Complexity

# Complexity and Resilience of Urban Network Structure

Lead Guest Editor: Jing-Hu Pan
Guest Editors: Liang Zhou, Xiuliang Yuan, and Miao Zhang

# Complexity and Resilience of Urban Network Structure

# Complexity and Resilience of Urban Network Structure

Lead Guest Editor: Jing-Hu Pan
Guest Editors: Liang Zhou, Xiuliang Yuan, and Miao Zhang

# Contents

Hindawi
Complexity
Volume 2022, Article ID 3667745, 12 pages
https://doi.org/10.1155/2022/3667745

*Research Article*

# Evolution of an Urban Network in the Yellow River Basin Based on Producer Services

**Weimin Gong, Zhibin Zhang, Long Chen, Xiaomin Ma, Xuewei Zhao, and Yan Guo**

*College of Geography and Environmental Science, Northwest Normal University, Lanzhou 730070, China*

Correspondence should be addressed to Zhibin Zhang; zbzhang@nwnu.edu.cn

Received 17 August 2021; Revised 4 October 2021; Accepted 19 March 2022; Published 11 April 2022

Academic Editor: Ali A. Minai



From the perspective of the producer services, a chain network model describes the evolution of an urban network in the Yellow River Basin and its mechanisms. The findings indicate the following phenomena: the spread of urban network connections in the basin and the proportional structure of cities at all levels show the characteristics of "flattening." The network relationship within the basin has a clear spatial and hierarchical orientation, presenting a "core-periphery" network structure and a "polarisation to trickle down" development model. The spatial organisation of the urban network societies in the river basin is remarkable. The closeness of the societies shows a trend at first decreasing and then increasing. The boundaries of the associations are consistent with the provincial administrative boundaries. The formation and evolution of the urban network in the basin have passed through the initial agglomeration stage, the hierarchical network formation stage, and the hierarchical network consolidation stage. From the perspective of influence mechanisms, proximity mechanisms, socio-economic development, and advances in communication technology have significantly impacted the formation of the urban network structure in the Basin, while the level of science and technology education and the degree of openness to the outside world have had less influence on the promotion of network connections between cities.

## 1. Introduction

Globalisation and informatisation are profoundly affecting and reconstructing the global urban system, as manifested by the intensification and enhancement of spatial differences and connections [1]. The global expansion of multinational companies has accelerated the continuous weakening of regional boundaries, with the regional spatial organisation model gradually shifting from central locations that emphasise centripetal spatial organisation and top-down single vertical connections to networks that emphasise node nature and horizontal connections [2, 3]. In this process, the traditional "local space" was gradually replaced by a new, nonregional "flow space," and the global economic organisation became an "open and multilateral network" [4, 5]. The formation of such a complex network produces a broader spatial influence that focuses cities of different natures, grades, and scales into the same networks and plays the functions and values of their respective nodes under the guidance and action of flow space [6].

As the best metaphor for the complex relationship between social actors in the new context, network research has become an increasingly important frontier of urban and economic geography [7]. In the early stages, the spatial image of the urban network was drawn mainly as a tangible physical network, with the traffic flow, such as highways, ports, railways, and aviation, as the focus of the research [8–10]. With the rise of the Internet, analysis of urban networks based on information flows such as mail, Internet traffic, and network broadband traffic began to emerge [11, 12]. The advent of the knowledge economy has led to the growth of innovative cyberspace with cities as the hubs, and the impact of intercity knowledge and technology flows on the urban network system gradually attracting the attention of the academic community [13–15]. In recent years, big data, represented by social networking site data [16], bus swiping data [17, 18], cell phone signalling data [19, 20], taxi GPS data [21], Weibo check-in data, and Baidu migration data [6, 22, 23], have provided more diversified support for modern urban network analysis. Under the influence of the central place theory, the hierarchical division of traditional urban systems has been mainly divided along with the perspectives of scale, grade, and urban function. The measurement method is mainly based on a large number of urban attributes and stresses urban internal characteristics and functions, ignoring the horizontal and mutual cooperation between cities [24]. Apart from this, some scholars have initiated discussions on the evolution of urban networks. With regard to the evolutionary mechanisms of urban networks, economic globalisation and regional economic integration have accelerated the industrial upgrading of central cities [25, 26]. As hubs of R&D institutions and headquarters, some cities have seen a huge rise in their role in regional development [27]. Overall, although these studies illustrate the external functions of producer services in the context of globalisation and informatisation in the formation of China's urban system, they have failed to further explore the evolution of urban networks from the perspective of productive service enterprises in the basin scale.

The cross-regional layout of corporate organisations and the networked development of organisational relationships have strongly promoted the formation of urban networks. The analysis of corporate microbehaviours with a worldwide layout has become an important entry point for studying world urban networks. In the new international division of labour, the disadvantages of the manufacturing industry for meeting the needs of structural transformation and product innovation have been prominent [28]. In this process, producer services have become an important driving force for the formation of metropolitan networks and an effective tool for explaining interurban connections. Research on the introduction of producer services into urban networks originates from Western scholars' attention to the world city system [29]. Relying on the production service company headquarters-branch network, Peter Taylor built the interlocking network model and took the lead in analysing the world city network with production service company data [30]. Western scholars' research on producer service city networks involves many aspects, such as world city networks and national and regional network patterns [31–33]. Their research perspectives are not limited to producer services themselves [11], but internal research on different industries is becoming increasingly abundant [34]. Influenced by Taylor's theory of world city networks, the research on city networks based on producer services is also gradually advancing in Chinese academic circles. Associated scholars mainly focus on the correlation between producer services, economic development, and the characteristics of urban networks. The research scope includes not only the national level [35] but also the urban system in the Yangtze River Delta [36], the urban network in Chengdu-Chongqing region, and the urban system in the Pearl River Delta [37, 38]. These empirical studies enrich the research results of China's urban regional scale and also strongly explain the outward function of producer services in the formation of the urban system in China against the background of globalisation and informatisation.

China upgraded ecological protection and high-quality development of the Yellow River Basin to a major national strategy in September 2019, setting a new goal for the development of the basin. The essential connotation of high-quality development is high efficiency, fair, green, and sustainable development is based on meeting the people's ever-increasing need for a better life [39]. As a spatial and regional unit comprising economic elements and social and economic activities, a watershed's high-quality development is not only the spatial implementation of high-quality economic development but also the external manifestation of coordinated social and economic development. Due to the existence of natural links, the interests of the development within the basin often form consistency, so promoting the integrated development of the basin often becomes the internal requirements of the basin development [40]. The Yellow River Basin is an important economic zone in China, and the strategic value of producer services for industrial transformation and coordinated economic development in the basin is increasingly prominent. From the perspective of productive service enterprises, this paper discusses the evolutionary characteristics, patterns, and mechanisms of urban networks in the Yellow River Basin. We hope to enrich the case study of urban networks of the Yellow River Basin, deepen the theoretical understanding of the spatial structure of it, and provide scientific basis for realising its high-quality development.

## 2. Data Sources and Methods

*2.1. Overview of the Study Area.* The Yellow River originates from the "water tower of China" in the source of the three rivers (Yellow River, Yangtze River, and Lancang River), as it passes through the Qinghai-Tibet Plateau, Inner Mongolia Plateau, Loess Plateau, North China Plain, and other topographic areas, and crosses the plateau mountains, temperate continental and monsoon climate zones, and arid, semiarid, and semihumid precipitation types. The comprehensive effects of natural elements such as topography and climate have given birth to the unique drainage structure of the Yellow River Basin: the upper and middle reaches of the river are ecologically fragile, and the environment has become a key constraint on the development of the basin. There are few large-scale tributaries, and the downstream section of the suspended river is linear. These fail to form a complete river network system, thus blocking the social and economic connections between the upstream and the downstream and between the basin and the outside. Based on the National Ministry of Water Resources of the Yellow River Water Resource Commission designated natural range, from its economic standpoint, the integrity of the basin to the eight provinces of Qinghai, Gansu, Ningxia, Inner Mongolia, Shaanxi, Shanxi, Henan, and Shandong and the statistical calibre of consistency, data availability, and comparability, 65 district cities were eventually chosen as research subjects.

*2.2. Data Source and Processing.* First, we used the regional keyword query function provided by the 11315 National Enterprise Credit Information System (https://www.11315.com/infnews/) and entered such keywords as "subsidiary" and "branch" to obtain the names of the internal branches in the research area and their corresponding company headquarter directories. Second, we logged on to the website of the State Administration for Industry and Commerce (https://www.gsxt.gov.cn/index.html) and use the company directory to perform a second query to supplement the company-attributed information. A total of 1,025 pieces of corporate data in 65 prefecture-level cities in the basin were obtained through sorting. Based on the obtained spatial locations of corporate headquarters and branches, only samples of corporate headquarters and branches in different places were kept. Finally, the sample enterprises were classified according to time. Those from 2000, 2010, and 2020 were screened out, and a total of 6953 valid samples were finally obtained.

*2.3. Research Methods*

*2.3.1. Chain Network Model.* In this paper, the connection strength of the network mainly refers to the connection strength between cities. The comprehensive network connection strength between city $a$ and city $b$ can be expressed as follows:

$$N_{ab} = L_{ab} + L_{ba}, \tag{1}$$

where $N_{ab}$ represents the strength of the connection between cities $a$ and $b$, $L_{ab}$ is the total number of enterprises with their headquarters in city $a$ and corresponding branches of each headquarters in city $b$, and $L_{ba}$ is the opposite vector value. The higher the total connectivity of a city, the better it can be integrated into the whole network of producer services.

The relative network connectivity for a single city can be expressed as follows:

$$P_a = \frac{N_a}{N_h}, \tag{2}$$

where $N_a$ represents the total connectivity between city a and other cities in the network, and $N_h$ is the city with the highest total network connectivity.

*2.3.2. Community Detection.* Community discovery can reflect the local characteristics of individuals in a network and their relationships and analyse and predict the interactions between elements of the entire network. Among them, the module-degree calculation method proposed by Newman is often used to measure the division quality of online communities [41]. The formula is as follows:

In the formula, $Q$ is the function of module degree, which quantitatively measures the result of community division. $M$ is the total number of network edges, while $A$ is the adjacency matrix corresponding to the network. When $A_{ij} = 1$, there is an edge between points $i$ and $j$. Otherwise, there is no edge. $K_i$ is the degree of node $i$; $C_i$ is the label where node $I$ belongs to a community.

*2.3.3. QAP Analysis.* This paper uses the Quadratic Assignment Procedure ($QAP$) as a nonparametric test method to quantitatively reveal the influence mechanism of the evolution of the urban agglomeration community structure. Compared with conventional statistical testing and regression methods, its advantage is that it does not need to meet the conditions of independence or noncollinearity between independent variables. The results obtained are more reliable, which is very suitable for the analysis of relational data. There are three specific analysis steps: first, select a number of influencing factors based on related theories, and research and build a model with the index network as the independent variable and the production service company's associated network as the dependent variable; second, use $QAP$ correlation analysis to test the correlation between the company's associated network and the index network; and finally, eliminate the factors with insignificant coefficients in the $QAP$ correlation analysis. The remaining factors and dependent variables are subjected to $QAP$ regression analysis to obtain the regression coefficients and test the indicators of each variable.

## 3. Analysis of the Urban Network Structure of the Yellow River Basin

*3.1. The Evolution of the Hierarchical Spatial Structure of Network Nodes.* The spatial level of nodes can better reflect the status and role of the city in the watershed than the traditional classification of the city hierarchy. In this paper, the score of the relative connectivity $P_a$ between different cities is used as the measurement for the classification of the city hierarchy. Comprehensively use the Q-type clustering method of SPSS19.0 hierarchical clustering to cluster 65 cities into 5 categories (Table 1).

From the hierarchical structure of urban network nodes, the network of producer service enterprises in the Yellow River Basin presents a clear diffusion trend, with the proportion structure of cities at all levels presenting a "flattening" feature. In this process, the number of cities in the first level of network connectivity rose from one in 2000 to four in 2020, as provincial capitals such as Taiyuan and Xi'an in the middle and upper reaches of the basin gradually became the central nodes of regional network connection. The reason may be that the middle and upper reaches of the Yellow River Basin are rich in natural resources and are home to a large number of traditional industries. In the process of China's industrial transformation and upgrading, modern emerging service industries with high knowledge-intensive characteristics have achieved rapid development in the middle and upper reaches of the river basin. Especially since 2010, with the continuous improvement of the development level of the producer services industry, those companies have gradually expanded in the central and western regional cities. Some provincial capitals in the middle and upper reaches of the basin, such as Taiyuan and Xi'an, have also joined the ranks of the first tier, as the network structure of cities in the basin has been further optimised. From the perspective of five different levels, the status of nodes in the intracity network of some cities is inconsistent with the characteristics of the traditional city hierarchy. For example, prefecture-level cities such as Tai'an, Jining, and Weifang have assumed the role of regional

Table 1: Hierarchical distribution of urban nodes based on producer services.

| Level | 2000 | 2010 | 2020 |
|---|---|---|---|
| The first | Jinan (1) | Jinan (1) | Jinane, Zhengzhou, Taiyuan, Xi'an (4) |
| The second | Taiyuan (1) | Weifang, Taian, Zhengzhou, Taiyuan, Xi'an, Lanzhou (6) | Jining, Weifang, Lanzhou (3) |
| The third | Zhengzhou, Huhehot (2) | Dongying, Zibo, Jining, Liaocheng, Jiaozuo, Linfen, Huhehot, Xining (8) | Zibo, Taian, Liaocheng, Huhehot, Yinchuan, Xining (6) |
| The fourth | Xi'an, Yinchuan, Lanzhou (3) | Heze, Shangluo, Shangqiu, Xinxiang, Jincheng, etc. (21) | Dongying, Binzhou, Heze, Shangqiu, Luoyang, Yuncheng, etc. (12) |
| The fifth | Weifang, Zibo, Luoyang, Baiyin, etc. (58) | Binzhoui, Kaifeng, Luoyang, Shuozhou, Yulin, etc. (29) | Puyang, Kaifeng, Xinxiang, Yulin, Guyuan, Dingxi, Haidong, etc. (40) |

Note: The number in brackets is the number of cities.

network contact centres, whereas provincial capital cities such as Xi'an, Yinchuan, Xining, and Lanzhou are no longer as prominent in the urban network. To a certain extent, this reflects the reality that the expansion of the urban network based on producer services does not completely follow the existing administrative hierarchy system based on geographical space.

From the perspective of the spatial distribution of urban network nodes in the Yellow River Basin, the city network's level distribution at the three time nodes all presented a "core-peripheral development model" (Figure 1). Before 2000, the distribution of high-level network nodes in the basin was relatively scattered and mostly located in the middle and lower reaches of the river basin, while the central cities in the upstream region were relatively low in the node level of the urban network. This further shows that the development of the producer service industry in the basin during this period was incomplete, leaving an extremely uneven intensity of urban connections between the upper and lower reaches of the basin. The distribution of urban network nodes in the river basin in 2010 changed considerably compared with 2000. The scale of producer services and the strength of intercity connections in some provincial capital cities, such as Xi'an, Lanzhou, and Xining, located in the middle and upper reaches of the river, have grown rapidly, with the Yellow River Basin initially showing a phenomenon of the "strong provincial capital." In this process, the middle and lower reaches of the river basin have become high-level and relatively concentrated network nodes, showing obvious spatial agglomeration effects. The reason may be that the producer services in this region not only have enjoyed a relatively high-level development but also have a sound basis for external output and play an important role as a network hub. In addition, Taiyuan, Linfen, and Xi'an in the middle reaches of the river basin, along with Lanzhou, Xining, and Huhehot in the upper reaches, also showed certain characteristics of agglomeration of producer services. The provincial capital cities in the basin showed a strong "leading" trend heading into 2020, with a development trend of "strong provincial capital." There was also a certain degree of differentiation within provincial capital cities. Provincial capital cities such as Zhengzhou, Xi'an, and Lanzhou grew rapidly compared with other prefecture levels in their provinces, producing a certain "siphon effect."

*3.2. Hierarchical Structure Evolution of Network System.* In order to further clarify the role of node cities in the river basin network system and the laws of their spatial connections, the natural breakpoint method is used in the ArcGIS software to divide the strength of intercity connections into three levels, creating a city network hierarchy diagram (Figure 2).

Through analysis, it is found that the urban network of the three time periods in the basin presents a continuous hierarchical structure, with its agglomeration and evolution showing obvious spatial orientation and "path dependence." In 2000, the overall network connection between cities in the basin was weak, with a network density of 0.017. A complete urban connection network had not yet been formed. The first-level city network was limited to the downstream areas, forming a four-pair association sequence of Jinan-Laiwu, Jinan-Liaocheng, Jinan-Weifang, and Jinan-Jining, with Jinan at the core. The second-level network was mainly distributed in the middle and upper reaches of the river basin, including Taiyuan-Jincheng, Hohhot-Ordos, and Hohhot-Bayannur. It can be seen that the network connection at this level is mainly within the province, and the spatial coordination between regions needs to be strengthened. The three-tier city network connections are mainly concentrated in some neighbouring cities with Lanzhou and Xining as the core, and Yinchuan and Xianyang, Lanzhou and Xining, and Xining and Yinchuan also forming spatial connections, with cross-provincial connections beginning to appear. In 2010, network connections between cities in the basin were significantly enhanced compared with the previous period. Its network density reached 0.091, and the radiation range of regional central cities increased, forming a multicentre network connection expansion structure with provincial capital cities as the core. Among them, the first-level network expanded to the middle and upper reaches, forming a total of 45 pairs of association sequences represented by Jinan-Weifang, Zhengzhou-Shangqiu, Taiyuan-Linfen, Xi'an-Lanzhou, and Lanzhou-Xining. The secondary city network has formed 10 pairs of sequences. In the connection direction, it shows that the provincial capital city is the radiating centre and spreads to the neighbouring cities. At the same time, the interprovincial city connections gradually strengthened. The three-tier city contact network during this period increased from 17 pairs in 2000 to 45 pairs, and the main framework of the network

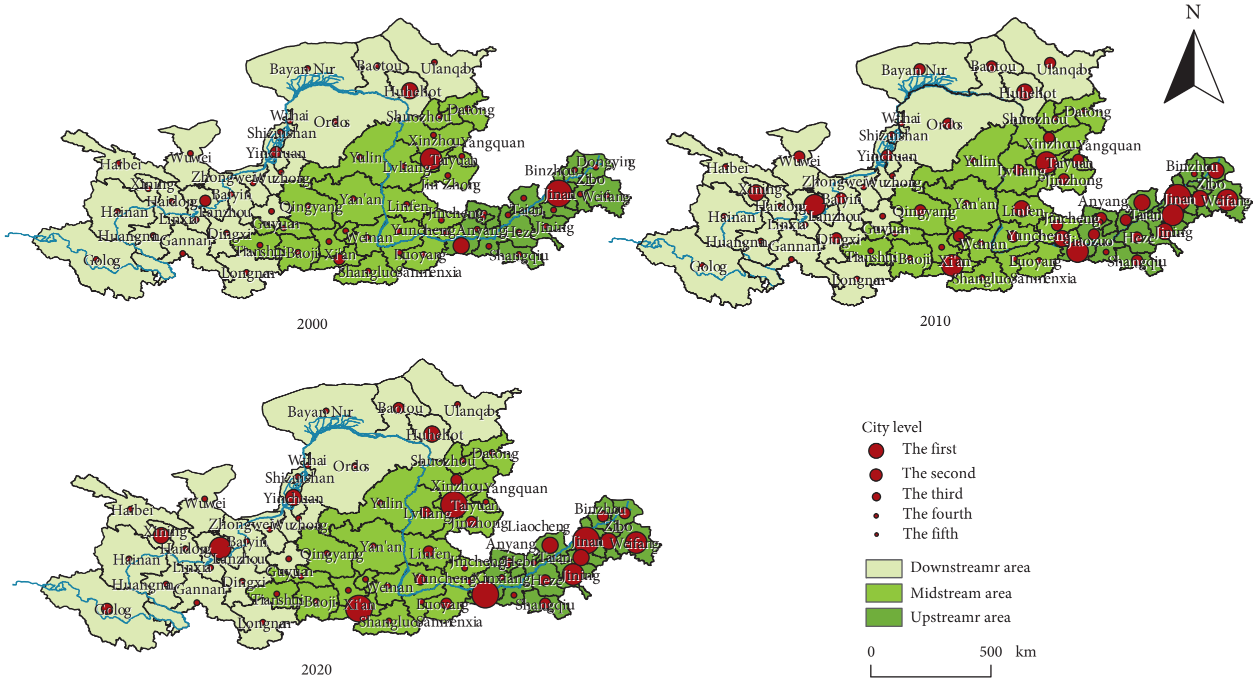



Figure 1: Hierarchical distribution map of river basin urban network nodes.

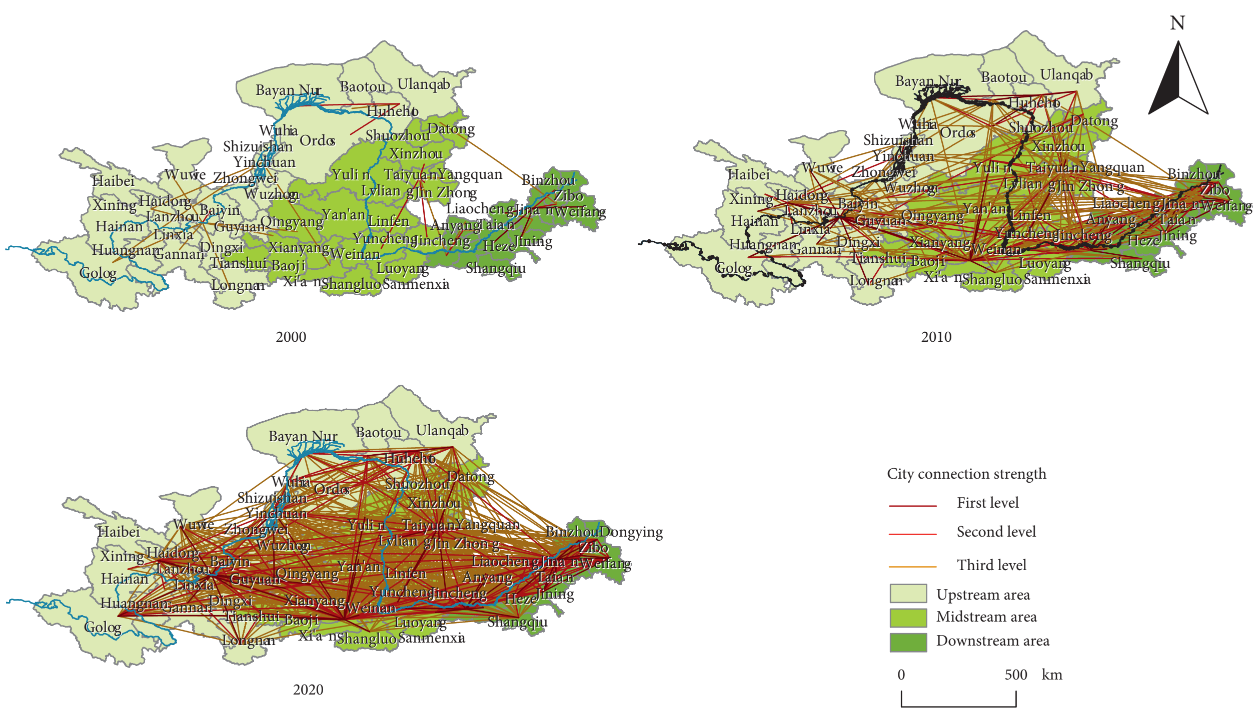



Figure 2: Hierarchical distribution of urban network in river basin.

structure between cities in the river basin was basically formed. In 2020, the overall network density in the river basin will reach 0.229, and the degree of urban network connections will be further improved, tightening the connections between cities. Through analysis, it is found that, compared with the previous period, the high-level urban network connection during this period has developed rapidly in the basin, while the urban network structure has basically remained stable, with the traditional core cities still playing leading roles.

*3.3. The Evolution of Network System Community Structure.* The *UCINET* software is used to calculate the urban modularity of the Yellow River Basin at three time nodes, and the network communities are divided accordingly. The number of splits of the association network was set as 2~15 times, and the modularity of different splits was calculated to determine the best classification of the community structure. The splitting times of the three time nodes are 7, 8, and 10, respectively, with the corresponding modularity the largest, which means that the division effect of urban communities is the most ideal. Through measurement, it was found that the maximum modularity of the three time nodes reached 0.781, 0.520, and 0.864, respectively, indicating that the urban association network of the Yellow River Basin is highly differentiated and presents a highly significant characteristic of the spatial organisation mode of community. On the whole, in the research period, the closeness of associations among communities in the basin showed a trend of first decreasing and then increasing, with the internal contact density of various associations increased.

By analysing the spatial organisational characteristics of various communities in the basin, their boundaries were highly consistent with the provincial administrative limits, denoting their status in the basin as a process of dynamic adjustment (Figure 3). Through analysis, it is found that seven community groups with Jinan, Zhengzhou, Taiyuan, Xi'an, Yinchuan, Hohhot, and Lanzhou as the core were formed in the basin in 2000, and the hierarchical gap between the community groups was relatively small during this period. Compared with the previous period in 2000, the producer service industry in 2010 has developed significantly in the upstream central cities. During this process, Xining City played an increasingly prominent role in regional connections and the formation of new community groups, leading to the growing status of the Xi'an and Lanzhou communities in the basin. In 2020, the Xi'an community in the upstream region continued the previous arrangement. Its status has been continuously promoted to become the highest-ranking community in the network. With the deepening of the degree of urban network connections in the river basin, the network connections within the community have differentiated, leaving some prefecture-level cities to gradually separate from the original community system and form new community groups.

## 4. Division and Model Induction of the Evolutionary Stages of Urban Network in the Yellow River Basin

Combining the regional development stage and related theories of industrial development, the formation and evolution process of the urban network in the Yellow River Basin during the study period is presented in three stages. In order to effectively reveal the evolution patterns and types of the urban network structure of the Yellow River Basin from the perspective of the producer service industry, the network structure patterns formed in three different stages are summarised thusly (Figure 4):

(1) The budding stage of hierarchical networks (pre-2000). The process of industrialisation is often accompanied by the continuous evolution of the internal structure of the producer service industry. Since various activities of the producer service industry were mainly provided by manufacturing enterprises in the early stages of industrialisation, the manufacturing industry in the Yellow River Basin during this period was largely dominated by labour and capital-intensive manufacturing, leaving the industrial structure low. Therefore, when the overall demand for production materials in the manufacturing industry in the basin is not strong, its producer service industry is also at a relatively low level of development. At this stage, the existing service activities in the Yellow River Basin were mainly concentrated in the regional central cities, and most of the service activities were rarely linked between provinces. Limited by the level of transportation infrastructure and communication technology at this stage, contacts between producer service companies occurred mainly by sending staff to contact local departments. The expansion of the companies within the region was not obvious. In this process, although the service function of regional central cities in the basin was strengthened, the driving effect on the surrounding cities was not strong and the service function differences were prominent. The whole region presented a system model of single centre development.

(2) Hierarchical network formation stage (2000-2010). With the refinement of the industrial division of labour and the development of economies of scale, especially the needs of manufacturing service model innovation and industrial transformation and upgrading, the producer service industry began to separate and grow independent from the manufacturing sector. From the perspective of enterprise space expansion, during this period, producer service enterprises in the Yellow River Basin increased their demands for information acquisition in the process of scale growth, the strength of network connections between cities increased, and a horizontal flow trend appeared. The producer service industries of cities with lower levels in the basin gradually began to develop, and the trend of multicentre development initially appeared. In the process, the spatial diffusion of producer service enterprises in the basin was enhanced, and the hierarchical characteristics of urban networks grew more and more obvious. On the whole, the connections between node cities in the basin during this period showed a trend of interweaving vertical and horizontal connections, while its spatial structure showed flattened characteristics.

(3) Consolidation stage of the hierarchical network (2010-present). After entering the middle stage of industrialisation, as China's economy gradually shifted from a manufacturing-oriented to a service-oriented economy, the producer service industry gradually became an important force in instigating urban economic growth and a major player in urban

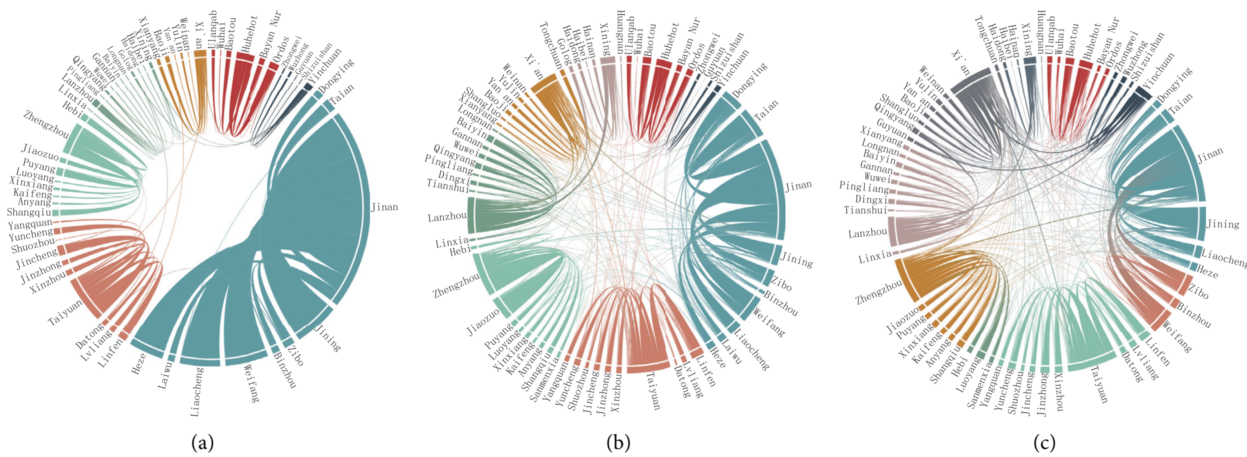



FIGURE 3: Evolution of urban network structure in the Yellow River Basin. (a) 2000. (b) 2010. (c) 2020.

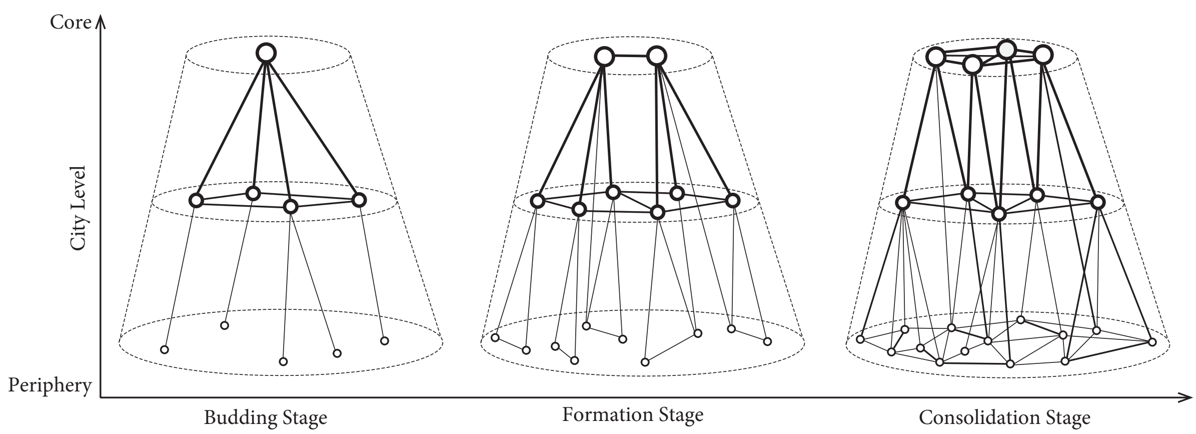



FIGURE 4: Model induction of the urban network of the Yellow River Basin in each evolutionary stage.

economic functions. Under the post-ford-style, flexible production organisation system, the international division of labour began to transform from product division to element division, and the organisation and management of enterprises in the basin gradually shifted from a functional layout structure to a multisegment structure. As the organisation and management of each branch were relatively complete and self-contained, this organisational structure also became more diffuse. At the same time, diseconomies of scale led to the transfer of some elements from high-level cities to low-level cities. Some nonadministrative central cities within the river basin enhanced the agglomeration and radiation capabilities of the elements, the horizontal connections became more significant and the urban network further consolidated. During this period, the hierarchical characteristics of the urban network in the basin became more and more obvious, the horizontal connections grew closer as the connection directions became more diverse, and the gap in service function connections narrowed.

## 5. The Influence Mechanism of the Evolving Urban Network Structure in the Yellow River Basin

*5.1. Variable Selection.* Based on the macroinfluencing factors and mechanisms in the existing literature, and taking into account the microbehaviour characteristics of producer service companies, the following variables are selected:

(1) Nearby mechanism variables. ① The proximity of policy systems. Considering that there are multiple administrative system environments in the river basin, different environment systems may have an impact on the development of enterprises. In this paper, 1 and 0 are used to indicate that the two cities belong to the same province or different provinces, reflecting the system proximity effect produced by interprovincial administrative divisions; ② Geographical proximity. Calculate the physical distance between cities based on the straight-line distance between them. Considering that the actual geographic distance between cities

may be too large and may cause errors in the estimation results, the actual distance is processed based on the research results of relevant scholars. ③ Cultural proximity. Local-scale economic connections are usually achieved through face-to-face informal communications. The cultural similarity is conducive to communication and coordination, which is of positive significance for the development of the producer service industry. Dialects are an important cultural representation in China. Among them, there are the many regional types of dialects in the Yellow River Basin, involving seven dialect areas. In view of this, use 1 and 0, respectively, were used to indicate that the two cities belonged to the same dialect area and different dialect areas, respectively, to reflect cultural proximity [42]. ④ Industrial proximity. In order to reflect the impact of the degree of similarity of similar industrial structures between cities in the river basin on urban economic relations, according to the industry classification in the city statistical yearbook, the industrial structure similarity coefficient is used to reflect the industrial proximity.

(2) Other variables. ① Social and economic development status. The population size is represented by the total population of the municipal area. The economic scale is represented by the GDP of the municipal districts. ② The level of science and technology education. The Intellectual capital is represented by the product of the number of scientific research and technical services of 10,000 people in the municipal area and the city's per capita local financial science and technology expenditures. The producer services is knowledge, technology, and capital-intensive industry. The easy access to required human resources is an important key factor in the location selection of producer service enterprises. Therefore, this article uses the average years of education to reflect the distribution of human resources. ③ Progress in technological exchanges. Select the accessibility index to characterise the progress of transportation technology. The total volume of post and telecommunication businesses is selected to reflect the progress of communication technology. ④ Openness to the outside world. The degree of openness to the outside world is an important indicator to measure the development level and scale of a country or region in international economic activities and its dependence on the global economy. This article mainly uses two indicators to characterise foreign trade dependence and foreign investment dependence [43]. Among them, the degree of dependence on foreign trade focuses on measuring the extent of the export-oriented national economy of prefecture-level units; the degree of dependence on foreign investment mainly reflects the degree of capital openness of prefecture-level units and the size of the potential for foreign economic development.

*5.2. Regression Analysis.* In QAP, correlation analysis examines whether the variables have collinearity, and regression analysis examines whether the influencing factors are significant. It was found through measurement (Table 2) that the correlation between most factors and the associated network of producer service companies over the course of the three years passed the 1% significance level test. With $R^2$ between 0.5 and 0.6, the model has achieved a relatively good fitting effect.

The results confirm that the proximity mechanism had a significant impact on the formation of the urban network structure in the Yellow River Basin and that the degree of influence on the urban network structure varies at different stages. Among them, the proximity of the administrative system based on the division of provincial administrative units is significantly positive in the three periods, indicating that administrative proximity strengthened the connection between the units. Although the regression coefficient of geographical proximity has passed the significance test in different periods and had a positive effect, with the development of transportation and information technology, its influence on the network relationship of producer service enterprises has gradually weakened. Cultural proximity is significantly positive in different periods, indicating that the smaller the regional dialect and cultural differences, the stronger the cultural identity between microindividuals, and the less likely it is to produce cultural conflicts, which are beneficial to the formation of regional internal connections. Although industrial proximity contributes to the improvement of regional specialisation and the exertion of the agglomeration effect, the positive influence on the formation of regional network connections is weak. In addition, because the influence of the dimension is eliminated, the size of the regression coefficient reflects the strength of the influence. Through measurement, it is found that institutional proximity has the largest coefficient in both the proximity mechanism variable group and all factors. The above results show that the formation of the urban network organisation relationship in the provincial administrative boundary region bears an important influence as the biggest obstacle to the regional network connection.

In the socio-economic development variable group, the two variables have passed the significance test in different periods, indicating that the radiation effect brought by the city's socio-economic development has promoted the network connection between cities. In the variable group of science and technology education levels, the regression results of knowledge capital and human resource distribution are not significant, indicating that a significant role for human and knowledge capital in the development of the producer service industry has not yet appeared. The reason may be that long-term, low-end manufacturing in the Yellow River Basin has dampened the development level of

Table 2: QAP analysis results of the evolution of urban network structure in the Yellow River Basin.

| Variable group | Variable | Network formation stage | | Hierarchical network formation stage | | Network form consolidation period | |
|---|---|---|---|---|---|---|---|
| | | Related analysis | Regression analysis | Related analysis | Regression analysis | Related analysis | Regression analysis |
| Proximity mechanism | System proximity | $0.186^{***}$ | $0.027^{*}$ | $0.083^{***}$ | $0.032^{**}$ | $0.137^{***}$ | $0.053^{***}$ |
| | Geographical proximity | $-0.195^{***}$ | $-0.253^{**}$ | $-0.059^{***}$ | $-0.176^{*}$ | $-0.106^{***}$ | $-0.065^{***}$ |
| | Cultural proximity | $0.057^{***}$ | $0.068^{**}$ | $0.074^{**}$ | $0.053^{**}$ | $0036^{***}$ | $0.050^{***}$ |
| | Industry proximity | $0.859^{***}$ | $0.032^{**}$ | $0.153^{**}$ | $0.095^{***}$ | $0.217^{***}$ | $0.062^{**}$ |
| Socio-economic development | Economic scale | $0.118^{***}$ | $0.109^{***}$ | $0.017^{**}$ | $0.083^{***}$ | $0.023^{***}$ | $0.136^{***}$ |
| | Population size | $0.291^{***}$ | $0.057^{***}$ | $0.156^{***}$ | $0.085^{***}$ | $0.142^{**}$ | $0.181^{***}$ |
| Science and technology | Knowledge capital | $0.075^{**}$ | $0.109^{**}$ | $0.108^{**}$ | $0.168^{*}$ | $0.095^{***}$ | $0.049^{**}$ |
| | Human resource distribution | $0.094^{**}$ | $0.072^{***}$ | $0.155^{**}$ | $0.129^{**}$ | $0.106^{***}$ | 0.083 |
| Communicating technological progress | Progress in transportation technology | $0.075^{***}$ | $0.127^{***}$ | $0.130^{***}$ | $0.019^{***}$ | $0.083^{***}$ | $0.228^{***}$ |
| | Communication technology progress | $0.263^{**}$ | $0.159^{***}$ | $0.020^{*}$ | $0.144^{**}$ | $0.152^{***}$ | $0.094^{*}$ |
| Degree of openness | Foreign trade dependence | $0.165^{**}$ | 0.052 | $0.065^{*}$ | $0.148^{*}$ | $0.179^{**}$ | $0.157^{**}$ |
| | Foreign capital dependence | $0.152^{**}$ | 0.049 | $0.084^{**}$ | $0.024^{**}$ | $0.218^{***}$ | $0.039^{*}$ |
| | Sample size | 65 | 65 | 65 | 65 | 64 | 64 |
| | $R^2$ | — | 0.60 | — | 0.58 | — | 0.52 |

Note: $^{***}$, $^{**}$, and $^{*}$ are significant at the statistical levels of 1%, 5%, and 10%, respectively.

the supporting producer service industry and that the two have fallen into a potential internal low-end lock-in cycle. This leads to the insignificant role of regional knowledge capital and human resources in the development of the producer service industry and the promotion of the value chain. In the communication technology progress variable group, both variables passed the significance test, indicating that the space-time compression effect produced by transportation technology progress promotes the connection between cities in the basin. The production service industry in areas with a higher level of information technology will have more information resources and lower information transmission costs, thereby attracting more producer service companies to deploy here. Through measurement, it is found that the variable of openness to the outside world has not passed the significance test. The reason may be that the overall globalisation process of the Yellow River Basin is relatively slow, the export-oriented economy is not prominent, and it has not had a significant effect on the network connections between cities in the basin.

## 6. Conclusions

Economic globalisation and the rapid development of information technology have strongly promoted the formation of urban networks. As the main driving force of structural transformation and an important source of product innovation, the key role of producer service enterprises in this process has received greater attention. Based on headquarters-branch data of producer service enterprises in different years, this paper uses the network analysis method to analyse the urban spatial structure of the Yellow River Basin. The main conclusions are as follows:

(1) The urban network in the Yellow River Basin has hierarchical stratification and path dependence under the influence of urban rights, presenting a "core-periphery" network structure and a "polarisation to trickle" development model. With the transformation of the organisation and operation mode of producer service enterprises to networks, the network connection among cities with producer service enterprises as the link has an obvious diffusion trend, and the proportional structure of cities at all levels demonstrates the characteristics of "flattening." In addition, the urban association network of the Yellow River Basin has formed a substantial spatial network of community groups. The status of each community in the basin is in a dynamic adjustment process, and the closeness of associations between communities has shown a trend of first decreasing and then increasing.

(2) Based on the urban administrative hierarchical system of geographic space, careful consideration is given to the spatial expansion process of producer service enterprises. Through inductive analysis, the formation and evolution of the city network based on the producer service industry are summarised as three stages, namely, the initial agglomeration stage, the network formation stage, and the network consolidation stage. At present, the overall urban

network of the Yellow River Basin, based on producer service enterprises, is at a point of consolidation of the hierarchical network. Its main characteristics are that the urban network has become more obvious, the horizontal connections are increasingly close, the connection directions are more diverse, and the service function connection gap is narrowing.

(3) From the perspective of influence mechanism, geographical, cultural, and administrative proximity have all cast a significant impact on the formation and evolution of urban networks in the basin, with administrative proximity having the greatest impact. Advances in communication technology and socioeconomic development have also carried a significant impact on the internal and external relations of the community. Under the combined effect of many factors, the urban network structure of the Yellow River Basin is becoming more and more refined from the perspective of producer service enterprises, with the horizontal connections between cities constantly increasing.

## 7. Discussion

The formation and evolution of the urban network structure denote a coupling process of two basic forms of movement, namely, spatial aggregation and diffusion. With the further division of labour, enterprises demonstrate multisectoral and multiregional tendencies, whereas changes in enterprise networks exert a direct influence on the spatial evolution of urban networks. Moreover, the most direct embodiment of the role of enterprise networks in urban networks is that the choice of an enterprise to enter or exit a city can result in a change in its position in the urban network. As a crucial driver of economic growth and industrial chain upgrading in cities, developing producer services displays a profound significance in strengthening intercity exchanges and cooperation and in promoting the efficient and orderly growth of regional urban systems. The relationship between enterprises and cities involves among different locations of enterprises and their branches, choice of enterprises regarding new investment sites and efforts of cities to attract enterprises by making themselves more competitive. Governments play a critical role in the clustering of enterprises and the development of industries. Furthermore, they drive the expected returns of enterprises. Therefore, they intervene in the decisions of enterprises by formulating policies and plans that facilitate industrial development, optimising support conditions, and improving supporting infrastructure. Additionally, the decisions of enterprises are largely influenced by the conditions of their cities (i.e., scale characteristics, such as population size, economic characteristics, infrastructure construction, informatisation level, human capital, and innovation factors) and their position in the urban system.

In the Yellow River Basin, the urban network system connected by producer services is expected to become increasingly *flat* with the diversification of production activities and the networking of producer services. This notion suggests the increased connectivity of urban networks. Interprovincial interaction and networking are currently mainstream in the Yellow River Basin, with central node cities exerting the siphon effect on the urban network. Hence, coordinated development between regions within the Basin to improve its urban network should be emphasised. At the national level, the Yellow River Basin can improve the quality of regional cooperation, except for internal cooperation, by increasing links with its neighbours and developed regions at home and abroad based on the local conditions. Additionally, it should aim to address imbalances and inadequacies in regional development, formulate relevant strategies that are more interconnected and holistic, and enhance its coordination and integrity. At the local level, efforts should be undertaken to improve synergies between cities in the Basin with core cities radiating and promoting the surrounding areas. Presently, core cities exhibit a major agglomeration effect but a poor positive spillover effect. This tendency inspires the breakdown of *barriers* between cities, strengthens synergy, and leverages the radiating and driving role of central cities. Furthermore, cities along the Yellow River Basin, especially those in the middle and lower reaches, should utilise their ecological resources to realise industrial complementarity between provinces and regions within the Basin. In this manner, they are "optimising the stock and improving the increment." This initiative is expected to promote the transformation of industrial structure from low-end to high-end. Moreover, it is expected to build an industrial cooperation alliance along the Yellow River Basin with the Yellow River as the link, cities as the carrier, and mutually beneficial cooperation as the premise. As a result, the restraints of the administrative region economy can be overcome, and a joint force for overall regional performance can be created.

## Data Availability

The basic data come from the 11315 National Enterprise Credit Information System (https://www.11315.com/infnews/) and the China State Administration for Industry and Commerce website (https://www.gsxt.gov.cn/index.html). The final data used to support the findings of this study contain confidential trade information, while the research findings are commercialised. Requests for sample data, 6 months after publication of this article, will be considered by the corresponding author.

## Conflicts of Interest

The authors declare that they have no conflicts of interest.

## Acknowledgments

The authors thank the Key Program of National Natural Science Foundation of China (No. 41961029) and "Innovation Star" Program for Outstanding Postgraduates in Gansu Province (2021CXZX-187).

## References

[1] E. Sheppard, "The spaces and times of globalization: place, scale, networks, and positionality," *Economic Geography*, vol. 78, no. 3, pp. 307–330, 2002.

[2] P. J. Taylor, M. Hoyler, and R. Verbruggen, "External urban relational process: introducing central flow theory to complement central place theory," *Urban Studies*, vol. 47, no. 13, pp. 2803–2818, 2010.

[3] D. F. Batten, "Network cities: creative urban agglomerations for the 21st century," *Urban Studies*, vol. 32, no. 02, pp. 313–3271, 1995.

[4] J. V. Beaverstock, R. G. Smith, and P. J. Taylor, "World city network:a new metageography?" *Environment & Planning A*, vol. 31, pp. 187–192, 2000.

[5] V. O. Frank, B. Martijn, and R. Otto, "On the economic foundation of the urban network paradigm:spatial integration,functional integration and economic complementarities within the Dutch Randstad," *Urban Studies*, vol. 47, no. 4, pp. 725–748, 2010.

[6] M. Xu, Z. Huang, and F. Cao, "The network structure of urban tourist destination and its evolution mode based on big data analysis: taking the data of Sina weibo signin as an example," *Geographical Research*, vol. 38, no. 4, pp. 937–949, 2019.

[7] P. Dicken, P. Kelly, K. Olds, and H. W.-C. Yeung, "Chains and networks, territories and scales: towards a r-elational framework for analysing the global economy," *Global Networks*, vol. 1, no. 2, pp. 89–112, 2001.

[8] K.-H. Shin and M. Timberlake, "World cities in Asia: cliques, centrality and connectedness," *Urban Studies*, vol. 37, no. 12, pp. 2257–2285, 2000.

[9] H. Matsumoto, "International urban systems and air passenger and cargo flows: some calculations," *Journal of Air Transport Management*, vol. 10, pp. 239–247, 2004.

[10] W. Wu, Y. Cao, and W. Cao, "On the patterns of integrated transportation accessibility in the Yangtze River Delta under opening conditions," *Geographical Research*, vol. 26, no. 2, pp. 391–402, 2007.

[11] E. J. Malecki, "The economic geography of the Internet's infrastructure," *Economic Geography*, vol. 78, no. 4, pp. 399–424, 2002.

[12] M. Wang and Y. Ning, "The network advantage of cities: an analysis of spatial structure and node accessibility of Internet backbones in China," *Geographical Research*, vol. 25, no. 2, pp. 193–203, 2006.

[13] C. Liu, M. Guan, and D. Duan, "Spatial pattern and influential mechanism of interurban technology transfer network in China," *Acta Geographica Sinica*, vol. 73, no. 8, pp. 1462–1477, 2018.

[14] T. Wang, S. Hennemann, and I. Liefner, "Spatial structure evolution of knowledge network and its impact on the NIS: Case study of biotechnology in China," *Geographical Research*, vol. 30, no. 10, pp. 1861–1872, 2011.

[15] H. Ma, "Research progress and prospects of city networks based on knowledge flow in China," *Economic Geography*, vol. 36, no. 11, pp. 207–213, 2016.

[16] Z. Feng, B. Wang, and Y. Chen, "China's city network characteristics based on social network space: an empirical analysis of Sina micro-blog," *Acta Geographica Sinica*, vol. 67, no. 8, pp. 1031–1043, 2012.

[17] Y. Long, Y. Zhang, and C. Cui, "Identifying commuting pattern of Beijing using bus smart card data," *Acta Geographica Sinica*, vol. 67, no. 10, pp. 1339–1352, 2012.

[18] Y. Liu, L. Chen, and Z. An, "Research on Job-Housing and Commuting in Wuhan Based on Bus Smart Card Data," *Economic Geography*, vol. 39, no. 2, pp. 93–102, 2019.

[19] L. Ding, X. Niu, and X. Song, "Measuring the employment center system in Shanghai central city: a study using mobile phone signaling data," *Acta Geographica Sinica*, vol. 71, no. 3, pp. 484–499, 2016.

[20] W. Zhong, D. Wang, and D. Xie, "Dynamic characteristics of Shanghai's population distribution using cell phone signaling data," *Geographical Research*, vol. 36, no. 5, pp. 972–984, 2017.

[21] Y. Han and W. A Wang, "On the spatial and temporal distribution of resident trip based on taxi GPS data," *Geomatics & Spatial Information Technology*, vol. 41, no. 2, pp. 87–89, 2018.

[22] F. Zhangxian, Y. Zhang, and W. Ye, "Spatial-Temporal pattern and dynamic mechanism of population flow of Changchun City during Chunyun period based on Baidu migration data," *Economic Geography*, vol. 39, no. 5, pp. 101–109, 2019.

[23] W. Liu and E. Shi, "Spatial pattern of population daily flow among cities based on ICT: a case study of "Baidu Migration"," *Acta Geographica Sinica*, vol. 71, no. 10, pp. 1667–1679, 2016.

[24] C. Wang, Y. Cao, and W. Song, "Research progress in urban network construction from the perspective of producer services industry," *Progress in Geography*, vol. 32, no. 7, pp. 1051–1059, 2013.

[25] F. Yu, C. Gu, and H. Wu, "Structure and transformation of urban function in China," *Urban Planning Forum*, vol. 65, no. 5, pp. 13–21, 2006.

[26] W. Houkai, "The new industrial division and conflict management in metropolitan region——based on the perspective of industrial- chain division," *China Industrial Economy*, vol. 27, no. 2, pp. 30–36, 2007.

[27] S. Li and C. Gu, "Empirical study on the intra-product specialization of Yangtze River Delta:A case study of auto manufacturing industry," *Geographical Research*, vol. 29, no. 12, pp. 2132–2142, 2010.

[28] C. Wang, Y. Cao, and G. Chen, "Study on urban network of Yangtze River Delta region based on producer services," *Geographical Research*, vol. 33, no. 2, pp. 323–335, 2014.

[29] S. Sassen, *The global city: New York, London, Tokyo*, Vol. 75, Princeton University Press, Princeton, NJ, USA, 2001.

[30] P. J. Taylor, G. Catalano, and D. R. F. Walker, "Measurement of the world city network," *Urban Studies*, vol. 39, no. 13, pp. 2367–2376, 2002.

[31] B. Derudder and P. J. Taylor, "Central flow theory: comparative connectivities in the world-city network," *Regional Studies*, vol. 52, pp. 1–14, 2018.

[32] P. J. Taylor, *World Cities Network: A Global Urban analysis*, Vol. 135, Routledge, London, UK, 2004.

[33] M. Zhao, B. Derudder, and J. Huang, "Examining the transition processes in the Pearl River Delta polycentric mega-city region through the lens of corporate networks," *Cities*, vol. 60, pp. 147–155, 2017.

[34] E. C. Rossi and P. J. Taylor, "Banking networks across Brazilian cities: Interlocking cities within and beyond Brazil," *Cities*, vol. 22, no. 5, pp. 381–393, 2005.

[35] M. Zhao and L. Zheng, "Research on China's city network based on production service industry," *City Planning Review*, vol. 36, no. 9, pp. 23–28, 2012.

[36] C. Wang and Y. Cao, "The evolution mode and mechanism of urban network from the perspective of producer services

industry: a case of the Yangtze River Delta," *Scientia Geographica Sinica*, vol. 39, no. 2, pp. 285–293, 2019.

[37] Y. Tan, Y. Yang, and B. Leng, "Urban network system in Chengdu- Chongqing region in the perspective of advanced producer service," *Progress in Geography*, vol. 30, no. 6, pp. 724–732, 2011.

[38] Y. Zhong and M. Huang, "Study on diversity of factors affected service indus-try develo-pment under different urban hierarchy: panel date analysis of the Pearl Riv-er Delta," *Industrial Economic Review*, vol. 4, no. 2, pp. 14–22, 2013.

[39] J. Zhao, D. Shi, and Z. Deng, "A framework of China's high-quality economic development," *Research on Economics and Management*, vol. 40, no. 11, pp. 15–31, 2019.

[40] X. Li, Y. Wen, and Y. Li, "High-quality development of the Yellow River Basin from a perspective of economic geography:Man-land and spatial coordination," *Economic Geography*, vol. 40, no. 4, pp. 1–10, 2020.

[41] L. Yan, Y. Sun, and S. Yao, "Urban network linkages of Chinese urban agglomerations based on an analysis of the Fortune China Top 500 Enterprises," *Geographical Research*, vol. 39, no. 7, pp. 1548–1564, 2020.

[42] P. Gao, D. He, and Y. Ning, "Community structure and proximity Mechanism of city clusters in middle reach of the Yangtze River: based on producer s-ervice firms'network," *Scientia Geographica Sinica*, vol. 39, no. 4, pp. 578–586, 2019.

[43] Y. Zhang and J. Pan, "Spatio-temporal pattern of economic opening rate at prefect-ure level cities or above in China," *Arid Land Geography*, vol. 41, no. 5, pp. 1123–1131, 2018.

Hindawi
Complexity
Volume 2022, Article ID 9854057, 1 page
https://doi.org/10.1155/2022/9854057

*Corrigendum*

# Corrigendum to "The Spatial Agglomeration and Industrial Network of Strategic Emerging Industries and Their Impact on Urban Growth in Mainland China"

**Jiaming Li, Peiyuan Zhang, Jialin He, and Dongqi Sun**

*Institute of Geographic Sciences and Natural Resources Research (CAS), 11A, Datun Road, Chaoyang District, Beijing 100101, China*

Correspondence should be addressed to Dongqi Sun; sundq@igsnrr.ac.cn

Received 22 January 2022; Accepted 22 January 2022; Published 28 February 2022



In the article titled "The Spatial Agglomeration and Industrial Network of Strategic Emerging Industries and Their Impact on Urban Growth in Mainland China" [1], the Acknowledgments section should read as follows.

## Conflicts of Interest

The authors declare that they have no conflicts of interest.

## Acknowledgments

The research has been supported by the Second Tibetan Plateau Scientific Expedition and Research, under Grant no. 2019QZKK0406.

## References

[1] J. Li, P. Zhang, J. He, and D. Sun, "The spatial agglomeration and industrial network of strategic emerging industries and their impact on urban growth in mainland China," *Complexity*, vol. 2021, Article ID 7199044, 12 pages, 2021.

Hindawi
Complexity
Volume 2022, Article ID 5975866, 13 pages
https://doi.org/10.1155/2022/5975866

*Research Article*

# Passenger Behavior Simulation in Congested Urban Rail Transit System: A Capacity-Limited Optimal Strategy Model for Passenger Assignment

**Kai Lu and Nan Cao**

*Traffic Control Technology Co., Ltd., Beijing 100070, China*

Correspondence should be addressed to Kai Lu; lukai_bjtu@163.com

Received 27 July 2021; Accepted 23 November 2021; Published 6 January 2022

Academic Editor: Jing-Hu Pan



Optimal strategy, one of the main transit assignment models, can better demonstrate the flexibility for passengers using routes in a transit network. According to the basic optimal strategy model, passengers can board trains based on their frequency without any capacity limitation. In the metropolitan cities such as Beijing, Shanghai, and Hong Kong, morning commuters face huge transit problems. Especially for the metro system, there is heavy rush in metro stations. Owing to the limited train capacity, some passengers cannot board the first coming train and need to wait for the next one. To better demonstrate the behavior of passengers pertaining to the limited train capacity, we consider capacity constraints for the basic optimal strategy model to represent the real situation. We have proposed a simulation-based algorithm to solve the model and apply it to the Beijing Subway to demonstrate the feasibility of the model. The application of the proposed approach has been demonstrated using the computational results for transit networks originating from practice.

## 1. Introduction

Providing passenger-oriented service is the most important and core target for every metro system. To support passenger-friendly service planning and modification, it is necessary to have a passenger assignment model, which can demonstrate the passenger behavior and the distribution of passengers. Based on the assignment results, it would be clear to determine if the service is passenger friendly and how to provide a better service. Because of this, the assignment models are wildly used in trip planning process and the passenger departure time decision [1–3]. At the same time, when the disturbance or disruption happened, the operation scheme and the timetable could be adjusted according to the assignment results, such as retiming and providing short loop trains. In this way, the disorder could be solved with a more passenger-friendly solution [4–6].

Service frequency, which has high relation with system service quality and efficiency, is the key component for metro systems. The frequency is determined with the passenger demand, the passenger distribution, and operation costs. In some exurban places or low-demand areas, the transit service is quite poor and the train frequency is at the low level. Passengers must be informed with the accurate timetable, and they can schedule their departure time to catch the specific train service at the station. To demonstrate this precise passenger behavior and the interaction between passengers and trains, a schedule-based transit assignment approach is needed.

The schedule-based assignment is modeled based on a time-space network according to a timetable [7–10]. Similar to the frequency-based assignment, congestion is also an issue in the schedule-based network. Poon et al. [11] used a time-increment simulation to obtain the arrival and departure time of passengers to predict dynamic queuing delays and update the shortest path for the next simulation run. Nuzzolo et al. [12] defined a dynamic loading process on each transit run according to the user's choice and the residual capacity of vehicles arriving at the stops. Considering the differential comfort level experienced by sitting and standing passengers, Hamdouch and Lawphongpanich [13] and Hamdouch et al. [14] proposed equilibrium conditions as variational inequalities that involve a vector-valued function of the expected strategy cost. The method of successive averages (MSA) was applied to solve the model. In addition to the capacity limitation, the supply is not stable. Hamdouch et al. [15] considered uncertainties in supply and proposed an analytical model to record the stochastic nature of transit schedules. The uncertainties are considered as the covariances of travel time between the links in a space-time graph. An MSA-based algorithm was used to solve the aforementioned problem. However, demand and supply uncertainties significantly affected the passenger behavior. Zhang et al. [16] modeled this phenomenon using the in-vehicle congestion parameter. A heuristic MSA-based algorithm was applied, and the results showed that the risk-taking attitude significantly impacts the travel mode and departure-time selection of passengers.

The schedule-based assignment could explain the passenger behavior in detail, but it is relatively complex and time consuming owing to the involvement of time-space path search and dynamic simulation. However, in metropolitans such as Beijing and Tokyo, when passengers planning their trip, they do not have to target to the specific train service because the train frequency is relatively high. Especially in the peak hour, the minimal headway is reduced to 90 s. Meanwhile, there is only one kind of service on the transit line, which means all trains on the same line stop at the same stations. In this situation, passengers do not have to consider timetables or train departure times before trying to catch the trains. Moreover, the train arrival times are announced beforehand for passengers to plan their travel. Passengers can switch their previous paths at some stations to travel more quickly or comfortably. This kind of behavior correlates with the frequency-based assignment assumption.

The frequency-based assignment is commonly used for transit and urban network planning, where the frequency is fixed and high [17]. It is assumed that the passengers may board the first attractive train when they arrived at the stop, which is named the optimal strategy [18] or shortest hyperpath [19]. However, owing to busy transit networks, systems get overloaded and highly congested. As a result, passengers who have missed the first train may have to wait for the second or third attractive train. To demonstrate this phenomenon, De Cea and Fernández [20] proposed the *effective frequency*, which is linearly related to equivalent average waiting time index, to correctly model the impact of congestion. Since then, several researchers have tried their best to focus on the frequency assignment with capacity constraints [21]. Cepeda et al. [22] considered that travel time is related to the passenger flow. They formulated an equivalent optimization problem with a new characterization of equilibria, which vanished in the computable gap function. Schmöcker et al. [23] modeled the capacity constraint with the "fail-to-board" probability and searched the hyperpath. The Markov network loading process was proposed for passengers who failed to board the first train and reconsidered their selection of routes. Disabled people and pregnant women cared more about the travel experience. Schmöcker et al. [24] introduced a "fail-to-sit" probability, according to which passengers would follow the priority rule that standing on-board passengers would occupy any available seats of alighting passengers before newly boarding passengers do. Passenger behavior is sensitive to perturbations when it comes to running times or service frequencies [25], and some systems provide online information on predicted arrival times. The frequency-based transit assignment model that considers online information can significantly reduce the overall travel time [26]. They tried to proposed capacity limitation with the mathematical formula and transform the model into a new optimization model.

Compared with the schedule-based assignment, the frequency-based assignment could solve the problem more quickly and also demonstrate passenger behavior. It would be interesting if we can combine the merits of the realism of schedule-based methods and the simplicity of frequency-based methods in a single framework to find a balance between accuracy and computational efficiency [27, 28]. At the same time, it is necessary to consider the capacity limitation during the model and make the model into a more practical way.

Therefore, this study considered capacity constraints for the classic optimal strategy model. Moreover, a CVX [29] (for Disciplined Convex Programming) simulation algorithm, which simulates the passenger behavior based on the timetable, has been proposed to solve the problem. The paper is arranged as follows: Section 2 describes the parameters and basic terminologies that will be used in the study. Moreover, a service network has been proposed as the base network for this study. Section 3 proposes the optimal strategy model with the capacity constraints, and the effective frequency has been used to enable the model. Section 4 proposes the CVX simulation algorithm to solve the model proposed in Section 3. Section 5 describes the application of the optimal strategy model and the CVX algorithm to the Beijing Subway Network to verify the feasibility of the model. Section 6 concludes the paper and discusses future related studies.

## 2. Notation and Service Network

*2.1. Notation.* We first list all the involved notations and decision variables (also mentioned in the aforementioned discussion) in Table 1.

*2.2. Service Network.* The transit network is the base of the transit assignment process. In the frequency-based assignment model, the network is based on physics lines, which means the physics track. Todiscribe passenger behavior in detail, we demonstrate a service network that comprises passenger behavior and train service links.

To simplify the network illustration, we consider a single direction. As shown in Figure 1(a), there are two physics lines: Line 1 and Line 2. Line 1 serves stations A, B, and E. Line 2 serves stations B, C, and D. Station B is a transfer

Table 1: Notations used in this study.

| Notation | Explanation |
|---|---|
| $G = (E, I, T)$ | Network $G$ with nodes $I$, link $E$, and time slice set $T$ |
| $K$ | Passenger set |
| $E^+_{a,t}$ | Links from node $a$ at time $t$ |
| $\overline{E}^*_t$ | Optimal strategy at time $t$ |
| $v^t_a$ | Passenger volume at node $a$ and time $t$ |
| $v^t_a$ | The total incoming passengers at node $a$ |
| $\mathrm{In}_a$ | The entry passenger at node $a$, which generated at node $a$ |
| $\alpha$ | Calibration parameters, which depends on the distribution assumed for buses headways and passenger arrival times. |
| $c^t_{(a,b)}$ | Link $(a, b)$ travel time at time $t$ |
| $f^t_{(a,b)}$ | Service frequency for link $(a, b)$ at time $t$ |
| $v^t_{(a,b)}$ | Passenger volume for link $(a, b)$ at time slice $t$ |
| $\overline{f}^t_{(a,b)}$ | Train valid frequency |
| $\mathrm{Cap}^t_{(a,b)}$ | Capacity at link $(a, b)$ at time slice $t$ |
| $\eta$ | Maximal loading factor, which is usually set as 120% or 130% |
| $c^t_e$ | Best value for link $e$ at time $t$ |
| $u_k$ | Total travel time for passenger $k$ from automated fare collection (AFC) data |
| $c^t_{e0}$ | Minimal travel time for link $e$ at time $t$ |
| $b^{k,t}$ | Total waiting time calculated from the AFC data with a given path for passenger $k$ at time $t$ |
| $x^t_{(a,b)}$ | Binary variable, whether link $(a, b)$ is recorded in the optimal strategy or not |
| $\delta^{k,t}_e$ | Binary variable, whether link $e$ is selected by passenger $k$ at time $t$ |
| $\gamma^{k,t}_e$ | Binary variable, whether link $e$ is the none-waiting link or not |

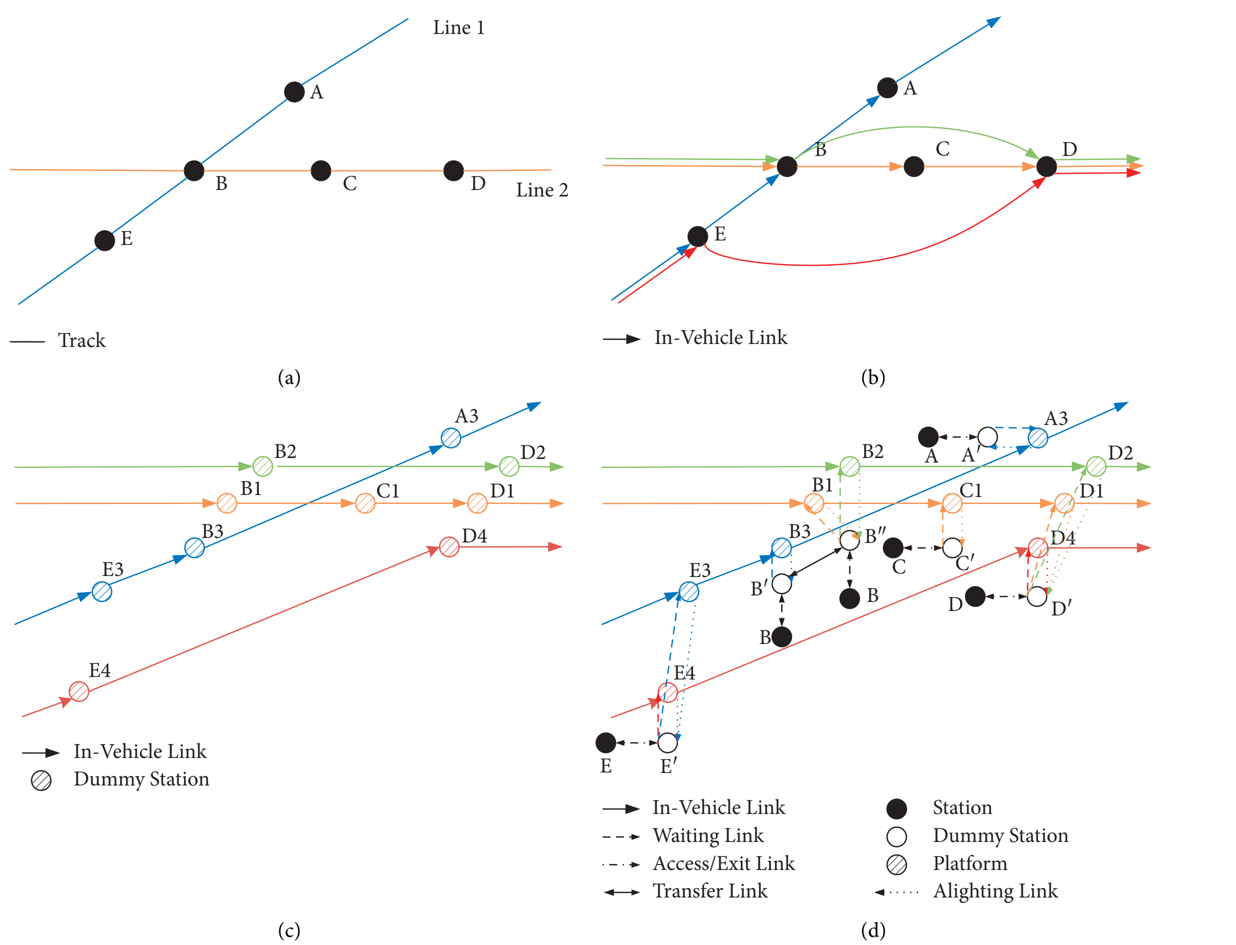



Figure 1: Transit service topology network: (a) physical network, (b) operation service, (c) frequency-based network, and (d) service network.

station. There are four kinds of transit services in the demo network, which are as follows (Figure 1(b)): stop-stop services denoted by blue and orange colors, skip-stop service denoted by green, and cross-line service denoted by red. In the frequency-based network, the stations are duplicated according to the transit services shown in Figure 1(c). Considering passenger behaviors such as waiting, boarding, and alighting, more links are added to the frequency-based network, which is defined as the service network here. There are five kinds of links: entry-egress links, waiting links, boarding-alighting links, in-vehicle links, and transfer links (Figure 1(d)).

The service network can detail the behavior of passengers from entrance to exit. We take the movement of a passenger from Station C to Station D as an example. The passenger reaches node C and walks to the dummy node C′, and he waits for trains at the waiting link C′–C1. When the train arrives, he can take the in-vehicle link C1–D1. He/she gets off when the train arrives at D1, following the alighting link D1–D′, and walks out of the station following D′–D.

*2.3. Link Cost and Frequency.* Each link has two parameters: link frequency and link cost. Link frequency represents the service frequency in which a passenger has to wait for the link. The link cost is the travel time or walking time spent on this link. The service topology network can be classified into two categories: the fixed link frequency and the flexible link frequency. The access and exit, boarding-alighting, in-vehicle, and transfer links belong to the fixed link frequency, where passengers can be served immediately when they reach to the start point of these link. The frequency was set to be infinite. The cost of transfer links, access, and exit links are based on the link walking time, which could be calibrated by field research. The cost of in-vehicle links comes from the timetable, which is the running time between the successive stations. We consider the boarding and alighting can be finished in a quick time and these link costs are set to be 0. For waiting links, the waiting time depends on the arrival distribution of trains and passengers. When the frequency of passengers follows a uniform distribution and the trains arrive evenlyand without capacity limitation, the waiting link frequency is half of the train frequency. The waiting link cost is 0.

## 3. Optimal Strategy with Capacity Constraints

In the normal optimal strategy model [2], it is assumed that there is no capacity limitation for the service network. However, in Beijing and Tokyo, the metro systems serve over 10 million passengers every day. In peak hours, the train capacity limitation is a huge obstacle for the passenger boarding process. Based on this, we have generated a time-based transit network $G = (E, I, T)$, where $T$ denotes the set for time slice $t$. We have introduced the capacity limitation constraint, shown in the last equations in F1, which is to minimize the total travel time of all passengers for all links. The train capacity changes or gets updated based on time and stations. The updated optimal strategy model (F1) with capacity constraint variables is shown below:

$$(F1)\ \mathrm{Min} \sum_{(a,b)\in E_t} c^t_{(a,b)} v^t_{(a,b)} + \sum_{a\in I} \alpha \frac{v^t_a}{\sum_{(a,b)\in E^+_{a,t}} f^t_{(a,b)} x^t_{(a,b)}},$$

$$\text{s.t.} \begin{cases} v^t_{(a,b)} = \dfrac{f^t_{(a,b)} x^t_{(a,b)}}{\sum_{(a,b)'\in E^+_{a,t}} f^t_{(a,b)'} x^t_{(a,b)'}} v^t_a, \ (a,b) \in E^+_{a,t}, \quad a \in I, \\ v^t_a = \sum_{(a,b)\in E^-_{a,t}} v^t_{(a,b)} + \mathrm{In}_a, \quad a \in I, \\ x^t_{(a,b)} = \begin{cases} 0, & \text{if } (a,b) \notin \overline{E}^*_t, \\ 1, & \text{if } (a,b) \in \overline{E}^*_t, \end{cases} \\ v^t_a \geq 0, \quad a \in I, \\ v^t_{(a,b)} \leq \eta \mathrm{Cap}^t_{(a,b)}. \end{cases} \tag{1}$$

We have applied effective frequency $\overline{f}^t_{(a,b)}$, which is related to the train capacity, to simplify the model. For instance, we assume that the passenger follows a uniform distribution, the train arrives evenly, and the service frequency is 3 min. When the train capacity is not limited, the waiting time of the passenger is 1.5 min. However, when there are many passengers on the platform and the train capacity is limited, a passenger may wait for the following trains to board on. In this condition, the waiting time is 4.5 min, and the valid train frequency is 9 min. Considering the effective frequency, the new model (F2) can be rewritten as

$$(F2)\ \text{Min} \sum_{(a,b)\in E_t} c^t_{(a,b)} v^t_{(a,b)} + \sum_{a\in I} \alpha \frac{v^t_a}{\sum_{(a,b)\in E^+_{a,t}} \overline{f}^t_{(a,b)} x^t_{(a,b)}},$$

$$\text{s.t.} \begin{cases} v^t_{(a,b)} = \dfrac{\overline{f}^t_{(a,b)} x^t_{(a,b)}}{\sum_{(a,b)'\in E^+_{a,t}} \overline{f}^t_{(a,b)'} x^t_{(a,b)'}} v^t_a,\ (a,b) \in E^+_{a,t}, \quad a \in I, \\ v^t_a = \sum_{(a,b)\in E^-_{a,t}} v^t_{(a,b)} + \text{In}_a, \quad a \in I, \\ x^t_{(a,b)} = \begin{cases} 0, & \text{if } (a,b) \notin \overline{E}^*_t, \\ 1, & \text{if } (a,b) \in \overline{E}^*_t, \end{cases} \\ v^t_a \geq 0, \quad a \in I. \end{cases} \tag{2}$$

The effective frequency has relaxed the transportation capacity constraints. The reformed model shares the same structure as the initial optimal strategy model, which implies that the model can be solved using the searching algorithm proposed by Spiess and Florian [18]. As the effective frequency cannot be directly obtained from the timetable, we have proposed a simulated CVX algorithm to obtain the effective frequency for each waiting link.

## 4. Passenger Simulation Process and Valid Train Frequency Updating Model

*4.1. Simulated CVX Algorithm.* The algorithm flow chart is shown in Figure 2. We cut the simulation time slice by 15 min. In the simulated CVX algorithm, there are two stop criteria. When the simulation time converges to the actual travel time, it implies that the simulation can represent the real situation. Thus, we can stop the simulation process. When the waiting link cost converges compared with the last iteration, the optimal strategy converges, and the simulation stops. The simulation will stop when it meets any of the aforementioned criteria.

The simulated CVX algorithm contains two important parts. The first one is the multi-agent simulation for the optimal strategy, which simulates the passenger behavior based on the optimal strategy and train capacity constraints. The second is the link-updating process to find the optimal value for the waiting link.

*4.2. Simulation Using the Capacity-Limited Optimal Strategy.* Agent-based simulation for transit services is not a new method [30–32]. The passenger behavior and timetable are set in the agent. For the previous research, the passenger path is assigned to each passenger agent. The passenger agent will follow the given path to finish its trip in the network. Based on the boarding, alighting, and transfer behaviors, the passenger agents and vechicle agents interact with each other.

Based on the previous research, we have designed the passenger and vehicle agents, denoted by red (left part) and blue (right part) in Figure 3. In this simulation process, passengers at each station first search for the optimal strategy. The passengers will follow the strategy when moving toward the destination. They can board the first train or transfer to the following trains according to the optimal strategy. When transferring to other train services, the train capacity should be considered to determine whether the passengers can board the train.

*4.3. Link-Updating Model for CVX.* A closed automated fare collection (AFC) system obtained an accurate total travel time for each passenger, which can help approach the real-time effective frequency. The least-squares link-updating model P1 is proposed as follows. The objective of P1 is to minimize the gap between the total calculated travel time and the total actual travel time.

$$\begin{aligned} &(P1)\ \min \sum_{k=1}^{K} \left( \sum_{e\in p(\overline{E}^t_i)} c^t_e - u_k \right)^2, \\ &\text{s.t.} \quad c^t_e \geq c^t_{e0}, e \in p\left(\overline{E}^t_i\right), \end{aligned} \tag{3}$$

where $\sum_{e\in p(\overline{E}^t_i)} c^t_e$ is the sum of successive links that the passenger used in the network to finish their trip and $\delta^{k,t}_e$ is the selecting index that represents whether the link is selected. P1 can be rewritten as P2, as given below:

$$\begin{aligned} &(P2)\ \min \sum_{K} \left( \sum_{e\in E} \delta^{k,t}_e c^t_e - u_k \right)^2, \\ &\text{s.t.} \begin{cases} c^t_e \geq c^t_{e0}, & e \in E, \\ \delta^{k,t}_e = \begin{cases} 1, & e \in p\left(\overline{E}^t_i\right), \\ 0, & \text{Otherwise}, \end{cases} \end{cases} \end{aligned} \tag{4}$$

where $\sum_{e\in E} \delta^{k,t}_e c^t_e$ defines the total travel time simulated for passenger $k$. As discussed before, all the link costs, except for waiting links, are constant. The total travel time gap can be set as the time gap between the simulation waiting time and real waiting time. Thus, the model can be rewritten as P3.

$$(P3)\min \sum_{K} \left( \sum_{a\in E^{\text{Boarding}}} \delta^{k,t}_e c^t_e - \left( u^{k,t} - \sum_{e'\in\left\{p\left(\overline{E}^t_i\right) - E^{\text{Boarding}}\right\}} \gamma^{k,t}_{e'} c^t_{e'} \right) \right)^2,$$

$$\text{s.t.} \begin{cases} c^t_e \geq 0, \quad e \in E^{\text{Boarding}} \\ \delta^{k,t}_e = \begin{cases} 1, & e \in \left\{p\left(\overline{E}^t_i\right) \cap E^{\text{Boarding}}\right\}, \\ 0, & \text{Otherwise}, \end{cases} \\ \gamma^{k,t}_e = \begin{cases} 1, & e \in \left\{p\left(\overline{E}^t_i\right) - E^{\text{Boarding}}\right\}, \\ 0, & \text{Otherwise}. \end{cases} \end{cases} \tag{5}$$

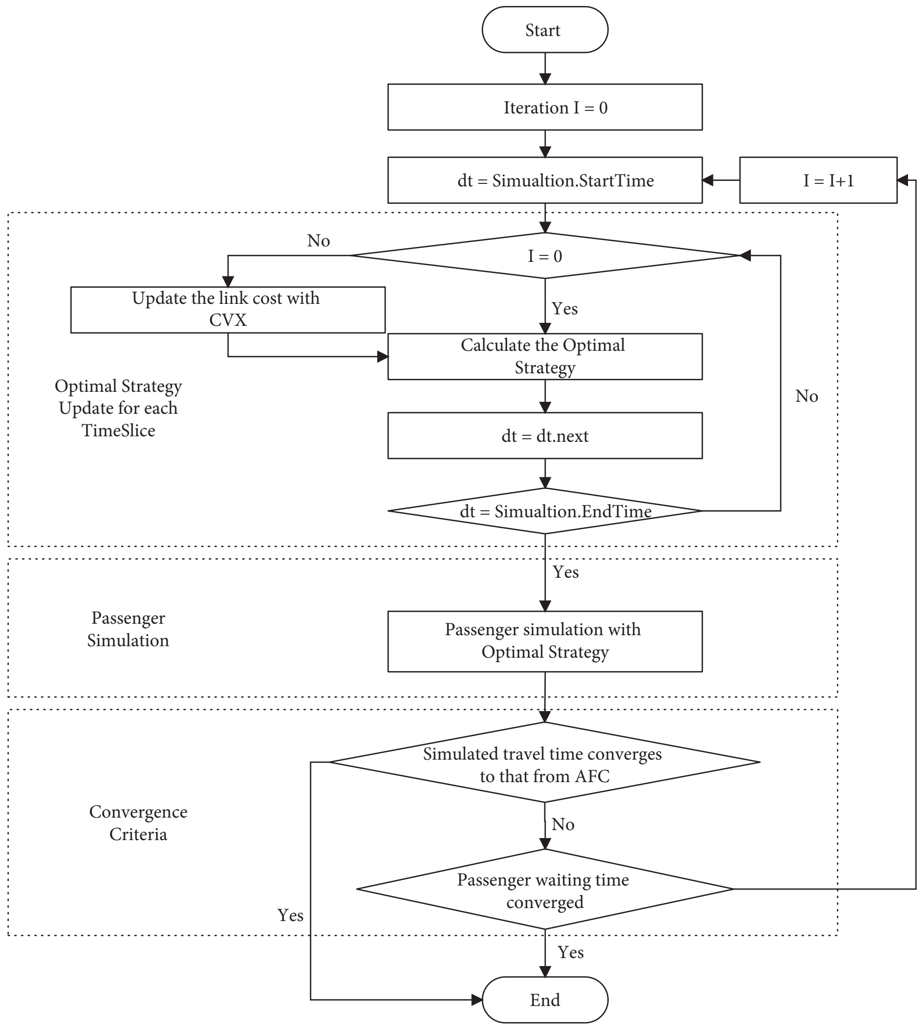



Figure 2: Dynamic simulation process for Beijing subway based on the optimal strategy.

The model complexity decreases from all the links to waiting links, and only the non-negative constraint remains. To simplify the model, $b^{k,t}$ is introduced, which is the total waiting time calculated based on the AFC data with a given path.

$$b^{k,t} = u^{k,t} - \sum_{e' \in \left\{p\left(\overline{E}_i^t\right) - E^{\text{Bording}}\right\}} \gamma_{e'}^{k,t} c_{e'}^{t}. \tag{6}$$

The matrix formulation of P3 is shown below:

$$\begin{aligned} &(P4)\min\left(\|AC - B\|_2\right)^{1/2}, \\ &\text{s.t.} C \geq \overrightarrow{0}. \end{aligned} \tag{7}$$

where $C$ is the column for all waiting links in the network and $A$ is the $k \times n$ matrix. For each element in $A$, $a_{i,j}$ is a dual variable, 1 denotes passenger $i$, and waiting link $j$ is used in this simulation. $B$ is the travel time gap for $k$ passengers. Considering the monotonicity of 2-norm, $\min(\|AC-B\|_2)^{1/2}$ equals $\min\|AC - B\|_2$. The current waiting link update model is shown as P5.

$$\begin{aligned} &(P5)\min\|AC - B\|_2, \\ &\text{s.t.} C \geq \overrightarrow{0}. \end{aligned} \tag{8}$$

The new model has a convex optimization problem and can be solved using CVX, which was proposed by Prof. Stephen Boyd and Dr. Michael Grant from Stanford University [22]. The link-updating process is shown in Figure 4.

## 5. Application for Beijing Subway

*5.1. Data in Beijing Subway.* Data used in the empirical study are listed in Table 2. We have used one-week AFC data (December 2016) of the Beijing Subway. During that time, there were 17 lines serving more than 10 million passengers every day with more than 8000 train services. The majority of line headways ranged from 2 to 5 min. In the peak hour, the headway could reach 90 s. After applying the service topology construction method, we gathered 96974 links, and every link is labeled with cost and frequency. Based on the annual report of the Beijing Subway, the on-time rate is >99.9%, and we use the timetable as the base document to initialize the train frequency.

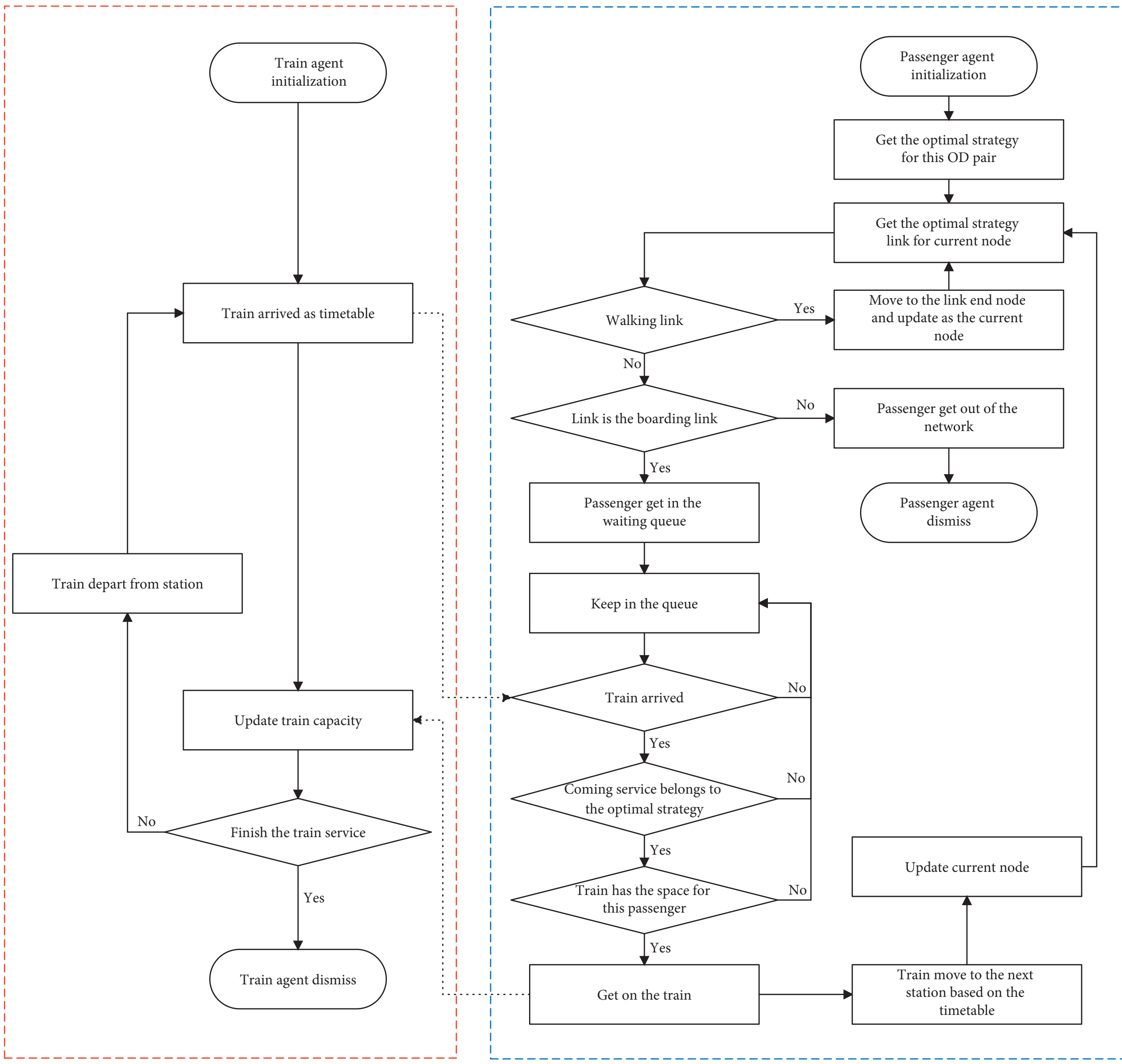



Figure 3: Interaction between passenger and vehicle agent.

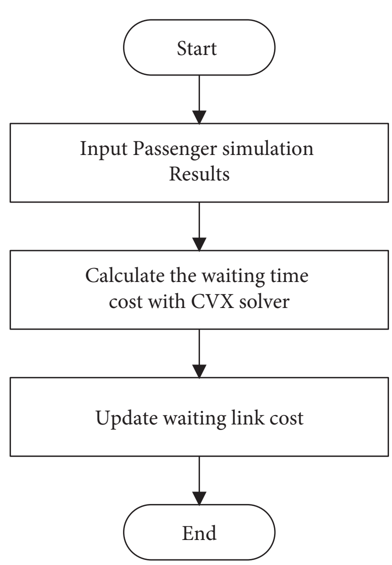



Figure 4: Updating process for waiting link.

Table 2: Dataset in the empirical study.

| Dataset | Fields | Explanation |
|---|---|---|
| AFC | Card ID | Unique number that could be taken as the passenger ID |
| | O station | Boarding station ID |
| | Entry time | Access time to the boarding station |
| | D station | Alighting station ID |
| | Exit time | Exit time from the alighting station |
| Link | Link ID | Unique number that could be taken as the link ID |
| | LinkType | Represent the link belongs to the walking/transfer/boarding or alighting link |
| | Link cost | Link travel time |
| | Link frequency | Service frequency for the link |
| | Day of week | Workday or weekend |
| | Time of day | Peak hour and off-peak hour |
| Timetable | Train no. | Train number |
| | StationID | Served station for a specific train, ordered by service sequence |
| | Train loop | Two terminal stations of the loop |
| | Arriving time | Arriving time at each station |
| | Departure time | Departure time at each station |

*5.2. Results Validation.* We focus on passenger behavior during the peak hours, 7:00–9:00 AM. The simulation starts at 6:00 AM to warm up the system and ends at 10:00 AM. The simulation process converges to the waiting time criterion after 23 iterations. The validation results are analyzed as follows:

*5.2.1. Verification of Total Travel Time.* Test samples are simulated results of 7:30 to 8:30 AM. For every 15 min, we have compared the real travel time with the simulated travel time using mean relative error (MRE) and mean absolute relative error (MARE).

$$\mathrm{MRE} = \frac{1}{n}\sum_{i=1}^{n}\frac{\left(T_i^{s-\mathrm{Exit}} - T_i^{a-\mathrm{Entry}}\right) - \left(T_i^{a-\mathrm{Exit}} - T_i^{a-\mathrm{Entry}}\right)}{\left(T_i^{s-\mathrm{Exit}} - T_i^{a-\mathrm{Entry}}\right)},$$

$$\mathrm{MARE} = \frac{1}{n}\sum_{i=1}^{n}\frac{\left|\left(T_i^{s-\mathrm{Exit}} - T_i^{a-\mathrm{Entry}}\right) - \left(T_i^{a-\mathrm{Exit}} - T_i^{a-\mathrm{Entry}}\right)\right|}{\left(T_i^{s-\mathrm{Exit}} - T_i^{a-\mathrm{Entry}}\right)}. \tag{9}$$

where $T_i^{s-\mathrm{Exit}}$ is the simulated exit time for passenger *i*. $T_i^{a-\mathrm{Exit}}$ and $T_i^{a-\mathrm{Entry}}$ are the exit and entry times recorded in the AFC system, respectively. The MRE and MARE results are listed in Table 3.

Based on the test period, MRE is ~7%, and MARE is ~10%. Considering the fluctuation in walking time of different passengers, we accept the simulation results.

*5.2.2. Verification of Typical Origin-Destination (OD) Pair.* Considering the diverse travel times among different OD pairs, we have selected typical OD pairs, which have a high passenger volume of 250–400 in 15 min. These OD pairs comprise large residential areas, e.g., Huilongguan and Huoying, and commercial areas, e.g., Xierqi and Fengtai Science Park. To balance the sample OD spatial distribution, we selected some other OD pairs located in the southeast part of the network.

We applied the *F*-test and *T*-test for the total travel time obtained from these OD pairs, as listed in Table 4. The result indicates that more than 99% of the OD pairs passed the test. Less than 1% of the OD pairs failed the test because the OD volume is somewhat minimal, which means that the less travel time sample is not very stable. Thus, the proposed model and algorithm can represent passenger behavior in the Beijing Subway network. Furthermore, we can analyze the passenger behavior using the optimal strategy.

*5.3. Passenger Behavior Analysis Using the Optimal Strategy*

*5.3.1. Waiting Time during the Peak Hours in the Morning.* Owing to the station entrance limitation, operation plan, and total travel demand, passengers get delayed while reaching the platform and need to wait for trains during the peak hours in the morning. To analyze the waiting time at the platform, we select the stations that has a top entry-in or transfer passenger volume, including the direction. The selected stations and their waiting times are shown in Figure 5.

For the first 15 min, from 7:30 to 7:45 AM, the average waiting time (AWT) is 157 s. The AWT decreases to 123 s in the next 15 min. This represents the crowd dispersed over time. Considering the Xingong station on Line 4 as an example, the train headway is 120 s. From 7:30 to 7:45 AM, the AWT increases to 297 s, which implies that passengers need to wait for three headways until they board the arriving train. After 30 min, the AWT deceases to 74.3 s, approximately half of the headway, which implies that the passenger can board the next train when it arrives. The condition of the Shuangjing station denoted by Line 10 is the opposite. The headway is 120 s. In the first 15 min, the AWT is 84 s, and it increases to 158.7 s after 30 min, which means that passengers have to wait for another train. This could be summarized as the rule that most of the commuters live in a rural area and work downtown. During the peak hours in the morning, the main stations, e.g., Tiangngyuan, Xingong, and Shahe, pertaining to waiting

Table 3: MRE and MARE results for every 15 min.

| Period (AM) | MRE% | MARE% |
|---|---|---|
| 7:30–7:45 | −6.094 | 9.5164 |
| 7:45–8:00 | −7.148 | 10.2293 |
| 8:00–8:15 | −7.026 | 10.4602 |
| 8:15–8:30 | −6.133 | 10.0834 |

Table 4: *T*-test and *F*-test for OD travel time.

| Origin | OD pairs destination | 7:30–7:45 AM | | 7:45–8:00 AM | | 8:00–8:15 AM | | 8:15–8:30 AM | |
|---|---|---|---|---|---|---|---|---|---|
| | | *V* | *P* | *V* | *P* | *V* | *P* | *V* | *P* |
| TiantongyuanNorth | Huixinxijie Beikou | 88 | Y | 117 | Y | 187 | Y | 122 | Y |
| Xingong | Xuanwumen | 54 | Y | 77 | Y | 52 | N | 61 | Y |
| Tuqiao | Sihui | 223 | Y | 227 | Y | 184 | Y | 145 | Y |
| Shahe University Park | Xierqi | 412 | Y | 428 | Y | 437 | Y | 352 | Y |
| Huoying | Wudaokou | 162 | Y | 215 | Y | 200 | Y | 160 | Y |
| Shilihe | Jintaixizhao | 99 | Y | 196 | Y | 269 | Y | 271 | Y |
| Changyang | Fengtai Science Park | 140 | Y | 204 | Y | 354 | Y | 281 | Y |
| Caofang | Chaoyangmen | 145 | Y | 169 | Y | 141 | Y | 82 | Y |
| Libafang | Fengtai Science Park | 89 | Y | 73 | Y | 150 | Y | 92 | Y |
| Songjiazhuang | Chaoyangmen | 105 | Y | 95 | Y | 93 | Y | 77 | Y |

*V*: OD volume; *P*: pass the test; Y: yes; N: no.

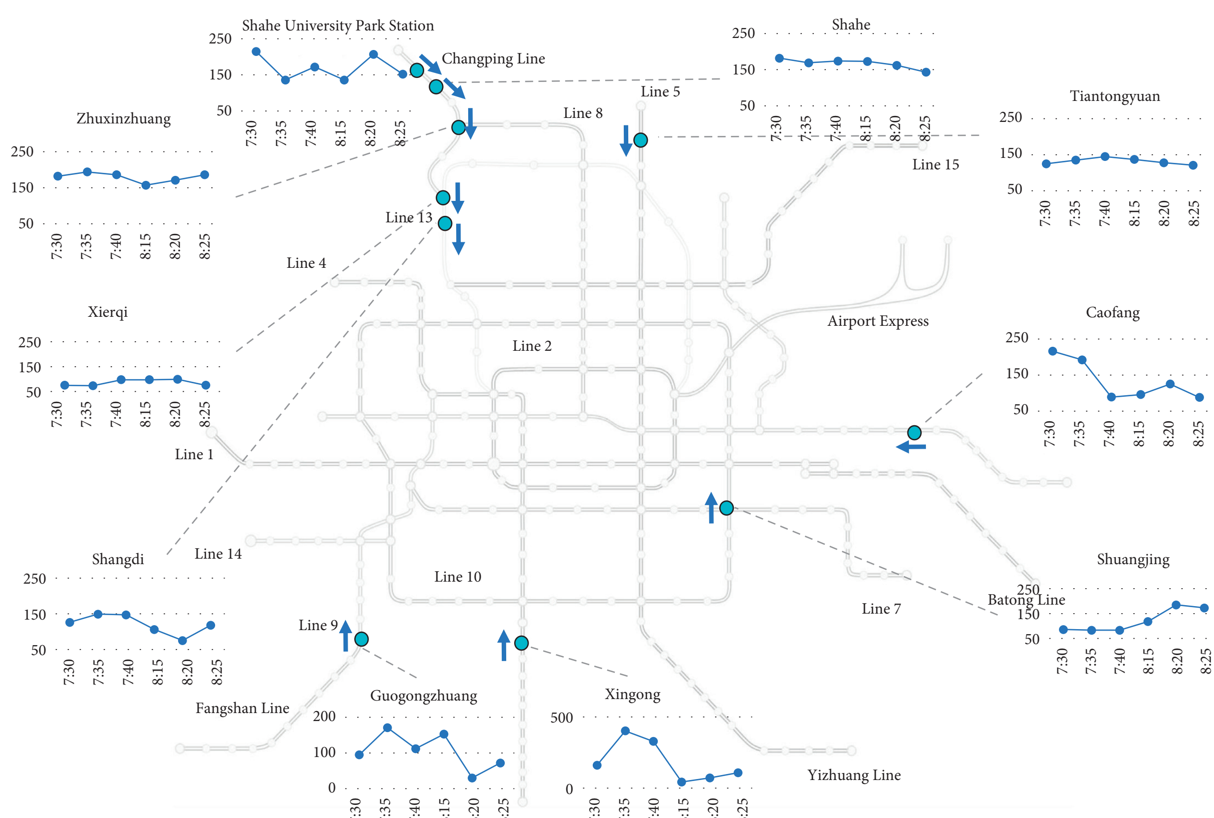



Figure 5: Waiting time analysis for stations.

time are distributed in rural lines. Over time, the crowd moves to the central part of the network and downtown stations.

*5.3.2. Loading Factor for Different Loops on the Same Line.* Line 4-daxing has two loops (Figure 6). The long loop is from Tiangongyuan to Anheqiaobei; both are terminal stations.

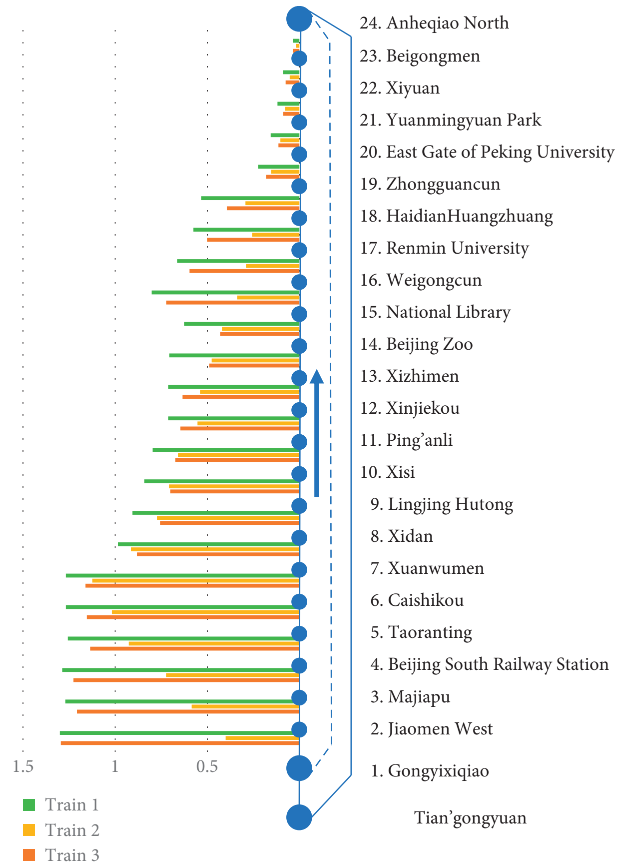



Figure 6: Operation loop of line 4-daxing in Beijing.

The shorter loop (dashed line) is from Gongyixiqiao to Anheqiaobei (Figure 6). To simplify the presentation, stations from Tiangongyuan to Gongyixiqiao have been excluded.

In the morning, many passengers travel from south to north. According to the optimal strategy, the service links from long and short loops are recorded, and passengers can choose a long loop or a short loop. We have selected three successive long-short loop train services. Train 1 and Train 3, i.e., long-loop services, arrive at Gongyixiqiao at 7:29:06 and 7:32:06 AM, respectively (Figure 6). Train 2, i.e., the short-loop service, arrives at Gongyixiqiao at 7:30:51 AM (Figure 6). The loading factors of the trains at each section from Gongyixiqiao to the terminal station are denoted by bar charts in Figure 6.

The loading factor in the long loop is greater than that in the short loop. A huge difference is observed between Gongyixiqiao and Taoranting (Section ID 1–5). In the stations in this section, passengers can switch between long-loop and short-loop trains. Owing to the low frequency of the short loop, which is 10 min on an average, passengers who board the long-loop train will not switch to the short loop train, even if the short loop has a lower loading factor. This leads to an imbalance between the long and short loops. For future operations, we should apply skip-stop plans with flexible loop operations to balance the passenger flow and reduce their total travel time.

*5.3.3. Transfer Volume in the Network.* The average transfer index in Beijing is 1.8, which means every passenger has to be transferred 1.8 times before arriving at the destination. Figure 7 represents the top eight transfer stations ordered by the transfer volume. For every 15 min, the transfer volume was over 4000 in average. Similar to the waiting time distribution, the transfer volume decreases with time in the rural area, while the transfer volume increases in the station that is located in the center of the network. In the future operation design, thorough services for some transit lines should be considered to reduce the transfer and total travel time, with the constraints of turn-back and rolling stocks.

From simulation results, it could be found that, in the morning peak hour, both the long waiting time stations and large transfer volume stations are not in the central of the network, which are specific evidences that most of the commuters are living in the rural area. Second, from the time series of waiting time and transfer volume, the peak of the morning peak hour is very intense. Because of the super high demand and the limited capacity, passengers need to wait for the second even third successive train. To better satisfy the passenger demand, the flexible operation solution such as the long-short loop was provided. Compared with the single operation scheme, the long-short loop relieved some operation pressure, but it could be better if other operation scheme could be applied, such as stop-skip pattern. Actually, the Beijing metro operation company keeps trying to

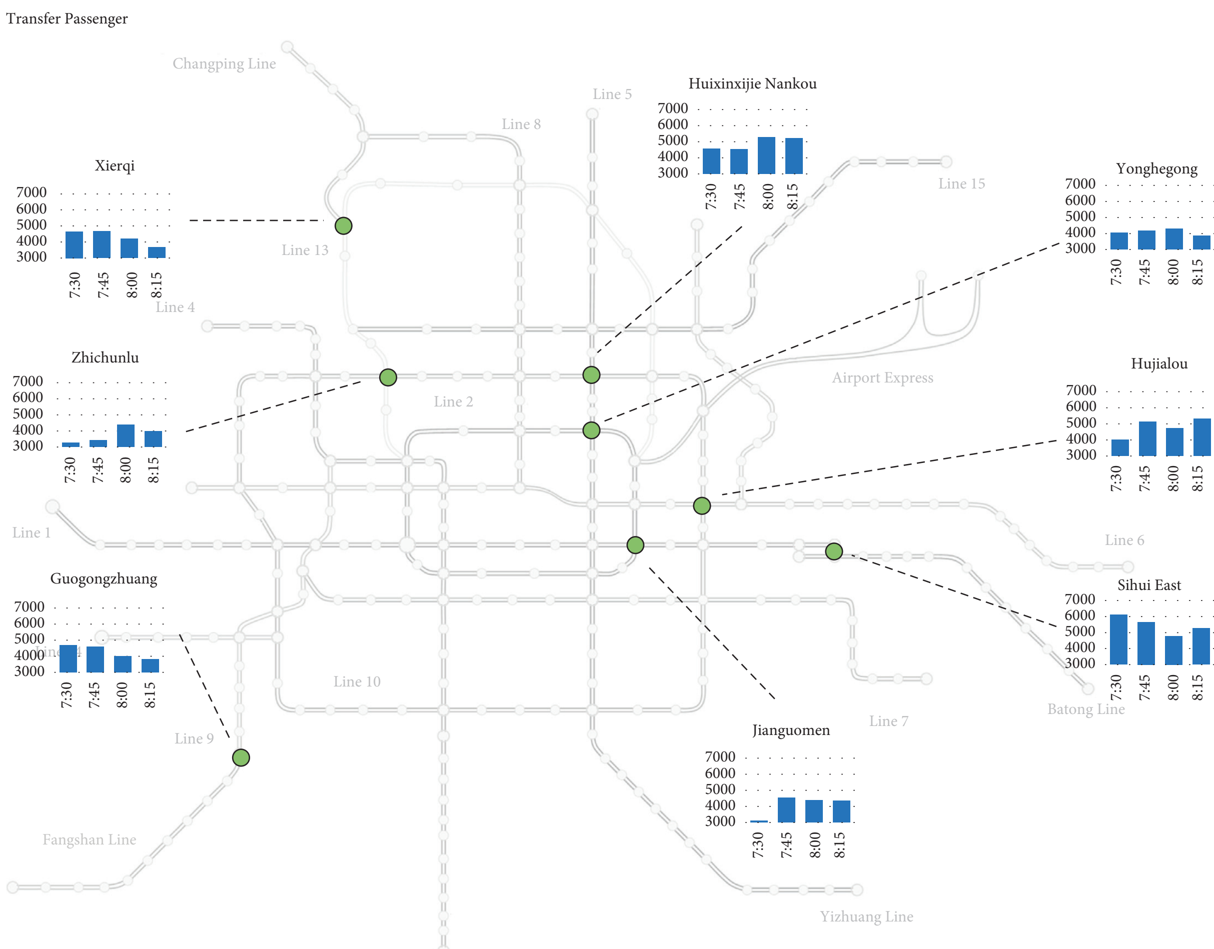



Figure 7: Locations of top eight transfer stations during the peak hours in the morning.

improve the system capacity. During the COVID-19 period in 2020, the more flexible and strong timetables were provided, such as the asymmetry timetable and skip-stop timetable to better serve the tidal passenger flow. At the same time, in order to reduce the passenger transfer time, the trough services between Line 1 and Line Batong are provided. Because of the lack of data, the performance of the new timetables was not analyzed in this research.

## 6. Conclusion

The Beijing Subway has a high and fixed transit service frequency based on the assumption analysis of the frequency-based transit assignment method and optimal strategy. Considering the train capacity limitation in the metro system, this study introduces the capacity constraint for the optimal strategy. To overcome the capacity limitation, a CVX simulation algorithm has been proposed. The empirical study demonstrates that the updated model and algorithm can solve the transit assignment problem for busy transit systems. The results detail the passenger behavior in the network. It is possible to analyze the passenger behavior under complex conditions to provide better solutions.

The main aim of the optimal strategy is to represent passenger behavior. However, some passengers preferred to determine the path in advance. In that case, they seldom change their path during the trip, and the logit model can demonstrate their behavior better than the optimal strategy. Every model has its own assumptions and characteristics. It is better to analyze passengers based on their perspectives.

Meanwhile, the objective of the optimal strategy model is to minimize the total travel time of passengers. However, the decision-making process considers more than total travel time. Moreover, they would consider discomfort or transfer time. Therefore, the transfer time and transfer penalty should be added to the objective function, thereby making the model closer to reality.

## Data Availability

The AFC and timetable used to support the findings of this study were supplied by the project: the metro network operation supervision and forecasting, provided by Transportation Operations Coordination Center and Beijing Jiaotong University, cannot be made freely available. The service link is from https://map.bjsubway.com, which is open to public.

## Conflicts of Interest

The authors declare no conflicts of interest.

## Authors' Contributions

Conceptualization, K.L. and N.C.; methodology, K.L. and N.C.; formal analysis, K.L; writing—original draft preparation, K.L.; resources, K.L. and N.C.; supervision, N.C.

## Acknowledgments

The authors would like to thank Dr. Alireza Khani from the University of Minnesota and Professor Zijia Wang from Beijing Jiaotong University for helping us edit and Polish the language of this paper. The authors would like to thank Editage (http://www.editage.cn) for English language editing. This work is supported by National Key R&D Program of China (ID: 2020YFB1600702), Beijing Postdoctoral Research Foundation (Award Number: ZZ2019-118), and Beijing Science and Technology Project (ID: Z191100002519003).

## References

[1] F. Chen, H. Yang, and Y. Yang, "Mas-based model for evaluating train timetables to minimise the waiting time," *KSCE Journal of Civil Engineering*, vol. 23, no. 3, pp. 1360–1368, 2019.

[2] Y. Zhou, Y. Wang, H. Yang, and X. Yan, "Last train scheduling for maximizing passenger destination reachability in urban rail transit networks," *Transportation Research Part B: Methodological*, vol. 129, pp. 79–95, 2019.

[3] Y. Chen, B. Mao, Y. Bai, T. K. Ho, and Z. Li, "Timetable synchronization of last trains for urban rail networks with maximum accessibility," *Transportation Research Part C: Emerging Technologies*, vol. 99, pp. 110–129, 2019.

[4] Y. Zhu and R. M. P. Goverde, "Dynamic passenger assignment for major railway disruptions considering information interventions," *Networks and Spatial Economics*, vol. 19, no. 4, pp. 1249–1279, 2019.

[5] L. L. Cheung, Shalaby, A. S. Amer, and Shalaby, "System optimal re-routing transit assignment heuristic: a theoretical framework and large-scale case study," *International Journal of Transportation Science and Technology*, vol. 6, no. 4, pp. 287–300, 2017.

[6] K. Gkiotsalitis and O. Cats, "Timetable recovery after disturbances in metro operations: an exact and efficient solution," *IEEE Transactions on Intelligent Transportation Systems*, pp. 1–11, 2020.

[7] N. Papola, F. Filippi, G. Gentile, and L. Meschini, "Schedule-based transit assignment: new dynamic equilibrium model with vehicle capacity constraints," *Schedule-Based Modeling of Transportation Networks*, vol. 46, pp. 1–26, 2010.

[8] A. Nuzzolo and U. Crisalli, "The schedule-based approach in dynamic transit modelling: a general overview," *Schedule-Based Dynamic Transit Modeling: Theory and Applications*, vol. 28, pp. 1–24, 2004.

[9] A. Nuzzolo, F. Russo, and U. Crisalli, *Transit Network Modelling. The Schedule-Based Dynamic Approach*, Collana Trasporti; Franco Angeli, Milan, Italy, 2003.

[10] A. Nuzzolo, F. Russo, and U. Crisalli, "A doubly dynamic schedule-based assignment model for transit networks," *Transportation Science*, vol. 35, no. 3, pp. 268–285, 2001.

[11] M. H. Poon, S. C. Wong, and C. O. Tong, "A dynamic schedule-based model for congested transit networks," *Transportation Research Part B: Methodological*, vol. 38, no. 4, pp. 343–368, 2004.

[12] A. Nuzzolo, U. Crisalli, and L. Rosati, "A schedule-based assignment model with explicit capacity constraints for congested transit networks," *Transportation Research Part C: Emerging Technologies*, vol. 20, no. 1, pp. 16–33, 2012.

[13] Y. Hamdouch and S. Lawphongpanich, "Schedule-based transit assignment model with travel strategies and capacity constraints," *Transportation Research Part B: Methodological*, vol. 42, no. 7-8, pp. 663–684, 2008.

[14] Y. Hamdouch, H. W. Ho, A. Sumalee, and G. Wang, "Schedule-based transit assignment model with vehicle capacity and seat availability," *Transportation Research Part B: Methodological*, vol. 45, no. 10, pp. 1805–1830, 2011.

[15] Y. Hamdouch, W. Y. Szeto, and Y. Jiang, "A new schedule-based transit assignment model with travel strategies and supply uncertainties," *Transportation Research Part B: Methodological*, vol. 67, pp. 35–67, 2014.

[16] Y. Zhang, W. H. K. Lam, A. Sumalee, H. K. Lo, and C. O. Tong, "The multi-class schedule-based transit assignment model under network uncertainties," *Public Transport*, vol. 2, no. 1-2, pp. 69–86, 2010.

[17] L. Meschini, G. Gentile, and N. Papola, "A frequency based transit model for dynamic traffic assignment to multimodal networks," in *Proceedings of the 17th International Symposium on Transportation and Traffic Theory*, vol. 2, no. 3, pp. 2-3, London, 2007 July.

[18] H. Spiess and M. Florian, "Optimal strategies: a new assignment model for transit networks," *Transportation Research Part B: Methodological*, vol. 23, no. 2, pp. 83–102, 1989.

[19] S. Nguyen and S. Pallottino, "Equilibrium traffic assignment for large scale transit networks," *European Journal of Operational Research*, vol. 37, no. 2, pp. 176–186, 1988.

[20] J. De Cea and E. Fernández, "Transit assignment for congested public transport systems: an equilibrium model," *Transportation Science*, vol. 27, no. 2, pp. 133–147, 1993.

[21] R. Cominetti and J. Correa, "Common-lines and passenger assignment in congested transit networks," *Transportation Science*, vol. 35, no. 3, pp. 250–267, 2001.

[22] M. Cepeda, R. Cominetti, and M. Florian, "A frequency-based assignment model for congested transit networks with strict capacity constraints: characterization and computation of equilibria," *Transportation Research Part B: Methodological*, vol. 40, no. 6, pp. 437–459, 2006.

[23] J.-D. Schmöcker, M. G. H. Bell, and F. Kurauchi, "A quasi-dynamic capacity constrained frequency-based transit assignment model," *Transportation Research Part B: Methodological*, vol. 42, no. 10, pp. 925–945, 2008.

[24] J.-D. Schmöcker, A. Fonzone, H. Shimamoto, F. Kurauchi, and M. G. H. Bell, "Frequency-based transit assignment considering seat capacities," *Transportation Research Part B: Methodological*, vol. 45, no. 2, pp. 392–408, 2011.

[25] K. Nökel and S. Wekeck, "Boarding and alighting in frequency-based transit assignment," *Transportation Research Record: Journal of the Transportation Research Board*, vol. 2111, no. 1, pp. 60–67, 2009.

[26] N. Oliker and S. Bekhor, "A frequency based transit assignment model that considers online information," *Transportation Research Part C: Emerging Technologies*, vol. 88, pp. 17–30, 2018.

[27] E. Cascetta and P. Coppola, "Assessment of schedule-based and frequency-based assignment models for strategic and operational planning of high-speed rail services,"

*Transportation Research Part A: Policy and Practice*, vol. 84, pp. 93–108, 2016.

[28] B. Otto, "Aggregation techniques for frequency assignment in public transportation," *Public Transport*, vol. 11, no. 1, pp. 51–87, 2018.

[29] http://cvxr.com/cvx/.

[30] E. Adamey, A. Kurt, and U. Ozguner, "Agent-based passenger modeling for intelligent public transportation," in *Proceedings of the 16th International IEEE Conference on Intelligent Transportation Systems (ITSC 2013)*, IEEE, The Hague, Netherlands, Oct 2013.

[31] T. Schelenz, Á. Suescun, L. Wikström, and M. Karlsson, "Application of agent based simulation for evaluating a bus layout design from passengers' perspective," *Transportation Research Part C: Emerging Technologies*, vol. 43, pp. 222–229, 2014.

[32] H. Yin, B. Han, D. Li, J. Wu, and H. Sun, "Modeling and simulating passenger behavior for a station closure in a rail transit network," *PLoS One*, vol. 11, no. 12, Article ID e0167126, 2016.

Hindawi
Complexity
Volume 2021, Article ID 5888549, 12 pages
https://doi.org/10.1155/2021/5888549

*Research Article*

# Research on the Coordinated Development of Global Urban Economic Competitiveness: Based on a Sample of 1007 Cities

**Xiaonan Liu,[1,2] Pengfei Ni,[3] Fangqu Niu ,[4] Bo Li,[5] and Qihang Li[6]**

[1]*Tehua Postdoctoral Programme, Beijing 100029, China*
[2]*Institute of Finance & Banking, CASS, Beijing 100710, China*
[3]*National Academy of Economic Strategy, CASS, Beijing 100006, China*
[4]*Institute of Geographic Sciences and Natural Resources Research, CAS, Beijing 100101, China*
[5]*School of Management, Tianjin University of Technology, Tianjin 300384, China*
[6]*Economic Research Center, Shandong University of Finance and Economics, Jinan 250014, China*

Correspondence should be addressed to Fangqu Niu; niufq@lreis.ac.cn

Received 28 July 2021; Revised 29 September 2021; Accepted 13 November 2021; Published 7 December 2021

Academic Editor: Jing-Hu Pan



Based on the global urban economic competitiveness data in 2017, this study conducts coupling analyses of the competitiveness indicator system. The comprehensive study on the coupling coordination degree among explanatory indexes of urban economic competitiveness concludes that the city with higher economic competitiveness rankings has a higher degree of coupling coordination (DCC); the city ranked lower in the economic competitiveness has a lower DCC. The cities with higher DCC are mainly those global cities or metropolis known for financial and technological innovations, while cities with bare coupling coordination are mainly in underdeveloped countries in Asia, Africa, and Latin America. Based on the findings, the paper employs a model that combines linear regression and quantile regression to identify the specific driving factors that affect the cities' competitiveness around the world. Therefore, every city should act according to local conditions, focus on the key drivers of urban development, and address the inadequacies to balance the economic development so as to enhance its competitiveness.

## 1. Proposition and Literature Review

*1.1. Proposition.* As urban population continues to grow after entering the new century, urban competitiveness is becoming more and more prominent in the development of human society. The improved transportation and advanced Internet and other information technologies have greatly shortened the distance between people and facilitated the interactions between cities with growing mutual influences. This has also aggravated the competitions for resources and businesses among the cities. With deepening economic globalization, the lagging development of developed economies, and the rapid rise of emerging economies, cities around the world face more opportunities and challenges. As the competition among cities is increasingly intensified, urban competitiveness has become a key issue concerning how a country or a region can survive and develop under the pressure of global competition. The World Economic Forum continues to foster international exchanges and cooperation to address the issue. Urban competitiveness, as a topic of scientific researches, has caused widespread concerns among experts and scholars in various disciplines, including economy, geography, and urban planning, and has become a major national policy agenda.

On the other hand, there is a wooden barrel effect in the actual economic and social development. It is very important to study whether there is a barrel effect in urban competitiveness and promotion strategies. According to the "Global Urban Competitiveness Report 2017–2018" jointly published by the Chinese Academy of Social Sciences and UN-HABITAT, the top ten cities in the global competitiveness are cities in developed countries, except Shenzhen in China. Over half of these cities are in the United States. The global economic competitiveness is highly concentrated in a few cities, a seriously unbalanced development among the cities worldwide. Cities in developed countries such as the United States, Germany, and the United Kingdom generally have relatively strong economic competitiveness, while cities in developing countries such as China and India make up a center-periphery pattern, in which the center cities have much stronger competitiveness than other cities in the country. To focus on priorities, address inadequacies, and shore up points of weakness are currently important starting points to enhance the urban economic competitiveness.

The DCC is to measure the synergy between systems and analyze the degree of coordinated development of systems. Therefore, the coupling coordination degree model can not only reflect the degree of interaction between systems, but also reflect the level of coordinated development of systems. In studying competitiveness, to analyze and compare the DCC in the urban competitiveness is very important for the cities to enhance their respective competitiveness. Such studies shed light on the future development and deepen the understanding of each city's specific internal structure and competitiveness. By analyzing the relative indexes of economic competitiveness, such as technology innovation, financial services, industrial systems, human resources, business environment, business cost, infrastructure, and living standards, the factors that are staying in line with, moving ahead of, or lagging behind the development of economic competitiveness can be identified. By addressing the key factors that restrict the DCC in city's competitiveness, adjusting the measures to suit local conditions, and concentrating on the factors that can significantly boost competitiveness, including technology, finance, local demand, business environment, and infrastructure, the economic competitiveness of cities around the world can be enhanced.

This study is intended to analyze the DCC and urban competitiveness of cities across the world and examines the influencing factors of urban competitiveness, which have not been fully and accurately analyzed due to the limitations of data collection and statistical range. Therefore, this paper fills in the research gap on global urban competitiveness and provides reference points in improving the economic competitiveness of global cities.

*1.2. Literature Review.* Coupling coordination theory was first proposed by Haken, a German physicist, in the 1970s [1]. It was originally used in the field of laser physics. Prior to this, in the 1930s, the Austrian-American biologist Bertalanffy first proposed the general system theory in the field of biology [2]. The theory of coupling coordination was developed by later generations and consistently applied in the field of economics [3–6]. There are few literatures on the DCC, and there are even fewer studies on coupling coordination between the various indicators of urban competitiveness and the cities' competitiveness. So far, there is no study on the DCC between various urban factors, which has important significance to the analysis and studies on the DCC between cities. This study, based on 1,007 sample cities, constructs a global competitiveness indicator system that comprises 10 Level-2 indicators and 24 Level-3 indicators to analyze the coupling coordination between economic competitiveness and the explanatory indicators, such as technological innovation, financial services, industrial systems, business environment, operational cost, and living standards. The derived DCC are classified into four tiers, to which the sample cities are assigned accordingly. This model shows the distribution of these cities' DCC and indicates a strong correlation to the urban competitiveness. The following analysis will be focused on the coupling degrees between the factors of urban competitiveness and the characteristics of economic competitiveness itself.

The multidimensional nature of urban competitiveness is the premise and basis for a possible coupling analysis. Most scholars emphasize that the factors affecting urban competitiveness are complex and multidimensional, different from the single factor for business competitiveness, the economic performance [7–11]. Lever (1993, 1999, 2002) believes that what makes a successful city should be a combination of economic development, sustainability, and quality of life [12–14]. Cities should pay more attention to "soft" assets such as innovation environment, corporate relations, resident's expectation, institutional capacity, and quality of education and research. He also points out that the supporting elements for successful competition are a series of multidimensional factors, including aspiring and people-oriented leadership; flexible and adaptive workforce; quick and responsive public administration; efficient corporate partnerships; and entrepreneurial environments. The Beacon Hill Institute has been evaluating the long-term competitiveness of each state of the United States since 2001 [15, 16]. In his study, the long-term competitiveness is attributed to eight factors: local government and taxation policies, security, infrastructure, human resources, technology, business incubation, inclusiveness, and environmental policy.

Kresl and others believe that regional competitiveness has a self-enhancement (inner coupling) mechanism [17–20]. This mechanism is first reflected in the fundamental aspects of the region (education, corporate culture, public infrastructure and services, institutional composition and trends, policy regime, and cultural atmosphere) and the external economic engagement in the region (specialized labor pool, specialized supply and service network, knowledge dissemination and technology transfer, specialized systems, and dedicated capital markets; the second is the regional competitiveness (innovation, investment, technology, business, networks, diversified and specialized economy, quality of life, and strategic policies). The third is output, that is, the external characteristics of competitiveness (productivity, employment, wages, and per capita GDP). The input and direct contribution of the output also become part of the driving force, which, in turn, contributes to the output; and then the output also reacts with the input and the impetus [21, 22]. This circular and accumulated self-enhancement process constitutes the coupling characteristic of the multidimensional indicators of the urban competitiveness. The pyramid model of urban competitiveness proposed by Begg (1999) also embodies the coupling characteristics [23]. This model is composed of four levels. The bottom level is the various environmental factors that affect urban competitiveness, including the top-down institutional and macro environment, company characteristics, business environment, innovation, and learning capacity. In this regional competitiveness assessment model, labor productivity and employment rate are regarded as dominant factors of urban competitiveness. Other environmental factors are explanatory factors.

Hao and Ni (1998, 2001) established a "Bow-String-Arrow-Target" model of urban competitiveness [24, 25]. The bow and string represent the hard and soft elements of competitiveness, while the arrow represents the competitive industry, the value of competitiveness. This objectively requires that a city has bow and string (hard factor and soft environment, input) to determine the arrow (industry, process) and then to form a multielement coupling system with decisive (value, economic output) and reciprocal feedbacks from opposite directions. At the same time, the coupling competitiveness also comes from common changes beyond the economic indicators. Ivan and Turok (2004) believe that competitiveness can be measured by the economic growth in foreign investment, the existing local businesses, and the start-ups [26]. The single index such as the gross added value of per capita GDP and per worker GDP is most commonly used to rank the overall economic growth and productivity. However, more complex measurements such as fairness, distribution efficiency, and indicators of sustainable economic growth should be included, because local governments usually adopt competition strategies that only hold on to labor costs and direct taxing that sacrifices the environment and cuts welfare [27, 28]. Bruneckiene et al. (2010) believe that the current elements of urban competitiveness are used as input to create future factor output, which then becomes an input in the cyclical process that forms new urban competitiveness. Of course, the whole process can also be run in reverse direction and is a continuous and circular cycle in improving urban competitiveness [29]. The strategic decisions are based on the latest measurement results of a city's competitiveness and potential. In other words, there is a correlation between the urban competitiveness in the economic sense and the urban competitiveness in the noneconomic sense. To coordinate and promote development and to promote coordination through development are a collective expression of the multidimensional and multifaceted urban competitiveness [30–33].

Ni (2006, 2017) defines the urban competitiveness as a city's capacity to attract, compete for, own, control, and transform resources, and to compete for, to seize, and to control markets, to create value so as to benefit its people, in comparison to other cities in the course of competition and development [34, 35]. Based on the model to measure a country's competitiveness developed by Porter et al. [36, 37], Ni et al. (2015) establish a model to study the factors of urban competitiveness [38]. Their model includes six latent variables: $UC = f$ (EQ, LE, LD, LC, GC, SE, HE). Among them, UC represents the input of urban competitiveness. EQ refers to the quality of the business entity; LE refers to the main supply of local factors; LD is the demand of the local market; LC reflects the internal networking and aggregation of the city; GC represents the city's external connection via external factors and markets; SE and HE, on the other hand, represent the institutional rules, environment and the local infrastructure, and environments that interact with the subject. In addition, the study of Ni et al. also uses quantitative methods to measure the competitiveness index of 500 cities around the world and divide these cities into 7 distinct categories by applying dynamic clustering analysis; with stepwise regression analysis, their study finds that different types of cities have different determinants of competitiveness: technological innovation, global connections, and international brands are the determinants to become the world's top cities; the business and living environment and the size of the population are the determinants for the fast growing emerging central cities; for the less developed cities, the determinants are infrastructure, wages, and living standards. Therefore, different strategies and measures should be adopted to enhance urban competitiveness.

## 2. Data and Research Methods

*2.1. Indicator System of Urban Competitiveness.* **A** sample of 1007 cities with a population of 500,000 or more across the world are selected based on the "2015 World Urbanization Prospects" issued by the UN Department of Economic and Social Affairs. These sample cities are distributed in 136 countries or regions in 6 continents. Among them, there are 566, 126, 131, 102, 75, and 7 in Asia, Europe, North America, Africa, South America, and Oceania, respectively. The reason for the selection of these cities is that they are very representative in terms of city sizes, the level of development, spatial distribution, etc.

Urban competitiveness is fundamentally a regional competitiveness. According to the theoretical framework, in terms of output, performance, and interpretation, urban competitiveness is a city within its spatial scope to create value and obtain the scale, level, and growth of economic rent. According to the principle of minimization of indexes, economic density (GDP per land area) is the proper indicator of the efficiency and level of value creation, while economic growth (the difference between GDP of the current year and GDP of the previous year) is an indicator of the scale and growth rate of value creation. These two indicators can be combined as an index to better reflect the explanatory variables of economic competitiveness and to express the overall long-term growth and the overall economic efficiency of the urban economy. To measure by using five consecutive years of average GDP growth and GDP density, we establish the model for urban economic competitiveness as follows:

$$\mathrm{GUCI}_i = \left(\frac{(\mathrm{GDP}_i/\mathrm{IDEA}_i) - (\overline{\mathrm{GDP}_i}/\mathrm{IDEA}_i)}{\partial^2}\right) \times \left(\frac{(\mathrm{GDP}_i/\mathrm{POP}_i) - (\overline{\mathrm{GDP}_i}/\mathrm{POP}_i)}{\partial^2}\right) + \left(\frac{\left(\mathrm{GDP}_i^t/\mathrm{GDP}_i^{t-1}\right) - \left(\overline{\mathrm{GDP}_i^t}/\mathrm{GDP}_i^{t-1}\right)}{\theta^2}\right), \quad (1)$$

where $\mathrm{GUCI}_i$ represents the index of the global urban competitiveness of city $i$.

The left side of the plus sign in formula (1) is the standardized economic density index of city $i$ after being adjusted with per capita GDP.

$((\mathrm{GDP}_i^t/\mathrm{GDP}_\mathrm{i}^{t-1}) - (\overline{\mathrm{GDP}_i^t}/\mathrm{GDP}_\mathrm{i}^{t-1})/\theta^2)$ is a standardized index of economic growth of the city $i$, which uses the average growth of 5 executive years of GDP to measure the comprehensive long-term growth. According to the theoretical framework, the relationship between competitive input and process and competitiveness should be the relationship between the explanatory variables and the explained variables interpretation. By making reference to the macroeconomic cycle theory and Michael Porter's competitive advantage theory, the model of the factors affecting competitiveness is as follows:

$$\mathrm{EEC} = F(\mathrm{FE}, \mathrm{TI}, \mathrm{IS}, \mathrm{HR}, \mathrm{LD}, \mathrm{CC}, \mathrm{SE}, \mathrm{IN}, \mathrm{LE}). \quad (2)$$

The input of urban competitiveness refers to the factors and environments of the city, including financial services, technological innovation, industrial system, human capital, local demand, business costs, institutional costs, infrastructure, living cost, and global connections. The output of all factors ultimately manifests itself as the city's competitiveness. Based on the above theoretical analysis, this paper constructs the following explanatory model of economic competitiveness:

Economic Competitiveness (EEC) is the explained variable; it is also called explicit variable. All variables in parentheses on the right side of formula (2) are explanatory variables, including Industrial System (IS), Technology Innovation (TI), Financial Services (FE), Human Resources (HR), Local Demand (LD), Corporate Cost (CC), System Environment (SE), Infrastructure (IN), and Living Environment (LE). All the data comes from the database of the Chinese Academy of Social Sciences City and Competitiveness Index. It should be pointed out that these indicators are compound indicators comprised of a number of specific indicators. See Table 1 for the specific indicator system.

*2.2. Measurements.* This section mainly introduces the degree of coupling and coordination among multiple variables.

Multivariable Coupling Coordination is as follows:

$$\mathrm{DC}_m = \left\{\frac{\prod_{i=1}^{n} g_i(x)}{\left[\prod_{i=1,i<j}^{n}\left(\left(g_i(x) + g_j(x)\right)/2\right)\right]^{2/n}}\right\}^{1/n}, \quad (3)$$

$$T = \sum_{i=1}^{n} \lambda_i g_i(x), \quad (4)$$

$$\mathrm{DCCI} = \sqrt{\mathrm{DC}_m * T}. \quad (5)$$

In formulas (3) to (5), where $g_i(x)$ represents the value of the $i$ variable of a city, $\mathrm{DC_m}$ is the degree of coordination between $n$ variables in a city, $\lambda_i$ is the weight given to the $i$ variable in the evaluation system, and DCCI is the degree of coupling and coordination between $n$ variables in a city; that is to say, it is the DCC between $n$ variables.

Coupling degree only reflects the degree of coordination between systems or elements, It focuses on emphasizing coordination, failure to reflect the interactive development of each other. Sometimes, there could be pseudocoordination, that is, coordination under a low level of development. Coupling degree and DCC are two different concepts. The degree of coupling and coordination emphasizes overall coordinated development, learning from each other's strengths, and avoiding shortcomings. High coupling degree between variables does not necessarily imply strong coupling coordination, because the coupling might occur with low coordination; however, the high DCC indicates higher coupling degree between systems, because it gives the weighted coefficient between the systems. The levels and categories of DCC are shown in Table 2.

How extensive and representative the samples are accounts for the accuracy and value of the study. This paper selects a sample of cities with a population of more than 500,000 across the world from the "2015 World Urbanization Prospects" issued by the UN Department of Economic and Social Affairs. Combined with China's urban context, a total of 1007 sample cities are selected from the global community. In terms of the spatial distribution, these 1,007 sample cities are distributed in 6 continents and 136 countries or regions; among them, 566 are in Asia, 126 in Europe, 131 in North America, 102 in Africa, 75 in South America, and 7 Oceania cities. They basically represent the different urban development in different regions around the world today. It should be noted that this article uses the metropolitan area (MSA) caliber. The data are drawn from the City and Competitiveness Index database of the Chinese Academy of Social Sciences.

Table 1: Explanatory indicators of the economic competitiveness of cities across the globe.

| Categories of indicators | Indicators | Categories of indicators | Indicators |
|---|---|---|---|
| Financial services (FE) | Bank index<br>Bank branches index<br>Exchange index | Operation cost (CC) | Loan interest rate<br>Tax to GDP ratio<br>Per capita income/benchmark price |
| Technology innovation (TI) | Patent index<br>Research paper index | System environment (SE) | Convenience of doing business<br>Level of economic freedom |
| Industry system (IS) | Production services business index<br>Technology enterprise index | Infrastructure (IN) | Convenience of transportation<br>Broadband users |
| Human resources (HR) | Workforce population (15–59)<br>The proportion of youth population<br>College index | Living environment (LE) | The logarithm of the number of airlines and airport distance<br>PM 2.5<br>Crime rate |
| Local demand (LD) | Total disposable income | | |

Table 2: The levels and categories of DCC.

| | Levels and categories of DCC | | | |
|---|---|---|---|---|
| Coordination degrees | 0–0.5 | 0.5–0.6 | 0.6–0.8 | 0.8–1 |
| Coordination categories | Scant coordination | Limited coordination | Good coordination | Strong coordination |
| Coordination levels | 4 | 3 | 2 | 1 |

## 3. Empirical Analysis

The empirical study first examines the DCC of cities around the world and then presents its spatial distribution and hierarchical structure. In the end, regression analysis is used to verify the impact of coupling coordination on urban competitiveness.

*3.1. The Coupling Coordination of Cities Worldwide.* Figure 1 is a kernel density graph of the DCC of cities across the world. It shows the distribution characteristics. Compared to a standard normal distribution, this pattern is a bit to the left. Such a distribution suggests that the DCC differ greatly with only a few cities reaching a high level, while many others have lower DCC and need improvement. Specifically, the global average DCC is 0.481, the median 0.515, the variance 0.0395, and the coefficient of variation 0.391; 37.94% of the cities are below average; those above average concentrate in a few developed countries. Currently, New York City of the United States has the highest DCC. Among the top 20 cities, seven are in the United States, three in China, two in Japan, and Australia, Germany, Russia, France, South Korea, Canada, Singapore, and the United Kingdom each have one city among them, indicating a notable imbalance in the distribution.

Figure 2 is a scatter plot of the DCC and economic competitiveness of cities around the world. It shows the correlation between the two. Cities ranked in the middle range by competitiveness have DCC between 0.4 and 0.6; the correlation coefficient between these cities' economic competitiveness and the DCC is 0.69, indicating very strong correlation between the two. There is a strong correlation between the degree of urban economic competitiveness and the DCC; i.e., cities with higher economic competitiveness rankings show higher DCC between the overall explanatory indicators and economic competitiveness, while cities ranked lower by economic competitiveness have lower DCC between overall explanatory indicators and economic competitiveness.

*3.2. Spatial Distribution of the DCC of Cities across the World.* In order to facilitate the analysis of the spatial distribution of the DCC of cities across the world, we categorize the 1,007 sample cities into four tiers: tier 1 includes cities of excellent coordination with coupling degrees between 0.8 and 1; tier 2 includes cities of good coordination with coupling degrees between 0.6 and 0.8; tier 3 covers cities of slight coordination with coupling degrees between 0.5 and 0.6; and tier 4 covers cities of coupling coordination in barely balanced condition with coupling degree at 0.5 or under. We use bubble chart to visualize the coupling coordination of the 1007 sample cities in this paper. Figure 3 shows the distribution of the coupling coordination of these cities.

The spatial distribution of the DCC shows that cities with excellent or good DCC account for less than a quarter of the 1,007 sample cities and that cities with DCC representing bare balance or severe imbalance account for 2/7 and 3/7 of the 1,007 sample cities, respectively. By continent, nearly 92% of the first two tiers are cities in Europe, North America, and Asia, with the three continents, respectively, accounting for 27.24%, 31.34%, and 32.84% of the total, while up to 86.33% of the latter two tiers are cities in Asia, Africa, and South America, of which cities in Central Asia, Africa, and South America account for about three-quarters of the total. Asian cities see larger gaps in their DCC. Developed cities in Europe, North America, and Asia have higher DCC,

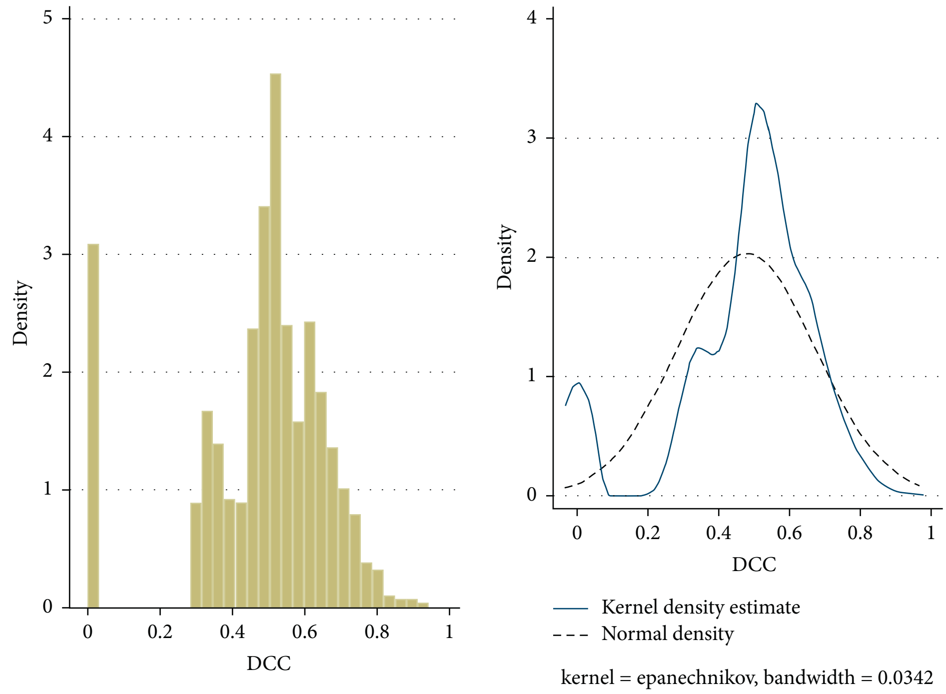



Figure 1: Histogram and kernel density distribution of the DCC.

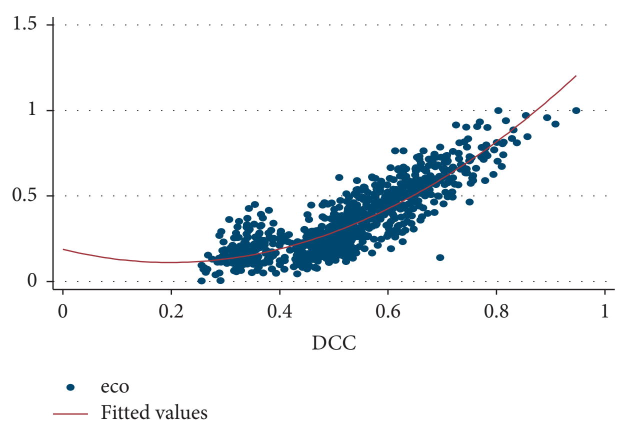



Figure 2: Scatter plot of the DCC of cities worldwide.

indicating better-coordinated development of various urban facets. Less developed cities in Asia and Africa, in contrast, have lower DCC. Among the 1,007 sample cities, the DCC of cities in BRICS countries such as China, Russia, and India generally have half of their cities in the latter two tiers, showing bare balance or severe imbalance of development. On the contrary, about half of the cities in G7 countries, such as the United Kingdom, France, and the United States, are in the first two tiers. So, overall, the DCC of cities in G7 countries are generally higher than those of cities of BRICS countries.

*3.3. Hierarchical Structure by DCC.* In order to conduct a more comprehensive analysis of DCC of cities across the world, this study ranks the city by its economic competitiveness and calculates the average, variance, coefficient of variation, and other statistical data of the cities' coupled coordination within each tier of the ranking. Table 3 shows that the cities ranked the top 20, the top 100, and the 200th have overall better coupling coordination than the cities ranked between the 200the and the 500th cities. The cities ranked below the 500th have imbalance of development. Table 4 shows that the DCC of these sample cities' make up a pyramid-like structure. First, the cities with the highest degrees of coordinated development make up the smallest group, with only 17 cities. Among them, except Sydney, which is an Oceania metropolis, others are the global financial and technology innovation hubs in Europe, America, and Asia. Second, 251 cities are ranked as cities with good coordination. These cities are mostly the financial and technological innovation centers within a region. Third,

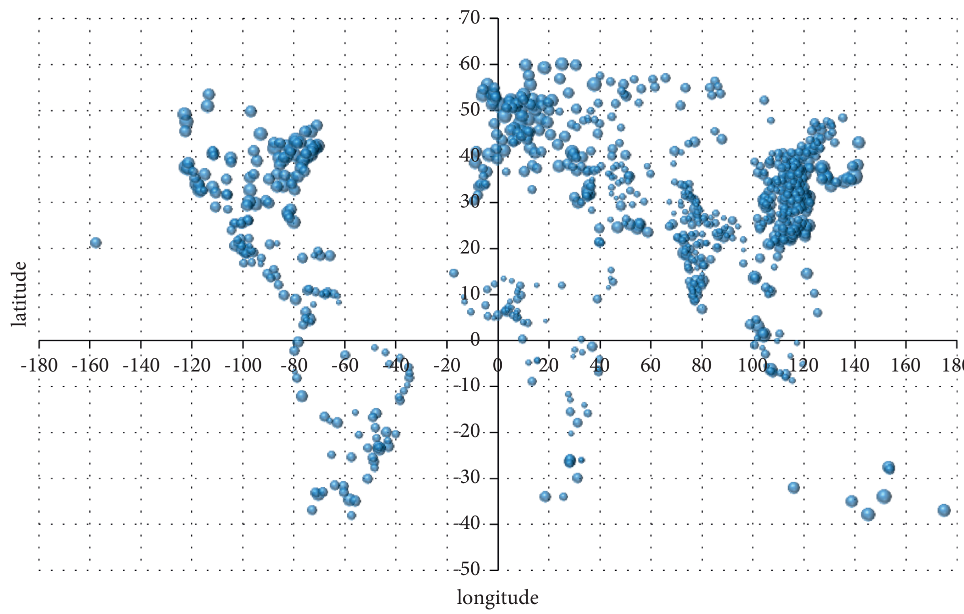



Figure 3: Spatial distribution of DCC of the cities around the world.

Table 3: Data of DCC of cities worldwide.

| DCC | Top 20 | Top 100 | 101st–200th | 201st–300th | 301st–500th | 501st–800th | 800th–1007th | All |
|---|---|---|---|---|---|---|---|---|
| Means | 0.810 | 0.727 | 0.639 | 0.586 | 0.499 | 0.400 | 0.317 | 0.477 |
| Median | 0.808 | 0.723 | 0.641 | 0.601 | 0.530 | 0.473 | 0.354 | 0.513 |
| Categories | Strong coordination | Good coordination | Good coordination | Limited coordination | Scanty coordination | Scanty coordination | Scanty coordination | Scanty coordination |
| Levels | Level I | Level II | Level II | Level III | Level IV | Level IV | Level IV | Level IV |
| Variance | 0.004 | 0.005 | 0.003 | 0.008 | 0.019 | 0.032 | 0.032 | 0.038 |
| Standard deviation | 0.062 | 0.069 | 0.056 | 0.089 | 0.139 | 0.178 | 0.178 | 0.196 |
| Coefficient of variation | 0.076 | 0.095 | 0.088 | 0.153 | 0.279 | 0.445 | 0.563 | 0.411 |

Table 4: Hierarchy of DCC of cities worldwide.

| Levels DCC | Categories DCC | Number of the cities | Means | Top 10 cities |
|---|---|---|---|---|
| Level I | Strong coordination | 16 | 0.839 | New York, tokyo, London, Seoul, Singapore, Beijing, Hong Kong, Shanghai, San Francisco, Chicago, etc. |
| Level II | Good coordination | 248 | 0.667 | Osaka, Washington, DC, Atlanta, Houston, Frankfurt, Seattle, Istanbul, Madrid, Zurich, Philadelphia, etc. |
| Level III | Limited coordination | 301 | 0.538 | Gebze, Leon, Allen, Hermosillo, Palermo, San Juan, Krakow, Hsinchu, Liege Albuquerque |
| Level IV | Scanty coordination | 442 | 0.317 | Karaj, Hengshui, Salem, Hebi, Voronezh Benxi, Palembang, Jingdezhen, Fuyang, Zigong Wait |

306 cities are ranked as cities with limited coordination. These cities have a certain level of coupling coordination and are mostly cities with average levels of economic development of a country, especially those new-economy countries. Fourth, 433 cities are ranked as cities with scanty/insufficient coordination. Most of them are located in Africa, South Asia, and South and East Europe. Table 3 shows the top 10 cities in each ranking category, and most of the cities with remarkable coordination are distributed in China and the United States. The United States has the most cities among the top 10 in good coordination ranking, while China has the most cities among the top 10 in the limited coordination ranking. The United States has more cities with stronger coupling coordination; China has fewer cities with strong coupling coordination such as Hong Kong, Shanghai, and Shenzhen. Chinese cities are mostly distributed in the middle ranges of the ranking and need improvement in coordinated development.

*3.4. Analysis of the Relative Index of Coupling Coordination.* The ratio of the explanatory indicator to the explained indicator is called the relative development index. That is to say, the ratio of each of the nine explanatory indicators, such as financial services, scientific and technological innovation, industrial systems, and human capital, to economic competitiveness index is relative development index. The DCC reflects the degree of coupling and coordinated development between the explanatory indicators of urban economic competitiveness, but it cannot reflect the relative development degree of each explanatory indicator and urban economic competitiveness. To this end, Table 5 and Figure 4 both analyze the relative development of the nine explanatory indicators and economic competitiveness of these cities by continents. Generally speaking, the relative index range between 0.8 and 1.2 indicates a balanced development. Below 0.8 or above 1.2 suggests a lagging development or advanced development.

Table 5 shows that the 1007 sample cities' relative indexes of operation cost, system environment, infrastructure, and living environment are generally higher than 2, indicating more development than urban economic competitiveness; these cities' relative indexes of financial services, technological innovation, and industrial system are generally low, indicating less developed than urban economic competitiveness; their relative index of human resource is around 1.2, suggesting coordinated development with the economic competitiveness. Specifically, the relative indexes of Africa are much higher than the relative development levels of other continents. In Africa, except in the areas of financial services, technological innovation, and the industrial system, the relative development indexes of the human resources, local demand, operation environment, system environment, infrastructure, and living environment are all between 2.7 and 6, way more advanced than the development level of economic competitiveness. In addition, the study found the following: first, in the relatively lower reference indexes of financial services, technological innovation, and industrial systems, sub-Africa and South America see much higher relative development level of financial services than Europe, North America, and Oceania, indicating that the relative development level of financial services in these continents is relatively lagging; Europe, North America, and Oceania have higher level of development in technological innovation than Asia, Africa, and South America; Europe, Oceania, and Africa have higher relative development index of industrial systems than North America, South America, and Asia; it shows that the relative development level of technological innovation in these continents is lagging but higher the world's average. Second, in the mid-range relative development levels of human capital and local demand, Europe, North America, and Oceania see lower indexes than Asia, Africa, and South America, and slightly lower than the world's average, lagging behind the development of economic competitiveness. Third, in the relative higher development levels of operation costs, system environment, infrastructure, and living environment, Asia and Africa have much higher relative indexes than Europe, America, and Oceania. The above conclusions show that the relative development of the nine explanatory indicators of urban economic competitiveness in different continents is different.

*3.5. Regression Analyses of the Factors Affecting Coupling Coordination of the Cities around the World.* In order to verify that the city's DCC is a key factor in urban competitiveness, Tables 6 and 7 are the benchmark regression analysis and factorial regression analysis on the DCC and economic competitiveness.

In Table 6, regression (1) represents the regression results of the economic competitiveness index alone and the DCC; regressions (2)-(5) represent the regression results of the economic competitiveness index and the DCC with additional control variables. Based on the benchmark regression analysis, we found that as the explanatory variables increase, the DCC is consistent with the significance level of other explanatory variables and economic competitiveness. This shows that the regression results are robust. In the regression analysis of (1)-(5), both the economic competitiveness index and the DCC are positively correlated at a 1% level, indicating a significant positive correlation between the two. The regression coefficient is 0.108. Table 6 shows that, in addition to the living environment, other indicators such as coupling coordination, financial services, technological innovation, industrial system, human capital, business cost, and business environment are all significantly correlated to economic competitiveness. In addition, the coefficients of each explanatory variable in the regression analysis (5) show that the infrastructure has the strongest impact on economic competitiveness, with a coefficient as high as 0.334, followed by financial services, technological innovation, and coupling coordination with a coefficient of 0.305, 0.305, and 0.108, respectively. The least influential are operation cost, industrial system, business environment, and human capital with coefficients 0.252, 0.0945, 0.0895, and 0.0703, respectively. In general, the DCC, financial services, technological innovation, and economic competitiveness are relatively robust and significant variables.

In order to further study the impacts of coupling coordination and other explanatory variables on the economic competitiveness of cities, we conducted a quintile regression analysis as shown in Table 7. We selected five representative quintiles of 0.1, 0.25, 0.5, 0.75, and 0.9, on which the indication and significance levels of the influencing factors are basically the same. However, the coefficient size and significance level display different patterns of change. First, the contribution from coupling coordination and technological innovation to the urban competitiveness is more concentrated near the middle quintiles, indicating that the two factors contribute more to cities with mid-to-upper levels of economic competitiveness. Second, the contribution from financial services to economic competitiveness gradually expands with the increase of urban economic competitiveness, and its significance also gradually increases. Third, the contribution from the industrial system is significant at the 0.1 quintile and 0.75 quintile, but not at other points, indicating that the factor contributes more to the cities with

Table 5: Relative Indexes of coupling coordination of the cities worldwide.

| Means | Relative index | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| | FE | TI | IS | HR | LD | CC | SE | In | LE |
| North America | 0.457 | 0.638 | 0.18 | 0.726 | 1.285 | 1.628 | 1.795 | 1.218 | 1.381 |
| Oceania | 0.432 | 0.704 | 0.387 | 0.67 | 1.111 | 1.125 | 1.521 | 0.953 | 1.194 |
| Africa | 0.906 | 0.34 | 0.322 | 2.709 | 2.841 | 5.002 | 4.678 | 3.582 | 6.408 |
| South America | 0.661 | 0.405 | 0.161 | 1.051 | 1.57 | 1.541 | 1.977 | 1.442 | 1.785 |
| Europe | 0.53 | 0.738 | 0.252 | 0.764 | 1.448 | 1.71 | 2.2 | 1.632 | 2.062 |
| Asia | 0.606 | 0.427 | 0.108 | 1.276 | 1.445 | 2.353 | 2.574 | 2.104 | 2.679 |
| Worldwide | 0.61 | 0.485 | 0.163 | 1.263 | 1.572 | 2.375 | 2.585 | 2.02 | 2.73 |

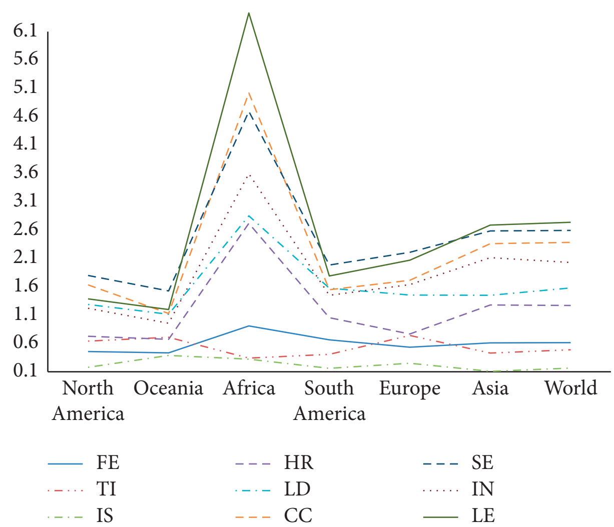



Figure 4: Relative indexes of DCC of the cities across the world.

Table 6: Benchmark regression analysis on the DCC and economic competitiveness.

| | Regression (1) | Regression (2) | Regression (3) | Regression (4) | Regression (5) |
|---|---|---|---|---|---|
| Variables | Economic competitiveness index | Economic competitiveness index | Economic competitiveness index | Economic competitiveness index | Economic competitiveness index |
| DCC | 0.684*** (30.01) | 0.177*** (7.35) | 0.165*** (7.59) | 0.115*** (5.38) | 0.108*** (4.95) |
| FE | | 0.508*** (8.88) | 0.430*** (6.03) | 0.295*** (4.26) | 0.305*** (4.32) |
| TI | | 0.550*** (21.46) | 0.381*** (14.78) | 0.303*** (11.78) | 0.305*** (11.71) |
| IS | | | 0.138*** (3.33) | 0.0975** (2.45) | 0.0945* (2.36) |
| HR | | | 0.0837*** (2.79) | 0.0706** (2.44) | 0.0703* (2.40) |
| CC | | | 0.320*** (15.20) | 0.252*** (11.68) | 0.252*** (11.61) |
| SE | | | | 0.0874*** (3.15) | 0.0895** (3.15) |
| IN | | | | 0.331*** (9.67) | 0.334*** (9.48) |
| LE | | | | | −0.00318 (−0.14) |
| Constant term | 0.0190* (1.65) | 0.0661*** (6.75) | −0.0963*** (−6.02) | −0.208*** (−10.81) | −0.206*** (−9.86) |
| Sample size | 1007 | 1007 | 1007 | 1007 | 1007 |

*Note.* Brackets indicate t-statistics, *indicates a significance level of 10%, **indicates a significance level of 5%, and ***indicates a significance level of 1%.

Table 7: Quantile regression analysis on economic competitiveness and coupling coordination.

| | OIS eco | 0.1 Eco | 0.25 eco | 0.5 eco | 0.75 Eco | 0.9 eco |
|---|---|---|---|---|---|---|
| DCC | 0.115*** (5.38) | 0.0758*** (2.90) | 0.0830*** (2.86) | 0.0798*** (2.71) | 0.146*** (5.01) | 0.112*** (2.67) |
| FE | 0.294*** (4.15) | 0.137 (1.58) | 0.182* (1.88) | 0.356*** (3.64) | 0.328*** (3.39) | 0.650*** (4.68) |
| TI | 0.304*** (11.70) | 0.197*** (6.16) | 0.302*** (8.52) | 0.350*** (9.73) | 0.374*** (10.51) | 0.327*** (6.42) |
| IS | 0.0977** (2.45) | 0.144*** (2.93) | 0.0712 (1.31) | −0.00391 (−0.07) | 0.167*** (3.05) | 0.0990 (1.26) |
| Human resources | 0.0704** (2.41) | 0.0952*** (2.65) | 0.0925** (2.32) | 0.0926** (2.29) | 0.0281 (0.70) | 0.00370 (0.06) |
| Operation cost | 0.252*** (11.63) | 0.225*** (8.46) | 0.240*** (8.12) | 0.247*** (8.26) | 0.234*** (7.89) | 0.262*** (6.17) |
| Business environment | 0.0878*** (3.09) | 0.174*** (4.97) | 0.131*** (3.38) | 0.0576 (1.47) | 0.0446 (1.15) | 0.0350 (0.63) |
| Infrastructure | 0.332*** (9.45) | 0.485*** (11.25) | 0.452*** (9.44) | 0.346*** (7.13) | 0.234*** (4.86) | 0.310*** (4.50) |
| Living environment | −0.00171 (−0.07) | −0.00702 (−0.25) | −0.00874 (−0.28) | 0.0362 (1.14) | −0.00542 (−0.17) | −0.0863* (−1.92) |
| _cons | −0.208*** (−9.95) | −0.371*** (−14.46) | −0.318*** (−11.16) | −0.226*** (−7.82) | −0.0947*** (−3.31) | −0.0504 (−1.23) |
| N | 1007 | 1007 | 1007 | 1007 | 1007 | 1007 |
| adj. $R^2$ | 0.790 | | | | | |

higher or lower economic competitiveness than to the cities with median competitiveness. Fourth, the contribution from operating cost to the cities' competitiveness is more even and with little difference at different quintiles. Fifth, the contribution from the business environment is more significant at the lower quintiles than at the mid-to-high quintiles, indicating that the factor has a significantly positive contribution to cities with lower economic competitiveness but a less significant contribution to the cities with higher economic competitiveness. From the above results, we can see that the technological innovation and financial services are the key to enhance urban economic competitiveness.

## 4. Conclusion and Discussions

Urban competitiveness is a huge project; it involves many discipline knowledge, including economy, geography, and urban planning especially, the coordinated development of global urban competitiveness is more systematic and complex, and it is changing dynamically. This article inevitably has shortcomings. The full text establishes a scientific evaluation index system by determining the research logic and technical route. The index system of urban economic competitiveness evaluation is constructed from the perspective of input and output. Generally speaking, there is a certain proportion of the allocation of various elements. Although only one element input will increase the output, with the increase of the input of this element, when the quantity reaches a certain value, the marginal output it brings is in a reduced state, which will cause a waste of resources. Since the output results brought by the input of various elements are also different, unreasonable element allocation will lead to low production efficiency [39, 40], which requires from us to determine different resource element allocation mode according to the productivity of different elements, to maximize realize the coordinated allocation of resources. Therefore, this article also introduces the degree of coupling and coordination (DDC) among multiple explanatory variables. We use empirical analysis to evaluate and analyze the influencing factors of urban economic competitiveness, DCC, and global urban economic competitiveness in the world.

Based on the 1,007 sample cities with a population over 500,000 across the globe, the study got the following conclusion: the global urban economic competitiveness and the degree of coupling and coordination (DCC) are unevenly distributed; the city with higher economic competitiveness rankings has a higher DCC; on the contrary, the opposite is true; the cities with higher DCC are mainly those global cities or metropolis known for financial and technological innovations, such as New York, Tokyo, London, Beijing, and Hong Kong, while cities with lower coupling coordination are mainly in underdeveloped countries in Asia, Africa, and Latin America. This is mainly because those global cities or metropolis have a good social and economic foundation and are in a leading position in financial services, technological innovation, infrastructure construction, and other aspects.

The DCC is pivotal to the cities' economic competitiveness. Improving the coupling and coordination of cities is of great significance to promoting the common prosperity of global cities. In the future, global cities should regard improving the DCC as an important strategic policy. For cities in different regions and cities at different stages of development, differentiated policies should be formulated based on relevant factors that affect the improvement of the cities' economic competitiveness.. Overall, they could adopt the principle of "strengthening the strengths to compensate for the weaknesses, laying the foundation and grasping the key" to enhance the competitiveness of global cities. While consolidating existing advantages, they should focus on improving financial service levels, improving technological innovation capabilities, strengthening human resource investment, upgrading urban infrastructure levels, and optimizing the living environment, so that all the people in the world can live a happy life.

In addition, this study uses cross-sectional data, and the evaluation is only based on the current state, while the competitiveness of cities is dynamically changing. In the future, panel data will be used to evaluate the dynamic changes in the competitiveness of world cities. The process is further research work.

## Data Availability

National Academy of Economic Strategy, CASS, and UN Habitat jointly released a "Global Urban Competitiveness Report (2017-2018)." All the data in this article comes from the data of the report 2017. Ownership of the data does not belong to us. Because the data is obtained for a fee. So, it cannot be provided.

## Conflicts of Interest

The authors declare that they have no conflicts of interest.

## References

[1] H. Haken, "Circuits & devices magazine," *IEEE*, vol. 28, no. 9, pp. 412–414, 1997.

[2] L. V. Bertalanffy, *Basics, Development and Application of General System Theory*, Tsinghua University Press, Beijing, CHN, 1987.

[3] G. Bi, Q. Yang, and S. Liu, "Coupling and coordinative development of China's provincial ecological civilization construction and urbanization," *Economic Geography*, vol. 37, no. 01, pp. 50–58, 2017.

[4] C. liao, "The quantitative evaluation and classification system of the coordinated development of environment and economy: taking the urban agglomeration of the pearl river delta as an example," *Tropical Geography*, vol. 2, pp. 76–82, 1999.

[5] D. Wu, L. Cao, and L. Chen, *Principles and Application of Synergetics*, Huazhong University of Science and Technology Press, Wuhan, CHN, 1990.

[6] X. Liu and P. Ni, "Analysis of coupling and coordination degree of competitiveness of Chinese large and medium cities," *Journal of Hebei University of Economics and Business*, vol. 40, no. 2, pp. 57–64, 2019.

[7] M. Storper, "The city: centre of economic reflexivity," *Service Industries Journal*, vol. 17, no. 1, pp. 1–27, 1997.

[8] J. Shen, "Urban competitiveness and urban governance in the globalizing world," *Asian Geographer*, vol. 23, no. 1-2, pp. 19–36, 2004.

[9] R. Martin and J. Simmie, "Path dependence and local innovation systems in city-regions," *Innovation*, vol. 10, no. 2-3, pp. 183–196, 2008.

[10] E. J. Malecki, "Hard and soft networks for urban competitiveness," *Urban Studies*, vol. 39, no. 5, pp. 929–945, 2002.

[11] M. Boddy, "Geographical economics and urban competitiveness: a critique," *Urban Studies*, vol. 36, no. 5-6, pp. 811–842, 1999.

[12] W. F. Lever, "Competition within the European urban system," *Urban Studies*, vol. 30, no. 6, pp. 935–948, 1993.

[13] W. F. Lever, "Competitive cities in Europe," *Urban Studies*, vol. 36, no. 5, pp. 1029–1044, 1999.

[14] W. F. Lever, "The knowledge base and the competitive city," in *Urban Competitiveness: Policies for Dynamic Cities*, I. Begg, Ed., Bristol University Press, pp. 11–32, Bristol, UK, 2002.

[15] J. Haughton and V. Slobodyanyuk, *State Competitiveness Report 2001*, Beacon Hill Institute, Boston, 2001, http://www.beaconhill.org/CompetitivenessHomePage.html.

[16] D. G. Tuerck, J. Haughton, F. Conte et al., *State Competitiveness Report 2007*, Beacon Hill Institute, Boston, 2007, http://www.beaconhill.org/CompetitivenessHomePage. html.

[17] M. Kitson, R. Martin, and P. Tyler, "Regional competitiveness: an elusive yet key concept?" *Regional Studies*, vol. 38, no. 9, pp. 991–999, 2004.

[18] P. K. Kresl, *Planning Cities for the Future: The Successes and Failures of Urban Economic Strategies in Europe*, Edward Elgar Publishing, UK, 2007.

[19] B. Gardiner, R. Martin, and P. Tyler, "Head," *Regional Studies*, vol. 38, no. 9, pp. 1045–1067, 2004.

[20] S. Iyer, M. Kitson, and B. Toh, "Social capital, economic growth and regional development," *Regional Studies*, vol. 39, no. 8, pp. 1015–1040, 2005.

[21] R. Camagni, R. Capello, and P. Nijkamp, "Towards sustainable city policy: an economy-environment technology nexus: an economy-environment technology nexus," *Ecological Economics*, vol. 24, no. 1, pp. 103–118, 1998.

[22] J. Jacobs, *The Economy of Cities*, Vintage, New York, USA, 1969.

[23] I. Begg, "Cities and competitiveness," *Urban Studies*, vol. 36, no. 5-6, pp. 795–809, 1999.

[24] S. Hao and P. Ni, "Research on the competitiveness of Chinese cities-taking several cities as cases," *Economic Science*, vol. 20, no. 3, pp. 50–56, 1998.

[25] P. Ni, "Analysis paradigm and conceptual framework of Chinese urban competitiveness," *Economic Dynamics*, vol. 6, pp. 14–18, 2001.

[26] I. Turok, "Cities, regions, and competitiveness," *Regional Studies*, vol. 38, no. 9, pp. 1069–1083, 2004.

[27] W. J. Steinle, "Regional competitiveness and the single market," *Regional Studies*, vol. 26, no. 4, pp. 307–318, 1992.

[28] P. J. Taylor, P. Ni, B. Derudder, M. Hoyler, J. Huang, and F. Witlox, *Catalano G. Global Urban Analysis: A Survey of Cities in Globalization*, Routledge, London, UK, 2012.

[29] J. Bruneckiene, A. Guzavicius, and R. Cincikaite, "Measurement of urban competitiveness in Lithuania[J]," *Engineering Economics*, vol. 21, no. 5, pp. 493–508, 2010.

[30] R. J. Rogerson, "Quality of life and city competitiveness," *Urban Studies*, vol. 36, no. 5, pp. 969–985, 1999.

[31] E. D'Arcy and G. Keogh, "The property market and urban competitiveness: a review," *Urban Studies*, vol. 36, no. 5-6, pp. 917–928, 1999.

[32] M. M. Kritz and R. E. Bilsborrow, "Migration, urbanization, and development: new directions and issues," *Population and Development Review*, vol. 24, no. 3, pp. 639–642, 1998.

[33] B. Gardiner, R. Martin, P. Sunley, and P. Tyler, “Spatially unbalanced growth in the British economy,” *Journal of Economic Geography*, vol. 13, no. 6, pp. 889–928, 2013.
[34] P. Ni and P. K. Kresl, *Global City Competitiveness Report: Competitiveness Challenges the World City System*, Social Science Literature Press, Beijing, CHN, 2006.
[35] P. Ni, *China City Competitiveness Report NO. 17*, Social Science Academic Press, Beijing, CHN, 2017.
[36] M. Porter, “Competitive advantage of nations,” *Harvard Business Review*, vol. 2, 1990.
[37] T. Berger, “Concepts of national competitiveness,” *Journal of International Business and Economy*, vol. 9, no. 1, pp. 91–112, 2008.
[38] P. Ni, B. Zhao, and H. Wei, “Competitive index construction and factor Analysis: based on a sample of 500 global typical cities,” *Urban Development Studies*, vol. 20, no. 6, pp. 72–79, 2013.
[39] L. X. Yang, “A Re-examination of the causes of China’s overcapacity from the perspective of all factors: based on the decomposition research of resource allocation and factor production efficiency,” *Social Sciences in Guangdong*, vol. 187, no. 5, pp. 14–22, 2017.
[40] C. R. Sun, Q. Qin, and S. Wang, “Does centrality of cities promote the upgrading of urban export structures?: evidence from Chinese cities and urban groups,” *World Economy Studies*, vol. 323, no. 1, pp. 17–31, 2021.

Hindawi
Complexity
Volume 2021, Article ID 9351741, 15 pages
https://doi.org/10.1155/2021/9351741

*Research Article*

# The Core-Periphery Structure in the Yangtze River Delta: An Enterprise Linkage Perspective, 1978–2019

**Jieying Yang,[1] Jingxiang Zhang,[1] Jiayi Lu,[2,3] and Dongqi Sun[2]**

[1] *School of Architecture and Urban Planning, Nanjing University, Nanjing 210093, China*
[2] *Key Laboratory of Regional Sustainable Development Modeling, Institute of Geographic Sciences and Natural Resources Research, Chinese Academy of Science, Beijing 100101, China*
[3] *College of Resources and Environment, University of Chinese Academy of Sciences, Beijing 100049, China*

Correspondence should be addressed to Dongqi Sun; sundq@igsnrr.ac.cn

Received 20 April 2021; Revised 12 July 2021; Accepted 29 October 2021; Published 28 November 2021

Academic Editor: Marcus Aguiar



Metropolitan areas are important for engaging in fierce global competition. Cities in metropolitan areas in China are generally characterized by a core-periphery structure. The Yangtze River Delta metropolitan area (YRD) is a national strategic region in which Shanghai, as a central city, drives the economic growth of hinterland cities. Exploring the spatiotemporal characteristics and influencing factors between the central city and its hinterland cities in the YRD can provide a basis for promoting regional development. Based on the headquarters-branches and enterprise investment data from 1978 to 2019, this study analyzes the spatiotemporal characteristics and influencing factors of enterprise linkages between Shanghai and its hinterland cities in the YRD. Our results reveal the following: (1) the headquarters-branch linkages between Shanghai and its hinterland cities manifest polarization characteristics, with different polarized characteristics among the three provinces; (2) the enterprise investment linkages between Shanghai and its hinterland cities are getting closer, but the key cities for investment in each province are different; (3) economic strength is a core factor that affects whether Shanghai establishes enterprise branches in its hinterland cities; and (4) the enterprise investment linkages between Shanghai and its hinterland cities depend on whether a city has a comparative advantage.

## 1. Introduction

With the rapid development of globalization and informatization, many cities have come to rely on infrastructure networks to form highly integrated metropolitan areas to address fierce global competition [1]. Metropolitan areas are basic geographical units that participate in global competition, and their development is of great significance to the sustained stability of the regional economy [2]. The links between cities are growing closer, city coordination has become a general method to promote economic development, and competitiveness has also become an important indicator of internationalization for measuring metropolitan areas [3]. As Scott points out, due to the increasing integration of cities in the world economy and emerging internal functional connections, these densely populated urbanized areas have become strategically crucial in the global economy [4].

Currently, many scholars point out that metropolitan areas have core-periphery structures, featuring a central city and hinterland cities, among which the central city plays the role of a “knowledge portal,” connecting the hinterland cities to the global economic network. The central city’s connection with the city inside the region is often closer than its connection with the city outside the region [5]. Diffusion from the central city to its hinterland cities is an intentional choice [6, 7]. Understanding how the central city affects its hinterland cities and discussing its evolutionary laws and influencing factors are important tasks for promoting both the integrated development of metropolitan areas and regional economic development and have become an important part of the reform of the national economic system [8].

The Yangtze River Delta metropolitan area (YRD) is the largest metropolitan area in China, located in the eastern coastal area, with an outstanding location, strong economic basis, and well-developed infrastructural constructions. In 2016, the central government proposed “the Development Plan for the Yangtze River Delta Metropolitan Area,” which clarified the central position of Shanghai in this area. In 2018, the State Council issued “Opinions on Establishing a New Mechanism for More Effective Regional Coordinated Development,” which pointed out that the central city plays a vital role in promoting the high-quality development of metropolitan areas. Currently, the construction of the YRD has become a national strategy. To implement policies formulated by the central government, it is necessary to conduct further research on the spatial organizational relationships that characterize the area.

To study urban networks, many scholars judge the relationship between central and hinterland cities through the comparison of urban population and land area. However, with the development of globalization and information technology, the logic of the relationship between cities has replaced the location theory. The relationship between cities no longer depends solely on their own size, but more on the strength of their relationship. The research on urban relevance based on “flow space” provides ideas for analyzing the relationship between central and hinterland cities. As enterprise linkage data have easy access and are directional, they can better analyze core-periphery problems in metropolitan areas. Therefore, urban relationship research based on enterprise linkages has become one of the important means to understand the internal network analysis of metropolitan areas.

Although the city network research based on enterprise linkages has many advantages, it is currently based more on headquarters-branch enterprise linkage data and mostly focuses on the comparison between global city networks and city clusters. Research on the core and periphery structure is relatively scarce, and there is also a lack of systematic investigation into how the central city affects the spatial evolution of the periphery city. Due to difficulties in data acquisition, current research focusing on the internal evolution of metropolitan areas is mostly based on short-term data analysis. However, it takes a long time for the central city to spill over to the peripheral areas. Based on short-term analysis, it is difficult to fully and deeply analyze the central city’s spread to peripheral cities.

Therefore, this study selected the YRD as an empirical area and analyzed its spatiotemporal characteristics and the factors influencing them from 1978 to 2019. We focused on the following three issues: (1) the characteristics of the temporal and spatial evolution of the enterprise linkages between Shanghai and its hinterland cities; (2) the factors influencing enterprise linkages between Shanghai and its hinterland cities; and (3) the possible core-peripheral mechanisms within the YRD.

## 2. Literature Review

Research on metropolitan areas is prevalent in the fields of urban geography and urban planning. Various theories have been advanced under its umbrella, such as the central place theory and the core-periphery theory. In the early stages of documentation, research often uses data such as city size, population, and GDP [9] to classify a city’s hierarchical scale system to determine relationships between cities. This type of analysis focuses on the characteristics of the scale of the city [10] and cannot accurately reflect the strength of relationships between cities. In 1996, Castells proposed the concept of the space of flow, emphasizing the constructivist nature of cities [11]. To a certain extent, this concept overturned the traditional logic of regional analysis based on scale attributes, focusing more acutely on the strength of the connections between cities [12–14]. Since then, the space of flow theory has gradually replaced the central place theory, forming a new urban system paradigm [15].

Due to the difficulty of obtaining relational data between cities, scholars are committed to discovering relational data to represent the linkages between cities for empirical research [11]. Airline passenger data [16], railway networks [17], bus networks [18], telecommunication flows [19], Bluetooth data [20], knowledge collaboration networks [21, 22], and enterprise linkages [23, 24] have all become important means of providing new perspectives for research on urban networks at home and abroad. However, there are many limitations to these studies. For example, it is difficult for traffic data to distinguish the purpose of tourists’ travel and the linkages between cities for the development of the tourism industry may be magnified as a result. Derudder et al. analyzed the infrastructure network in South Asia and found that a city located between two important transportation hubs may have an overestimated transportation connectivity. This may be the result of the layout of the transportation network rather than its real connectivity [25]. The measurement method of monitoring telecommunication flows and knowledge collaboration networks encounters difficulty in distinguishing the starting point from the ending point between cities [17]. This method focuses more on judging the hierarchical status of the city in the network structure [21].

International studies have shown that the essence of the city network is the economic network, and enterprises are the agents of the city network [26]. The enterprise linkage method is conceptually closest to the actual process of a city network [27]. The enterprise linkage data can solve the problem of direction attributes in the analysis of the core-peripheral structure. At the global level, some scholars have analyzed the global city network and its core-peripheral structure by using the headquarters-branch enterprise data of advanced producer services (APS) or multinational corporations. Taylor analyzed the distribution of 175 APS enterprises in 138 cities and explored the home region and outreach of the region on a global scale [26]. Derudder and Taylor used the data of APS enterprises to analyze 157 cities and to discern the core-periphery model [28]. At the regional scale, some scholars have compared and analyzed the differences between various metropolitan areas. Zilai and Tao compared the YRD and the Middle Reaches of the Yangtze River and analyzed the network connection patterns of the two metropolitan areas in the global and national systems [29]. Lu et al. compared the different effects of central cities on hinterland cities in the Beijing-Tianjin-

Hebei region and the YRD [30]. However, only a few scholars have focused on the structural characteristics of metropolitan areas. Salder takes the Greater Birmingham area as an example to explore the relationship between the enterprise network and the city network, and in doing so, strengthened our understanding of the core-periphery relationship [31]. Yeh et al. and Zhang et al. employed the connection analysis of APS companies to conclude that the Pearl River Delta has a hierarchical structure with Guangzhou and Shenzhen as the central cities and used social network analysis to describe the characteristics of the cities [32, 33].

In general, research at the global level regards the metropolitan area as a whole and lacks any discussion of internal urban network structures. Regional-level research focuses on the analysis of urban hierarchical structures and character descriptions, and related studies mostly use short-term headquarters-branch data, lacking the support of long-term continuous data and ignoring the important indicator presented by enterprise investment data. The incompetency of data is not conducive to systematically explaining the evolution of metropolitan areas.

Currently, studies have limited explanatory power when it comes to the analysis of long-term spatiotemporal characteristics and their influencing factors amid the core-periphery structure of the YRD. Therefore, based on the perspective of enterprise linkages, this study uses headquarters-branch and enterprise investment data from 1978 to 2019 in the YRD to analyze its spatiotemporal characteristics and utilizes pool OLS models to explore the influencing factors of the connection between Shanghai and its hinterland cities. A complete and continuous data foundation can provide support for long-term evolution analysis and promote our understanding of diffusion characteristics between central and hinterland cities in the region [34].

## 3. Materials and Methods

*3.1. Study Area.* The YRD is a highly urbanized metropolitan area and one of the largest financial hubs in China. Over the years, the YRD has undergone tremendous changes in scope. In 2010, the "Regional Plan for the Yangtze River Delta Region (2010–2015)" proposed that the YRD include 16 cities. In 2016, the "Development Plan for the Yangtze River Delta Metropolitan Area" stated that the YRD, in fact, includes 26 cities. In 2019, the "Outline of the Yangtze River Delta Regional Integrated Development Plan" updated that number to 41. These include one province-level city (Shanghai) and three provinces (Jiangsu, Zhejiang, and Anhui Provinces), comprising a total of 40 cities. Shanghai is the central city, and the remaining 40 are hinterland cities. These hinterland cities include 13 prefecture-level cities in Jiangsu Province: Nanjing, Wuxi, Xuzhou, Changzhou, Suzhou, Nantong, Lianyungang, Huai'an, Yancheng, Yangzhou, Zhenjiang, Taizhou, and Suqian; 11 prefecture-level cities in Zhejiang Province: Hangzhou, Ningbo, Wenzhou, Jiaxing, Huzhou, Shaoxing, Jinhua, Quzhou, Zhoushan, Taizhou, and Lishui; and 16 prefecture-level cities in Anhui Province: Hefei, Wuhu, Bengbu, Huainan, Maanshan, Huaibei, Tongling, Anqing, Huangshan, Chuzhou, Fuyang, Suqian, Lu'an, Bozhou, Chizhou, and Xuancheng. This study adopts these 41 cities as its research scope (Figure 1), and the period under study begins in 1978.

*3.2. Basic Idea and Data Sources.* According to the core-periphery theory, metropolitan areas include the central city and the hinterland cities. Headquarters-branch enterprises build organizational connections between the central city and the hinterland cities; if a headquarters' enterprise is in Shanghai and its branch enterprise is in a hinterland city, this is regarded as a headquarters-branch enterprise linkage. Enterprise investment refers to the behavior of enterprises that are located in the central city and invest in hinterland cities; if an enterprise located in Shanghai is the investor and the invested enterprise is located in a hinterland city, this will be regarded as an enterprise investment linkage (Figure 2).

The spatiotemporal characteristics of the regional core-periphery structure and its influencing factors are complex. To start to make sense of this complexity, we conducted the following processes: (1) analyzing the spatiotemporal characteristics of the enterprise linkages between the central city and the hinterland cities and (2) discussing the factors influencing the establishment of enterprise linkages between the central city and hinterland cities.

The data source for this research is the enterprise registration information of the State Administration for Industry and Commerce (http://www.gsxt.gov.cn). It contains data on the number of headquarters-branch enterprises and the amount of enterprise investment between cities, from 1978 to 2019. We organized the data with Shanghai as the enterprise headquarters/investor and the 40 hinterland cities as the enterprise branches/investees. This study analyzes the spatiotemporal evolution of the dynamics between Shanghai and other cities in the YRD through the enterprise linkages method to reflect Shanghai's connection with other cities. Based on regression analysis, we inferred the factors that influence Shanghai's establishment of enterprise linkages with other cities.

According to Wei, globalization, marketization, and other factors have influenced China's regional economic development patterns [35]. We divided the period from 1978 to 2019 into four stages. China began to reform and open up in 1978, and it was not until 1992 that the socialist market economy system was initially established. Therefore, we set the first stage from 1978 to 1990, the period of the commodity economy, and the second stage from 1990 to 2000, the period of marketization. China's accession to the World Trade Organization in 2001 rendered it more closely connected to the world. Therefore, we marked the third stage from 2000 to 2010, the period of globalization. The last period was marked from 2011 to 2019. At this stage of the financial crisis and the adjustment of the global economic structure, the focus of China's economic development gradually turned to localization.

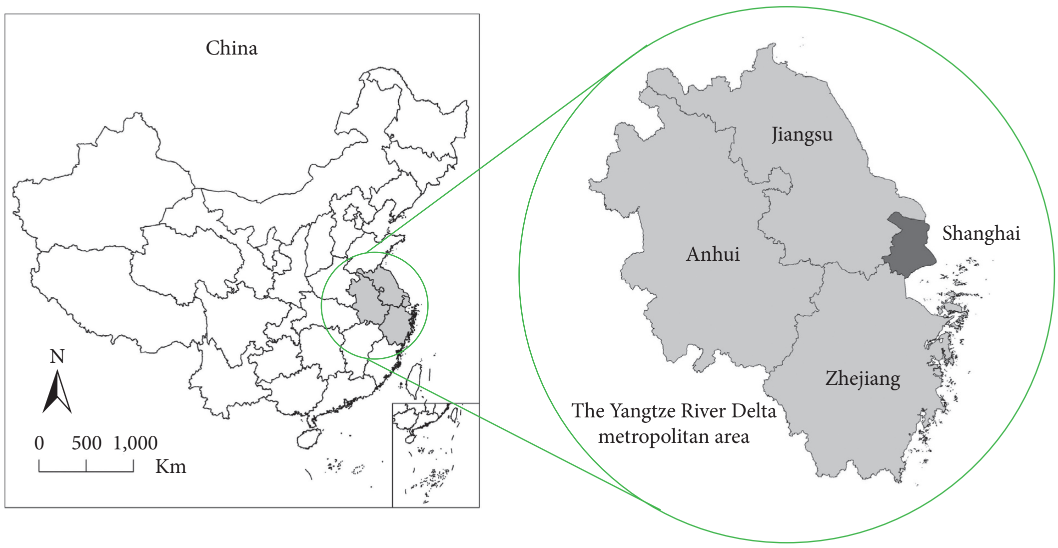



Figure 1: Map of the study area.

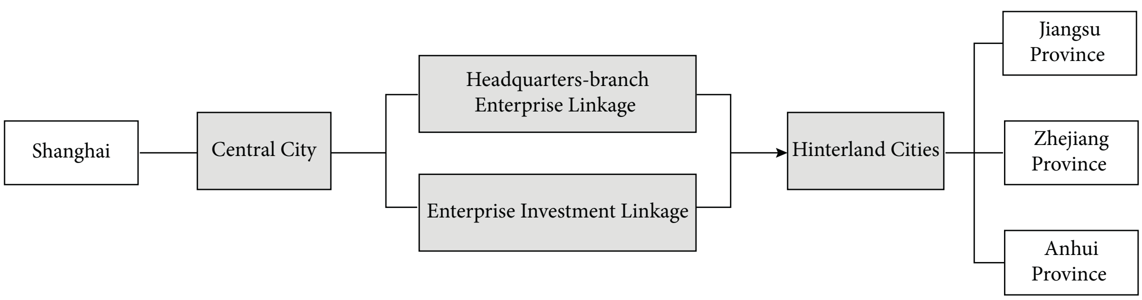



Figure 2: Enterprise linkages between the central city and hinterland cities.

*3.3. Research Method and Influencing Factors.* This article focuses on a comparative study of Shanghai's influencing factors in different provinces from 1978 to 2019. Through the rationality test, the test results based on $R^2$ are all greater than 0.88, this paper believes that the pool OLS model is suitable for analyzing Shanghai's influence factors on other hinterland cities.

Drawing on the practices observed in the previous literature [36–39] and combining them with the actual situation, we selected 10 indicators that have potentially impacted Shanghai's enterprise linkages in terms of capital input, labor input, urbanization, and level of industrialization (Table 1). The formula for the regression model is as follows:

$$Y_i = \beta_0 + \beta_1 \ln \mathrm{FAI}_i + \beta_2 \ln \mathrm{LR}_i + \beta_3 \ln \mathrm{UR}_i + \beta_4 \ln \mathrm{IDU}_i + \beta_5 \ln \mathrm{FRE}_i + \beta_6 \ln \mathrm{DP}_i + \beta_7 \ln \mathrm{PD}_i + \beta_8 \ln \mathrm{CGP}_i + \beta_9 \ln \mathrm{HM}_i + \beta_{10} \ln \mathrm{TS}_i + \varepsilon_i \, (i = 1, \ldots, n), \tag{1}$$

where $Y_i$ is the number of enterprise linkages established between city $i$ and Shanghai. (1) $\mathrm{FAI}_i$ refers to the capital input. Capital inputs can provide support for the development of enterprises. This study uses per capita investment in fixed assets to measure capital inputs. (2) $\mathrm{LR}_i$ refers to the labor input. Sufficient labor provides the necessary production factors for the development of enterprises and creates conditions for them to expand the scale of the market and form capital accumulation. This study takes the number of employees in each city to measure the input of urban labor. (3) $\mathrm{UR}_i$ refers to the urbanization. The ratio of the nonagricultural population to the total population is taken to indicate the level of urbanization. (4) $\mathrm{IDU}_i$ represents the level of industrialization. The level of industrialization is the ratio of the output value of the secondary industry to the total output value. (5) $\mathrm{FRE}_i$ refers to the marketization. The level of marketization refers to the degree to which the market plays a role in resource allocation. The higher the degree of marketization, the higher the degree of openness of the city. Since it is difficult to obtain complete indicators for state-owned enterprises over a long period of time, this study uses the ratio of fiscal budget revenue to fiscal budget expenditure indicating the level of marketization [40]. (6) $\mathrm{DP}_i$ refers to the capital status. Capital status indicates the wealth of the city, and the per capita savings deposit balance of urban and rural residents is used to find a city's capital status. (7) $\mathrm{PD}_i$ refers to the labor abundance. Labor abundance indicates the abundance of the population, and this study uses population density to measure it. (8) $\mathrm{CGP}_i$ refers to the market capacity. Market capacity represents the level of consumption in a city, which is determined by the total number of retail sales of consumer goods per capita. (9) $\mathrm{HM}_i$

Table 1: Regression variables and representational meaning.

| Independent variables | Interpretation of independent variables (unit of measurement) | Representational meaning |
|---|---|---|
| FAI | Per capita investment in fixed assets (yuan/person) | Capital input |
| LR | Number of employees (ten thousand people) | Labor input |
| UR | Ratio of nonagricultural population (%) | Urbanization |
| IDU | Ratio of the output value of the secondary industry in the total output value (%) | Industrialization level |
| FRE | Ratio of fiscal budget revenue to fiscal budget expenditure | Marketization |
| DP | Per capita savings deposit balance of urban and rural residents (yuan/person) | Capital status |
| PD | The population density (person/km$^2$) | Labor abundance |
| CGP | Total retail sales of consumer goods per capita (yuan/person) | Market capacity |
| HM | Highway mileage (km) | Traffic accessibility |
| TS | Number of fixed telephone users (ten thousand households) | Communication facility status |

refers to the traffic accessibility. The level of traffic accessibility is related to transportation costs. Theories in new economic geography dictate that underdeveloped areas will lead to high transportation costs due to backward transportation conditions. The highway mileage in the city was used as a measurement indicator. (10) $TS_i$ represents communication facility status. The status of communication facilities is part of infrastructure construction, which can help companies convert internal costs into social public costs. This study analyzed the number of fixed telephone users. $\beta_1 \sim \beta_{10}$ are the regression coefficients, $\varepsilon$ is the error disturbance term, and $n$ is the number of cities in the study area. The results of the White test and Variance Inflation Factor (VIF) test indicate that the multiple linear regression has no heteroscedasticity and multicollinearity problems.

The data sources for regression analysis include the yearbooks, statistical yearbooks, or statistical bulletins of 40 cities in the YRD from 1978 to 2019, such as "Nanjing Statistical Yearbook (1978–2019)," "Hangzhou Statistical Yearbook (1978–2019)," and "Hefei Statistical Yearbook (1978–2019)."

## 4. Results and Discussion

### *4.1. Spatiotemporal Characteristics of Enterprise Linkages in the YRD*

*4.1.1. Spatiotemporal Characteristics of Headquarters-Branch Enterprise Linkages.* In general, Shanghai enterprises have obvious polarization characteristics in the establishment of branch companies in hinterland cities and the gradient of hinterland cities is becoming increasingly obvious (Figure 3, Tables 2–4). In terms of different provinces, Shanghai has different polarization characteristics among the three provinces. (1) The headquarters-branch enterprise linkages in Jiangsu Province have "dual polarization" characteristics. Suzhou and Nanjing are the two cities with the most branches, and the proportion of their enterprise branches in Jiangsu Province continues to increase. Specifically, from 1978 to 2019, the total proportion of the two cities' branches in Jiangsu Province increased from 18.18% to 50.74%. Among them, the proportion of branches in Suzhou maintained a steady growth and the number of enterprise branches in Suzhou and Nanjing exhibited a widening trend. (2) The number of enterprise branches in Zhejiang Province is characterized by a dynamic of "core replacement." From 1992 to 2019, the ratio of Hangzhou's enterprise branches to the total number of branches in Zhejiang increased from 19.44% to 34.37%, while in Ningbo, it dropped from 36.11% to 21.41%. Hangzhou's core position has emerged, gradually replacing Ningbo as the city with the most branches. (3) The headquarters-branch enterprise linkages in Anhui Province exhibit significant "single-point polarization" characteristics. From 1992 to 2019, the proportion of the total number of branches in Hefei increased rapidly, from 17.65% to 43.30%. As the second-largest city in Anhui, Wuhu's proportion of the total number of enterprise branches in the period from 1992 to 2019 gradually decreased from 17.65% to 8.78%.

In terms of time, we have the following: (1) During the period of commodity economy, the establishment of enterprise branches was characterized by neighboring effects. Cities close to Shanghai were prioritized for the establishment of enterprise branches. Anhui Province is far from Shanghai. Compared with the Jiangsu and Zhejiang Provinces, there were fewer branches in the cities of Anhui. (2) During marketization, enterprise branches were concentrated in the three provincial capital cities, and the status of provincial capital cities was prominent. From 1992 to 2003, the growth rate of enterprise branches in provincial capital cities in their province was significantly higher than that in other periods and was also significantly higher than that of other cities in the same province. The proportion of Nanjing in Jiangsu Province increased by 8.27%, while the proportion of Suzhou only increased by 1.61% and the proportion of Hangzhou in Zhejiang Province increased by 5.33%, surpassing Ningbo to become the city with the largest number of branches in Zhejiang Province; the proportion of Hefei in Anhui Province also increased by 16.32%. (3) During the period of globalization, Hefei's influence in the YRD increased and its primacy also increased significantly. From 2003 to 2009, although Jiangsu and Zhejiang Provinces still far surpassed Anhui Province in the total number of enterprise branches, at the city level, the gap between Hefei and Suzhou/Nanjing/Hangzhou decreased. Wuhu, the second-largest city in Anhui Province, had a widening gap with Hefei. The gap widened from 17.84% in 2003 to 29.97% in 2009. (4) During the localization period, the three major cities of Suzhou, Hangzhou, and Nanjing established their dominant positions, becoming the cities with the closest

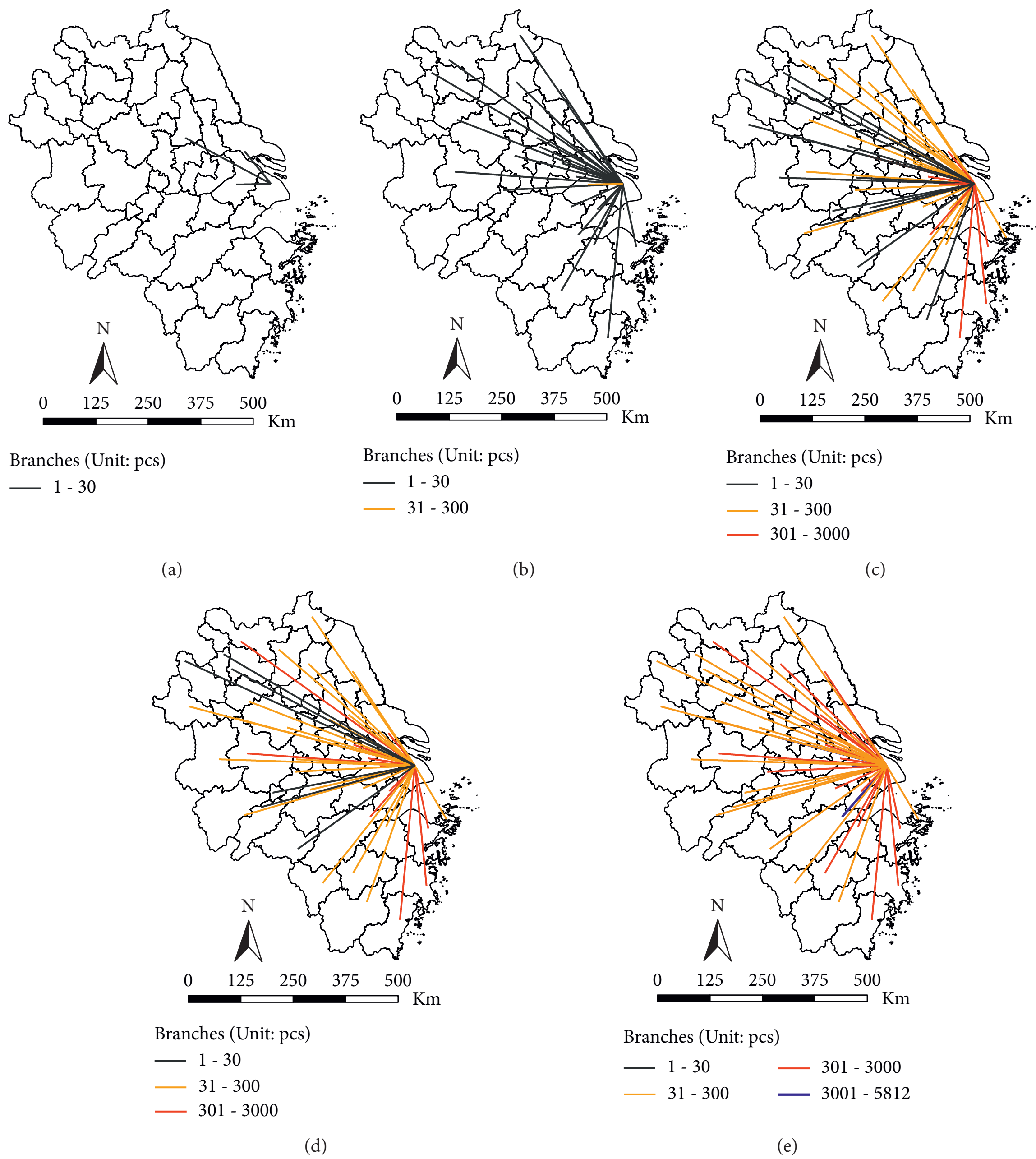



FIGURE 3: Headquarters-branch enterprise linkages between Shanghai and hinterland cities. (a) 1978. (b) 1992. (c) 2003. (d) 2009. (e) 2019.

enterprise branch linkages to Shanghai. In 2019, the total number of branches in the three cities accounted for nearly 40% of the total number of branches in the YRD.

*4.1.2. Spatiotemporal Characteristics of Enterprise Investment Linkages.* According to the classification of natural breaks in ArcGIS 10.3, the investment amount of enterprises can be divided into four types. These are low investments (1–1000), medium investments (1001–1000000), high investments (1000001–1000000000), and superhigh investments (1000000001–1016156440). There is a significant difference in the amount of investment represented by each type (Figure 4, Tables 5–7).

Overall, Shanghai's enterprise investment in hinterland cities is concentrated in some cities (Figure 4). In terms of the different provinces, we have the following: (1) Enterprise investment in Jiangsu Province is mainly concentrated in economically developed southern Jiangsu. Nanjing, Suzhou, Wuxi, and Changzhou were the key invested cities in Jiangsu Province. Over the years, these four cities have been in the top five positions of the investment rankings, and their total investment in Jiangsu Province has remained above 80%. (2) Enterprise investment in Zhejiang Province focuses on economic and characteristic resources. Shanghai's enterprise investments focused on two cities, Hangzhou and Ningbo. From 1978 to 2019, the total investment in the two cities

Table 2: Proportion of headquarters-branch enterprise linkages in cities in Jiangsu Province.

| | 1978 (%) | 1992 (%) | 2003 (%) | 2009 (%) | 2019 (%) |
|---|---|---|---|---|---|
| Nanjing | — | 10.99 | 19.26 | 20.62 | 18.80 |
| Wuxi | — | 10.44 | 16.76 | 14.85 | 13.11 |
| Xuzhou | — | 8.24 | 4.88 | 4.10 | 3.72 |
| Changzhou | — | 6.59 | 5.21 | 4.98 | 6.48 |
| Suzhou | 18.18 | 22.53 | 24.14 | 29.07 | 31.94 |
| Nantong | 72.73 | 9.89 | 7.93 | 7.07 | 8.41 |
| Lianyungang | — | 8.79 | 2.90 | 2.57 | 1.35 |
| Huai'an | — | 0.55 | 2.57 | 2.15 | 2.02 |
| Yancheng | — | 0.55 | 4.62 | 3.76 | 3.54 |
| Yangzhou | 9.09 | 6.04 | 3.40 | 3.65 | 3.85 |
| Zhenjiang | — | 6.04 | 3.79 | 2.85 | 2.80 |
| Taizhou | — | 9.34 | 3.79 | 3.31 | 2.78 |
| Suqian | — | — | 0.74 | 1.00 | 1.20 |

Table 3: Proportion of headquarters-branch enterprise linkages in cities in Zhejiang Province.

| | 1978 | 1992 (%) | 2003 (%) | 2009 (%) | 2019 (%) |
|---|---|---|---|---|---|
| Hangzhou | — | 19.44 | 24.77 | 26.17 | 34.37 |
| Ningbo | — | 36.11 | 20.16 | 24.21 | 21.41 |
| Wenzhou | — | 2.78 | 11.86 | 9.93 | 9.27 |
| Jiaxing | — | 2.78 | 8.24 | 9.87 | 8.68 |
| Huzhou | — | 8.33 | 4.08 | 2.88 | 3.64 |
| Shaoxing | — | 27.78 | 6.90 | 4.78 | 5.84 |
| Jinhua | — | 2.78 | 4.08 | 4.61 | 5.75 |
| Quzhou | — | — | 1.27 | 1.04 | 1.33 |
| Zhoushan | — | — | 2.65 | 2.67 | 1.99 |
| Taizhou | — | — | 15.03 | 12.90 | 6.55 |
| Lishui | — | — | 0.95 | 0.94 | 1.17 |

Table 4: Proportion of headquarters-branch enterprise linkages in cities in Anhui Province.

| | 1978 | 1992 (%) | 2003 (%) | 2009 (%) | 2019 (%) |
|---|---|---|---|---|---|
| Hefei | — | 17.65 | 33.97 | 38.94 | 43.30 |
| Wuhu | — | 17.65 | 16.13 | 8.97 | 8.78 |
| Bengbu | — | 23.53 | 8.54 | 6.24 | 5.29 |
| Huainan | — | — | 5.69 | 5.55 | 3.73 |
| Maanshan | — | — | 4.93 | 3.95 | 3.97 |
| Huaibei | — | 41.18 | 3.80 | 1.60 | 1.88 |
| Tongling | — | — | 1.52 | 2.28 | 2.05 |
| Anqing | — | — | 7.40 | 5.86 | 4.28 |
| Huangshan | — | — | 2.09 | 2.28 | 2.34 |
| Chuzhou | — | — | 5.31 | 5.70 | 4.54 |
| Fuyang | — | — | 5.31 | 4.56 | 5.40 |
| Suzhou | — | — | 1.14 | 2.28 | 3.13 |
| Liuan | — | — | 0.38 | 3.80 | 3.86 |
| Bozhou | — | — | 0.19 | 1.98 | 2.49 |
| Chizhou | — | — | 0.38 | 1.98 | 1.37 |
| Xuancheng | — | — | 3.23 | 4.03 | 3.57 |

remained above 55% in Zhejiang Province. In other years, Shanghai also invested in other cities, such as Shaoxing (in 1992), Wenzhou (in 2009), and Jiaxing (in 2019), but the proportion of investment was lower than that seen in Hangzhou and Ningbo. (3) Enterprise investment in Anhui Province has gradually decentralized. The amount of Shanghai's investment in Hefei dropped sharply. From 1992 to 2019, the proportion of Shanghai's total investment in Hefei in Anhui Province decreased from 99.88% to 26.23%, making its proportion closer to that of Wuhu.

In terms of time, Shanghai's enterprise investment linkages in hinterland cities have become increasingly close and tend to increase in general (Figure 4). (1) During the period of the community economy, the investment scope of Shanghai enterprises expanded from two cities in 1978 to 25 in 1992. The investment scope was relatively scattered, and the main levels of investment seen were "low" and "medium" investments. (2) During the period of marketization, the investment scope of Shanghai enterprises expanded to all cities in the YRD. "Medium investment" cities accounted for

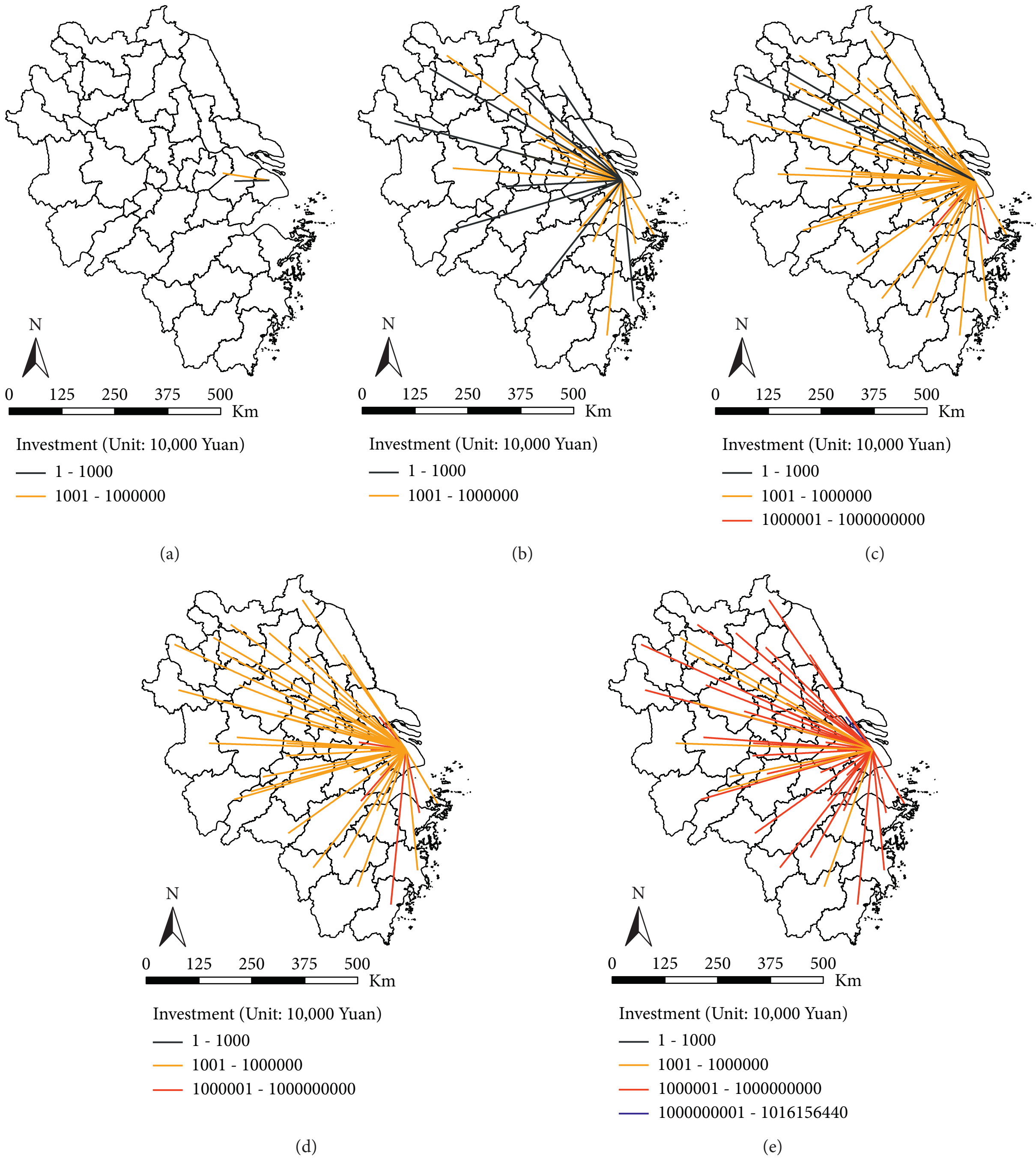



Figure 4: Enterprise investment linkages between Shanghai and hinterland cities. (a) 1978. (b) 1992. (c) 2003. (d) 2009. (e) 2019.

85%, and "high investment" cities accounted for 10%. (3) During the period of globalization, the amount invested by Shanghai enterprises increased. The proportion of "high investment" cities increased slightly, reaching 17.50%, and the rest were "medium investment" cities. (4) During the localization period, the investment linkages between Shanghai and its hinterland cities were upgraded and 80% of those cities reached "high investment" status.

Table 5: Proportion of enterprise investment in cities in Jiangsu Province.

| | 1978 (%) | 1992 (%) | 2003 (%) | 2009 (%) | 2019 (%) |
|---|---|---|---|---|---|
| Nanjing | — | 76.85 | 35.25 | 37.86 | 2.66 |
| Wuxi | 96.73 | 2.24 | 5.05 | 11.04 | 1.13 |
| Xuzhou | — | 1.06 | 1.81 | 1.88 | 0.32 |
| Changzhou | — | 1.50 | 4.25 | 3.69 | 0.68 |
| Suzhou | 3.27 | 16.36 | 44.55 | 27.58 | 3.91 |
| Nantong | — | 0.64 | 1.21 | 7.74 | 89.20 |
| Lianyungang | — | — | 0.46 | 1.50 | 0.18 |
| Huai'an | — | 0.05 | 1.01 | 0.83 | 0.14 |
| Yancheng | — | 0.00 | 0.47 | 1.70 | 0.44 |
| Yangzhou | — | 0.85 | 4.39 | 2.47 | 0.34 |
| Zhenjiang | — | 0.40 | 0.87 | 1.92 | 0.56 |
| Taizhou | — | 0.04 | 0.59 | 1.50 | 0.33 |
| Suqian | — | — | 0.10 | 0.30 | 0.11 |

Table 6: Proportion of enterprise investment in cities in Zhejiang Province.

| | 1978 | 1992 (%) | 2003 (%) | 2009 (%) | 2019 (%) |
|---|---|---|---|---|---|
| Hangzhou | — | 25.01 | 36.99 | 42.56 | 29.71 |
| Ningbo | — | 52.50 | 35.89 | 22.66 | 28.20 |
| Wenzhou | — | 2.69 | 5.08 | 10.93 | 4.78 |
| Jiaxing | — | 1.78 | 5.42 | 8.28 | 20.52 |
| Huzhou | — | 0.09 | 5.76 | 4.35 | 3.60 |
| Shaoxing | — | 15.71 | 2.91 | 3.20 | 3.88 |
| Jinhua | — | — | 1.99 | 1.72 | 2.11 |
| Quzhou | — | 0.01 | 0.26 | 0.80 | 0.77 |
| Zhoushan | — | 2.16 | 4.13 | 3.76 | 3.52 |
| Taizhou | — | 0.06 | 1.54 | 1.72 | 2.47 |
| Lishui | — | — | 0.04 | 0.03 | 0.45 |

Table 7: Proportion of enterprise investment in cities in Anhui Province.

| | 1978 | 1992 (%) | 2003 (%) | 2009 (%) | 2019 (%) |
|---|---|---|---|---|---|
| Hefei | — | 99.88 | 55.36 | 46.61 | 26.23 |
| Wuhu | — | 0.02 | 24.74 | 18.93 | 18.74 |
| Bengbu | — | — | 1.30 | 1.91 | 4.16 |
| Huainan | — | — | 0.58 | 0.78 | 2.30 |
| Maanshan | — | — | 3.20 | 5.53 | 4.95 |
| Huaibei | — | 0.05 | 0.01 | 0.45 | 2.37 |
| Tongling | — | — | 0.40 | 0.46 | 1.77 |
| Anqing | — | 0.02 | 0.69 | 1.05 | 2.98 |
| Huangshan | — | — | 5.22 | 3.29 | 3.95 |
| Chuzhou | — | — | 2.11 | 3.97 | 4.16 |
| Fuyang | — | 0.03 | 2.62 | 1.04 | 2.61 |
| Suzhou | — | — | 0.74 | 0.97 | 1.44 |
| Liuan | — | — | 0.30 | 9.19 | 1.74 |
| Bozhou | — | — | 0.03 | 1.95 | 5.25 |
| Chizhou | — | — | 1.80 | 1.86 | 1.14 |
| Xuancheng | — | — | 0.88 | 2.01 | 16.21 |

### *4.2. Results of Enterprise Linkages in the YRD*

*4.2.1. Results of Headquarters-Branch Enterprise Linkages.* According to Table 8, the overall comparison of the influencing factors of headquarters-branch enterprise linkages in the three provinces shows that capital status (DP) and communication facilities status (TS) had a significant positive impact on the establishment of enterprise branches. This shows that economic strength and the degree of informatization are the key factors that affect whether Shanghai sets up an enterprise branch in the city. Jiangsu Province and Zhejiang Province/Anhui Province have different demands for labor reserves. Jiangsu places more emphasis on the total amount of labor, while Zhejiang and

Anhui place more emphasis on the distribution density of labor capital. This may be because Jiangsu pays more attention to the distribution of labor on a larger scale, and the establishment of branch connections by enterprises focuses more on prefecture-level city units, while Zhejiang and Anhui Province pay more attention to labor in small-scale areas. The lower the industrialization level of Zhejiang and Anhui Provinces, the more the branches set up by local enterprises. This may be because places with a higher proportion of the tertiary industry are more attractive for enterprises to set up branches. Compared with Zhejiang, Shanghai pays more attention to the local market capacity when setting up business branches in Jiangsu and Anhui, which shows the positive significance of urban residents' consumption power for business branches. For provinces that are less economically underdeveloped, the establishment of branches by enterprises is more inclined to be in areas with high levels of urbanization.

In Jiangsu Province, capital status (DP) and communication facility status (TS) had a significant positive impact on enterprise branch establishment. Suzhou and Nanjing are the cities where Shanghai has the most branches in Jiangsu Province. Suzhou has strong economic strength and is close to Shanghai. It can quickly pass through the construction of national economic development zones and provincial development zones and build advanced manufacturing bases around Shanghai, forming a close connection with the city. Large-scale infrastructure construction and industrial clusters have attracted many people to work. The rapid development and upgrading of the industry have laid a good foundation for the establishment of branch enterprises in Shanghai. With a closer connection with Shanghai in the later period, relying on the overflow of resources in Shanghai, Suzhou's economic level and the level of supporting infrastructure have been continuously improved, forming a virtuous economic cycle [41]. Relying on the high-tech development zone established in the early years, Nanjing has fully enhanced the competitive advantages of leading industries such as petrochemicals, automobile manufacturing, and electronic complete machines and components and formed an industrial scale. The development of the industry has promoted the accumulation of considerable laborers, and Nanjing has also continued to attract Shanghai to set up enterprise branches. Capital input (FAI) had a significant negative impact on the establishment of enterprise branches. A possible reason may be that the market-oriented reform in Jiangsu Province was early, and the economically developed regions attracted many people based on their own advantages, forming an intensive development path. Per capita investment in fixed assets is relatively small in economically developed areas, so Shanghai's enterprise branch establishments choose cities with low capital input. This indicates that Shanghai's establishment of branch enterprises in Jiangsu Province tended to be concentrated in areas with affluent residents and complete infrastructure and that enterprise branch establishment did not depend on capital investment.

In Zhejiang Province, communication facility status (TS), capital status (DP), labor abundance (PD), and traffic accessibility (HM) all had positive impacts on enterprise branch establishment. However, the industrialization level (IDU) had a significant negative correlation with it. This indicates that the establishment of branch enterprises in Shanghai in Zhejiang Province tended to be in cities with well-developed infrastructure, high levels of residential income, and sufficient labor and cities which were not dominated by industrial enterprises. Relying on the advantages of the provincial capital, Hangzhou possesses economic advantages and labor resources. From 2009 to 2019, Shanghai set up more enterprise branches in Hangzhou than Ningbo. During this period, Hangzhou prioritized the service industry and digital economy to drive the development of tertiary industry as its main goal, actively undertaking the international service industry, and the proportion of the secondary industry continued to decline. In 2014, the Hangzhou Municipal Party Committee deliberated and approved the "Several Opinions on Accelerating the Development of the Information Economy" and created a strategic plan for the development of the information economy. Hangzhou officially launched the "No. 1 Project" of "Developing the Information Economy and Promoting Smart Applications." After that, Hangzhou used the strength of the whole city to promote the priority development of the information economy. In 2018, the proportion of Hangzhou's secondary industry dropped to 31.7% and the proportion of tertiary industry reached nearly 66.2%. The development of the tertiary industry has promoted the improvement of infrastructure facilities and attracted a large flow of labor. Ningbo has always set industrial development as its goal. In the early years, Ningbo relied on the advantages of its seaports to vigorously promote economic development and formed a strong enterprise relationship with Shanghai. Currently, the Ningbo Municipal Government has determined the strategic goal of building a strong industrial city, and the secondary industry accounts for more than 50% of its economy in Ningbo. When the economy enters the advanced stage, the higher the level of industrialization, the lower the proportion of the tertiary industry and the lesser the advantage to the city when it establishes a branch in Shanghai.

In Anhui Province, marketization (FRE), communication facilities (TS), labor abundance (PD), urbanization (UR), capital status (DP), and market capacity (CGP) had a positive and significant impact on the establishment of enterprise branches. Traffic accessibility (HM) was a necessary element for the establishment of enterprise branches, but it was not a key element. However, the industrialization level (IDU) had a significant negative impact on enterprise branch establishment. In general, Shanghai's establishment of branch companies in Anhui mainly considers the city's openness, economic strength, and labor level, and it prefers cities that do not rely on the secondary industry. Anhui Province is a relatively underdeveloped area, and many regions in the province are still in the initial stage of urbanization development, with low levels of urbanization and marketization. Problems such as the large proportion of the agricultural population and the shortage of urban labor have restricted the development of cities. Anhui Province concentrates its

Table 8: Regression results of headquarters-branches enterprise linkages.

| Independent variable | YDR | Jiangsu | Zhejiang | Anhui |
|---|---|---|---|---|
| ln FAI | −0.044 (−0.368) | −0.643*** (−2.915) | −0.091 (−0.839) | 0.172 (1.121) |
| ln LR | 0.337*** (2.825) | 0.323* (1.940) | 0.070 (1.444) | 0.137 (0.797) |
| ln UR | 0.650*** (4.212) | 0.514 (0.928) | 0.096 (0.894) | 0.589*** (2.860) |
| ln IDU | −0.666 (−1.617) | 0.127 (0.171) | −0.351*** (−3.510) | −2.143*** (−5.608) |
| ln FRE | 1.311*** (3.693) | −0.588 (−0.712) | −0.053 (−0.772) | 2.054*** (5.044) |
| ln DP | 0.606*** (3.320) | 1.126*** (4.176) | 1.127*** (5.692) | 0.367** (2.294) |
| ln PD | 0.408*** (3.032) | 0.104 (0.164) | 0.640*** (4.954) | 0.460*** (4.323) |
| ln CGP | 0.234 (1.167) | 0.448* (1.763) | 0.081 (0.302) | 0.522** (2.191) |
| ln HM | 0.282** (2.190) | 0.152 (0.508) | 0.249*** (4.029) | 0.287* (1.804) |
| ln TS | 0.044*** (3.143) | 0.053*** (2.954) | 0.443*** (6.204) | 0.354*** (4.903) |
| Constant | −10.298*** (−6.208) | −9.641** (−2.321) | −13.775*** (−12.318) | −7.917*** (−5.284) |
| $N$ | 40 | 13 | 11 | 16 |
| $R^2$ | 0.893 | 0.898 | 0.954 | 0.992 |

Note: the symbols *, **, and *** represent the significance levels of 10%, 5%, and 1%, respectively; the values in parentheses are the corresponding $t$-values.

advantageous resources in the provincial capital city Hefei through infrastructure construction and the provision of public products. After the reform of the tax-sharing system, local governments' dependence on local revenues has increased, which has weakened the ability of local governments to redistribute resources to achieve fairness goals [42]. Hefei's first position in Anhui Province has been continuously strengthened. It has advantages in marketization, infrastructure, and labor and has become an important city in the central region to undertake the gradient transfer of industries in the developed coastal areas of China.

*4.2.2. Results of Enterprise Investment Linkages.* According to Table 9, comparing the factors influencing enterprise investment in the three provinces, capital status (DP) had a positive impact on Shanghai's enterprise investment in all the three provinces. In Jiangsu and Zhejiang Provinces, the higher the market capacity, the more they can attract enterprise investment. This may be because Shanghai's investment in these two provinces relies on the support of their local high-consumption markets, while its investment in Anhui Province does not. The relationship between enterprise investment in Zhejiang and the level of industrialization is inversely correlated, but the investment of enterprises in Jiangsu and Anhui Provinces is positively correlated with the level of industrialization. A possible reason for this may be that Shanghai's investment in Zhejiang Province is biased toward cities whose industrial industries have transformed into tertiary industries. Enterprise investment in Jiangsu Province depends on the construction of communication infrastructure, while in Anhui Province, it is more inclined to areas with underdeveloped communication infrastructure. This may be related to the difference between the advantageous industries in Jiangsu and Anhui. Shanghai is more willing to invest in places with good communication infrastructure in Jiangsu Province, while investment in Anhui Province relies on its natural base advantages and its information infrastructure tends to be imperfect. The capital investment (FAI) and investment choices of enterprises in Jiangsu and Zhejiang Provinces had a significant negative impact. A possible reason may be that Shanghai's choice of investment locations in Jiangsu and Zhejiang Provinces is biased toward cities with earlier market development.

In Jiangsu Province, capital investment (FAI) had a significant negative correlation with enterprise investment. Capital status (DP), labor input (LR), communication facility status (TS), industrialization level (IDU), and market capacity (CGP) were positively correlated with enterprise investment. This shows that investors in Jiangsu were more inclined to choose cities with economic strength, high income levels, and the strong consumption power of residents. The status of the urban industry and the adequacy of labor were also considerations that supported investment choices. Since the 1970s, southern Jiangsu has accumulated wealth through the development of townships and village enterprises. It has a strong economy, high population density, and complete support facilities, which give southern Jiangsu a competitive advantage in the YRD. The convenient transportation linkages between southern Jiangsu and Shanghai have encouraged the transfer of Shanghai's traditional manufacturing industries to southern Jiangsu. Many iron and steel, metallurgy, building material, petrochemical, and energy industries have been established along the transportation line, that is, the Yangtze River. Urban infrastructure construction guaranteed the expansion and development of enterprises. In Kunshan, Suzhou, multiple small- and medium-sized foreign companies in the IT industry have gathered. Six of Taiwan's top ten notebook computer manufacturers gathered and brought hundreds of Taiwanese-supporting enterprises to settle in Kunshan, rendering the area the world's largest notebook computer production base. Due to the existence of economies of scale and scope, specialized regions have further attracted related enterprises to join, forming an evolutionary process of path dependence [43], and perfect basic supporting facilities can provide a guarantee for industrial expansion. With the improvement of the market mechanism, the establishment of branches of enterprises in Shanghai no longer relies on the promotion of the government, instead it depends more on the market promotion.

Table 9: Regression results of enterprise investment.

| Independent variable | YDR | Jiangsu | Zhejiang | Anhui |
|---|---|---|---|---|
| ln FAI | 0.257 (1.272) | $-0.932^{***}$ (−3.219) | $-0.583^{**}$ (−2.626) | −0.221 (−0.339) |
| ln LR | $0.411^{**}$ (2.079) | $0.567^{***}$ (2.930) | 0.071 (0.622) | $1.461^{*}$ (2.002) |
| ln UR | $0.664^{**}$ (2.575) | 0.008 (0.013) | 0.236 (1.397) | 0.367 (0.418) |
| ln IDU | 0.241 (0.346) | $2.109^{**}$ (2.150) | $-0.536^{***}$ (−3.452) | $3.949^{**}$ (2.429) |
| ln FRE | $1.280^{**}$ (2.130) | −1.348 (−1.223) | $-0.296^{*}$ (−1.796) | 1.940 (1.121) |
| ln DP | $1.128^{***}$ (3.639) | $2.082^{***}$ (5.753) | $1.426^{***}$ (4.512) | $1.183^{*}$ (1.738) |
| ln PD | 0.281 (1.254) | −0.081 (−0.118) | $1.098^{***}$ (4.280) | −0.015 (−0.034) |
| ln CGP | 0.157 (0.462) | $0.715^{**}$ (2.035) | $0.833^{*}$ (1.845) | 0.461 (0.455) |
| ln HM | 0.239 (1.112) | $0.603^{*}$ (1.864) | 0.054 (0.388) | 0.352 (0.521) |
| ln TS | −0.002 (−0.078) | $0.067^{***}$ (2.728) | −0.087 (−0.774) | $-1.225^{***}$ (−3.990) |
| Constant | $-12.461^{***}$ (−4.423) | $-20.667^{***}$ (−3.903) | $-9.324^{***}$ (−4.164) | $-24.407^{***}$ (−3.830) |
| $N$ | 40 | 13 | 11 | 16 |
| $R^2$ | 0.884 | 0.937 | 0.927 | 0.916 |

Note: the symbols $^{*}$, $^{**}$, and $^{***}$ represent the significance levels of 10%, 5%, and 1%, respectively; the values in parentheses are the corresponding $t$-values.

In Zhejiang Province, capital status (DP), labor abundance (PD), and market capacity (CGP) had positive effects on enterprise investment. However, the industrialization level (IDU) had a significant negative impact on enterprise investment. Capital investment (FAI) and marketization (FRE) had a negative effect on enterprise investment, but they were not decisive factors. This shows that in Zhejiang Province, Shanghai investors were more inclined to choose hinterland cities with strong economic growth among residents, high residential spending power, and dense populations as investment locations. The tertiary industry-led cities represented by Hangzhou are hot cities for investment in Shanghai. In addition, many cities in Zhejiang have formed an industrial structure dominated by light industry through the development of family workshops and family factories. Due to the promotion of industrialization in these cities over the years, the proportion of tertiary industries has continuously increased and their urban service capabilities have improved. The establishment of a characteristic market can form a competitive advantage nationwide and has a profitable and abundant labor market. It has an absolute advantage in undertaking Shanghai's capacity overflow. The development of clothing and consumer goods as the leading industries also depends on the residents' high consumption power. Examples of this include the Yiwu small commodity market, Haining leather market, and Shaoxing textile market. Nowadays, the Shaoxing textile market is developing rapidly, attracting a large amount of investment and transforming into the professional textile wholesale market with the largest scale, the highest turnover and the most variety in Asia. The textile and garment industry clusters in Jiashan, Zhejiang, and other places have developed rapidly during this period, and the employment density of the textile and garment manufacturing industry has also increased significantly.

In Anhui Province, industrial level (IDU), labor input (LR), and capital status (DP) had positive effects on enterprise investment choices, but communication facility status (TS) had a significant negative correlation with enterprise investment. This indicates that Shanghai's investment in Anhui Province tended to be in cities with immature infrastructure. The economic strength of Anhui Province does not have comparative advantages in the YRD, but ecological environment resources and agricultural industry are some of this province's advantages. Due to the slow development of Anhui's urbanization process, cities that developed by relying on ecological industries are still in the development stage. These areas often have insufficient infrastructure and poor infrastructure facilities. Xuancheng is one of the cities with an excellent ecological environment in Anhui Province, with a forest coverage rate of 59.34%. There are four national ecological demonstration areas, one national nature reserve, and three provincial nature reserves. Outstanding location advantages and good natural conditions provide a strong guarantee for the development of modern agriculture. Xuancheng has become an important production and supply base for grain, oil, tea, poultry, and forest products in the YRD. The agricultural value of Xuancheng has significant advantages. In 2019, Shanghai Brightdairy Dairy Co., Ltd. cooperated with Xuancheng to invest 6 billion yuan to build a food production and processing base.

## 5. Conclusion and Discussion

This study used headquarters-branch enterprise linkage and enterprise investment data from 1978 to 2019 to explore the enterprise linkages between Shanghai and its hinterland cities in the YRD. It expands the time range of the research area and improves the limitations of short and incomplete data in previous studies. At the same time, it also innovatively introduced a number of relevant factors, which further analyzed the influencing factors of the internal evolution mechanism of the metropolitan area and enriched our understanding of the temporal and spatial evolution of the core-peripheral structure of the YRD. The main conclusions of this study are as follows:

(1) In the YRD, the headquarters-branch enterprise linkages between Shanghai and hinterland cities have gradually become polarized, and they are increasingly concentrated in large cities with strong economic strength. Suzhou, Hangzhou, and Nanjing have become the cities with the highest number of

branches. In terms of different provinces, the headquarters-branch linkages between Jiangsu and Shanghai exhibit “dual polarization” characteristics. Suzhou and Nanjing have the highest number of branches in Jiangsu. The headquarters-branch linkages between Zhejiang and Shanghai demonstrate “core replacement” characteristics. Hangzhou has gradually replaced Ningbo as the city with the most branches in Zhejiang. The headquarters-branch linkages between Anhui and Shanghai manifest significant “single-point polarization” characteristics, and Hefei’s position in Anhui Province is prominent. The city’s economic strength has become a key factor in Shanghai’s choice of whether to establish a branch location for a company. Based on the regression results, in the three provinces of Jiangsu, Zhejiang, and Anhui, the capital status (DP) and the status of communication facilities (TS) positively impacted the construction of enterprise branches. The main reason is that cities with economic foundations can quickly respond to market development trends and provide construction capital to undertake enterprise spillovers in the central city.

(2) Enterprise investment linkages between Shanghai and its hinterland cities are getting closer, but the key cities for investment in each province are different. In the choice of investment locations for hinterland cities, Shanghai prefers cities that can represent the characteristics of the province and focus more on considering the city’s comparative advantages. Enterprise investment in Jiangsu Province is mainly concentrated in economically developed southern Jiangsu. Nanjing, Suzhou, Wuxi, and Changzhou are the core investment cities in Jiangsu. Enterprise investment in Zhejiang Province focuses on economically developed regions or cities with characteristic resources. The core investment cities mainly include Hangzhou and Ningbo. Enterprise investment in Anhui Province has been gradually decentralized, and the amount of Shanghai’s investment in Hefei has dropped sharply. Based on the regression results of enterprise investment linkages, the higher the market capacity in Jiangsu and Zhejiang provinces, the more it can attract enterprise investment. Enterprise investment in Jiangsu Province depends on the construction of communication infrastructure, while enterprise investment in Anhui Province, on the contrary, is more inclined to areas with underdeveloped communication infrastructure.

Research is of great significance to the future development of metropolitan areas. Research has shown that the enterprise branch linkages of the central city and the peripheral cities in metropolitan areas are increasingly concentrated in economically developed areas and their radiating effect on other cities is weaker. This view is consistent with the conclusions of Hymer who believe that enterprise locations tend to be concentrated in a few central cities [44]. This shows that the relationship between enterprises is based on economic relevance rather than geographic proximity [45]. This process often leads to greater development in already affluent areas, making the gap between rich and poor areas larger and larger [46]. Therefore, in order to reduce the gap within the metropolitan areas, the government can increase investment in infrastructure in economically backward areas in terms of policies and systems to provide support for undertaking the overflow of resources in central cities. In terms of enterprise investment, there are significant differences in the regression models of different provinces, reflecting the different ways of spreading Shanghai to different provinces. The characteristic industry of the city has the effect of attracting investment from the central city.

The interenterprise investment has the phenomenon of linkage in the same region, which will promote the phenomenon of investment convergence in the investment of enterprises in the same region. Therefore, cities need to fully integrate their own advantageous resources in planning and give full play to the natural geographical environment, advantageous industrial industries, and other supporting factors. Cities should seize the strategic opportunity for the integrated development of the Yangtze River Delta, strengthen cooperation with central cities and internal cooperation, strengthen the leading and radiating demonstration role of prominent industries, and use the contextual and spillover effects of leading industries to attract central cities to establish themselves. They must invest in links to promote their own economic development.

Due to data acquisition and space limitations, this study still has room for improvement. For example, the selected influencing factor data were not complete due to incomplete data of earlier years, so it was impossible to conduct a more detailed discussion of influencing factors by time. Due to space restrictions, no further detailed research has been conducted on the influencing factors of each city. More refined analysis can provide detailed support for our research. This article needs to further expand the data mining in this field in the future.

## Data Availability

The enterprise linkage data used to support the findings were supplied by the enterprise registration information of the State Administration for Industry and Commerce (http://www.gsxt.gov.cn).

## Disclosure

The funding sources had no role in the study design, data collection, analysis, interpretation, or writing of this manuscript.

## Conflicts of Interest

The authors declare that there are no conflicts of interest.

## Acknowledgments

This research was supported by the National Natural Science Foundation of China (Nos. 52078245 and 41971162).

## References

[1] C. L. Ross, *Megaregions: Planning for Global Competitiveness*, Island Press, Washington, DC, USA, 2009.

[2] J. Gottmann, "Megalopolis or the urbanization of the northeastern seaboard," *Economic Geography*, vol. 33, no. 3, pp. 189–200, 1957.

[3] F. van Oort, M. Burger, and O. Raspe, "On the economic foundation of the urban network paradigm: spatial integration, functional integration and economic complementarities within the Dutch randstad," *Urban Studies*, vol. 47, no. 4, pp. 725–748, 2010.

[4] N. Brenner and N. Theodore, "Cities and the geographies of "actually existing neoliberalism"," *Antipode*, vol. 34, no. 3, pp. 349–379, 2002.

[5] W. Anderson, *Economic Geography*, Routledge, New York, NY, USA, 2012.

[6] J. V. Beaverstock, M. A. Doel, P. J. Hubbard, and P. J. Taylor, "Attending to the world: competition, cooperation and connectivity in the world city network," *Global Networks*, vol. 2, no. 2, pp. 111–132, 2002.

[7] R. Yang, J. Zhanga, Q. Xub, and X. Luo, "Urban-rural spatial transformation process and influences from the perspective of land use: a case study of the Pearl River Delta Region," *Habitat International*, vol. 104, Article ID 102234, 2020.

[8] L. Zhou, X. Dang, H. Mu, B. Wang, and S. Wang, "Cities are going uphill: slope gradient analysis of urban expansion and its driving factors in China," *The Science of the Total Environment*, vol. 775, Article ID 145836, 2021.

[9] A. G. Champion, "A changing demographic regime and evolving poly centric urban regions: consequences for the size, composition and distribution of city populations," *Urban Studies*, vol. 38, no. 4, pp. 657–677, 2001.

[10] Z. P. Neal, "From central places to network bases: a transition in the U.S. Urban hierarchy, 1900-2000," *City & Community*, vol. 10, no. 1, pp. 49–75, 2011.

[11] M. Castells, *The Rise of Network Society*, Blackwell, Oxford, UK, 1996.

[12] J. V. Beaverstock, R. G. Smith, and P. J. Taylor, "World-city network: a new m," *Annals of the Association of American Geographers*, vol. 90, no. 1, pp. 123–134, 2000.

[13] P. J. Taylor, "Specification of the world city network," *Geographical Analysis*, vol. 33, no. 2, pp. 181–194, 2001.

[14] P. J. Taylor, *World City Network: A Global Urban Analysis*, Routledege, New York, NY, USA, 2004.

[15] P. J. Taylor, M. Hoyler, and R. Verbruggen, "External urban relational process: introducing central flow theory to complement central place theory," *Urban Studies*, vol. 47, no. 13, pp. 2803–2818, 2010.

[16] M. C. Mahutga, fnm Xiulian Ma, D. A. Smith, and M. Timberlake, "Economic globalisation and the structure of the world city system: the case of airline passenger data," *Urban Studies*, vol. 47, no. 9, pp. 1925–1947, 2010.

[17] W. Zhang and B. Derudder, "Approximating actual flows in physical infrastructure networks: the case of the Yangtze River Delta high-speed railway network," *Bulletin of Geography. Socio-Economic Series*, vol. 31, no. 31, pp. 145–160, 2016.

[18] X. Liu, B. Derudder, and K. Wu, "Measuring polycentric urban development in China: an intercity transportation network perspective," *Regional Studies*, vol. 50, no. 8, pp. 1302–1315, 2016.

[19] G. Krings, F. Calabrese, C. Ratti, and V. D. Blondel, "Urban gravity: a model for inter-city telecommunication flows," *Journal of Statistical Mechanics: Theory and Experiment*, vol. 2009, 2009.

[20] P.-A. Laharotte, R. Billot, E. Come, L. Oukhellou, A. Nantes, and N.-E. El Faouzi, "Spatiotemporal analysis of Bluetooth data: application to a large urban network," *IEEE Transactions on Intelligent Transportation Systems*, vol. 16, no. 3, pp. 1439–1448, 2015.

[21] Y. Li and N. A. Phelps, "Articulating China's science and technology: knowledge collaboration networks within and beyond the Yangtze River Delta megalopolis in China," *Chinese Geographical Science*, vol. 28, no. 2, pp. 247–260, 2018.

[22] Y. Li and N. A. Phelps, "Knowledge polycentricity and the evolving Yangtze River Delta megalopolis," *Regional Studies*, vol. 51, no. 7, pp. 1035–1047, 2017.

[23] M. Zhao, X. Liu, B. Derudder, Y. Zhong, and W. Shen, "Mapping producer services networks in mainland Chinese cities," *Urban Studies*, vol. 52, no. 16, pp. 3018–3034, 2015.

[24] B. Derudder, Z. Cao, X. Liu et al., "Changing connectivities of Chinese cities in the world city network, 2010-2016," *Chinese Geographical Science*, vol. 28, no. 2, pp. 183–201, 2018.

[25] B. Derudder, X. Liu, C. Kunaka, and M. Roberts, "The connectivity of South Asian cities in infrastructure networks," *Journal of Maps*, vol. 10, no. 1, pp. 47–52, 2014.

[26] P. J. Taylor, B. Derudder, M. Hoyler, and P. Ni, "New regional geographies of the world as practised by leading advanced producer service firms in 2010," *Transactions of the Institute of British Geographers*, vol. 38, no. 3, pp. 497–511, 2013.

[27] F. Pan, W. Bi, J. Lenzer, and S. Zhao, "Mapping urban networks through inter-firm service relationships: the case of China," *Urban Studies*, vol. 54, no. 16, pp. 3639–3654, 2017.

[28] B. Derudder and P. Taylor, "Change in the world city network, 2000–2012," *The Professional Geographer*, vol. 68, no. 4, pp. 624–637, 2016.

[29] T. Zilai and L. Tao, "A comparative analysis of urban system in the Yangtze Delta Region and the Middle Yangtze Region: an approach of firm-based interlocking network," *Urban Planning Forum*, vol. 2, pp. 24–31, 2014.

[30] J. Y. Lu, D. Sun, J. Yu, J. Li, and F. Niu, "Local versus nonlocal enterprise linkages of global cities: a comparison between beijing and Shanghai, China," *Complexity*, vol. 2020, Article ID 8918534, 13 pages, 2020.

[31] J. Salder, "The networked economy of firms in city-region peripheries," *European Urban and Regional Studies*, vol. 28, pp. 195–212, 2020.

[32] A. G. Yeh, F. F. Yang, and J. Wang, "Producer service linkages and city connectivity in the mega-city region of China: a case study of the Pearl River Delta," *Urban Studies*, vol. 52, no. 13, pp. 2458–2482, 2015.

[33] X. Zhang, Q.-e. Guo, D. M.-w. Cheung, and T. Zhang, "Evaluating the institutional performance of the Pearl River Delta integration policy through intercity cooperation network analysis," *Cities*, vol. 81, pp. 131–144, 2018.

[34] E. Meijers, "From central place to network model: theory and evidence of a paradigm change," *Tijdschrift voor Economische en Sociale Geografie*, vol. 98, no. 2, pp. 245–259, 2007.

[35] Y. D. Wei, "Decentralization, marketization, and globalization: the triple processes underlying regional development in China," *Asian Geographer*, vol. 20, no. 1-2, pp. 7–23, 2001.

[36] Y. D. Wei, "Spatiality of regional inequality," *Applied Geography*, vol. 61, pp. 1–10, 2015.

[37] G. Li and C. Fang, "Spatial econometric analysis of urban and county-level economic growth convergence in China," *International Regional Science Review*, vol. 41, no. 4, pp. 410–447, 2018.

[38] A. Frenkel, "Why high-technology firms choose to locate in or near metropolitan areas," *Urban Studies*, vol. 38, no. 7, pp. 1083–1101, 2001.

[39] P. Guimaraes, O. Figueiredo, and D. Woodward, "Industrial location modeling: extending the random utility framework," *Journal of Regional Science*, vol. 44, no. 1, pp. 1–20, 2004.

[40] Z. Xuebo, Y. Wei, Z. Yali, S. Jinping, and W. Zhenbo, "Spatio-temporal differentiation and its influencing factors ofregional economic growth in Beijing-Tianjin-Hebei region," *Acta Geographica Sinica*, vol. 73, no. 10, pp. 1985–2000, 2018.

[41] L. Zhou, X. Dang, Q. Sun, and S. Wang, "Multi-scenario simulation of urban land change in Shanghai by random forest and CA-Markov model," *Sustainable Cities and Society*, vol. 55, Article ID 102045, 2020.

[42] F. H. F. Liao and Y. D. Wei, "Dynamics, space, and regional inequality in provincial China: a case study of Guangdong province," *Applied Geography*, vol. 35, no. 1-2, pp. 71–83, 2012.

[43] R. Martin and P. Sunley, "Path dependence and regional economic evolution," *Journal of Economic Geography*, vol. 6, no. 4, pp. 395–437, 2006.

[44] S. Hymer, *The Multinational Corporation and the Law of Uneven Development: Perspectives on World Politics*, Routledge, London, UK, 2006.

[45] P. Marek, M. Titze, C. Fuhrmeister, and U. Blum, "R&D collaborations and the role of proximity," *Regional Studies*, vol. 51, no. 12, pp. 1761–1773, 2017.

[46] E. Wang, "Fiscal decentralization and revenue/expenditure disparities in China," *Eurasian Geography and Economics*, vol. 51, no. 6, pp. 744–766, 2010.

Hindawi
Complexity
Volume 2021, Article ID 5514135, 14 pages
https://doi.org/10.1155/2021/5514135

*Research Article*

# Determinants of Intercity Air-Passenger Flows in the "Belt and Road" Region

**Lu Zhang**,[1] **Yannan Zhao**,[2] **Bart Dessein**,[3] **and Philippe De Maeyer**[4]

[1]*Beijing Technology and Business University, School of International Economics and Management, Beijing 100048, China*
[2]*Institute of Geographic Sciences and Natural Resources Research, Chinese Academy of Sciences, Beijing 100101, China*
[3]*Faculty of Arts and Philosophy, Ghent University, Ghent 9000, Belgium*
[4]*Department of Geography, Ghent University, Ghent 9000, Belgium*

Correspondence should be addressed to Yannan Zhao; zhaoyn@igsnrr.ac.cn

Received 8 February 2021; Revised 27 June 2021; Accepted 13 October 2021; Published 11 November 2021

Academic Editor: Francesco Lo Iudice



Air-passenger flow, denoting intercity connections, has been a focal point of studies pertaining to urban networks. While most existing studies include only the geoeconomic characteristics of nodes as explanatory variables, this analysis developed a gravity model by incorporating further factors (e.g., cultural disparity and institutional disparity) that might influence air-passenger flows in the "Belt and Road" region. The primary findings are as follows: (1) The cultural and institutional disparities correlate negatively with the air-passenger flows in this region; (2) air-passenger flows are positively related to border, population and economy size, and economic disparity; (3) flows tend to first increase and then subsequently decrease as geographical distance increases; (4) the impact of the factors on the flows varies by subregion. This study could serve as a reference for those interested in gaining a greater insight into air-passenger flows and could also help improve regional strategies for air-transportation development.

## 1. Introduction

In November 2013, the Chinese government implemented the "Belt and Road" initiative (BRI, hereafter) which aims to improve the connectivity and regional cooperation between the BRI countries. In March 2015, several government departments jointly released detailed plans for the BRI, encompassing the five major priority areas of policy coordination, facilities connectivity, unimpeded trade, financial integration, and people-to-people bonds [1]. Beyond these, a predominant focus has been on facilitating infrastructure connectivity which promotes the transport, energy, and communication infrastructure, forming a network connecting all the subregions within Asia and between the Asian, European, and African countries. Within this infrastructure connectivity, air-transportation infrastructure has linked even remote and inaccessible geographic locations and fostered intercity relations [2–4]. Thus, air transportation could immensely aid developing countries by unlocking the potential of trade and tourism. Air-transport demand has been of increasing interests to airlines, airports, government institutions, and scholars in recent years. The distribution and evolution law of traffic flow have been the core research objects of transportation geography.

Therefore, in the background of the BRI, what has been unclear is which factors have promoted or blocked the air-passenger flows in the BRI region. In sum, the previous studies have showed that factors influencing the formation of the transportation structure are complex since the cause comes from both internal network itself and external reasons. Internal factors, such as topological effects [5], have been proven to impact network formation. External factors, including social, economic, and geographic elements, impact the formation of the air-transport network structure and determine its functionality. Thus, these external factors are more pivotal to the actual evolution of the aviation network structure. Essentially, these factors are conducted from the node and edge aspects: (1) From the node aspect, population and economic size enormously influence the passenger flows, as a specific number of passengers and the ability of meeting demands are critical to the aviation network structure [6–9]; (2) from the edge aspect, geographical distance (which can trigger cost increments) and the border (which can cause the boundary effect) between nodes affect the formation of the air-transport network [10, 11]. However, apart from these factors, cultural and institutional factors have been largely ignored in existing literature. Previous studies have noted that cultural differences (e.g., common language) can both stimulate and hinder leisure travel [12, 13]. Regulation by complicated cognitive and normative institutions is a prominent part of the business flows in BRI countries and should hence be incorporated as key factors.

Thus, the aim of our analysis is twofold. First, we examine some emerging potential factors, such as cultural and institutional ones, to determine whether they influence air-passenger flows in the BRI region. Second, besides these potential factors, we reexamine some determinants, such as economic disparity, economic size, population, geographical distance, and border. This paper is divided into five main sections. The analysis commences with a section methodology that provides a rationale on some potential determinants that could influence air-passenger flows. Subsequently, the methodology section introduces the model we have adopted. The results section reports the parameter interpretation and model estimation results. Next, we discuss the primary findings as well as the limitations of this analysis. In the last section, we present the concluding remarks.

## 2. Literature Review and Hypothesis

In general, air-passenger flows can be broadly classified into two categories: business passengers and leisure passengers. Business passengers typically constitute a small percentage but are the main source of revenue in airlines. In comparison, the number of leisure passengers is higher [14]. Realizing the category by reviewing the literature is helpful to explain how potential determinants influence each passenger category. Previous studies have identified economy, population size, and geographical distance as the most crucial determinants of air-passenger flows [6, 10, 11, 15]. Given the vast expanse of the BRI region and the large disparity, we believe that, in addition to these determinants, cultural, institutional, and economic disparity must also be considered.

*2.1. Cultural Disparity and Air Passengers.* Although, during its implementation, BRI advocated for the promotion of cultural exchange [16], the manner and extent to which cultural disparity could influence the development of the BRI region remain unexplored. Before reviewing the literature, we propose the following definition for cultural disparity (CD): the cultural similarity between origins and destinations. A small CD indicates a cultural similarity between the two countries, while a large CD represents significantly different cultures. Regarding leisure passengers, several studies have explored the impact of CD on tourist flows [13, 17]. Mixed results were observed with respect to the relation between CD and passengers' destination selection [18, 19], indicating that CD could positively or negatively affect air-passenger flows. For instance, CD negatively impacts intercity air passengers, because passengers are more likely to visit destinations with cultures similar to their own [13].

In summary, CD affects both leisure and business behaviors and, thus, could be either an inhibitor or a motivator in air-passenger flows between origins and destinations. In reality, BRI countries have considerably different religions and languages (barriers), indicating large communication costs among these countries. Furthermore, most BRI countries are not primary trade partners. Thus, they are unwilling to overcome large communication costs for trade. Thus, we consider CD as a potential inhibitor in this analysis. Therefore, the following hypothesis was established:

*Hypothesis 1.* CD negatively affects air-passenger flows. A large CD between cities indicates small air-passenger flows.

*2.2. Institutional Disparity and Air Passenger.* In general, the institutional context varies significantly across countries [20]. Various institutions have invariably posed a special concern for multinational enterprises; their special regulatory environment may influence the foreign market access and entry mode, thereby inhibiting international business [21, 22]. For instance, a multinational enterprise may face legal restrictions on the number of equity shares to be bought in local businesses [23]. Differences in institutional arrangements may challenge the transfer of strategic organizational practices to their overseas subsidiaries and create hurdles in gaining legitimacy [24]. Moreover, the efficiency of the investment rules varies based on the infrastructure charge levied by private operators or the government [11]. Therefore, institutional differences may exacerbate the information asymmetry between the partners, the risk of partner opportunism, and the cost of doing business abroad [25]. While reviewing leisure passengers, multinational tour operators enjoy an advantage in attracting tourists to the countries they invest in due to their reputation [26]. Therefore, we assume that various institutions in the BRI countries could create obstacles to air passengers' exchange.

*Hypothesis 2.* ID negatively affects air-passenger flows. A large ID between cities indicates small air-passenger flows.

*2.3. Geographical Distance and Air Passengers.* Geographical distance (GD), an important concept in geography, has been identified by the location theory (proposed by Johann Heinrich von Thunen) and industrial location theory (proposed by Alfred Weber) has been identified as a valued variable used to analyze the distribution of agriculture, industry, and other economic activities. From a transportation-cost perspective, long-distance traveling generally leads to lesser demand for air travel, primarily owing to the high travel cost [10]. Leisure passengers are more sensitive to the price than business passengers [27]. From the viewpoint of cultural geography, residents of geographic locations that are in proximity to each other invariably have a similar language and religion [28], which stimulates air-passenger flows. Thus, we propose the third hypothesis as follows:

*Hypothesis 3.* GD has a complicated effect on air-passenger flows. The latter might be stimulated or limited by distance.

*2.4. Economic Disparity and Air Passengers.* Similar to the definition offered for CD, a large economic disparity (ED) indicates a large economic gap between the regions. It has been well acknowledged that income is the most common variable in studies on tourism [15]. Further studies have observed that airlines in developed countries have large capacity and demand [29, 30]. However, there is still a paucity of research on how the economic gap between the origin and destination affects the tourist demand. According to the World Bank statistics, the gap in GDP per capita between the BRI countries was up to 325 times in 2017. In general, the large economic gaps indicate differences in needs, thereby expanding the complementarity and promoting exchanges between the regions. Thus, the following hypothesis is proposed:

*Hypothesis 4.* A large ED has a positive effect on air-passenger flows.

## 3. Methodology

*3.1. Study Area.* A list of 198 cities from 66 BRI countries was compiled as per the following criteria: (1) The study area covered the six subregions in the Eurasian landmass (East Asia, South Asia, Central and Eastern Europe, West Asia, Southeast Asia, and Central Asia; see Figure 1), including 66 countries and 1,718 airports in 2018; (2) cities with airports operating more than 25 airlines were included; (3) the capital cities of all the 66 countries were represented. Data on weekly intercity flights helped map the intercity transport linkages. The dataset was gathered in the first week of August 2018 using Google Flight, with each record containing information on the flights, including schedule and the airline. Ideally, directed and annual data should be used in studies of this nature; however, owing to time constraints, our dataset was much smaller. Our overall goal was to estimate the weight differences in the edges. From the data collected on several weekly flights in August, we observed the following: (a) All binary connections were retrieved, and (b) demand did not show either a peak or a trough. Thus, this effect can be considered to be of minor relevance. Resultantly, data on the 147,970 nonstop flights linking 2,717 city pairs in one week were collected (Figure 2).

*3.2. Data Source.* The conceptual framework of this analysis is shown in Figure 3. The main aim of this analysis is to examine the influencing factors impacting intercity transport linkages in the BRI region. Since the BRI region includes six subregions and exhibits obvious subregional features in economic and flight distributions, the interpretation of the factors influencing air-passenger flows is distinguished by subregions.

*3.2.1. Cultural Disparity.* Two different measurement methods were adopted for calculating CD. In the first method, cultural difference indicators were constructed using databases, e.g., the European Values Survey and World Values Survey [32, 33]. However, the data on developing countries are not invariably available. The second method was modeled on Hofstede's cultural framework [34] which uses multidimensional indicators. In comparison, the second method offers comprehensive coverage and is thus extensively used.

Based on the comparison of the above-mentioned methods, we use Hofstede's cultural dimensions to construct a cultural integration index. The dimensions have considered the six indicators, namely, the power distance index (PDI), individualism versus collectivism (IDV), masculinity versus femininity (MAS), uncertainty avoidance index (UAI), long-term orientation (LTO), and indulgence versus restraint (IVR). The first four cultural indicators are widely applied to denote cultural status. Moreover, data on LTO and IVR are missing for most countries in the region. Thus, we selected the first four cultural indicators listed above to propose cultural blending variables curated for this analysis. The cultural scores of the 29 countries can be retrieved from Hofstede's homepage (https://www.hofstede-insights.com/). Missing data on some countries were collected from other studies that had calculated the same by applying the Hofstede method [35–38]. Finally, 59 countries were included in our CD analysis. It is calculated utilizing the following formula:

$$\mathrm{CD}_{ij} = \frac{\left(\sum_{m=1}^{4}\left[\left(C_{im} - C_{jm}\right)^2 / V_m\right]\right)}{4}, \tag{1}$$

where $i$ and $j$ represent the selected cities; $m$ represents the four indices reflecting the CD, while $V_m$ is the score variance of all involved countries on the $m^{\text{th}}$ dimension; $\mathrm{CD}_{ij}$ represents the cultural distance between cities $i$ and $j$; $C_{im}$ is Hofstede's score of the $m^{\text{th}}$ dimension of city $i$, while $C_{jm}$ is the same dimension's cultural score of city $j$. It is noteworthy that $\mathrm{CD}_{ij}$ equals 0 while cities $i$ and $j$ belong to the same country.

*3.2.2. Institutional Disparity.* The establishment of ID facilitates the calculation of the similarities and differences in institutions. The Worldwide Governance Indicators (WGI), which we retrieved from the World Bank dataset, are well acknowledged and were applied to calculate the distance [39–41]. The six-dimensional indices include the Control of Corruption (CC), Government Effectiveness (GE), Political Stability (PS), Regulatory Quality (RLQ), Rule of Law (RL), and Voice and Accountability (VA). ID was calculated using the following formula:

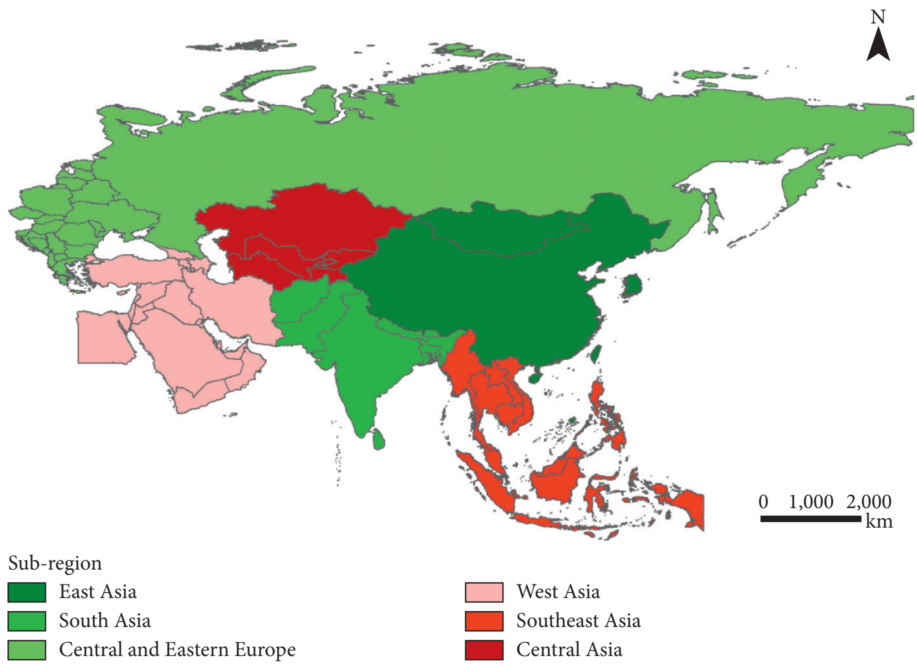



FIGURE 1: Scope of the "Belt and Road."

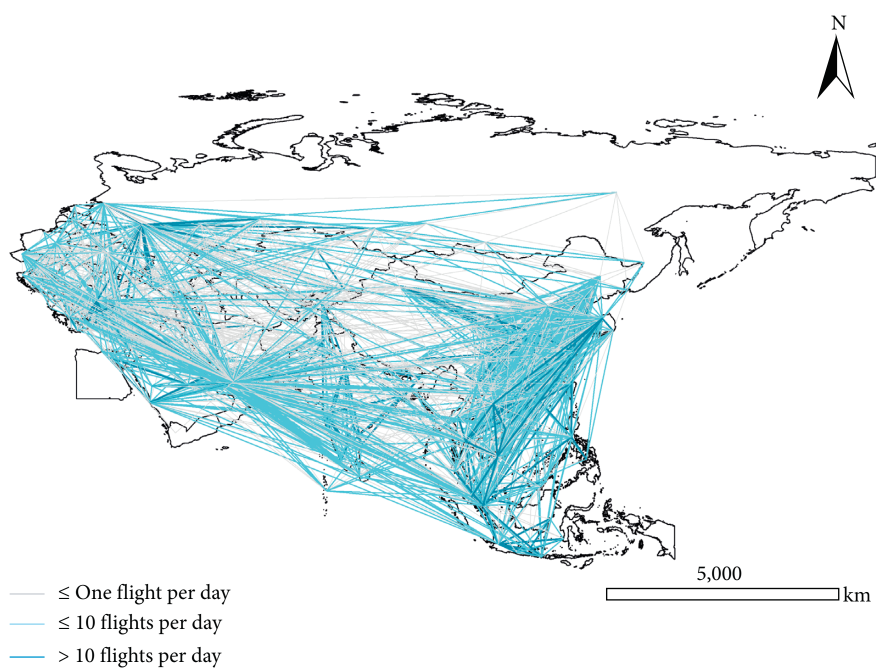



FIGURE 2: Flights in the BRI region [31].

$$\mathrm{ID}_{ij} = \frac{\left(\sum_{k=1}^{6}\left[\left(I_{ik} - I_{jk}\right)^2/V_k\right]\right)}{6}, \qquad (2)$$

where $i$ and $j$ represent the selected cities; $k$ represents the six dimensions of ID, while $V_k$ represents the score variance of all the involved countries in the $k^{\text{th}}$ dimension; $\mathrm{ID}_{ij}$ denotes the institutional distance between the two cities; $I_{ik}$ is the score of the $k^{\text{th}}$ dimension of city $i$, while $I_{jk}$ is the same dimension's score of city $j$. It is noteworthy that $\mathrm{ID}_{ij}$ equals 0 while cities $i$ and $j$ belong to the same country.

*3.2.3. Economic Disparity.* Economic size is derived from the Visible Infrared Imaging Radiometer Suite imagery nighttime light (NTL) data, and ED was calculated as follows:

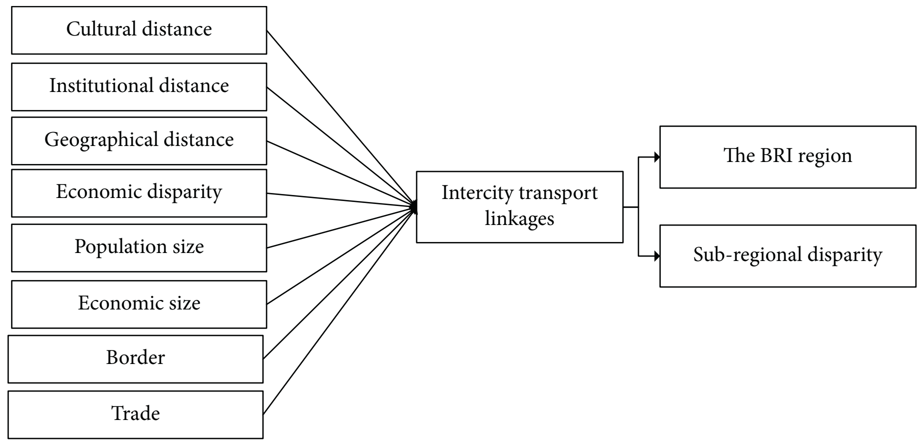



Figure 3: Conceptual framework.

$$\mathrm{ED}_{ij} = \frac{\left(L_i - L_j\right)^2}{L_i * L_j}, \tag{3}$$

where $L_i$ and $L_j$ represent NTL values of cities $i$ and $j$ and $\mathrm{ED}_{ij}$ denotes the economic disparity between the two cities.

*3.2.4. Other Factors.* The geographic distance between the cities is calculated in kilometers by applying the great circle method (Several methods are available to calculate the distance between two points on the surface. Among these, the great circle distance, calculated by the great circle method, is the shortest distance between two points on a sphere.). Both city boundary and population data were obtained from the website http://population.be/. The trade data were collected from the World Bank dataset.

*3.3. Model Specification.* A regression model will be estimated in this section with these factors as the independent variables and the air-passenger flows as the dependent variable. The gravity model has been widely used in explaining various interregional and international flows [42, 43]. Although the model has been criticized for its lack of theoretical foundation, several studies have investigated the relationship between economic theory and the gravity model [44, 45]. Thus, we applied the gravity model in this analysis. The ordinary least squares (OLS) method is adopted to perform estimations. Approximately, 183 cities from 56 countries were included, considering the data available in some countries while calculating the CD and ID.

The dependent variable in this analysis is the number of weekly flights between the selected cities. The independent variables include not only the multidistances (CD, ID, GD, and ED) that were initially developed but also some other variables. Specifically, the cities' economic size, population size, trade, and border (determining whether the flight is an intercity or an international one) were included. The gravity function was specified as follows:

$$\mathrm{Flow}_{ij} = f\left(\mathrm{CD}_{ij}, \mathrm{ID}_{ij}, \mathrm{ED}_{ij}, \mathrm{GD}_{ij}, E_i, E_j, \mathrm{POP}_i, \mathrm{POP}_j, \mathrm{trade}_i, \mathrm{trade}_j, \mathrm{border}_{ij}\right). \tag{4}$$

To estimate the coefficient elasticities, the regression function was calculated based on

$$\begin{aligned}\mathrm{Flow}_{ij} = {} & \alpha + \beta \mathrm{CD}_{ij} + \gamma \mathrm{ID}_{ij} + \theta \mathrm{ED}_{ij} + \varepsilon_1 \mathrm{GD}_{ij} \\ & + \varepsilon_2 \left(\mathrm{GD}_{ij}\right)^2 + \delta\left(E_i * E_j\right) + \varepsilon\left(\mathrm{POP}_i * \mathrm{POP}_j\right) \\ & + \varphi\left(\mathrm{trade}_i * \mathrm{trade}_j\right) + \vartheta \mathrm{border}_{ij},\end{aligned} \tag{5}$$

where $\mathrm{Flow}_{ij}$ is the number of direct flights in one week between cities $i$ and $j$; $\mathrm{CD}_{ij}$ represents the CD between cities $i$ and $j$; $\mathrm{ID}_{ij}$ denotes the ID between cities $i$ and $j$; $\mathrm{ED}_{ij}$ stands for the ED between cities $i$ and $j$. A quadratic function is applied in estimating GD's impact in which $\mathrm{GD}_{ij}$ demonstrates the GD between cities $i$ and $j$; $\mathrm{POP}_i$ and $\mathrm{POP}_j$ represent the populations of cities $i$ and $j$, respectively; $E_i$ and $E_j$ represent the economic sizes of cities $i$ and $j$, respectively; $\mathrm{trade}_i$ and $\mathrm{trade}_j$ show the trade values in the countries of cities $i$ and $j$; $\mathrm{border}_{ij}$ determines the city pair of a flight belonging to one country. Moreover, $\mathrm{border}_{ij} = 1$ if the city pair of one flight belongs to one country, while $\mathrm{border}_{ij} = 0$ if cities $i$ and $j$ belong to two different countries.

The above model was built to examine the influence of the distance of four dimensions as well as other determinants on air-passenger flows. Model 2, which was constructed based on the following equation, was to explore how each cultural factor (e.g., PDI, IDV, MAS, and UAI) and each institutional factor (e.g., CC, GE, RL, VA, PS, and RLQ) affect air-passenger flows:

$$\text{Flow}_{ij} = \alpha + \beta_1\text{PDI}_{ij} + \beta_2\text{IDV}_{ij} + \beta_3\text{MAS}_{ij} + \beta_4\text{UAI}_{ij} + \gamma_1\text{CC}_{ij} + \gamma_2\text{GE}_{ij} + \gamma_3\text{RL}_{ij} + \gamma_4\text{VA}_{ij} + \gamma_5\text{PS}_{ij} + \gamma_6\text{RLQ}_{ij} + \theta\text{ED}_{ij} + \varepsilon_1\text{GD}_{ij} + \varepsilon_2\left(\text{GD}_{ij}\right)^2 + \delta\left(E_i * E_j\right) + \varepsilon\left(\text{POP}_i * \text{POP}_j\right) + \varphi\left(\text{trade}_i * \text{trade}_j\right) + \vartheta\text{border}_{ij}, \tag{6}$$

where $\text{CC}_{ij}, \text{GE}_{ij}, \text{RL}_{ij}, \text{VA}_{ij}, \text{PS}_{ij}, \text{LQ}_{ij}$ represent the disparities of CC, GE, RL, VA, PS, and RLQ between cities *i* and *j*, respectively; $\text{PDI}_{ij}$ represents the PDI disparity between cities *i* and *j*; $\text{IDV}_{ij}$, $\text{MAS}_{ij}$, and $\text{UAI}_{ij}$ show the disparity of IDV, MAS, and UAI between cities *i* and *j*, respectively.

All the models were completed using the SPSS software.

## 4. Results

*4.1. Descriptive Statistics, Reliability, and Correlation.* Table 1 gives the definition of the variables. Table 2 presents the correlation results between two potential variables and the collinear diagnostic results. As shown, all variance inflation factor (VIF) values are below 4, and tolerance values are larger than 0.657, thereby indicating little cause for concern on the major multicollinearity among these independent variables.

*4.2. Regression Results in the BRI Region.* In Model 1, CD and ID are examined as determinants, whereas their original indices (e.g., PDI, IDV, and CC) are examined as determinants in Model 2. Here, the model is adopted to (1) test the impacts of CD, ID, and their original indices and (2) determine whether the various factors are stable in the different models. The results from the two models are tabulated (Table 3). Two original observations are drawn from these two models. First, the determination coefficients for the two models are high. The independent variables together account for more than 39.9% of the variations of each model. Second, the air-passenger flows (dependent indicator) significantly correlate with the independent variables, indicating that the regression results in the two models are satisfactory.

By examining the hypotheses proposed earlier, the following results were obtained:

(1) The hypothesis that CD has a negative impact on air-passenger flows is well supported. Specifically, the CD coefficient in Model 1 is −0.068, indicating a negative impact on air-passenger flows and the expected negative signs. While the regression analysis is traced back to the four indices that are used to calculate the CD, not all of them show the same negative impact. The results suggest that IDV and MAS negatively influence air-passenger flows. Specifically, the IDV coefficients in Model 2 amount to −0.055. The MAS coefficient in Model 2 is −0.033. The PDI and UAI demonstrate no significant effect.

(2) Hypothesis 2, in which ID negatively impacts air-passenger flows, is well supported. The ID coefficient in Model 1 measures −0.058, indicating a consistent and expected effect on air-passenger flows. While the regression analysis is traced back to the six original indices explored in Model 2 (used to calculate the ID), they indicate quite consistent effects. Only CC negatively impacts the air-passenger flows, with a coefficient of −0.172. In contrast, the RL (0.108) and VA (0.029) coefficients in Model 2 have a positive effect on air-passenger flows. Moreover, GE, PS, and RLQ in Model 3 show no significant impact on air-passenger flows.

(3) Hypothesis 3, concerning GD having a complicated impact on air-passenger flows, is well supported. The significant results in the quadratic function are observed, indicating an inverted U-shaped relation between GD and air-transport flows. The air-passenger flows in the BRI region are accompanied by a trend of geographical distance increasing first and decreasing afterward. Furthermore, GD has a more significant impact on the flows compared to CD and ID.

(4) Hypothesis 4, in which ED has a positive impact on air-passenger flows, is well supported. Both models demonstrate positive coefficients, indicating that city pairs with a larger economic disparity are more attractive than those with a small economic disparity. While comparing the effects shown in the two models, considerably similar coefficients (ranging from 0.022 to 0.03) are obtained.

Furthermore, economic and population sizes positively affect air-passenger flows. As indicated by the two models, city pairs with higher economic and population sizes are more attractive than those with lower sizes. Our results also indicate that the border has a significant impact on passenger flows, suggesting that the number of air passengers tends to be higher if a city pair is located in the same country. Moreover, the impact of trade on air passengers tends to be unstable. Specifically, while in Model 1 the impact of trade on air-transport passengers is negative, in Model 2 the impact becomes insignificant.

*4.3. Regression Results in the Subregions.* The regression in subregions is conducted under Model 1, and the adjusted R-squared is shown in Table 4. Three striking observations can be made here: (1) In general, the $R^2$ on the diagonal is higher than that on the nondiagonals, indicating that the model is relatively more applicable in explaining the flows within the subregions; (2) in the intrasubregional linkages, the regression model is more applicable in interpreting the flows within Southeast and South Asia, with $R^2$ being 0.553 and 0.543, respectively. Regarding the flows within East Asia and Central and Eastern Europe, $R^2$ is 0.387 and 0.356, respectively. The $R^2$ value in West Asia is the lowest, at 0.326; (3) while interpreting the flows between the subregions, the regression model is more applicable in interpreting those

Table 1: Description of variables.

| Index | Abbreviations | Description | Maximum values | Minimum values | Means | Standard deviations |
|---|---|---|---|---|---|---|
| Power distance index | PDI | In score | 104 | 13 | 71.1 | 20.62 |
| Individualism versus collectivism | IDV | In score | 80 | 14 | 36.37 | 15.55 |
| Masculinity versus femininity | MAS | In score | 112 | 9 | 49.88 | 22.54 |
| Uncertainty avoidance index | UAI | In score | 112 | 8 | 66.97 | 21.37 |
| Control of corruption | CC | In score | 2.07 | −1.67 | −0.23 | 0.78 |
| Government effectiveness | GE | In score | 2.21 | −1.82 | 0.03 | 0.78 |
| Political stability | PS | In score | 1.53 | −2.91 | −0.31 | 1.01 |
| Regulatory quality | RLQ | In score | 2.18 | −2.09 | 0.00 | 0.85 |
| Rule of law | RL | In score | 1.83 | −2.01 | −0.11 | 0.82 |
| Voice and accountability | VA | In score | 1.20 | −2.13 | −0.43 | 0.91 |
| Geographical distance | GD | In kilometers | 11567.28 | 224 | 4550.27 | 2619.27 |
| Population size | Pop | In thousand persons | 2449 | 1.2 | 254.16 | 336.26 |
| Economic size | $E$ | In nighttime light volume ($10^4$ nanowatts/cm$^2$/sr) | 915.59 | 0.00 | 9.55 | 30.78 |
| Trade | Trade | In US dollar thousand | 2395400 | 0 | 86629.24 | 285953.09 |

Table 2: Summary of the descriptive statistics and correlation of the variables.

| | CD | ID | ED | GD | E | Pop | Trade | Border | VIF | Tolerance |
|---|---|---|---|---|---|---|---|---|---|---|
| CD | 1.00 | | | | | | | | 1.52 | 0.658 |
| ID | 0.04 | 1.00 | | | | | | | 1.03 | 0.967 |
| ED | 0.03 | 0.00 | 1.00 | | | | | | 1.03 | 0.975 |
| GD | −0.15 | 0.04 | −0.09 | 1.00 | | | | | 1.42 | 0.707 |
| E | −0.09 | −0.02 | −0.09 | 0.05 | 1.00 | | | | 1.20 | 0.832 |
| Pop | −0.06 | 0.02 | −0.01 | −0.03 | 0.33 | 1.00 | | | 1.30 | 0.770 |
| Trade | 0.13 | 0.11 | 0.01 | 0.04 | 0.00 | −0.06 | 1.00 | | 2.42 | 0.414 |
| Border | 0.39 | 0.09 | 0.03 | 0.30 | 0.05 | −0.29 | 0.70 | 1.00 | 3.45 | 0.289 |

Table 3: Estimation results for the OLS model.

| Variables | Model 1 | Model 2 |
|---|---|---|
| CD | −0.068 (***) | |
| PDI | | NS |
| IDV | | −0.055 (***) |
| MAS | | −0.033 (***) |
| UAI | | NS |
| ID | −0.058 (***) | |
| CC | | −0.172 (***) |
| GE | | NS |
| RL | | 0.108 (***) |
| VA | | 0.029 (***) |
| PS | | NS |
| RLQ | | NS |
| ED | 0.03 (***) | 0.022 (***) |
| GD | 0.662 (***) | 0.669 (***) |
| GD$^2$ | −0.907 (***) | −0.912 (***) |
| Economic size | 0.116 (***) | 0.09 (***) |
| Pop | 0.187 (***) | 0.182 (***) |
| Border | 0.369 (***) | 0.385 (***) |
| Trade | −0.032 (***) | NS |
| $R^2$ | 0.399 | 0.408 |
| Significance | *** | *** |

***, **, and * indicate that the correlations are significant at 0.01, 0.05, and 0.1 levels, respectively. NS demonstrates that no significance has been found. $R^2$ in this analysis refers to adjusted $R^2$. The sample size $N = 16836$ in Models 1 and 2.

between South Asia and West Asia ($R^2 = 0.381$) and less applicable between Southeast Asia and Central and Eastern Europe ($R^2 = 0.034$). With regard to the flows between East Asia and the other subregions, the $R^2$ between Central and East Asia is the highest (0.342), while that between West and East Asia is the lowest (0.04). Regarding the flows between Central Asia and Southeast and South Asia, the regression models are insignificant. Taking into account the flows with Southeast Asia, East Asia displays the highest $R^2$ (0.265), followed by South Asia ($R^2 = 0.161$) and West Asia ($R^2 = 0.118$). Regarding the flows with South Asia, West Asia has the highest $R^2$ (0.381, as mentioned above), followed by Southeast Asia ($R^2 = 0.161$). The lowest $R^2$ occurs in flows with East Asia (0.037), while the value is insignificant in flows with Central Asia. With regard to the flows between Central and Eastern Europe and the other subregions, Central Asia has the best fit at 0.268, followed by East and West Asia, with $R^2$ being 0.102 and 0.074, respectively. In the flows between West Asia and the subregions, South Asia shows the best fit with $R^2$ equaling 0.381, followed by Central ($R^2 = 0.236$) and Southeast Asia ($R^2 = 0.118$). In comparison, the value of $R^2$ between West Asia and East Asia is 0.04, indicating the lowest value.

In general, the impact of CD on the air flows varies within the subregions (Table 5). Among them, the flows within South Asia have the largest negative coefficient, which is 0.265, followed by the flows between South and West Asia. The negative influence coefficient of the CD is 0.217. The coefficient of 0.196 has a negative impact on the flows within East Asia. Furthermore, the impact of the CD on the flows between the subregions is not invariably negative. Positive coefficients are generated within West Asia and between West Asia and Central and Eastern Europe, with coefficients of 0.098 and 0.07, respectively.

The impacts of the ID on the flows show immense disparity between and within the subregions (Table 6). Among all negative impacts, the coefficient between South Asia and West Asia is the largest, at 0.269. The negative coefficients between Southeast Asia and South Asia on the one hand and Central and Eastern Europe on the other are 0.149 and 0.101, respectively. Furthermore, the negative coefficient between East and West Asia is 0.096. In contrast, positive effects emerge in the flows within East Asia and between East Asia and Central and Eastern Europe, with $R^2$ equaling 0.244 and 0.058, respectively.

The impacts of ED on the flows within and between the subregions are primarily positive (Table 7). It is noteworthy that the flows between West Asia and the other subregions are positively affected by the economic disparity. The flows between West Asia and Central Asia have the largest positive impact coefficients, amounting to 0.235. While referring to the flows within South Asia and Central and Eastern Europe, the negative impact of the ID is observed.

The impacts of GD on the flows within and between the subregions adhere to quadratic equations (Tables 8 and 9). Specifically, the coefficients of $GD^2$ within all subregions, except Central Asia, are below 0 (Table 8). Furthermore, the coefficients of $GD^2$ between West Asia and South Asia and Central and Eastern Europe are below 0. Moreover, the coefficients between East Asia and Southeast Asia are also below 0. These negative values indicate that the flows initially increase and subsequently decrease as the distance augments. Contrarily, the coefficients of $GD^2$ between East Asia and Central Asia and Central and Eastern Europe are above 0, indicating that the flows decrease first and increase afterward as the distance increases.

The impacts of the economic and population size on the flows within and between the subregions are positive. As shown in Tables 10 and 11, all coefficients are above 0. Specifically, economic size has the most significant impact on the flows between East and Central Asia, with a coefficient of 0.325. In contrast, it has the lowest impact on the flows between East Asia and Central and Eastern Europe. In terms of the population size, the flows within Southeast Asia are influenced the most (with a coefficient of 0.526), whereas the flows between South Asia and Central and Eastern Europe (with a coefficient of 0.084) are affected the least.

While observing the impact of trade on the flows within and between the subregions (Table 12), it is noteworthy that the flows within West Asia and between West Asia and other subregions are positively influenced by trade, except the flows between East Asia and West Asia. Further, regarding the flows within and between East Asia and other subregions, trade has negative impacts.

## 5. Discussion

By constructing multidimensional distances, this analysis hypothesized about the impacts of CD, ID, GD, and ED on the air-passenger flow formation in the BRI region. Applying a gravity model, the test results of the hypotheses were obtained, which are described in Table 13 with several significant findings.

Previous studies have claimed that a large CD may promote complementary trade between countries [46–48]. The positive impact of CD on international trade could be explained from the perspective of firms being more sensitive to the benefits brought about by complementary trade. The comparative advantages offered by complementary trade invariably outweigh the extra cost of the cultural barriers. However, some evidence suggests that companies tend to expand their business to countries that are culturally similar to their existing business branches to reduce the associated uncertainty and risks [49]. Consistent with the conclusions drawn in a previous study [50], our analysis showed that cultural disparity has a significant negative effect on tourist flows. Specifically, IDV and MAS have negative impacts on air-passenger flows. Thus, air passengers in the BRI region tend to be generalized among cities with similar IDV and MAS. First, IDV scores are high in the European countries (e.g., Hungary [80], Latvia [70], and Poland [60]) compared to those in some Asian countries (e.g., China [20], India [48], Indonesia [14], Pakistan [14], South Korea [18], and Thailand [18]) [51, 52]. IDV refers to how people define themselves (in terms of "I" or "we") and their relationship with others. Thus, a low IDV score denotes more collective traits. IDV is significantly and negatively correlated to air-passenger flows, and the size effect is the largest (0.055 in

Table 4: Adjusted $R$-squared for the gravity model within and between subregions.

| | East Asia | Central Asia | Southeast Asia | South Asia | Central and Eastern Europe | West Asia |
|---|---|---|---|---|---|---|
| East Asia | 0.387 (***) | | | | | |
| Central Asia | 0.342 (***) | — | | | | |
| Southeast Asia | 0.265 (***) | NS | 0.553 (***) | | | |
| South Asia | 0.037 (***) | NS | 0.161 (***) | 0.543 (***) | | |
| Central and Eastern Europe | 0.102 (***) | 0.268 (***) | 0.034 (***) | NS | 0.356 (***) | |
| West Asia | 0.04 (***) | 0.236 (***) | 0.118 (***) | 0.381 (***) | 0.074 (***) | 0.326 (***) |

The number of samples from the five Central Asian countries is small, so there is no value for $R^2$ between the Central Asian cities; ***, **, and * indicate that the correlations are significant at 0.01, 0.05, and 0.1 levels, respectively. NS demonstrates that no significance has been found (also applied in Tables 5–12).

Table 5: Results of OLS regression analysis of the CD within and between subregions.

| | East Asia | Central Asia | Southeast Asia | South Asia | Central and Eastern Europe | West Asia |
|---|---|---|---|---|---|---|
| East Asia | −0.196 (**) | | | | | |
| Central Asia | −0.17 (***) | — | | | | |
| Southeast Asia | NS | NS | NS | | | |
| South Asia | −0.091 (**) | NS | NS | −0.265 (***) | | |
| Central and Eastern Europe | NS | −0.143 (*) | NS | NS | −0.089 (***) | |
| West Asia | NS | NS | −0.081 (**) | −0.217 (***) | 0.07 (***) | 0.098 (**) |

Table 6: Results of OLS regression analysis of the ID within and between subregions.

| | East Asia | Central Asia | Southeast Asia | South Asia | Central and Eastern Europe | West Asia |
|---|---|---|---|---|---|---|
| East Asia | 0.244 (**) | | | | | |
| Central Asia | NS | — | | | | |
| Southeast Asia | NS | NS | NS | | | |
| South Asia | NS | NS | −0.149 (***) | NS | | |
| Central and Eastern Europe | 0.058 (**) | NS | −0.101 (***) | NS | NS | |
| West Asia | −0.096 (***) | NS | NS | −0.269 (***) | NS | NS |

Table 7: Results of OLS regression analysis of the ED within and between subregions.

| | East Asia | Central Asia | Southeast Asia | South Asia | Central and Eastern Europe | West Asia |
|---|---|---|---|---|---|---|
| East Asia | NS | | | | | |
| Central Asia | NS | — | | | | |
| Southeast Asia | NS | NS | NS | | | |
| South Asia | NS | NS | 0.091 (**) | −0.089 (*) | | |
| Central and Eastern Europe | NS | NS | NS | NS | −0.051 (**) | |
| West Asia | NS | 0.235 (**) | 0.092 (**) | 0.101 (***) | 0.067 (***) | NS |

Table 8: Results of OLS regression analysis of the GD within and between subregions.

| | East Asia | Central Asia | Southeast Asia | South Asia | Central and Eastern Europe | West Asia |
|---|---|---|---|---|---|---|
| East Asia | 2.178 (***) | | | | | |
| Central Asia | −8.112 (***) | — | | | | |
| Southeast Asia | NS | NS | 1.224 (***) | | | |
| South Asia | −0.971 (*) | NS | NS | 2.116 (***) | | |
| Central and Eastern Europe | −4.465 (***) | −0.393 | NS | NS | NS | |
| West Asia | NS | NS | NS | 2.15 (**) | 0.498 (**) | 1.65 (***) |

Model 2) compared to the other three indices. Second, societies with high MAS scores are driven by achievements, competition, and success, while those with a low score value life quality and happiness. Passengers from a country with a low MAS score have more leisure time, thus leading to a high tourist demand [53]. MAS is significantly and negatively correlated to air-passenger flows, with a coefficient of 0.033. Thus, air passengers prefer destinations with similar MAS scores.

ID negatively impacts air-passenger flows, as shown in Models 1 and 2. Specifically, CC has a negative influence on air-passenger flows, while RL and VA exhibit a positive influence. However, PS, RLQ, and GE show no significant influence. CC indicates the public's view on the use of public

Table 9: Results of OLS regression analysis of the $GD^2$ within and between subregions.

| | East Asia | Central Asia | Southeast Asia | South Asia | Central and Eastern Europe | West Asia |
|---|---|---|---|---|---|---|
| East Asia | −2.122 (***) | | | | | |
| Central Asia | 7.696 (***) | — | | | | |
| Southeast Asia | −0.946 (*) | NS | −1.443 (***) | | | |
| South Asia | NS | NS | NS | −2.374 (***) | | |
| Central and Eastern Europe | 4.336 (***) | −0.393 | NS | NS | −0.682 (**) | |
| West Asia | NS | NS | NS | −2.527 (***) | −0.653 (**) | −1.775 (***) |

Table 10: Results of OLS regression analysis of the economic size within and between subregions.

| | East Asia | Central Asia | Southeast Asia | South Asia | Central and Eastern Europe | West Asia |
|---|---|---|---|---|---|---|
| East Asia | NS | | | | | |
| Central Asia | 0.325 (***) | — | | | | |
| Southeast Asia | 0.293 (***) | NS | 0.257 (***) | | | |
| South Asia | 0.071 (**) | NS | 0.229 (***) | 0.106 (**) | | |
| Central and Eastern Europe | 0.052 (***) | NS | NS | NS | NS | |
| West Asia | 0.126 (***) | 0.306 (***) | 0.161 (***) | 0.115 (***) | NS | 0.289 (***) |

Table 11: Results of OLS regression analysis of population size within and between subregions.

| | East Asia | Central Asia | Southeast Asia | South Asia | Central and Eastern Europe | West Asia |
|---|---|---|---|---|---|---|
| East Asia | 0.479 (***) | | | | | |
| Central Asia | 0.377 (***) | — | | | | |
| Southeast Asia | 0.500 (***) | NS | 0.526 (***) | | | |
| South Asia | 0.192 (***) | NS | 0.308 (***) | 0.188 (***) | | |
| Central and Eastern Europe | 0.225 (***) | 0.231 (**) | 0.154 (***) | 0.084 (**) | 0.341 (***) | |
| West Asia | 0.159 (***) | 0.327 (***) | 0.162 (***) | NS | 0.19 (***) | 0.308 (***) |

Table 12: Results of OLS regression analysis of trade within and between subregions.

| | East Asia | Central Asia | Southeast Asia | South Asia | Central and Eastern Europe | West Asia |
|---|---|---|---|---|---|---|
| East Asia | −0.125 (***) | | | | | |
| Central Asia | −0.118 (***) | — | | | | |
| Southeast Asia | −0.07 (**) | NS | NS | | | |
| South Asia | −0.162 (***) | NS | NS | NS | | |
| Central and Eastern Europe | NS | NS | NS | NS | NS | |
| West Asia | NS | 0.263 (**) | 0.116 (***) | 0.13 (***) | 0.096 (***) | 0.177 (***) |

Table 13: Test results of the hypotheses.

| | Hypotheses | Estimated results |
|---|---|---|
| H1 | CD has a negative effect on air-passenger flows. | Supported |
| H2 | ID has a negative effect on air-passenger flows. | Supported |
| H3 | GD has complicated effects on air-passenger flows. | Supported |
| H4 | ED has a negative effect on air-passenger flows. | Supported |

power to meet private interests. As shown by Randrianarisoa et al., a country with high corruption hinders airport efficiency [54]. Therefore, countries tend to build more air-transport linkages to countries with similar CC scores. RL captures the perceptions of the extent to which citizens have confidence in and abide by the rules of society. VA indicates the degree to which citizens could participate in governance, expressing freedom. High scores of RL and VA guarantee the development of air infrastructure to some extent, which facilitates cultural exchanges.

Both ED and economy size positively impact air-passenger flows in the BRI region. The results pertaining to economic size suggest that the air-transport demand has a high income elasticity and could be defined as luxury from an economic perspective [27, 55]; this furthermore aligns with the findings of previous studies [56]. Air passengers, especially in international flows, are driven by economic growth to a large extent [15, 57, 58]. Thus, airlines in more developed countries have immense capacity and demand [29, 30]. In 2017, the GDP per capita in the BRI region was

Table 14: Regression results among subregions.

| | East Asia | Central Asia | Southeast Asia | South Asia | Central and Eastern Europe | West Asia |
|---|---|---|---|---|---|---|
| East Asia | −CD, ID, $-GD^2$, Pop, −trade | | | | | |
| Central Asia | −CD, $GD^2$, *E*, Pop, −trade | — | | | | |
| Southeast Asia | $-GD^2$, *E*, Pop, −trade | — | $-GD^2$, *E*, Pop | | | |
| South Asia | −CD, *E*, Pop, −trade | — | −ID, ED, *E*, Pop | −CD, −ED, $-GD^2$, *E*, Pop | | |
| Central and Eastern Europe | ID, $GD^2$, *E*, Pop | −CD, $-GD^2$, Pop | −ID, Pop | Pop | −CD, ED, $-GD^2$, Pop | |
| West Asia | −ID, *E*, Pop | ED, *E*, Pop, trade | −CD, ED, *E*, Pop, trade | −CD, −ID, ED, $-GD^2$, *E*, trade | CD, ED, $-GD^2$, Pop, trade | CD, $-GD^2$, *E*, Pop, trade |

5,558 dollars, which increased by 4.39 times that in 2000. A previous study also presented the fact that the given percentage of growth in GDP is matched by an identical growth percentage in annual travel [59]. Regarding ED, cities from a lower level of economic development tend to schedule flights with cities having a higher economic position to achieve development.

In previous studies, the impact of GD was found to be positive within a certain distance [60]. After a certain threshold, however, GD starts to negatively impact flows [61]. In our analysis, GD proves to have a similar irregular impact on air-passenger flows. Within a certain distance, air-passenger flow increases in accordance with increasing distance. This is induced by its advantages compared to other traffic modes (e.g., automotive). However, once a specific distance is exceeded, air-passenger flows are unable to break through beyond the distance decay theory and thus decrease along with the increment of the GD. This can be attributed to, among other things, transportation costs rising with increasing GD. Geographical distance is considerably more important than cultural disparity [62]. Moreover, population has an immense impact on air-passenger flows, indicating that a large population produces more potential holiday travelers. Population's positive influence on air-passenger flows could be explained from the perspective of demand, which is supported by the tourist demand and the abilities of satisfying the same [6]. Apart from the factors mentioned above, air-passenger flows are also susceptible to their domestic or international nature. Domestic flow accounts for the majority of the air-passenger flows in the BRI region since visiting friends/relatives and leisure traffic also constitute a large part of domestic travel. It is noteworthy that flows between some subregions (East and Central Asia, East Asia, and Central and Eastern Europe) decrease first and increase subsequently as GD increases. This phenomenon is explained from the perspective of the urban cluster, as economy and population have been concentrated in East China. Resultantly, these cities have introduced more frequent flights to cities in Central Asia and Central and Western Europe compared to cities in Central and West China.

Moreover, the impact of these factors on air-passenger flows varies based on the subregions and is significantly different from the one in the BRI region as a whole (Table 14). In comparison, the regression model is more applicable in explaining the flows within subregions than between them. Specifically, CD negatively impacts flows in the BRI region within and between most subregions, indicating that cultural ties continue to dominate flight patterns. According to previous studies, international transactions are influenced by not only the costs of overcoming physical distances, such as those of transportation and tariffs, but also the costs associated with the collection and interpretation of the information required to impact such transactions [63]. Thus, CD significantly dissuades firms from investing in foreign countries [64]. Furthermore, it positively impacts the flows within West Asia and between West Asia and Central and Eastern Europe. This can be associated with the tendency of countries with smaller territories to have lower domestic aviation market shares. Therefore, cities tend to arrange flights connecting to foreign cities that display a large cultural disparity. However, negative influences are also observed within South Asia and Central and Eastern Europe. In the context of GD, typically, the flows (e.g., within East Asia, South Asia, and Southeast Asia) increase and then decrease as the distance augments. This is consistent with the findings of previous studies where a distance decay effect was observed [61]. However, in comparison, the flows between East and Central Asia and East Asia and Central and Eastern Europe first decrease and then increase as the distance augments. Furthermore, economic size significantly impacts the flows between East and Central Asia, while population size has a similar effect on the flows between East Asia and other subregions. With the common factor being the East Asian cities, it can be deduced that the growth of population and economy size in these subregions drives outbound tourism and, consequently, increases the number of international flows.

## 6. Conclusions

The objective of this paper was twofold. First, we estimated the cultural and institutional factors influencing air-passenger flows. We conclude that CD and ID correlate negatively with air-passenger flows in the BRI region. Second, we examined the significance of the most widely used factors, such as the ED, border, economy, and population size. We conclude that air-passenger flows in the BRI region have positive relations with ED, border, economy, and population size. Third, a complex GD elasticity is confirmed. In general, flows are first positively and then negatively influenced, with a proportionate increase of GD. Fourth, the impact of these factors on air-passenger flows varies by subregion. Specifically, CD has positive impacts on air-passenger flows within West Asia and between Central and Eastern Europe and West Asia. ID has positive impacts on air-passenger flows within East Asia and between East Asia and Central and Eastern Europe. ED has negative impact on flows within South Asia, while GD has a negative impact first and a positive impact subsequently on air-transport flows between East and Central Asia and East Asia and Central and Eastern Europe.

Theoretically, this paper contributes novel insights into understanding air-passenger flows from the perspective of cultural and institutional disparity. Specifically, it contributes to the existing literature by examining how cultural proximity, which can be formidable in terms of its operationalization and interpretation, influences the formation of air-passenger flows. Practically, this paper briefly illustrates that air passengers in the BRI cities target destinations with similar cultural backgrounds and state institutions. However, the results do not intend to eliminate these differences between the BRI countries. Instead, cultural diversity should be respected, and equal communication between different cultures should be advocated. In view of this, enhanced policy communication is necessary to overcome the barriers erected by institutional differences and to foster and enhance mutual trust among the BRI countries. Moreover, an approach to building a successful low-fare airline business model in the BRI region is also significant for the flow since current air passengers in the BRI region continue to be sensitive to transportation costs.

Nevertheless, our study contains some limitations which may offer various scopes for future research. One of the limitations concerns the reliability of the flight data utilized in this analysis. Instead of an accurate passenger flow, weekly flight data have been utilized in this study. Moreover, neither the size of the airplanes nor the origin-destination flow has been taken into account. As a result, the difference in transport passenger caused by seasonal travel is unrecognizable. Thus, even though the actual physical movements can be reflected by the weekly flights to some extent, accurate passenger flow data are more ideal by comparison to comprehend the social and economic processes [14]. A natural progression of this analysis could be to assess more factors since the future of interurban transport will be determined by the interaction of consumer preferences, bilateral aviation policy, technological developments, and the availability of resources to meet mobility needs [65]. Moreover, by considering cultural and institutional differences, this analysis assumes that cities in the same country exhibit no cultural and institutional disparity. On the contrary, there are many autonomous regions within several countries (e.g., Guangxi and Tibet in China). As a result, even cultures and institutions within a country may show disparities. Additionally, it is acknowledged that the 2019 novel coronavirus (COVID-19) has caused irreparable loss to air-transport [66–68]. Therefore, being sufficiently aware of how to develop efficient strategies under emergency events is instrumental for the air-transport industry.

## Data Availability

The data collected during the study are freely available from Google flight search.

## Conflicts of Interest

The authors declare no conflicts of interest.

## Acknowledgments

This work was supported by the National Natural Science Foundation of China (nos. 42101210 and 42101170), Beijing Philosophy and Social Science Foundation (no. 21JJC023), and China Postdoctoral Science Foundation (no. 2020M680660).

## References

[1] NDR Commission, "Vision and actions on jointly building silk road economic belt and 21st-century maritime silk road," 2015, http://en.ndrc.gov.cn/newsrelease/201503/t20150330_669367.html.

[2] J. H. Choi, G. A. Barnett, and B.-S. Chon, "Comparing world city networks: a network analysis of Internet backbone and air transport intercity linkages," *Global Networks*, vol. 6, no. 1, pp. 81–99, 2006.

[3] L. Zhang, H. Du, X. Zhang, P. De Maeyer, B. Dessein, and Y. Zhao, "Robustness and edge addition strategy of air transport networks: a case study of "the belt and road"," *IEEE Access*, vol. 7, pp. 96470–96477, 2019.

[4] L. Zhang, Y. Zhao, D. Chen, and X. Zhang, "Analysis of network robustness in weighted and unweighted approaches: a case study of the air transport network in the belt and road region," *Journal of Advanced Transportation*, vol. 2021, Article ID 8810254, 13 pages, 2021.

[5] S. Vinciguerra, K. Frenken, and M. Valente, "The geography of internet infrastructure: an evolutionary simulation approach based on preferential attachment," *Urban Studies*, vol. 47, no. 9, pp. 1969–1984, 2010.

[6] F. Dobruszkes, M. Lennert, and G. Van Hamme, "An analysis of the determinants of air traffic volume for European metropolitan areas," *Journal of Transport Geography*, vol. 19, no. 4, pp. 755–762, 2011.

[7] L.-Y. Chang, "International air passenger flows between pairs of APEC countries: a non-parametric regression tree approach," *Journal of Air Transport Management*, vol. 20, pp. 4–6, 2012.

[8] L. Zhang, Y. Zhao, X. Zhang, and H. Du, "Effect mechanisms of peasant relocation decision-making behaviours in the process of rural spatial restructuring: the case of Hotan Region, China," *Sustainable Cities and Society*, vol. 63, Article ID 102429, 2020.

[9] Y. Zhao, L. Zhang, and X. Zhang, "The spatial distribution of retail outlets in Urumqi: the application of points of interest," *Open Geosciences*, vol. 12, no. 1, pp. 1541–1556, 2020.

[10] Y. Eilat and L. Einav, "Determinants of international tourism: a three-dimensional panel data analysis," *Applied Economics*, vol. 36, no. 12, pp. 1315–1327, 2004.
[11] S.-i. Mun and S. Nakagawa, "Pricing and investment of cross-border transport infrastructure," *Regional Science and Urban Economics*, vol. 40, no. 4, pp. 228–240, 2010.
[12] Y. Zhang, "International arrivals to Australia: determinants and the role of air transport policy," *Journal of Air Transport Management*, vol. 44-45, pp. 21–24, 2015.
[13] S. I. Ng, J. A. Lee, and G. N. Soutar, "Tourists' intention to visit a country: the impact of cultural distance," *Tourism Management*, vol. 28, no. 6, pp. 1497–1506, 2007.
[14] Z. Neal, "The backbone of bipartite projections: inferring relationships from co-authorship, co-sponsorship, co-attendance and other co-behaviors," *Social Networks*, vol. 39, pp. 84–97, 2014.
[15] C. Lim, "Review of international tourism demand models," *Annals of Tourism Research*, vol. 24, no. 4, pp. 835–849, 1997.
[16] S. Cui, "Belt and road initiative is also about cultural exchanges, official says," 2017, http://www.chinadaily.com.cn/china/19thcpcnationalcongress/2017-10/20/content_33492115.htm.
[17] C. Vietze, "Cultural effects on inbound tourism into the USA: a gravity approach," *Tourism Economics*, vol. 18, no. 1, pp. 121–138, 2012.
[18] M. Jackson, "Cultural influences on tourist destination choices of 21 pacific rim nations," in *Proceedings of the 11th Australian Tourism and Hospitality Research Conference, CAUTHE 2001: Capitalising on Research*, p. 166, Canberra, Australia, 2001.
[19] Y. Yang, H. Liu, and X. Li, "The influences of national cultural distance on international destination choice of Japanese and Chinese residents," *Tourism Tribune*, vol. 31, no. 10, pp. 45–55, 2016.
[20] M. Demirbag, M. McGuinness, and H. Altay, "Perceptions of institutional environment and entry mode," *Management International Review*, vol. 50, no. 2, pp. 207–240, 2010.
[21] D. Ionascu, K. E. Meyer, and S. Estrin, "Institutional distance and international business strategies in emerging economies," in *William Davidson Institute Working*no. 728, , p. 728, William Davidson Institute at the University of Michigan Stephen M. Ross Business School, Ann Arbor, MI, USA, 2004.
[22] J.-H. Bae and R. Salomon, "Institutional distance in international business research," in *The Past, Present and Future of International Business & Management*, T. Devinney, T. Pedersen, L. Tihanyi, T. M. Devinney, T. Pedersen, and L. Tihanyi, Eds., pp. 327–349, Emerald Group Publishing Limited, Cheltenham, UK, 2010.
[23] T. Kostova and S. Zaheer, "Organizational legitimacy under conditions of complexity: the case of the multinational enterprise," *Academy of Management Review*, vol. 24, no. 1, pp. 64–81, 1999.
[24] S. Elsner, "Study 1 - effects of institutionalized entry modes on entry mode choices," in *Retail Internationalization*, pp. 41–74, Springer Fachmedien Wiesbaden, Wiesbaden, Germany, 2014.
[25] J.-F. Hennart, "The transaction costs theory of joint ventures: an empirical study of Japanese subsidiaries in the United States," *Management Science*, vol. 37, no. 4, pp. 483–497, 1991.
[26] J. Zhang and C. Jensen, "Comparative advantage: explaining tourism flows," *Annals of Tourism Research*, vol. 34, no. 1, pp. 223–243, 2007.
[27] M. Brons, E. Pels, P. Nijkamp, and P. Rietveld, "Price elasticities of demand for passenger air travel: a meta-analysis," *Journal of Air Transport Management*, vol. 8, no. 3, pp. 165–175, 2002.
[28] M. Domosh, R. P. Neumann, P. L. Price, and T. G. Jordan-Bychkov, *The Human Mosaic: A Cultural Approach to Human Geography*, Macmillan, New York, NY, USA, 2011.
[29] K. Button and S. Taylor, "International air transportation and economic development," *Journal of Air Transport Management*, vol. 6, no. 4, pp. 209–222, 2000.
[30] J. T. Bowen, "The economic geography of air transport," in *The Geographies of Air Transport*, A. R. Goetz and L. Budd, Eds., pp. 57–76, Routledge, London, UK, 2016.
[31] L. Zhang, H. Du, Y. Zhao, P. De Maeyer, B. Dessein, and X. Zhang, "Drawing topological properties from a multi-layered network: the case of an air transport network in "the Belt and Road" region," *Habitat International*, vol. 93, Article ID 102044, 2019.
[32] R. Inglehart and W. E. Baker, "Modernization, cultural change, and the persistence of traditional values," *American Sociological Review*, vol. 65, no. 1, pp. 19–51, 2000.
[33] R. Inglehart, *Islam, Gender, Culture, and Democracy: Findings from the World Values Survey and the European Values Survey*, Vol. 4, De Sitter Publications, , Whitby, Canada, 2003.
[34] G. Hofstede, *Culture's Consequences: International Differences in Work-Related Values*, Sage, Beverly Hills, CA, USA, 1984.
[35] I. Piniuta, "Belarusian mentality in the context of intercultural communication: grounds or pessimism and optimism," *Sustainable Multilingualism*, vol. 11, no. 1, pp. 34–53, 2017.
[36] A. Rodriguez and A. Brown, "Cultural differences: a cross-cultural study of urban planners from Japan, Mexico, the U.S., Serbia-Montenegro, Russia, and South Korea," *Public Organization Review*, vol. 14, no. 1, pp. 35–50, 2014.
[37] M. I. M. Irfan, "Cultural dimensions of Hofstede and their impact on organizational performance in Sri Lanka," *Imperial Journal of Interdisciplinary Research*, vol. 2, no. 10, pp. 1160–1169, 2016.
[38] C. Rarick, G. Winter, G. Falk, I. Nickerson, and C. Barczyk, "Afghanistan's younger, elite, and educated population: a cultural assessment and possible implications for the economic and political future of the country," *Global Journal of Management and Business Research*, vol. 14, no. 3, pp. 27–33, 2013.
[39] D. Kaufmann, A. Kraay, and M. Mastruzzi, "The worldwide governance indicators: methodology and analytical issues," *Hague Journal on the Rule of Law*, vol. 3, no. 2, pp. 220–246, 2011.
[40] M. Aleksynska and O. Havrylchyk, "FDI from the south: the role of institutional distance and natural resources," *European Journal of Political Economy*, vol. 29, pp. 38–53, 2013.
[41] L. Zhang and Z. Xu, "How do cultural and institutional distance affect China's OFDI towards the OBOR countries?" *Baltic Journal of European Studies*, vol. 7, no. 1, pp. 24–42, 2017.
[42] I.-H. Cheng and H. J. Wall, "Controlling for heterogeneity in gravity models of trade and integration," *Federal Reserve Bank of St. Louis Review*, vol. 87, no. 1, pp. 49–63, 2005.
[43] J. Khadaroo and B. Seetanah, "The role of transport infrastructure in international tourism development: a gravity model approach," *Tourism Management*, vol. 29, no. 5, pp. 831–840, 2008.
[44] J. E. Anderson and E. Van Wincoop, "Gravity with gravitas: a solution to the border puzzle," *The American Economic Review*, vol. 93, no. 1, pp. 170–192, 2003.
[45] J. E. Anderson, "The gravity model," *Annual Review of Economics*, vol. 3, no. 1, pp. 133–160, 2011.
[46] J. R. Markusen, "Factor movements and commodity trade as complements," *Journal of International Economics*, vol. 14, no. 3-4, pp. 341–356, 1983.

[47] M. Lankhuizen, H. L. F. De Groot, and G.-J. M. Linders, "The trade-off between foreign direct investments and exports: the role of multiple dimensions of distance," *The World Economy*, vol. 34, no. 8, pp. 1395–1416, 2011.

[48] M. B. M. Lankhuizen and H. L. F. de Groot, "Cultural distance and international trade: a non-linear relationship," *Letters in Spatial and Resource Sciences*, vol. 9, no. 1, pp. 19–25, 2016.

[49] B. Kogut and H. Singh, "The effect of national culture on the choice of entry mode," *Journal of International Business Studies*, vol. 19, no. 3, pp. 411–432, 1988.

[50] Y. Yang and K. K. F. Wong, "The influence of cultural distance on China inbound tourism flows: a panel data gravity model approach," *Asian Geographer*, vol. 29, no. 1, pp. 21–37, 2012.

[51] G. Hofstede, ""Hofstede Center," National Cultural Dimensions," 2010, http://geert-hofstede.com/.

[52] H. Orlady, *Human Factors in Multi-Crew Flight Operations*, Routledge, London, UK, 2017.

[53] B. M. Farahani and B. Mohamed, "The influence of national culture on tourists' behaviour towards environment," *WIT Transactions on Ecology and the Environment*, vol. 148, pp. 573–582, 2011.

[54] L. M. Randrianarisoa, D. Bolduc, Y. Y. Choo, T. H. Oum, and J. Yan, "Effects of corruption on efficiency of the European airports," *Transportation Research Part A: Policy and Practice*, vol. 79, pp. 65–83, 2015.

[55] R. R. Clewlow, J. M. Sussman, and H. Balakrishnan, "The impact of high-speed rail and low-cost carriers on European air passenger traffic," *Transport Policy*, vol. 33, pp. 136–143, 2014.

[56] E. Marrocu and R. Paci, "Different tourists to different destinations. evidence from spatial interaction models," *Tourism Management*, vol. 39, pp. 71–83, 2013.

[57] X. Fu, T. H. Oum, and A. Zhang, "Air transport liberalization and its impacts on airline competition and air passenger traffic," *Transportation Journal*, vol. 49, no. 4, pp. 371–390, 2010.

[58] D. Gillen, "Aviation economics and forecasting," in *Air Transport Management: An International Perspective*, L. Budd and S. Ison, Eds., pp. 95–121, Routledge, New York, NY, USA, 2016.

[59] Y. Crozet, "The prospects for inter-urban travel demand," *OECD/ITF Joint Transport Research Centre Discussion Papers*, no. 14, 2009.

[60] L. Zhang, H. Du, Y. Zhao, R. Wu, and X. Zhang, "Urban networks among Chinese cities along "the Belt and Road": a case of web search activity in cyberspace," *PLoS One*, vol. 12, no. 12, Article ID e0188868, 2017.

[61] B. McKercher, "The effect of distance decay on visitor mlx at coastal destinations," *Pacific Tourism Review*, vol. 2, no. 3-4, pp. 215–223, 1998.

[62] L. Håkanson and B. Ambos, "The antecedents of psychic distance," *Journal of International Management*, vol. 16, no. 3, pp. 195–210, 2010.

[63] B. Ambos and L. Håkanson, "The concept of distance in international management research," *Journal of International Management*, vol. 20, no. 1, pp. 1–7, 2014.

[64] H. Berry, M. F. Guillén, and N. Zhou, "An institutional approach to cross-national distance," *Journal of International Business Studies*, vol. 41, no. 9, pp. 1460–1480, 2010.

[65] G. De Rus, "Interurban passenger transport: economic assessment of major infrastructure projects," *OECD/ITF Joint Transport Research Centre Discussion Paper*, no. 18, 2009.

[66] F. Serrano and A. Kazda, "The future of airports post COVID-19," *Journal of Air Transport Management*, vol. 89, Article ID 101900, 2020.

[67] X. Sun, S. Wandelt, and A. Zhang, "How did COVID-19 impact air transportation? A first peek through the lens of complex networks," *Journal of Air Transport Management*, vol. 89, Article ID 101928, 2020.

[68] S. Maneenop and S. Kotcharin, "The impacts of COVID-19 on the global airline industry: an event study approach," *Journal of Air Transport Management*, vol. 89, Article ID 101920, 2020.

Hindawi
Complexity
Volume 2021, Article ID 9933582, 12 pages
https://doi.org/10.1155/2021/9933582

*Research Article*

# Exploring Coupling Relationship between Urban Connection and High-quality Development Using the Case of Lanzhou-Xining Urban Agglomeration

**Junfeng Yin**[1] **Haimeng Liu**[2] **Peiji Shi,**[1] **and Weiping Zhang**[1]

[1]*College of Geography and Environmental Science, Northwest Normal University, Lanzhou 730070, China*
[2]*Institute of Geographical Sciences and Natural Resources Research, Chinese Academy of Sciences, Beijing 100101, China*

Correspondence should be addressed to Haimeng Liu; liuhm@igsnrr.ac.cn

Received 27 August 2021; Revised 12 October 2021; Accepted 21 October 2021; Published 3 November 2021

Academic Editor: Ning Cai



Based on socioeconomic statistical data, transport data, and network big data, the urban connection index (UCI) was constructed in terms of industry, transportation, information, and innovation, and the high-quality development index (HDI) was established from five aspects: innovation, coordination, green development, openness, and sharing. Taking Lanzhou-Xining urban agglomeration as a case, the urban connection intensity and high-quality development level were measured to analyze the relationship between them. From 2012 to 2018, the UCI and HDI of each city showed different degrees of growth. Note that there exist significant regional differences, with Lanzhou and Xining having the largest difference. The biggest gap among cities lies in the innovative connection intensity. Moreover, urban external connections are closely related to high-quality development, especially innovation and green development. For every 1% increase in industrial and transport connection, the HDI will increase by 0.317% and 0.159%, respectively. This study provides a methodological reference for measuring urban connectivity and provides decision support for high-quality development in China and other countries.

## 1. Introduction

An urban system is characterized by openness, nonlinearity, diversity, fluctuation, and self-organization [1, 2]. The openness is directly revealed by the continuous exchange of matter, energy, and information between different cities. This external connection allows the urban system to constantly replenish energy, which is one of the main driving forces to keep its vitality and orderliness [3, 4]. In urban agglomerations, cities are particularly interconnected [5], and the closeness between cities is an important indicator in reflecting the development degree of urban agglomerations. In 2017, China proposed the concept of high-quality development, aiming to transform from seeking "growth" to seeking "better growth" with general improvement in socioeconomic development. High-quality development will be the theme of China's socioeconomic development in the 14th Five-Year Plan (2021–2025) and beyond. Following the high-quality development path means staying committed to the people-oriented approach and the concept of innovative, coordinated, green, open, and shared development [6]. Currently, urban agglomerations in China comprise 75% of the population and 80% of the gross domestic product (GDP), being major spatial carriers of socioeconomic development [7]. Therefore, it is necessary to explore the relationship between urban connection intensity and high-quality development level in typical urban agglomerations for regional sustainable development [8, 9].

How to measure urban connection intensity and high-quality development level? What is the relationship between these two systems? Existing studies have mainly focused on the quantitative measurement of urban networks and connections. Different sources of flows were used to reflect the functional linkages of cities, such as migration [10], transport [11, 12], trade [13], and information [14]. The POLYNET project measured the degree of cohesion of urban agglomerations in Europe by analyzing daily commuting flows [15]. The intensity, direction, network pattern, and hierarchical structure of the connections between cities have been discussed [16–19]. The measurement methods of urban connection intensity mainly include the gravity model [20, 21], urban flow intensity model [22], regional accessibility [23], and social network analysis [24, 25]. However, most studies focused only on one dimension of urban connections, lacking comprehensive analysis. In addition, the concept of high-quality development proposed by the Chinese government has only gained attention recently. High-quality development in the new era is expected to coordinate the relationship between speed and quality of development [26]. That is, total factor productivity needs to be increased and old growth drivers should be replaced [27]. Long et al. evaluated the high-quality development of the Yellow River basin from five aspects of society, economy, resources, ecology, and culture [28]. Although the influence of foreign direct investment characteristics on China's high-quality economic development has been studied [29], there is no consensus on the understanding and measurement of urban high-quality development.

Considering the shortcomings of current research, the urban connection index (UCI) and high-quality development index (HDI) were constructed based on the complexity theory. Taking Lanzhou-Xining urban agglomeration as a case, the connection intensity and high-quality development level of each city from 2012 to 2018 were evaluated. Then, we explained quantitatively and qualitatively the coupling and coordination relationship between UCI and HDI. Finally, the popularization significance and policy implications were discussed. This study provides a methodological reference for measuring urban connectivity and provides decision support for high-quality development in China and other countries.

## 2. Theoretical Basis

A city is a complex system like an organism, and its inflow and outflow of matter, energy, and information are metabolisms [30–32]. Complexity science believes that the openness of a system and the fluidity of internal system elements are the driving forces for the vitality and continuous evolution of all complex systems on the Earth: a basic law of nature [33]. Similarly, as an open complex organic regional system, the city also obeys this law. There is an uninterrupted material, people, energy, and information flow within a city or between different cities [34]. The various "flows" between cities are driven by active and passive forces, forming intricate network connections between cities. The key to assessing a city's metabolic vitality is to quantify the intensity of these flows. The stronger the interaction between cities is, the higher the socioeconomic output of the cities is [35]. Especially, in the Information Age characterized by a new network society and informational cities, this interaction is becoming more significant for urban development [36].

A two-way positive and negative feedback effect occurs between UCI and HDI (Figure 1). The strengthened external connections of the urban system promote and support the

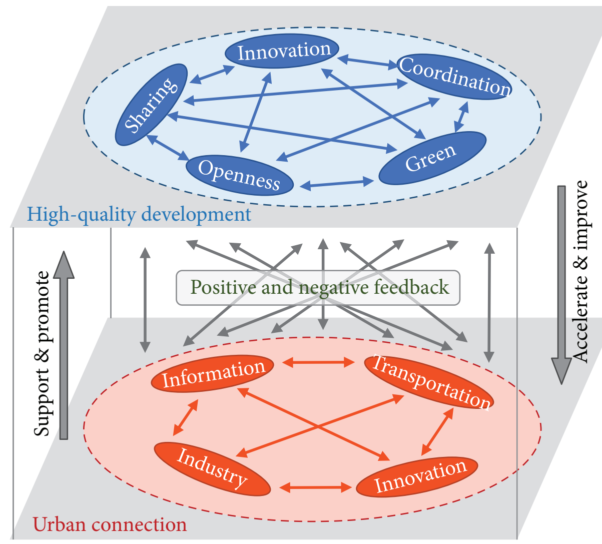



Figure 1: Relationship between urban connection and high-quality development.

high-quality development of a city. The urban high-quality development will in turn promote the openness of the city. At the city scale, the higher the openness of a city system is, the higher the mobility of various elements within the city is and the higher the livability and sustainability of the city are [37]. At the urban agglomeration scale, the higher the connections between cities in terms of industry, transport, information, and innovation are, the stronger the vitality and competitiveness of the urban agglomeration and the cities within it are and the greater the high-quality development of the region will be. Thus, we believe cities' external connections may be related to their high-quality development. Quantitative verification will be performed below.

## 3. Methodology and Data

*3.1. Study Area.* The Lanzhou-Xining urban agglomeration is one of the 19 important urban agglomerations proposed in the 13th Five-Year Plan of China (2016–2020). It is a typical urban agglomeration in the initial development stage in central and western China, and the Chinese government aims to build a growth pole to support the regional development. Located at the intersection of three natural regions in China, the eastern monsoon region, arid, and semiarid region, Qinghai-Tibet Plateau is an important transportation hub and economic channel connecting the central and western China with eastern China with numerous trunk railways and national highways (Figure 2). It covers an area of $9.75 \times 10^4\,km^2$. The total population was 11.94 million, and the urbanization rate was 49.79% by 2019. The GDP in 2019 was about 577.18 billion yuan, accounting for 49.40% of the sum of Gansu and Qinghai province.

*3.2. Urban Connection Intensity Measurement.* The flow of people, matter, and energy is mainly realized through industrial and traffic links between cities and the flow of information through the Internet, telecommunication

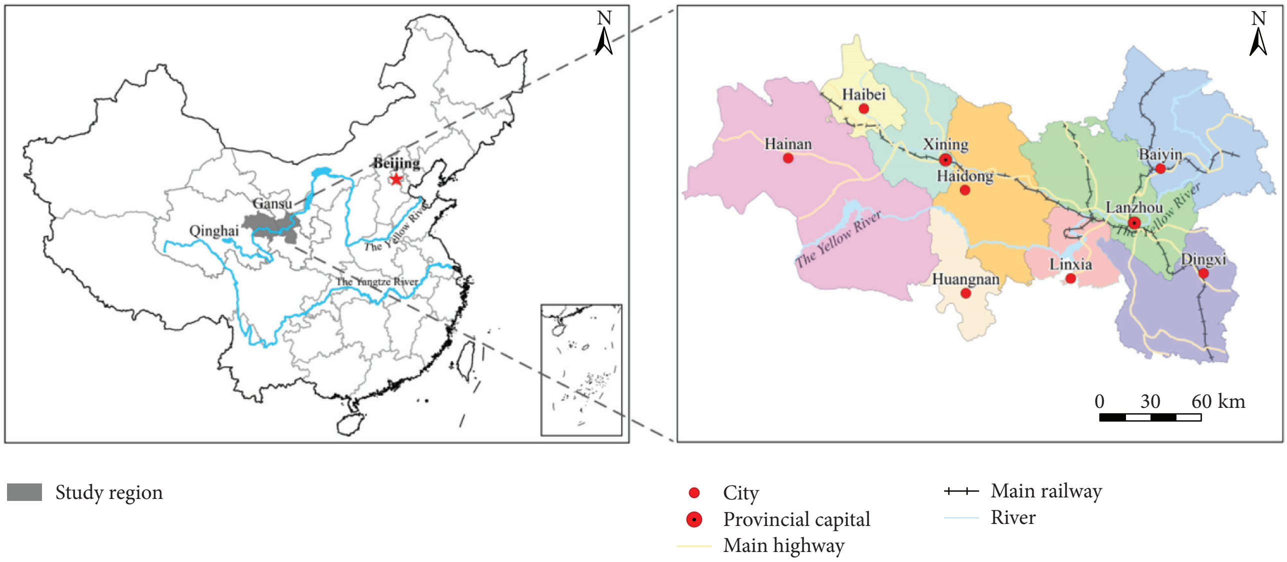



Figure 2: Lanzhou-Xining urban agglomeration.

networks, and other channels [38, 39]. Each city is one node in a complex network formed by various "flows" between cities [40]. Therefore, we chose four dimensions of industry, transportation, information, and innovation to comprehensively measure the external connection intensity of a city.

*3.2.1. Industrial Connection Intensity.* It is calculated based on the industrial influence coefficient and induction coefficient [41]. Influence coefficient refers to the amount of output that needs to be increased by other industrial sectors after one unit of the final product is added [42]. The formula is as follows:

$$Tt_j = \frac{\sum_{i=1}^{n} b_{ij}}{(1/n)\sum_{i=1}^{n}\sum_{j=1}^{n} b_{ij}}, \quad i, j = 1, 2, 3, \ldots, n, \tag{1}$$

where $Tt_j$ represents the influence coefficient of $j$ industrial sector, $b_{ij}$ is the coefficient of the $i$th row and $j$th column in the Leontief inverse matrix, and $n$ is the number of industries: $\sum_{i=1}^{n} b_{ij}$ is the sum of the $j$th column of the matrix and is the average of the sum of the columns of the matrix.

Induction coefficient refers to the amount of output required by a certain sector to meet the production demand of all sectors after each sector in the region adds one unit of the final product [43]. The formula is as follows:

$$St_i = \frac{\sum_{j=1}^{n} b_{ij}}{(1/n)\sum_{i=1}^{n}\sum_{j=1}^{n} b_{ij}}, \quad i, j = 1, 2, 3, \ldots, n, \tag{2}$$

where $St_i$ represents the induction coefficient of the sector $i$ affected by other industrial sectors. $\sum_{j=1}^{n} b_{ij}$ is the sum of the $i$th row of the matrix.

Drawing on previous studies [44, 45], we established industrial connection intensity based on "spatial connection" and "influence." The industrial connection intensity is expressed by actual connection, which is related to the industrial scale. The formula is as follows:

$$Ci_k = \sum_{i=1}^{18} St_i Tt_i Sc_i a_m, \tag{3}$$

$$Ci_{k1-k2} = \frac{\sqrt{C_{ik1} C_{ik2}}}{D_{k1-k2}}, \tag{4}$$

$$E_{p-IC} = \sum_{i=1}^{m} Ci_{k1-k2}, \tag{5}$$

where $Ci_k$ represents the industrial connection intensity of $n$ industries of city $k$ and its value can reflect the external radiation capacity of the city's industrial development, $Sc_i$ represents the scale of employment in each industry, $a_m$ represents the weight of each industry $i$ obtained by the analytic hierarchy process (AHP), $Ci_{k1\text{-}k2}$ indicates the intensity of industrial connections between cities $k_1$ and $k_2$, $D_{k1\text{-}k2}$ indicates the distance between two cities, calculated using the shortest expressway operating mileage between two cities, $E_{p-IC}$ indicates the total intensity of industrial connections in city $p$, and $m$ is the number of cities in the urban agglomeration.

*3.2.2. Transportation Connection Intensity.* The intensity of intercity transportation connection determines the connection degree of people, matter, and other flows [46]. Considering the density of train connections in Lanzhou-Xining urban agglomeration is low and many cities still rely on low-grade highways, the number of passenger buses running between cities is added into the model. The formula is as follows:

$$T_{pq} = aG_{pq} + bD_{pq} + cK_{pq} + dB_{pq}, \tag{6}$$

$$E_{pq} = \frac{T_{pq} + T_{qp}}{2}, \tag{7}$$

$$E_{p-TC} = \sum_{q=1}^{m} E_{pq}, \tag{8}$$

where $G_{pq}$, $D_{pq}$, $K_{pq}$, *and* $B_{pq}$, respectively, indicate the daily number of G-series high-speed railway, D-series high-speed railway, ordinary railway, and passenger buses from city $p$ to city $q$. $a$, $b$, $c$, and $d$ represent the coefficients of G-series high-speed railway, D-series high-speed railway, ordinary railway, and passenger buses. To scientifically quantify the connection intensity of different vehicles, the weight according to the speed of various vehicles was calculated. Taking the speed of G-series high-speed rail (speed of 300 km/h) as the benchmark, $a$ is taken as 1, correspondingly $b$ is 250/300 = 5/6, $c$ is 120/300 = 2/5, and $d$ is 80/300 = 4/15. $T_{pq}$ represents the transportation connection intensity from city $p$ to city $q$; $E_{pq}$ represents the average value of transportation connection between city $p$ and city $q$; $E_{p\text{-}TC}$ represents the total intensity of transportation connections in city $p$.

*3.2.3. Information Connection Intensity.* Using Baidu search index [47], we collected the number of searches for each other between the two cities from 2012 to 2018 in Lanzhou-Xining urban agglomeration. The spatial matrix of the Baidu search index for two cities was constructed, and then, the total amount of information connections is calculated by [48]

$$E_{p-NC} = \sum_{q=1}^{m} P_q \times Q_p, \tag{9}$$

where $P_q$ represents the number of searches of city $p$ to city $q$, $Q_p$ represents the number of searches of city $q$ to city $p$, and $E_{p\text{-}NC}$ indicates the information connection intensity of city $p$.

*3.2.4. Innovative Connection Intensity.* Using the Web of Science database, we adopted centrality to reflect the position of cities in the knowledge innovation network of the urban agglomeration [49]. We searched for the co-authored published papers between cities in the database during the research period. The formula is as follows:

$$E_{p-INC} = \sum_{q=1}^{m} \text{Citypair}_{qp}, \tag{10}$$

where $\text{Citypair}_{qp}$ represents the number of co-published papers between city $p$ and city $q$ and $E_{p\text{-}INC}$ represents the innovative connection intensity of city $p$.

*3.2.5. UCI.* We constructed the UCI based on the four indexes of $E_{p\text{-}IC}$, $E_{p\text{-}TC}$, $E_{p\text{-}NC}$, and $E_{p\text{-}INC}$. Using AHP [50], the weight of industry, transportation, information, and innovation are set as 0.283, 0.266, 0.234, and 0.217, respectively. The formula of UCI is as follows:

$$Z_{p-UCI} = \sum_{s=1}^{4} E_{ps} W_s, \tag{11}$$

where $Z_{p\text{-}UCI}$ represents the urban connection index of the $p$th city, $E_{ps}$ represents the standardized value of the connection intensity of the $s$th index of the $p$th city, and $W_s$ represents the weight of the $s$th index.

*3.3. A Comprehensive Index System for High-Quality Development.* High-quality development is one of China's top priorities in the new era [51]. The evaluation system of urban high-quality development is based on China's new development concept, namely, "innovation, coordination, green, openness, and sharing," the core of which is to balance the speed and quality of development [52]. Drawing on previous studies [28], combined with data availability and representativeness, we constructed the comprehensive index system. The index weight was determined by expert consultation [53] and the entropy method [54], and then, the average of the two was calculated as the comprehensive weight (Table 1).

The HDI is calculated as follows:

$$H\ DI = \sum_{j=1}^{m} W_j Y_{ij}, \tag{12}$$

where $Y_{ij}$ represents the standardized processing value of the $j$th index in the $i$th city, $W_j$ represents the weight of the $j$th index, and $m$ is the number of indexes.

*3.4. Coupling Degree and Coordination Degree.* The concept of "coupling" refers to the interaction and mutual influence of two or more systems [34]. We used the coupling degree and coordination degree to explore the possible relationship between UCI and HDI. The formula is as follows:

$$C = \sqrt{\frac{u_1 \times u_2}{(u_1 + u_2)^2}}, \tag{13}$$

where $C$ represents the coupling degree, $u_1$ represents the UCI, and $u_2$ is the HDI. The formula of coordination degree is as follows:

$$D = \sqrt{C \times T}, \tag{14}$$

$$T = \alpha u_1 + \beta u_2, \tag{15}$$

where $D$ represents the coordination degree, $T$ represents the comprehensive evaluation index of the two systems, and $\alpha$ and $\beta$ represent the undetermined coefficient (we assumed that the two systems are equally important; both $\alpha$ and $\beta$ were assigned 1/2). Referring to previous research [55], the coupling degree and coordination degree were classified, as listed in Table 2.

*3.5. Data Sources.* The socioeconomic data were derived from The Statistical Yearbook of Chinese Cities, Statistical Yearbook of Gansu and Qinghai Province, Industry Classification of National Economy (GB/T 4754-2017), and statistical yearbook of each city. The railway traffic data were obtained from the Ministry of Transport (https://www.mot.gov.cn/), the National Railway Administration (http://www.nra.gov.cn/), and the website of China Railway 12306 (https://www.12306.cn). Highway traffic operation data were obtained from The Atlas of China's Expressways and Urban and Rural Roads and the official websites (http://jtys.gansu.gov.cn/ and http://jtyst.qinghai.gov.cn/). Passenger

Table 1: Evaluation system of urban high-quality development.

| Goal layer | Criterion layer | Weight | Indexes | Index mark | Index attributes | Comprehensive weight |
|---|---|---|---|---|---|---|
| Urban high-quality development index system | Innovation development | 0.2 | Number of patent applications per 10,000 people | A1 | + | 0.072 |
| | | | R&D internal expenditure as a percentage of GDP | A2 | + | 0.059 |
| | | | R&D personnel per 10,000 people | A3 | + | 0.069 |
| | Coordination development | 0.2 | Urban-rural income ratio | B1 | − | 0.058 |
| | | | GDP per capita | B2 | + | 0.063 |
| | | | The proportion of urban population in total population | B3 | + | 0.079 |
| | Green development | 0.2 | $PM_{2.5}$ concentration | C1 | − | 0.069 |
| | | | Green coverage rate in built-up area | C2 | + | 0.072 |
| | | | Energy consumption per unit of GDP | C3 | − | 0.059 |
| | Openness development | 0.2 | The proportion of total import and export in GDP | D1 | + | 0.070 |
| | | | Actual foreign investment as a proportion of GDP | D2 | + | 0.068 |
| | | | Number of internet broadband access users per 10,000 people | D3 | + | 0.062 |
| | Sharing development | 0.2 | Number of beds in health institutions per 10,000 people | E1 | + | 0.062 |
| | | | Public library collections per 10,000 people | E2 | + | 0.065 |
| | | | Road area per capita | E3 | + | 0.073 |

Table 2: The classification of coupling degree and coordination degree.

| Coupling | Coupling phase | Coordination | Coordination level |
|---|---|---|---|
| $0.0 < C \leq 0.3$ | Low coupling stage | $0.00 \leq D < 0.19$ | Severe imbalance |
| $0.3 < C \leq 0.5$ | Antagonism stage | $0.20 \leq D < 0.39$ | Mild disorder |
| $0.5 < C \leq 0.8$ | Run-in stage | $0.40 \leq D < 0.59$ | General coordination |
| $0.8 < C \leq 1.0$ | Highly coupling stage | $0.60 \leq D < 0.89$ | Well coordination |
| | | $0.80 \leq D < 1.00$ | Advanced coordination |

frequencies were obtained from Qunar (https://www.qunar.com/) and Ctrip (https://www.ctrip.com/). Data related to Baidu Search Index were from its official website (http://index.baidu.com). The number of co-published papers was determined from the Web of Science database. $PM_{2.5}$ concentration data were obtained from the Dalhousie University Atmospheric Composition Analysis Group.

## 4. Results

### *4.1. UCI of Lanzhou-Xining Urban Agglomeration*

*4.1.1. Industrial Connection Intensity.* The intensity of industrial connections among cities in Lanzhou-Xining urban agglomeration has been rising, from an average of 226.08 in 2012 to 502.65 in 2018 (Figure 3(a)). The intensity of industrial connections among cities in the region varied greatly. Cities in the eastern part of the urban agglomeration (Lanzhou, Baiyin, Dingxi, and Linxia) were more closely connected with the external industry than cities in the western part (Xining, Haidong, Hainan, Haibei, and Huangnan). The industrial connection intensities in Hainan, Haibei, and Huangnan were all below 200. During the study period, Lanzhou was the strongest and Hainan the weakest of the urban agglomeration in industrial connection intensity, and the gap between the two widens year by year.

*4.1.2. Transportation Connection Intensity.* Lanzhou has always had the largest total transportation connections in the urban agglomeration, which increased from 54.27 in 2012 to 132.20 in 2018 (Figure 3(b)). Lanzhou, Xining, and Dingxi are important transportation hubs, and the intensity of transportation connections between the three cities has always been at the top. However, the transportation connection of Huangnan and Haibei is relatively weak, and the growth is slow. In the urban agglomeration, only Lanzhou and Dingxi have G-series high-speed trains, and Lanzhou, Xining, and Dingxi have D-series high-speed trains. Generally, cities located along the main railway lines have the closest transportation connections. The peripheral cities are weakly connected with the core cities.

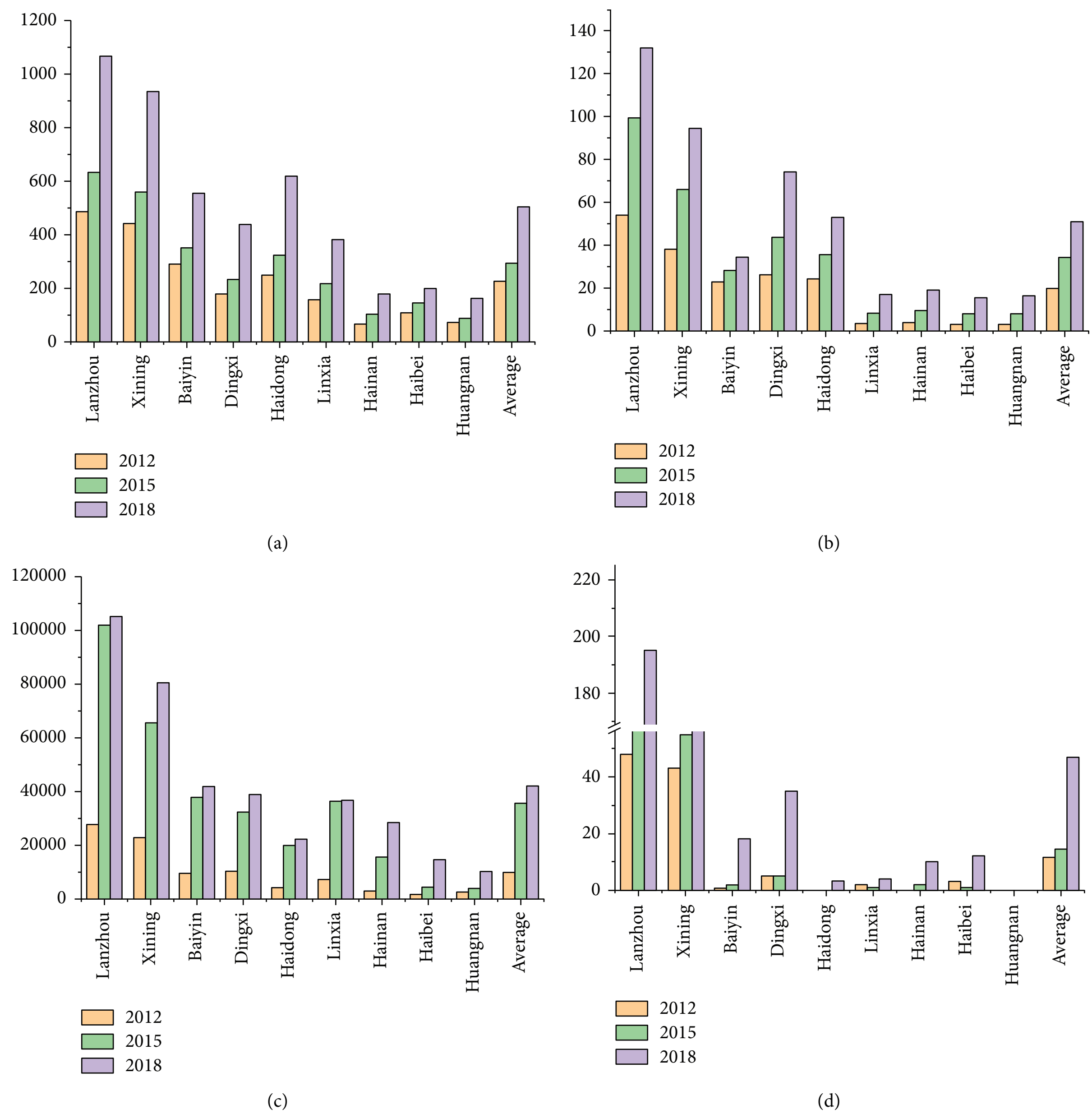



Figure 3: The four dimensions of urban connection indexes in Lanzhou-Xining urban agglomeration. (a) Industrial connection, (b) transportation connection, (c) information connection, and (d) innovative connection.

*4.1.3. Information Connection Intensity.* From 2012 to 2018, the information connection intensity of all cities in Lanzhou-Xining Urban agglomeration showed an increasing trend, but the rate of increase differed (Figure 3(c)). The intensity of information connection for Lanzhou and Xining has always been superior. The information connection intensity of Huangnan was the worst. Hainan's information links are growing at a rapid pace. Overall, the information connection intensity in Gansu province was much higher than that in Qinghai province.

*4.1.4. Innovative Connection Intensity.* The innovative connection intensity decreased from the core cities of the urban agglomeration such as Lanzhou and Xining to the surrounding cities, presenting a cliff-like decline; an obvious dual-center spatial structure was established (Figure 3(d)). Lanzhou and Xining have abundant higher education institutions and scientific research institutions, offering many high-level innovation platforms, which are conducive to the cities' innovative connections. The western region has poor conditions for scientific research and innovation, leading to the relatively low innovative connection intensity.

*4.1.5. UCI.* From 2012 to 2018, the UCI of each city in Lanzhou-Xining urban agglomeration showed an increasing trend, but the speed and amplitude of the rise differed. Cities with the fastest increase in the UCI were Huangnan and Hainan, with an average growth rate of over 6.60% and

1.21%, respectively (Figure 4). The UCI of Lanzhou was 2.84 times that of the average of the urban agglomeration, indicating that Lanzhou is closely connected with other cities and plays a leading role in the urban agglomeration. Xining had the second-highest UCI, and the UCI of Huangnan has always been at the bottom. The radiation and driving effects of Lanzhou and Xining on the surrounding cities are becoming stronger. The average UCI has been steadily increasing, reflecting that the spatial connection within urban agglomeration is getting closer.

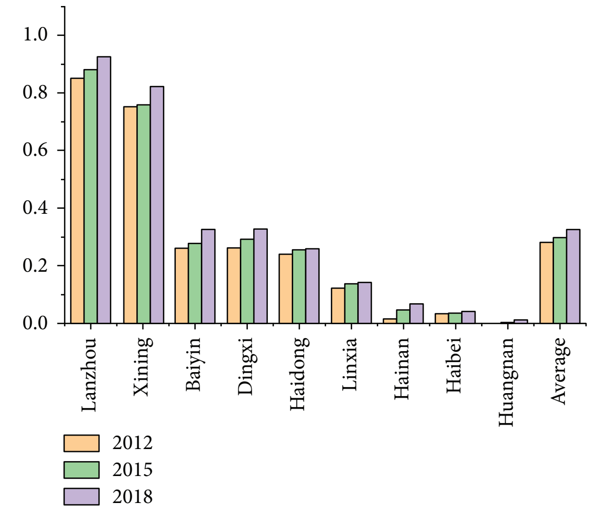



Figure 4: Urban connection index in Lanzhou-Xining urban agglomeration.

*4.2. The High-Quality Development of Lanzhou-Xining Urban Agglomeration.* There are obvious spatial differences in the HDI and four dimensions of all cities in the Lanzhou-Xining urban agglomeration (Figure 5). Lanzhou and Xining are significantly better than other cities in all dimensions. The innovation index of Lanzhou was about 5 times the average of the urban agglomeration. The gap in the green development index among cities was small, and Huangnan had the lowest score. For the dimension of openness development, the scores of Linxia, Hainan, Haibei, and Huangnan scored were lower than 0.01, nearly five times lower than the highest value. The variation in the shared development index was markedly different from that in other dimensions, showing the spatial characteristics of "low in the middle and high on both sides."

At the time scale, the HDI of each city showed an upward trend during the study period, but the growth rates differed. Lanzhou has always been the first in terms of the HDI, whereas Linxia and Haibei have always been at the bottom. The highest HDI value was 4.6 times higher than the lowest HDI. The largest gap was in innovation development, and the smallest gap was seen in green development and shared development. The variation coefficient (VC) of the HDI reflects the degree of imbalance in spatial development in the urban agglomeration. Table 3 shows that the imbalance reduced from 2012 to 2018, and the VC decreased from 0.847 to 0.697. Overall, the HDI of the eastern cities in the urban agglomeration was higher than that of the western cities.

### *4.3. Coupling Relationship between UCI and HDI*

*4.3.1. Pearson Correlation Analysis.* UCIs were significantly correlated with most indexes of high-quality development, especially innovation development and green development (Table 4), indicating that the enhanced connection of cities is likely to facilitate development. The urban connection was strongly correlated with the level of innovation development and green development. The three indexes of innovation development were significantly related to the three subsystems of urban connection; the proportion of actual use of foreign direct investment in GDP was only significantly related to industrial and information connection; the urban connection subsystem did not significantly correlate with the urban-rural income ratio and the number of books in public libraries per 10,000 people.

*4.3.2. Panel Regression Analysis.* Based on the panel data of UCI and HDI, combined with the data structure and the Hausman test results, we selected a random-effects regression model to test the relation between HDI and four UCIs. As shown in Table 5, lnIC, lnTC, lnNC, and lnINC represent the logarithmic form of industry, transportation, information, and innovation, respectively. The industrial and traffic connection had a significant positive impact on the HDI, while the information and innovative connection had no significant impact.

The elasticity coefficient between industrial connection and HDI was 0.317, indicating that, for every 1% increase in the degree of industrial connection of a city, its high-quality development level increased by about 0.317%. The elasticity coefficient of a city's external transportation connection and HDI was 0.159, indicating that every 1% increase in a city's transportation connection increases its high-quality development level by 0.159%. Increased transportation connections could accelerate the free flow of production factors such as population, logistics, and information between cities and increase in the openness and sharing of cities. The information connection and the innovation connection had no significant impact on the HDI, but the force was positive, indicating that an enhancement of the information and innovative connection may improve the urban high-quality development level to a certain extent.

*4.3.3. Coupling Degree and Coordination Degree.* The coupling degree of all cities showed an upward trend during 2012–2018 (Figure 6), but the magnitude of increase varied greatly. The coupling degree of Huangnan increased rapidly, from 0.071 in 2012 to 0.346 in 2018, but it was still at the bottom. The coupling degree of other cities varied less, with values in the range of 0.37–0.49, and maintained the same coupling stage during the study period. The coordination degree of Huangnan, Haibei, and Hainan was below 0.2, which represents a serious imbalance state; Lanzhou and

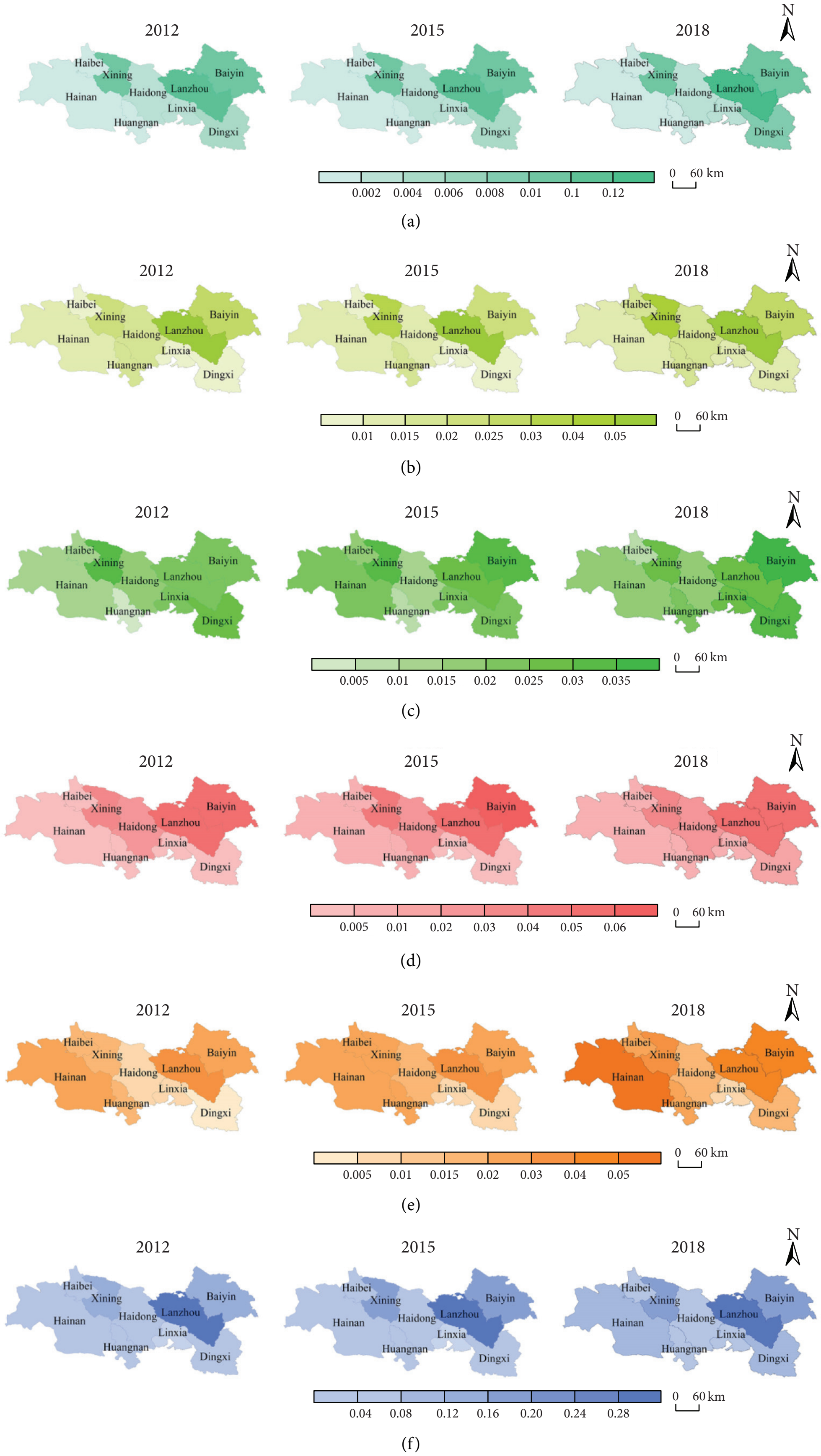



Figure 5: Spatial differentiation of the high-quality development level in Lanzhou-Xining urban agglomeration.

Table 3: The high-quality development index of cities, 2012–2018.

| Year | Lanzhou | Baiyin | Dingxi | Linxia | Xining | Haidong | Hainan | Haibei | Huangnan | Coefficient of variation |
|---|---|---|---|---|---|---|---|---|---|---|
| 2012 | 0.284 | 0.160 | 0.050 | 0.038 | 0.149 | 0.067 | 0.055 | 0.042 | 0.043 | 0.847 |
| 2015 | 0.289 | 0.175 | 0.050 | 0.047 | 0.179 | 0.069 | 0.064 | 0.052 | 0.059 | 0.781 |
| 2018 | 0.318 | 0.197 | 0.085 | 0.057 | 0.185 | 0.080 | 0.086 | 0.059 | 0.075 | 0.697 |

Table 4: Correlation test between UCI and HDI in Lanzhou-Xining urban agglomeration.


Notes: red color indicates significance at 1% level; green color indicates significance at 5% level; gray color indicates not significant.

Xining had a higher coordination degree. Baiyin, Dingxi, Linxia, and Haidong were in a state of mild disorder.

From 2012 to 2018, the degree of coupling and coordination between UCI and HDI of cities both increased; the values were [0.38, 0.44] and [0.24, 0.30], respectively, indicating that there is an antagonistic stage and a mild disorder between the external connection and urban high-quality development. The degree of coupling and coordination both increased, indicating that the degree of interaction and mutual influence between the two has deepened. In 2012, the general coordination areas included Lanzhou and Xining, while the severely imbalanced cities included Huangnan, Haibei, Hainan, and Linxia, and the mildly disordered cities included Baiyin, Dingxi, and Haidong. In 2018, only Huangnan and Haibei remained as the severely imbalanced cities, and the number of mildly disordered areas increased. The coordination degree between UCI and HDI of Lanzhou and Xining has always been high.

Table 5: Regression results of HDI and four UCIs.

| Variable | Coefficient estimates |
|---|---|
| lnIC | $0.317^{*}$ (0.172) |
| lnTC | $0.159^{**}$ (0.072) |
| lnNC | 0.013 (0.021) |
| lnINC | 0.023 (0.017) |
| Constant | $-4.989^{***}$ (1.072) |
| R-squared | 0.876 |
| City FE | Yes |
| Year FE | No |
| Number of observations | 27 |

Note: $^{***}$, $^{**}$, and $^{*}$ indicate significant correlation at 1%, 5%, and 10% confidence level. The numbers in parentheses indicate the standard error.

## 5. Discussion

*5.1. Research Implications.* The concept of high-quality development put forward by the Chinese government is consistent with the UN's 2030 Sustainable Development Goals (SDGs), representing the concrete implementation way of the SDGs in China [56]. It emphasizes both the speed and the quality of development, regards the people-oriented concept, and balances efficiency and equity. The evaluation system of urban high-quality development proposed in this paper may be a reference for studying other countries or regions.

Most previous studies have focused on urban connections or urban networks only considering individual aspects, such as knowledge collaboration [49, 57], traffic flow [46, 58], and network connection [59]. The UCI established in this study synthesized these dimensions. Considering the actual situation of the case, related methods have been improved. For example, we added the number of passenger buses in the transportation connection model to improve the accuracy; when constructing the information connection model, network big data were used to reflect the intensity of information flow between cities. This study provides a methodological reference for future research.

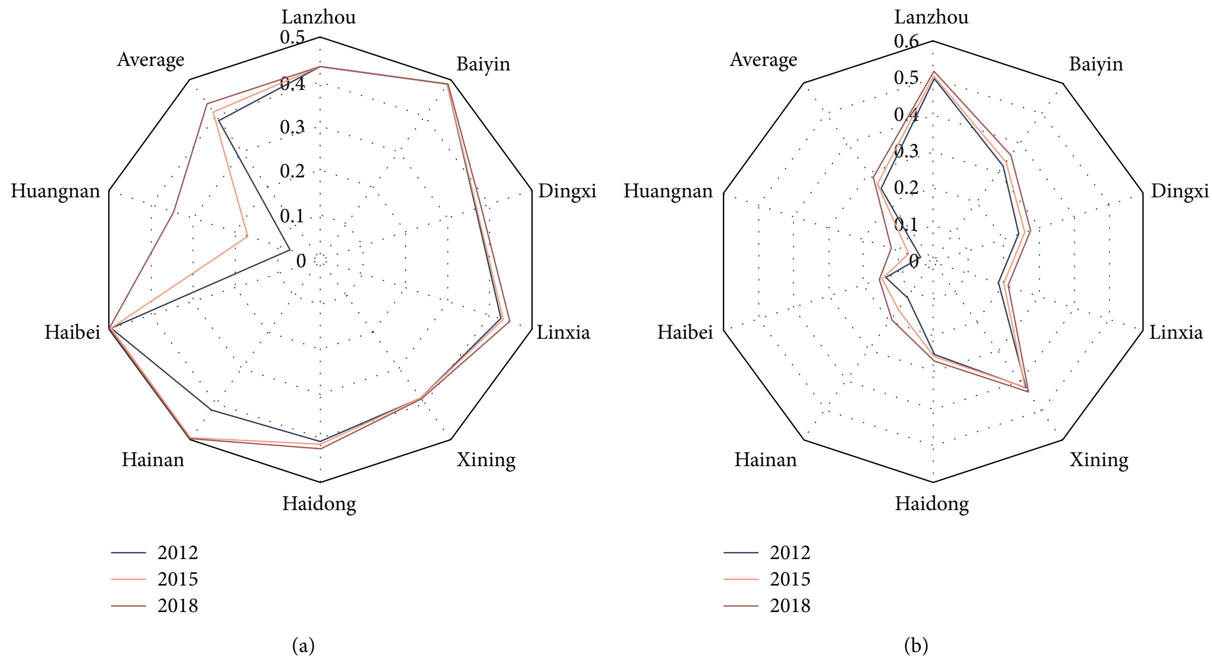



Figure 6: The degree of coupling and coordination between UCI and HDI in Lanzhou-Xining urban agglomeration. (a) Coupling degree and (b) coordination degree.

*5.2. Policy Implications.* This study proved that the cities' external connections have positive effects on urban development by quantitative analysis, providing a decision-making reference for other regions or countries. The multicenter and network of urban agglomerations help to achieve balanced and high-quality regional development. For industrial connections, it is important to further clarify the industrial function positioning of different cities so as to promote functional complementarity. Cities should develop in complementarity, barriers to the flow of elements should be broken, and an industrial chain cycle should be formed. For transportation connections, the traffic lines between major cities and key towns should be strengthened; the integrated ticket connection between different modes of transportation and the one-card interoperability should be promoted. For information connection, many measures can effectively promote network connections between cities such as improving the layout of high-speed information networks, accelerating the construction of 5G networks, and promoting the integration of e-government platforms. For innovative connection, cities should improve the market mechanism to transform scientific and technological achievements and promote the in-depth integration of production, education, and research of innovation subjects.

*5.3. Research Shortcoming and Prospects.* The present study only considered four dimensions of industry, transportation, information, and innovation in the urban connection. In the future, it is necessary to improve the evaluation index system and supplement the city's connection in the dimensions of politics and culture. Whether the relationship between UCI and HDI demonstrated in Lanzhou-Xining urban agglomeration is a universal need to be tested by further using other regions as examples. The interaction mechanism and quantitative relationship between urban connections and high-quality development need to be further explored by applying complexity science.

## 6. Conclusions

This study contributes to the scientific community in two points. The UCI and HDI were constructed from multiple dimensions based on big data and statistical data, and the coupling relationship between UCI and HDI was determined. The main findings are as follows.

From 2012 to 2018, the UCI of each city in Lanzhou-Xining urban agglomeration showed different degrees of growth. The fastest growing cities were Huangnan and Haibei. Lanzhou and Xining had the highest connection intensity in all dimensions. Urban high-quality development can be measured through five dimensions: innovation, coordination, green development, openness, and sharing. The overall high-quality development level of each city was on the rise, but there were regional differences. Lanzhou and Xining had the highest level of HDI, but the other cities were far behind. The biggest gap was in the innovation development level, and the gap in green and sharing development was small.

UCIs are significantly correlated with most indexes of high-quality development, especially innovative and green development. This means that the enhancement of the external connections and openness of an urban system is likely to facilitate high-quality development. Analysis of our case shows that, for every 1% increase in industrial and transportation connection, the urban high-quality development level will increase by 0.317% and 0.159%, while the influence of information and innovative connection is not significant. There is an antagonistic relation and a mild disorder between the UCI and HDI of the cities, but the trend is improving.

## Data Availability

All input data used in these analyses were derived from published sources cited within the article. Any other datasets generated in the current study are available from the corresponding author upon request.

## Conflicts of Interest

The authors declare no conflicts of interest.

## Acknowledgments

This research was funded by the National Natural Science Foundation of China (Grant no. 41771130) and the Research Funding for the Second Comprehensive Scientific Investigation of the Qinghai-Tibet Plateau (Grant no. 2019QZKK1005).

## References

[1] H. H. Shih-Kung Lai, *Urban Complexity and Planning: Theories and Computer Simulations*, Taylor and Francis, Milton Park, Abingdon, UK, 2016.

[2] Y. J. Shi, G. F. Zhai, L. H. Xu et al., "Assessment methods of urban system resilience: from the perspective of complex adaptive system theory," *Cities*, vol. 112, p. 13, 2021.

[3] M. Batty, "The size, scale, and shape of cities," *Science*, vol. 319, no. 5864, pp. 769–771, 2008.

[4] X. Cui, C. Fang, H. Liu, and X. Liu, "Assessing sustainability of urbanization by a coordinated development index for an urbanization-resources-environment complex system: a case study of Jing-Jin-Ji region, China," *Ecological Indicators*, vol. 96, pp. 383–391, 2019.

[5] M. J. Burger and E. J. Meijers, "Agglomerations and the rise of urban network externalities," *Papers in Regional Science*, vol. 95, no. 1, pp. 5–15, 2016.

[6] H. Xin, "Jinping Xi on China's high-quality development," 2021.

[7] C. Fang, "Progress and the future direction of research into urban agglomeration in China," *Acta Geographica Sinica*, vol. 69, no. 8, pp. 1130–1144, 2014.

[8] T. B. Levent, S. Kundak, and A. A. Gulumser, "City-to-city linkages in a mobile society: the role of urban networks in Eurocities and Sister Cities," *International Journal of Services Technology and Management*, vol. 10, no. 1, pp. 83–109, 2008.

[9] B. Blondé, "At the cradle of the transport revolution?" *The Journal of Transport History*, vol. 31, no. 1, pp. 89–111, 2010.

[10] W. Zheng, A. Kuang, X. Wang, and J. Chen, "Measuring network configuration of the Yangtze river middle reaches urban agglomeration: based on modified radiation model," *Chinese Geographical Science*, vol. 30, no. 4, pp. 677–694, 2020.

[11] T. Wang, W. Yue, X. Ye, Y. Liu, and D. Lu, "Re-evaluating polycentric urban structure: a functional linkage perspective," *Cities*, vol. 101, 2020.

[12] H. Yang, M. Dijst, P. Witte, H. van Ginkel, and J. E. Wang, "Comparing passenger flow and time schedule data to analyse high-speed railways and urban networks in China," *Urban Studies*, vol. 56, no. 6, pp. 1267–1287, 2019.

[13] S. Hallegatte, "Storm damages and inter-city trade," *Nature Sustainability*, vol. 3, no. 8, pp. 577-578, 2020.

[14] D. Li, Y. D. Wei, and T. Wang, "Spatial and temporal evolution of urban innovation network in China," *Habitat International*, vol. 49, pp. 484–496, 2015.

[15] P. Hall, *The Polycentric Metropolis: A Western European Perspective on Mega-City Regions, Governance and Planning of Mega-City Regions*, Routledge, Milton Park, Abingdon, UK, 2010.

[16] C. L. Fang, X. H. Yu, X. L. Zhang, J. W. Fang, and H. M. Liu, "Big data analysis on the spatial networks of urban agglomeration," *Cities*, vol. 102, p. 16, 2020.

[17] H. Tong, P. Shi, J. Luo, and X. Liu, "The structure and pattern of urban network in the Lanzhou-Xining urban agglomeration," *Chinese Geographical Science*, vol. 30, no. 1, pp. 59–74, 2020.

[18] R. Capello, "The city network paradigm: measuring urban network externalities," *Urban Studies*, vol. 37, no. 11, pp. 1925–1945, 2000.

[19] L. Wang, W. Xiao, and C. Wang, "Evolution of hierarchical structure and spatial pattern of coastal cities in China-based on the data of distribution of marine-related enterprises," *Polish Maritime Research*, vol. 24, no. s3, pp. 58–64, 2017.

[20] Y. Chen, "Urban gravity model based on cross-correlation function and Fourier analyses of spatio-temporal process," *Chaos, Solitons & Fractals*, vol. 41, no. 2, 2008.

[21] Y. Zhao, G. Zhang, and H. Zhao, "Spatial network structures of urban agglomeration based on the improved gravity model: a case study in China's two urban agglomerations," *Complexity*, vol. 2021, Article ID 6651444, 17 pages, 2021.

[22] X. Jin, G. Hu, H. Ding, S. Ye, Y. Lu, and J. Lin, "Evolution of spatial structure patterns of city networks in the Yangtze River Economic Belt from the perspective of corporate pledge linkage," *Growth and Change*, vol. 51, no. 2, 2020.

[23] L. Zhang and Y. Lu, "Regional accessibility of land traffic network in the Yangtze River Delta," *Journal of Geographical Sciences*, vol. 17, no. 3, pp. 351–364, 2007.

[24] C. Gan, M. Voda, K. Wang, L. Chen, and J. Ye, "Spatial network structure of the tourism economy in urban agglomeration: a social network analysis," *Journal of Hospitality and Tourism Management*, vol. 47, pp. 124–133, 2021.

[25] X. Cao, G. Zeng, L. Lin, and L. Zou, "Hierarchical characteristics and proximity mechanism of intercity innovation networks: a case of 290 cities in China," *Complexity*, vol. 2021, Article ID 5538872, 14 pages, 2021.

[26] Y. Lichi, "Evaluation of urban development quality based on the connotation of high-quality development: a case study of guangdong province," *Journal of Landscape Research*, vol. 11, no. 2, pp. 83–89, 2019.

[27] L. Chen, W. Ye, C. Huo, and K. James, "Environmental regulations, the industrial structure, and high-quality regional economic development: evidence from China," *Land*, vol. 9, no. 12, p. 517, 2020.

[28] J. Long, Z. Qiting, M. Junxia, and Z. Zhizhuo, "Evaluation and prediction of the level of high-quality development: a case study of the Yellow River Basin, China," *Ecological Indicators*, vol. 129, 2021.

[29] A. Jahanger, "Influence of FDI characteristics on high-quality development of China's economy," *Environmental Science and Pollution Research*, vol. 28, no. 15, pp. 18977–18988, 2021.

[30] A. Huerta-Barrientos, *Scaling in Urban Complex Systems: Mexico City Metabolism*, Intech Open Science, London, UK, 2018.

[31] H. M. Liu, C. L. Fang, and Q. Gao, "Evaluating the real-time impact of COVID-19 on cities: China as a case study," *Complexity*, vol. 2020, Article ID 8855521, 11 pages, 2020.

[32] L. Inostroza and H. Zepp, "The metabolic urban network: urbanisation as hierarchically ordered space of flows," *Cities*, vol. 109, 2021.

[33] K. Mainzer and R. Landauer, "Thinking in complexity: the complex dynamics of matter, mind, and mankind," *American Journal of Physics*, vol. 63, no. 8, 1995.

[34] H. Liu, C. Fang, and K. Fang, "Coupled Human and Natural Cube: a novel framework for analyzing the multiple interactions between humans and nature," *Journal of Geographical Sciences*, vol. 30, no. 3, pp. 355–377, 2020.

[35] L. M. A. Bettencourt, "The origins of scaling in cities," *Science*, vol. 340, no. 6139, pp. 1438–1441, 2013.

[36] M. Castells, "Space of flows, space of places: materials for a theory of urbanism in the information age," *The City Reader*, pp. 263–274, Routledge, Milton, UK, 2015.

[37] A. Bassolas, H. Barbosa-Filho, B. Dickinson et al., "Hierarchical organization of urban mobility and its connection with city livability," *Nature Communications*, vol. 10, no. 1, p. 4817, 2019.

[38] Y. Wei, W. Song, C. Xiu, and Z. Zhao, "The rich-club phenomenon of China's population flow network during the country's spring festival," *Applied Geography*, vol. 96, pp. 77–85, 2018.

[39] W. Zhang, B. Derudder, J. Wang, and F. Witlox, "An analysis of the determinants of the multiplex urban networks in the Yangtze river delta," *Tijdschrift voor Economische en Sociale Geografie*, vol. 111, no. 2, pp. 117–133, 2020.

[40] P. J. Taylor, "Leading world cities: empirical evaluations of urban nodes in multiple networks," *Urban Studies*, vol. 42, no. 9, pp. 1593–1608, 2005.

[41] B. Trinh and K. Kobayashi, "Measuring the effective rate of protection in Vietnam's economy with emphasis on the manufacturing industry: an input-output approach," 2010.

[42] L. Lai, X. Huang, and W. Liu, "Adjustment for regional ecological footprint based on input-output technique: a case study of Jiangsu Province in 2002," *Acta Ecologica Sinica*, vol. 26, no. 4, pp. 1285–1292, 2006.

[43] L. Xie, Q. Xue, and Z. Yuan, "Composition and spatial difference of agro-industry carbon footprint in Hebei province, North China," *Ecological Indicators*, vol. 97, pp. 141–149, 2019.

[44] G. H. M. Sonis, *Fields of Influence and Extended Input-Output Analysis: A Theoretical Account*, Avebury, Aldershot, UK, 1991.

[45] M. Sonis and G. J. D. Hewings, "Economic complexity as network complication: multiregional input-output structural path analysis," *The Annals of Regional Science*, vol. 32, no. 3, pp. 407–436, 1998.

[46] F. Zhang, Y. Ning, and X. Lou, "The evolutionary mechanism of China's urban network from 1997 to 2015: an analysis of air passenger flows," *Cities*, vol. 109, 2021.

[47] K. Liu, T. Wang, Z. Yang et al., "Using Baidu search index to predict dengue outbreak in China," *Scientific Reports*, vol. 6, no. 1, pp. 38040–38049, 2016.

[48] Z. Zhang and W. Tang, "Analysis of spatial patterns of public attention on housing prices in Chinese cities: a web search engine approach," *Applied Geography*, vol. 70, 2016.

[49] Y. Li and N. A. Phelps, "Megalopolis unbound: knowledge collaboration and functional polycentricity within and beyond the Yangtze River Delta Region in China," *Urban Studies*, vol. 55, pp. 433–460, 2014.

[50] Q. Wang, R. Han, Q. Huang et al., "Research on energy conservation and emissions reduction based on AHP-fuzzy synthetic evaluation model: a case study of tobacco enterprises," *Journal of Cleaner Production*, vol. 201, pp. 88–97, 2018.

[51] S. Mei, H. Xuguang, Z. Liyan et al., "Transformation performance and subsystem coupling of resource-based cities in China: an analysis based on the support-pressure framework," *Integrated Environmental Assessment and Management*, pp. 1–14, 2021.

[52] B. Zhu, M. Zhang, Y. Zhou et al., "Exploring the effect of industrial structure adjustment on interprovincial green development efficiency in China: A novel integrated approach," *Energy Policy*, vol. 134, Article ID 110946, 2019.

[53] L. Ran, X. D. Tan, Y. Xu et al., "The application of subjective and objective method in the evaluation of healthy cities: a case study in Central China," *Sustainable Cities and Society*, vol. 65, 2021.

[54] D. Luo, L. W. Liang, Z. B. Wang, L. K. Chen, and F. M. Zhang, "Exploration of coupling effects in the economy-society-environment system in urban areas: case study of the Yangtze river delta urban agglomeration," *Ecological Indicators*, vol. 128, 2021.

[55] L. Ding, W. Zhao, Y. Huang, S. Cheng, and C. Liu, "Research on the coupling coordination relationship between urbanization and the air environment: a case study of the area of Wuhan," *Atmosphere*, vol. 6, no. 10, 2015.

[56] Y. Cheng, H. Liu, S. Wang, X. Cui, and Q. Li, "Global action on SDGs: policy review and outlook in a post-pandemic era," *Sustainability*, vol. 13, no. 11, 2021.

[57] Z. Ba, Y. Ma, and Z. Liang, "Exploring the effect of city-level collaboration and knowledge networks on innovation: evidence from energy conservation field," *Journal of Informetrics*, vol. 15, no. 3, 2021.

[58] X. Yang, R. Wang, D. Guo, and W. Sun, "The reconfiguration effect of China's high-speed railway on intercity connection—a study based on media attention index," *Transport Policy*, vol. 95, pp. 47–56, 2020.

[59] H. Wang, W. Wang, Y. Meng, and Y. Zhang, "Degree of user attention to a webpage based on Baidu index: an alternative to page view," *Journal of Experimental & Theoretical Artificial Intelligence*, vol. 26, no. 2, 2014.

Hindawi
Complexity
Volume 2021, Article ID 2855542, 10 pages
https://doi.org/10.1155/2021/2855542

*Research Article*

# On Time Effect of Preschool Education: Social Analysis Based on CUCDS

**Bohan Yan** [1] **Yongjun Feng,** [2] **and Ning Cai** [3,4]

[1] School of Sociology, Henan University of Economics and Law, Zhengzhou 450046, China
[2] School of Education, Shaanxi Normal University, Xi'an 710062, China
[3] School of Artificial Intelligence, Beijing University of Posts & Telecommunications, Beijing 100876, China
[4] Key Laboratory of China's Ethnic Languages and Information Technology for Ministry of Education, Northwest Minzu University, Lanzhou 730000, China

Correspondence should be addressed to Bohan Yan; yanbh12@tsinghua.org.cn

Received 12 August 2021; Revised 29 August 2021; Accepted 1 September 2021; Published 14 October 2021

Academic Editor: Jing-Hu Pan



Cognitive ability is an important aspect of children's development, but there is still room for discussion about the impact of preschool education on children's cognitive ability. Based on the data of China Urbanization and Children Development Survey (CUCDS) of Tsinghua University, this paper categorizes cognitive ability into Chinese language cognition and mathematical cognition. It is discovered that the impact of preschool education on children's cognitive development differs depending on the cognitive ability and the length of time. In particular, preschool education has both short-term and long-term effects on children's Chinese cognitive ability, while there is only a short-term effect on the development of children's mathematical cognitive ability without long-term effect.

## 1. Introduction

Cognitive ability is a very important aspect of development for children. It can in a sense even be claimed that cognitive ability constitutes the basis of children's upward mobility. Due to the close relationship between education and ultimate class status acquisition, the fundamental role of cognitive ability in the promotion of status is prominently manifested as its influence on education acquisition. Relevant research was first conducted through the Wisconsin model, with the mediating role of "intelligence" and other social-psychological variables in the causal chain of "family background-education acquisition" being found [1]. Since then, with the improvement of measurement technology and the abundant accumulation of data, more and more research studies on intelligence and cognitive ability have been conducted. It has become a consensus that the basic role of cognitive ability accounts for improving status. It was proved not only to be an effective indicator for predicting academic achievement and educational acquisition [2–4] but also to have significant effect on job acquisition, career performance, and economic income in the labor market [4–7]. Moreover, a recent study combined cognitive ability with family background for analysis and found that cognitive ability can compensate the disadvantage of family background and ultimately help children from those disadvantaged families to gain higher economic status [8]. In view of the increasing recognition of the important role of cognitive ability, some scholars even put forward the concept of "cognitive capital" to represent a kind of accumulated asset that can be employed to create and grasp opportunity and extend well-being to cope with environmental challenges and pressures [9].

The above empirical studies on the importance of cognitive ability regard it as an independent variable, while actually cognitive ability is also a resulting variable impacted by various sociological factors such as education. Preschool/early education is one of the possible factors. Preschool education experience may affect cognitive development from two aspects—educational opportunity and quality.

There are still significant rural-urban and regional differences in preschool education opportunities of Chinese children. Therefore, this paper focuses on the impact of preschool education opportunities and analyzes the impact of preschool education experience on children's cognitive ability development by exploring the national sample data of China, which is relatively rare in existing literature. This paper endeavors to answer the following questions. (1) Does preschool education opportunity have significant impact on children's cognitive development? (2) If the first question is true, is the influence different upon different levels or aspects of cognitive ability? (3) Considering ageing, is the effect short term, long term, or both?

## 2. Literature Review and Assumptions

Although theories affecting cognitive development are full of contradictions and controversies ("genetic determinism" and "environmental determinism"), there is a consensus on the main aspect, that is, changes in cognitive environment and stimuli will lead to differences in cognitive ability. Biological basis is the fact that a growing brain is plastic. During the growth period, especially in the process of brain development, different stimuli may lead to different cognitive levels.

In short, the development of cognitive ability bears sensitive and critical period. The early childhood (0–6 years old) is usually regarded as the sensitive and critical period of growth, not only because many brain structures and biochemical pathways are developed at this stage but also because the brain development speeds at this stage [10, 11]. In this critical period, implementation of appropriate mild intervention will help to improve health and well-being, education effectiveness, skill potential, employment status, and quality of life. Oppositely, negative stimuli will instead lead to depreciation of cognitive capital, damage of physical and mental health, and reduction of education effectiveness and life opportunities [12]. The idea of "critical period" is the foundation of the implementation of preschool education. Because the cognitive ability and behavior of children are more malleable than adults, investment in early childhood education has a higher return on investment than compensation education in middle and late stages [11]. The World Bank estimates that the return on investment in early education is about 7% to 18%, much higher than the return on financial capital. The return on investment of early education in China is also within this range, roughly between 7% and 15% [13]. In view of this, investing in early childhood education is regarded as the most effective intervention to help disadvantaged families/children to break the intergenerational transmission of poverty [14, 15].

There are two senses for preschool education, broad and narrow. In broad sense, preschool education refers to all forms of education that preschool children receive, such as school, family, and society. In the narrow sense, it only refers to the formal education implemented by specialized preschool education institutions. This paper only concentrates on the narrow sense. It deeply impacts the growth of a child whether or not he/she receives formal and standardized preschool education at the right age and whether or not the quality of preschool education is sound, especially the cognitive development. Existing empirical data from countries outside of China almost unanimously confirm that the cognitive level or academic achievement of children with formal preschool education is generally higher than those without [16, 17]. The role of preschool education, especially higher quality preschool education, in reducing the impact of risk (such as poverty) on children's cognitive development/academic achievement is also firmly validated [18]. Empirical studies from China have similar findings, but there is a lack of national samples. Chen and Liu, applying the survey data of the Program for International Student Assessment (PISA) in Shanghai, found that preschool education has dual effect of "cultivating excellence" and "making up the gap," which can promote not only academic achievement but also educational equity among students [19]. But the findings were limited to Shanghai, China. Luo et al. conducted a study in six state-level poverty-stricken counties in Shanxi, Gansu, and Henan provinces. Through the analysis of 505 children aged 4–5, they found that there was a significant correlation between cognitive ability of children and formal preschool education experience [20].

As far as the duration of the effect of preschool education on cognitive development is concerned, the literature consistently shows that preschool education has a short-term effect on children's cognitive development [21]. However, it is still controversial whether preschool education has a long-term effect. Some studies have found that there is truly long-term effect [22], while some do not support this [23]. The literature within China also consistently confirms the short-term effects of preschool education, but conclusions towards long-term effects are not uniform. Two studies using the data of "China Education Tracking Survey" (CEPS) found that preschool education has a long-term effect on the development of cognitive ability [24, 25]. Another analysis using the data of "China Family Tracking Survey" (CFPS) found that there was no significant correlation between preschool education and children's cognitive ability, without long-term effect from preschool education to cognitive ability [26]. Both CEPS and CFPS are national data, but there is a problem of insufficient coverage of children's age range. The baseline data of CEPS only cover two cohorts of students in grade 7 and grade 9, while Gong and others only used the data of children aged 11–15 in CFPS data.

In addition, most relevant studies simply treat children's cognitive ability as a "whole" for analysis, whereas cognitive ability is actually of multi-level and multi-facet. Some tests may put emphasis on measuring cognitive ability in learning, memory, comprehension, and classification abilities, while some other focus on measuring cognitive ability in reasoning and judgment, logical thinking, abstract thinking ability, etc. [27]. This overall analysis method that does not distinguish between cognitive abilities may cover up some real and interesting information, thereby biasing the conclusions.

Based on the above analysis, via mining the data of China Urbanization and Children Development Survey (CUCDS) of Tsinghua University, a large national sample that covers a wider range of children's ages, this paper endeavors to explore the influence of preschool education experience on different levels of cognitive ability of Chinese children and the duration of effect by dividing the cognitive ability into two categories: the Chinese language cognition and mathematical cognition, and the children into two age groups: 3–10 and 11–15. The Chinese test of CUCDS data focuses on the measurement of common sense, vocabulary, classification, understanding, and reasoning ability. These abilities reflect individual learning and memory, reasoning and judgment, understanding, and comprehensive conceptual thinking ability. Mathematics test pays more attention to the measurement of calculation, problem-solving, and reasoning ability, which reflects an individual's logical and abstract thinking and reasoning ability. The two tests not only measure cognitive abilities but also hold particular emphasis. In terms of time division, the reason explaining why children are divided into two age groups is based on previous practice, with limited data. The maximal age surveyed by CUCDS is 15. Considering the external stimulus of cognitive ability, duration, and preschool education, one can put forward the hypothesis that cognitive ability in different aspects and levels will be affected by preschool education, with both short-term and long-term effects. This can be summarized into the following two assumptions.

*Assumption 1.* Preschool education experience has both short-term and long-term effects on children's Chinese language cognitive development.

*Assumption 2.* Preschool education experience has both short-term and long-term effects on children's mathematics cognitive development.

## 3. Data, Variables, and Model Setting

*3.1. Data.* The data employed in this paper are from China Urbanization and Children Development Survey (CUCDS) of Tsinghua University. The survey was conducted in 2012. With multi-stage sampling scheme and PPS sampling method, adults and children from 28 provinces, 147 districts and counties, and 500 villages in mainland China except Qinghai, Tibet, and Hainan were randomly selected for the interview. The weighted valid sample size for the question "whether or not they received preschool education" was 4,963.

The test of children's cognitive ability is the main content of module for children in the questionnaire. The 3–12-year-old part of the "children's ability test" was compiled by Hou-Can Zhang in Beijing Normal University, and the 13–15-year-old part was designed by Jean Yeung in National University of Singapore, with reference to the PISA test and the cognitive scale of the module for children in the American Income Tracking Survey. All measuring tools are suitable for Chinese children aged 3–15. The children on the test are divided into 4 age groups: 3–6 years old, 7–8 years old, 9–12 years old, and 13–15 years old. Each age group has a corresponding subtest (local language, mathematics, and English) with the English test being only applicable for the two oldest age groups. Because this paper focuses on the duration of the effect of preschool education on cognitive development and comparison should be conducted among different ages, only the data of children's Chinese and mathematical cognition test are concerned.

As for the test reliability, in view of the non-one-dimensional structure of Chinese and mathematics tests, the alpha coefficient cannot be taken as the most ideal indicator to measure the stability of the test results. Considering that the subtests are basically arranged in a way from easy to difficult, the odd-even method is adopted to judge the stability of test tools by calculating the split half reliability. However, this method of halving the test length will underestimate the test reliability. In order to compensate the error, the Spearman–Brown formula is usually used to correct the split half reliability. Table 1 shows the split half and calibrated reliability of Chinese and mathematics proficiency tests for children of all ages. It can be seen that the tests have a high stability, with the calibrated reliability above 0.84. Especially for mathematics tests, the lowest value of the calibrated reliability also reaches 0.90.

### 3.2. Variables

*3.2.1. Dependent Variables.* In this paper, the Chinese language and mathematics cognitive abilities are dependent variables, based on corresponding subtests. The content, difficulty, and duration of subtests were different in different age groups. Test difficulty increases with the age of children. For young children aged 3–6 years, the assessment time takes about 20 minutes, while for older children aged 9–12 and 13–15 years, it takes 30 minutes. The Chinese and mathematics tests for children of different age groups and the time limits for completion are shown in Table 2.

The full score of Chinese and mathematics subtest is 50. Descriptive statistics of Chinese and mathematics cognitive score variables are shown in Table 3.

*3.2.2. Independent Variables.* The core independent variable of this paper is preschool education. According to the index "whether or not ever been to kindergarten or preschool," the variable discretely classifies the children who have been to kindergarten or preschool as those received preschool education. In the preschool education choice model, preschool education is the dependent variable, with individual, family, and regional variables being the independent ones. Specifically, individual level variables include children's gender and age; family level variables include the parents' educational levels to measure family cultural capital, the number of children aged 0–15 years to measure family structure, and the birthplace of children to measure family economic status; regional level variables include regions (eastern, central, and western) and rural-urban areas (urban or rural). Table 4 lists the descriptive statistics of the main independent variables in this paper.

Table 1: Split half reliability and calibrated reliability of the test.

| Test name | Age group (years) | Split half reliability | Calibrated reliability of Spearman–Brown formula |
|---|---|---|---|
| Chinese test | 3–6 | 0.76 | 0.86 |
| | 7-8 | 0.82 | 0.90 |
| | 9–12 | 0.78 | 0.88 |
| | 13–15 | 0.73 | 0.84 |
| | 3–6 | 0.86 | 0.92 |
| Mathematics test | 7-8 | 0.88 | 0.94 |
| | 9–12 | 0.91 | 0.95 |
| | 13–15 | 0.82 | 0.90 |

Table 2: Subtest and time limit (minutes) of children of different age groups.

| Age group (years) | Chinese | Mathematics |
|---|---|---|
| 3–6 | 10 | 10 |
| 7-8 | 12 | 13 |
| 9–12 | 12 | 18 |
| 13–15 | 15 | 15 |

Table 3: Descriptive statistics of Chinese cognitive ability and Mathematics cognitive ability.

| Variable | Sample size | Minimum value | Maximum value | Mean | Standard deviation |
|---|---|---|---|---|---|
| Chinese | 4938 | 1.00 | 50.00 | 26.04 | 11.04 |
| Mathematics | 4798 | 1.00 | 50.00 | 21.18 | 11.82 |

Table 4: Descriptive statistics of main variables.

| Variables | Weighted mean |
|---|---|
| Child's gender (1 = male) | 0.54 |
| Child's age (years) | 9.02 (3.84) |
| Father's education level | |
| Primary school and below | 0.24 |
| Junior middle school | 0.46 |
| Senior high school and equivalent | 0.21 |
| University or above | 0.08 |
| Mother's educational level | |
| Primary school and below | 0.36 |
| Junior middle school | 0.41 |
| Senior high school and equivalent | 0.15 |
| University or above | 0.07 |
| Birthplace (1 = hospital) | 0.77 |
| Number of children aged 0–15 years | 1.64 (0.84) |
| Region | |
| Western | 0.31 |
| Central | 0.33 |
| Eastern | 0.36 |
| Urban and rural (1 = urban) | 0.39 |
| Received preschool education or not (1 = yes) | 0.70 |

*Note*. Standard deviation is given in brackets. Due to missing values, the percentage sum of some variables may not be equal to 100%.

*3.3. Model Setting.* We adopt the propensity score matching (PSM) method to analyze the effect of preschool education on children's cognitive development. Most of the existing research studies on the influence of preschool education on children's cognitive development are mainly based on simple comparison between children receiving or not receiving preschool education, often lacking consideration of endogenous problems. Furthermore, children receiving and not receiving preschool education may be two groups with systematic differences. If endogenous problems are never addressed to eliminate possible confounding variable effects, the comparison results cannot be definitely attributed to the influence of preschool education but may be the effect of other confounding variables such as family. Therefore, in order to analyze the "net effect" of preschool education, we need to introduce a method that can eliminate the selective error of confusing variables and solve the endogenous problem. One of the methods is propensity score matching, which is based on counterfactual causality. In this paper, the steps of this method are as follows: the confounding variables that lead to the imbalance between children receiving (intervention group) and not receiving preschool education (control group) are included in the logit regression model, and the propensity score is calculated on this basis. According to the common support domain of propensity score, the intervention and the control groups were matched to find the ideal counterfactor. The Average Effect of Treatment on the Treated (ATT) of the intervention group was calculated for the matched samples.

## 4. Short-Term Effect of Preschool Education on Children's Cognitive Development

*4.1. Establishing Preschool Education Choice Model (3–10 Years Old).* With reference to the existing literature and combined with the characteristics of the research, this paper selects the independent variables at three levels of individual, family, and region to establish the preschool education choice model. Considering that the main task of the study is to analyze the "net effect" of preschool education experience to children's cognitive ability, the independent variables included in the model need to meet the requirements to be correlated to both children's access to preschool education opportunities and cognitive ability. The individual level variables ultimately included gender and age. The family level variables are education level of parents, the number of children aged 0–15 in the family, and the birthplace. The current economic status is not suitable because only the family economic status before receiving preschool education can affect the choice of children. According to research of Xu and Xie [28], we indirectly measure the economic status before children's preschool education by their birthplace (born in hospital or at home). At the regional level, regional (eastern, central, or western) and rural-urban (urban or rural) variables are selected. This is because there is evident imbalance in regional advance of preschool education in China, not only between urban and rural areas, but also between eastern, central, and western provinces. Statistics released by the Ministry of Education show that the national gross enrollment rate for preschool education has been steadily increasing, from 56.6% in 2010, 67.5% in 2013, to 81.7% in 2018. However, the growth cannot cover up the significant differences between rural/urban areas and among regions. Taking the year 2013 as example, the average gross enrollment rate of preschool education in China was 67.5%, accounting for 45.0% in urban areas and only 22.5% in rural areas. The situation of discrepancy is extensive. In terms of regional differences, taking the gross enrollment rate of preschool education as index, the value in Tibet is only 52.0%, and that in Yunnan Province is 54.0%, while that in Zhejiang, Jiangsu, Fujian, and Guangdong provinces is higher than 95.0%, with this index in Shanghai even being 100.0%.

Let preschool education be a binary variable (1 = received; 0 = not received) and the aforementioned individual, family, and regional level variables be independent variables; then, fit the logit model (Table 5). The results show that gender and education level of fathers have no significant influence to whether children receive preschool education, whereas the lack of preschool education experience is more likely attributed to low educational level of mothers, multiple children, poor family economic status, and regional factors. The significant effect of age may indicate that some younger children simply do not receive preschool education at the right age.

*4.2. Matching Samples and Balance Test (3–10 Years Old).* Based on the establishment of preschool education choice model, the matching score of preschool education is calculated, and then one-fourth of the standard deviation of this score is adopted as the caliper. The nearest neighbor matching within the scope of caliper is taken as method for sample matching to find the relatively ideal "counterfactor" between the children receiving and not receiving preschool education. Because the cases outside the common support domain (the overlapping part of the probability density distribution) will be excluded when the propensity score is matched, the samples are effectively utilized with larger common support domain of the intervention and the control groups. Figure 1 shows that the common support domain of the two groups is large, which meets the basic condition of propensity score matching.

As for the effect of matching, a balance test is required, that is, to test whether there are systematic differences in confounding variables between the matched intervention and control groups. The test method is determined according to the measurement level of confounding variables. Generally, chi-square analysis and analysis of variance are applied to test categorical and continuous variables, respectively. Table 6 shows the balance test results. Before matching, there were significant differences between the two groups of children in other confounding variables except gender. After matching, the systematic differences between the two groups of children in other variables other than region disappeared, indicating that the overall matching sample passed the balance test well, and then one can use the matched samples for impact effect analysis.

*4.3. On Short-Term Effect.* The effect of preschool education on children's cognitive ability was calculated by using matched samples. It is found that the Chinese language cognitive ability ($N = 1463$) and mathematics cognitive ability ($N = 1425$) of children who have received preschool education are significantly higher than those without preschool education, with a difference of 1.79 points in Chinese language and 2.13 points in mathematics cognitive ability (Table 7). Since the analysis is for children aged 3–10 years who are receiving or just completed preschool education, this result indicates that preschool education has immediate or short-term effects on cognitive development.

## 5. Long-Term Effect of Preschool Education on Children's Cognitive Development

Through the same logic, this paper analyzes the influence of preschool education on the cognitive development of children aged 11–15 years. Firstly, the choice model of preschool education is established to calculate the propensity to receive preschool education. Then, the nearest neighbor method within the range of caliper is taken for sample matching and balance test. Finally, ATT is calculated by employing the matched samples.

*5.1. Establishing Preschool Education Choice Model (11–15 Years Old).* The variables included in the model are the same as those for children aged 3–10 years with conclusions being basically similar (Table 8). In addition to the child's gender and father's education level, the variables that have no significant impact on preschool education also include the child's age. Still, those children with higher education level of mothers, fewer siblings, better family economic status, birthplace in eastern or Central China, and in urban areas are more likely to receive preschool education.

Table 5: Logit model for predicting the tendency to receive preschool education (3–10 years old).

| Variables | Coefficient |
|---|---|
| Child's gender (1 = male) | −0.11 (0.09) |
| Child's age (years) | 0.22 (0.02)*** |
| Father's education level (reference group: primary school and below) | |
| Junior middle school | 0.14 (0.12) |
| Senior high school and equivalent | 0.22 (0.16) |
| University or above | 0.08 (0.26) |
| Mother's education level (reference group: primary school and below) | |
| Junior middle school | 0.29 (0.11)* |
| Senior high school and equivalent | 0.36 (0.17)* |
| University or above | 0.71 (0.28)* |
| Number of children aged 0–15 years | −0.25 (0.06)*** |
| Birthplace (1 = hospital) | 1.24 (0.13)*** |
| Region (reference group: western) | |
| Central | 0.50 (0.11)*** |
| Eastern | 0.73 (0.12)*** |
| Urban or rural (urban = 1) | 0.73 (0.11)*** |
| Intercept | −1.89 (0.25)*** |
| N | 3008 |
| Log likelihood | −1469.77*** |
| Virtual $R^2$ | 14.8% |

*Note.* * $p < 0.05$, ** $p < 0.01$, and *** $p < 0.001$; standard error is given in brackets.

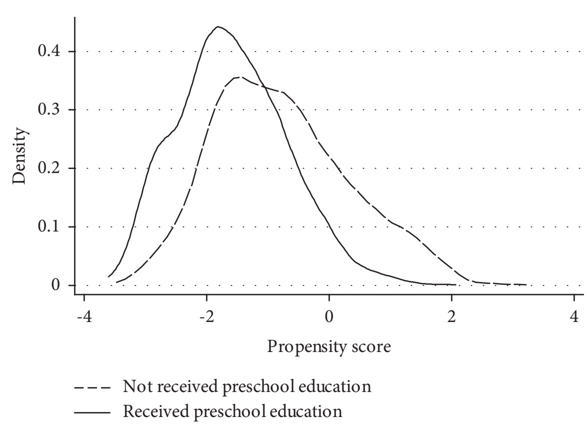



Figure 1: Probability density distribution of propensity score matching between intervention group and control group (3–10 years old).

*5.2. Matching Samples and Balance Test (11–15 Years Old).* Figure 2 shows that the common support domain of intervention and control groups is large, which is suitable for propensity score matching.

One-fourth of the standard deviation of the propensity score is still utilized as the caliper standard for nearest neighbor matching. After matching, the samples pass the balance test. The test results are shown in Table 9.

*5.3. On Long-Term Effect.* Table 10 shows that the Chinese language ($N = 729$) and mathematics cognitive ability ($N = 705$) of children who have received preschool education are both higher than those without receiving preschool education, where the Chinese cognitive ability is 2.07 points higher and the mathematics is 0.86 points higher. However, the difference in mathematics cognitive level between the two groups of children does not reach the significance level of 0.05. In other words, preschool education has a long-term effect on children's Chinese cognitive development than mathematical ability. It can be interpreted by the theory on fluid intelligence and crystal intelligence, which holds that the general factor of human cognitive ability can be attributed to fluid and crystallized intelligence. Fluid intelligence is based on individual physiological conditions, depending on innate endowment, and is not or less affected by education and life experience, whereas crystallized intelligence is acquired through learning, which is deeply influenced by education and daily experience [29]. Mathematics cognitive ability belongs to the former. Therefore, preschool education experience has no significant long-term impact on it. Of course, this conclusion still deserves further testing.

Table 6: Balance test of intervention group and control group (3–10 years old).

| Variables | Before matching (F or $\chi^2$) | After matching (F or $\chi^2$) |
|---|---|---|
| Child's gender | 0.51 | 1.22 |
| Child's age | $111.25^{***}$ | 1.12 |
| Father's education level | $67.94^{***}$ | 5.09 |
| Mother's education level | $69.26^{***}$ | 10.65 |
| Number of children aged 0–15 years | $77.56^{***}$ | 1.96 |
| Birthplace | $144.60^{***}$ | 4.82 |
| Region | $64.25^{***}$ | $22.74^{***}$ |
| Urban or rural | $76.62^{***}$ | 5.75 |

*Note.* $^{**}p < 0.01$; $^{***}p < 0.001$.

Table 7: Effect of preschool education on children's cognitive development (3–10 years old).

| | Intervention group (mean) | Control group (mean) | ATT |
|---|---|---|---|
| Chinese | 26.08 | 24.29 | $1.79\ (0.56)^{**}$ |
| Mathematics | 22.23 | 20.10 | $2.13\ (0.63)^{***}$ |

*Note.* $^{**}p < 0.01$; $^{***}p < 0.001$; standard error is given in brackets.

Table 8: Logit model for predicting the tendency to receive preschool education (11–15 years old).

| Variables | Coefficient |
|---|---|
| Child's gender (1 = male) | −0.14 (0.11) |
| Child's age (years) | −0.06 (0.04) |
| Father's education level (reference group: primary school and below) | |
| Junior middle school | −0.19 (0.13) |
| Senior high school and equivalent | −0.00 (0.19) |
| University or above | −0.25 (0.38) |
| Mother's education level (reference group: primary school and below) | |
| Junior middle school | $0.73\ (0.13)^{***}$ |
| Senior high school and equivalent | $0.99\ (0.23)^{***}$ |
| University or above | $1.12\ (0.42)^{**}$ |
| Number of children aged 0–15 years | $-0.27\ (0.07)^{***}$ |
| Birthplace (1 = hospital) | $0.64\ (0.12)^{***}$ |
| Region (reference group: western) | |
| Central | $0.50\ (0.13)^{***}$ |
| Eastern | $1.24\ (0.14)^{***}$ |
| Urban or rural (urban = 1) | $0.70\ (0.13)^{***}$ |
| Intercept | 0.44 (0.58) |
| N | 1452 |
| Log likelihood | $-1043.74^{***}$ |
| Virtual $R^2$ | 16.9% |

*Note.* $^{*}p < 0.05$, $^{**}p < 0.01$, and $^{***}p < 0.001$; standard error is given in brackets.

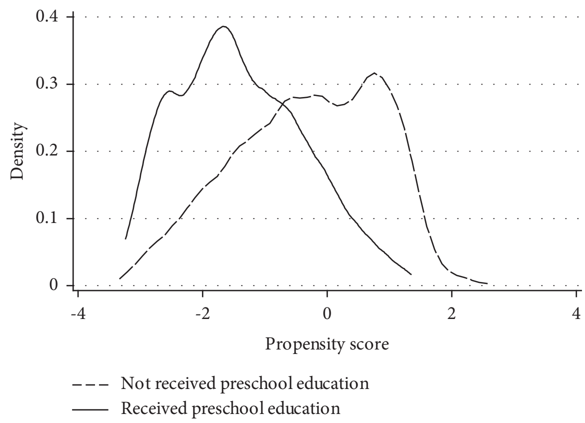



Figure 2: Probability density distribution of propensity score matching between intervention group and control group (11–15 years old).

Table 9: Balance test of intervention group and control group (11–15 years old).

| Variables | Before matching (F or $\chi^2$) | After matching (F or $\chi^2$) |
|---|---|---|
| Child's gender | 0.07 | 0.05 |
| Child's age | 1.73 | 2.93 |
| Father's educational level | $76.80^{***}$ | 0.56 |
| Mother's educational level | $136.24^{***}$ | 4.41 |
| Number of children aged 0–15 years | $68.35^{***}$ | 2.00 |
| Birthplace | $172.53^{***}$ | 0.05 |
| Region | $113.07^{***}$ | 7.33 |
| Urban or rural | $92.57^{***}$ | 1.37 |

*Note.* $^{**}p < 0.01$; $^{***}p < 0.001$.

Table 10: Effect of preschool education on children's cognitive development (11–15 years old).

| | Intervention group (mean) | Control group (mean) | ATT |
|---|---|---|---|
| Chinese | 27.19 | 25.12 | $2.07\ (0.85)^{*}$ |
| Mathematics | 20.97 | 20.11 | 0.86 (0.92) |

*Note.* $^{*}p < 0.05$, $^{**}p < 0.01$, and $^{***}p < 0.001$; standard error is given in brackets.

## 6. Conclusion, Suggestion, and Limitation

In this paper, based on CUCDS data, the selection tendency for preschool education was analyzed through logit regression, and PSM was employed to estimate the impact of preschool education on children's cognitive development. The following conclusions were reached:

(1) Preschool education experience has important influence on children's cognitive development, but this influence varies with different levels, facets, and time effects of cognitive ability. Specifically, preschool education has both short-term and long-term effects on the development of children's Chinese language cognitive ability, while there are only short-term effects on the development of children's mathematical cognitive ability. Previous studies simply claim that preschool education has significant impact on children's cognitive ability but without temporal scales. There are only research studies on the duration of the influence effect, either finding that preschool education has both short-term and long-term effects on cognitive development or finding that preschool education only has short-term but no long-term effects on cognitive ability. Evidently, the conclusion of this paper is distinct from previous studies because it subdivides cognitive ability into different aspects.

(2) Family background exerts important influence on whether children have access to preschool education as well as their cognitive development. It can be attributed to three aspects of capital, namely, family literacy, structure, and economy. In families with rich literacy capital, parents with high educational level tend to pay more attention to education and have higher expectations for the development of children. This kind of attention and expectation will be transformed into children's enthusiasm for learning, which is conducive to the development of cognitive ability; moreover, parents with high education level are willing to provide financial support and especially attention for cognitive development. More importantly, they are able to provide guidance being more appropriate and effective and assistance to the cognitive and academic development, and this

aspect is especially prominent for mothers since they play the primary role of caregivers in most families. The internal logic of the influence of family structure is "resource dilution theory" [30], that is, with prescribed family resources (including the attention resources of main caregivers, especially parents, besides economic conditions), greater number of children is a disadvantage to cognitive development because less resources can be allocated to each child. Family economic capital has impact on children's cognitive development by means of resource transformation, such as attending high-quality kindergartens, attending tutorial classes, and purchasing learning materials. The region where children are located has an important influence on whether children receive preschool education and their cognitive development. This is mainly due to the evident imbalance in the regional development of preschool education in China. Distinctions among different stages of education, rural-urban areas, and eastern, central, and western provinces are undoubtedly the main content of the unbalanced and insufficient development of Chinese education. The fundamental reason for this difference lies in the dualistic division of Chinese education system—the division between rural and urban areas and the internal division of all stages of education, from kindergarten to university. The former leads to prominent differences in education between urban and rural regions, while the latter leads to the distinction between the key and ordinary schools. The essence of dualistic division is that limited education funds are allotted to urban schools and key schools [31].

Based on the above analysis, policy recommendations are given in the subsequent paragraphs. Children from disadvantaged families are more likely to be at significant disadvantage in cognitive level due to the lack of preschool education and other reasons. Such families are mainly distributed in rural and western regions, and because the important influence of preschool education and family cultural capital were marked by parents' educational level on cognitive development, in order to block the severe circle of intergenerational transmission of poverty and to promote social mobility, especially the realization of upward mobility of children from disadvantaged strata, primary areas and appropriate strategies should be determined for intervention, in addition to certain specific assistance means. There is no doubt that education should be the focus of intervention. In particular, we should pay attention to the education problems in rural, western, and other underdeveloped areas. We should not only develop and popularize preschool education and improve the quality of compulsory education to "directly intervene" children's cognitive development but also vigorously develop high school education in underdeveloped areas, improve the enrollment rate of universities, and increase the average years of education in the region to promote children's cognitive development indirectly.

In addition, considering that long-term effect of preschool education only applies to Chinese cognitive ability, the enlightenment to relevant assistance departments and personnel may be that for those children without preschool education experience, targeted reinforced training in Chinese may be more important in the process of academic assistance and improvement.

Admittedly, there are still some limitations. First, cognitive development is affected not only by the environment acquired but also by natural biological factors. As the acquired environment affects the degree, genetic factors also bring limitations to the scope of cognitive development. Due to the lack of collection of genetic information data, it is impossible to identify the involvement of genes and environment in cognitive test results. This also affects the accuracy of the evaluation of the effects of preschool education on cognitive ability. Second, although propensity score matching is an effective method to confirm causality, it is only an alternative choice in the absence of tracking data, which has the limitations of losing sample size and changing research conclusions due to different confounding variables included in the model. Third, cognitive ability is extremely complex, and relevant research to date has still been in primitive stage. Therefore, the division of cognition by Chinese language and mathematics in this paper cannot cover all of its contents. However, since cognition has different levels and aspects, the measurement results of Chinese and mathematical cognition with different emphases are bounded to reflect these differences. In addition, existing studies have confirmed that the supply of education welfare in China has a ladder-like decline pattern of "good in the east and bad in the west," and education welfare is a manifestation of the problem of educational balance in terms of quality. Compared with preschool education opportunities, the impact of the quality of preschool education on cognitive development should be a more interesting issue. Due to data limitations, the task to offer solutions can be left to the future.

## Data Availability

The data used to support the findings of the study are available from the corresponding author upon reasonable request.

## Conflicts of Interest

The authors declare that they have no conflicts of interest.

## Acknowledgments

This study was supported by the National Social Science Foundation of China (grant no. 16CSH008), Training Plan for Young Backbone Teachers in Colleges and Universities in Henan Province in 2020, General Program of Shaanxi Normal University High-Level Talents Research Startup

Project (grant no. 1110011175), and Fundamental Research Funds for the Central Universities (grant no. 21SZYB11).

## Supplementary Materials

I use the data of Tsinghua University which has been uploaded in supplementary material. (*Supplementary Materials*)

## References


[1] W. H. Sewell, R. M. Hauser, P. Blau, and O. D. Duncan, "The influence of the American occupational structure on the Wisconsin model," *Contemporary Sociology*, vol. 21, no. 5, pp. 598–603, 1992.

[2] R. L. Linn and C. N. Hastings, "A meta analysis of the validity of predictors of performance in law school," *Journal of Educational Measurement*, vol. 21, no. 3, pp. 245–259, 1984.

[3] U. Neisser, G. Boodoo, T. J. Bouchard et al., "Intelligence: knowns and unknowns," *American Psychologist*, vol. 51, no. 2, pp. 77–101, 1996.

[4] N. R. Kuncel, S. A. Hezlett, and D. S. Ones, "Academic performance, career potential, creativity, and job performance: can one construct predict them all?," *Journal of Personality and Social Psychology*, vol. 86, no. 1, pp. 148–161, 2004.

[5] J. Cawley, J. Heckman, and E. Vytlacil, "Three observations on wages and measured cognitive ability," *Labour Economics*, vol. 8, no. 4, pp. 419–442, 2001.

[6] L. S. Gottfredson, "Where and why g matters: not a mystery," *Role of General Mental Ability in industrial, Work, and Organizational Psychology*, vol. 15, no. 1/2, pp. 25–46, 2002.

[7] J. W. Boudreau, W. R. Boswell, T. A. Judge, and R. D. Bretz, "Personality and cognitive ability as predictors of job search among employed managers," *Personnel Psychology*, vol. 54, no. 1, pp. 25–50, 2001.

[8] R. I. Damian, R. Su, M. Shanahan, U. Trautwein, and B. W. Roberts, "Can personality traits and intelligence compensate for background disadvantage? predicting status attainment in adulthood," *Journal of Personality and Social Psychology*, vol. 109, no. 3, pp. 473–489, 2015.

[9] M. Richards and I. Deary, "Cognitive capital in the British birth cohorts: an introduction," *Longitudinal &Life Course Studies*, vol. 1, no. 3, pp. 197–200, 2010.

[10] J. J. Heckman, J. Stixrud, and S. Urzua, "The effects of cognitive and noncognitive abilities on labor market outcomes and social behavior," *Journal of Labor Economics*, vol. 24, no. 3, pp. 411–482, 2006.

[11] J. J. Heckman and D. V. Masterov, "The productivity argument for investing in young children," *Review of Agricultural Economics*, vol. 29, no. 3, pp. 446–493, 2007.

[12] D. J. Noble, K. Martin, L. Qin et al., "What could cognitive capital mean for China's children?," *PsyCh Journal*, vol. 6, no. 2, pp. 153–160, 2017.

[13] L. Wang et al., "Benefit-cost calculations of government investment into early childhood development in rural China," *Journal of East China Normal University*, vol. 37, no. 3, pp. 118–128, 2019, in Chinese.

[14] M. Woodhead, "Pathways through early childhood education in Ethiopia, India and Peru: rights, equity, and diversity," *Childhood Poverty Multidisciplinary Approaches*, vol. 5, pp. 1327–1333, 2012.

[15] J. J. Heckman and L. K. Raut, "Intergenerational long-term effects of preschool-structural estimates from a discrete dynamic programming model," *Journal of Econometrics*, vol. 191, no. 1, pp. 164–175, 2016.

[16] W. S. Barnett, "Benefits of compensatory preschool education," *Journal of Human Resources*, vol. 27, no. 2, pp. 279–312, 1992.

[17] Y. P. Tang, W. Y. Wang, and Y. M. Hu, "Pre-primary education and future academic achievement: empirical research based on PISA2012," *Studies in Foreign Education*, vol. 43, no. 5, pp. 99–109, 2016, in Chinese.

[18] J. Hall, K. Sylva, E. Melhuish, P. Sammons, I. Siraj-Blatchford, and B. Taggart, "The role of pre-school quality in promoting resilience in the cognitive development of young children," *Oxford Review of Education*, vol. 35, no. 3, pp. 331–352, 2009.

[19] C. J. Chen and Q. Liu, "A quasi-experiment study of impact of early childhood education on students' academic achievements when they were 15 years old based on the Shanghai data from PISA," *Studies in Early Childhood Education*, vol. 1, pp. 3–12, 2017, in Chinese.

[20] R. F. Luo, Q. R. Zhao, and M. He, "The empirical analysis of early childhood education of poor rural China," *Studies in Early Childhood Education*, vol. 1, pp. 7–10, 2009, in Chinese.

[21] K. Burger, "How does early childhood care and education affect cognitive development? An international review of the effects of early interventions for children from different social backgrounds," *Early Childhood Research Quarterly*, vol. 25, no. 2, pp. 140–165, 2010.

[22] J. J. Heckman, *Giving Kids a Fair Chance*, The MIT Press, Cambridge, MA, USA, 2013.

[23] K. G. J. Duncan and G. J. Duncan, "Can early childhood interventions decrease inequality of economic opportunity?," *RSF: The Russell Sage Foundation Journal of the Social Sciences*, vol. 2, no. 2, pp. 123–141, 2016.

[24] H. M. Wang, Y. X. Wu, and C. Huang, "Family socioeconomic status, preschool education and the adolescents' cognitive & non-cognitive abilities," *Youth Studies*, vol. 6, pp. 46–57, 2017, in Chinese.

[25] L. Zheng, Q. Y. Weng, and X. Gong, "Does preschool attendance matter for the urban-rural cognition gap of middle school students? Evidence from CEPS," *Sociology Study*, vol. 34, no. 3, pp. 122–145, 2019, in Chinese.

[26] X. Gong, D. Xu, and W.-J. Han, "The effects of preschool attendance on adolescent outcomes in rural China," *Early Childhood Research Quarterly*, vol. 37, pp. 140–152, 2016.

[27] R. J. Gregory, *Psychological Testing: History, Principles, and Applications*, Machinery Industry Press, Beijing, China, 2012.

[28] H. Xu and Y. Xie, "The causal effects of rural-to-urban migration on children's well-being in China," *European Sociological Review*, vol. 31, no. 4, pp. 502–519, 2015.

[29] R. B. Cattell, *Intelligence: Its Structure, Growth and Action*, North-Holland, Amsterdam, Netherlands, 1987.

[30] J. Blake, "Family size and the quality of children," *Demography*, vol. 18, no. 4, pp. 421–442, 1981.

[31] Y. L. Zhang, "Education in China: the expansion of inequality and its driving force," *Popular Tribune*, vol. 3, pp. 25–29, 2013, in Chinese.

Hindawi
Complexity
Volume 2021, Article ID 4923532, 18 pages
https://doi.org/10.1155/2021/4923532

*Research Article*

# Spatiotemporal Characteristics and Resilience of Urban Network Structure during the Spring Festival Travel Rush: A Case Study of Urban Agglomeration in the Middle Reaches of Yangtze River in China

**Shimei Wei and Jinghu Pan**

*College of Geography and Environmental Science, Northwest Normal University, Lanzhou 730070, China*

Correspondence should be addressed to Jinghu Pan; panjh_nwnu@nwnu.edu.cn

Received 5 May 2021; Revised 26 August 2021; Accepted 7 September 2021; Published 28 September 2021

Academic Editor: Alejandro F. Villaverde



With the increasing trend of globalization, large-scale and diffuse population flow have become vital carriers characterizing users' spatial behaviors. Network analysis provides a new perspective to uncover the topology and evolution of the population flow and understand its influence on regional development. By gathering the Autonavi migration index during the Spring Festival travel rush (SFTR) in 2019, 2020, and 2021, the population flow networks among 31 cities of urban agglomeration in the middle reaches of the Yangtze River were constructed to analyze spatiotemporal dynamic characteristics and explore the structure resilience. Results show that although the changing trends of population flow during the 40-day SFTR of 2019, 2020, and 2021 are consistent, the population floating scale in 2020 and 2021 shows remarkable abnormalities before and after the Spring Festival due to the need for prevention and control of COVID-19. The intensity of population floating of the regional urban network in 2020 was the weakest, and Changsha became the focus of most population flow, while Wuhan was the most advantageous city in 2019 and 2021. As the third core city in the regional network, the siphon effect of Nanchang was still weak. A situation of tripartite confrontation in the region is formed. However, the higher intensity of population flow in 2021 increased the instability of the regional urban network, potentially exposing the region to higher risks and pressures. Therefore, it is necessary to pay more attention to the peripheral cities to improve regional resilience.

## 1. Introduction

With the increasing trend of globalization and the promotion of technology, the improvement in high-speed connection infrastructures (aviation, high-speed rail, telecommunications, Internet, etc.) has strengthened the connections between cities. The hierarchical space of cities dominated by administrative divisions tends to change into the functional space of the relational network. Especially in the recent years, with the rapid development of society and economy, the connections among cities tend to be complicated and diversified, and the dynamic urban network in economy, technology, and information-driven by population flow has become a hot research topic. The "space of flows" theory proposed by Castells [1] can describe vividly the phenomenon of spatial connection in city networks. The urban network is usually represented by a network diagram, which is composed of nodes and links (edges). In a regional urban network, nodes and links correspond to cities and connections between cities, respectively. Flow elements, population flow, information flow, capital flow, technology flow, etc., gradually get rid of the limitation of geographical space, and communicate and propagate in larger regions. With the rise of flow space theory, regional urban networks were constructed based on the data of various resource flows among cities, and the special characteristics of their spatial structure were studied, which provide scientific support for the strategies of regional development.

As one of the largest and most far-reaching geographic processes since China's reform and opening up, population migration, which is an important carrier of flow elements among cities, is regarded as an activity in which essential productive factors are reconfigured in space [2, 3]. The Chinese Lunar New Year, the Spring Festival, is the most celebratory time of the year in China, during which a massive human migration takes place as individuals travel back to their hometowns [4]. China's Spring Festival travel rush (SFTR) is described as rarely seen population floating in the world and is the largest and periodically floating population in human history due to the particularity of its population flow scale. National Development and Reform Commission of China stated that the annual spring migration (called Chunyun in China) often lasts for 40 days. During the SFTR in 2019, 2020, and 2021 (from January 21, 2019 to March 1, 2019; from January 10, 2020 to February 18, 2020; from January 28, 2021 to March 8, 2021), there had been about 3 billion, slightly higher than 3 billion, and 1.7 billion passengers between urban and rural areas, respectively. The large-scale and diffuse population flow promotes the reaggregation and diffusion of social and economic factors to some extent. Meanwhile, it is, however, a challenge and examination to the transportation infrastructure systems [5].

It is generally known that the Corona Virus Disease 2019 (COVID-19) was first identified and reported in Wuhan of China, the capital city of Hubei province with 11 million inhabitants and the most significant transport hub in Central China [6]. Due to its characteristics of person-to-person contact, extensive population flow has substantially increased social contacts in public, which caused COVID-19 to reach essentially everywhere [7]. On December 31, 2019, before the Chinese Lunar New Year (began on January 24, 2020), about 5 million people left Wuhan [7]. Further spatial spread of COIVD-19 was of great concern because of the upcoming Spring Festival Holiday (from January 24, 2020 to January 31, 2020). To control the rapid spread of COVID-19, Chinese government ordered lockdown policies in late January 2020 [8]. Particularly in Wuhan, where the largest number of infected people live, the Municipal Government completely closed the city on January 23, 2020. Nevertheless, the global multipoint outbreak of COVID-19 has evolved from a public health emergency into a nontraditional disaster. It directly exposes that without the support of a solid healthy and safe urban development environment, the urban prosperity and vitality may strengthen the vulnerability of the city, and even cause short-term "shock" and long-term "sequelae" of the city. Additionally, in the process of coping with the COVID-19, China's highly effective coordination mechanism among cities, nodes control of urban traffic network, construction of interregional public health networks, and regional coordinated production of antiepidemic materials, to a certain extent, reflect the coordinated cooperation among cities in response to disasters and crises, which will form a benign network synergy effect. Therefore, faced with long-term chronic pressures and short-term impacts, the underlying questions are how to investigate spatiotemporal characteristics of population mobility networks and how to measure the capacity of regional resilience.

Regional resilience is a rising concept in international research fields. It is usually measured by urban network structure resilience, focusing on the capacity of a city system to restore, maintain, or improve the original network characteristics and important functions in response to regional shocks [9]. With the acceleration of the process of urbanization, the uncertainties and unknown risks facing urban systems are also increasing [10]. Especially in the recent years, the sustainable development of cities has been seriously threatened by natural disasters, human disasters, the economy, the environment, safety, transportation, and society, which directly affect the life safety and quality of life of urban residents. In these cases, a system can fail, leading to a major reduction or complete loss in performance with respect to some or all measures. Moreover, if the function of a city fails, other cities in the region may no longer operate normally, but consider and select new connection objects. Accordingly, the pressures of other cities will increase, which bring some challenges to the sustainable development of regions. As a typical embodiment of regional spatial characteristics, the urban network structure has an impact on regional sensitivity to shocks, adaptability, and the ability to develop new growth paths [11]. Since Holling (1973) [12] first introduced the concept of resilience into ecosystem studies, it has been routinely used in research in disciplines ranging from environmental research to materials science and engineering, psychology, sociology, and economics [13, 14]. Recent studies have shown that the concept and definition of resilience can also be applied to regional networks [15]. In previous studies on the resilience of regional networks, some scholars focused on the resilience characteristics of finance networks [16], infrastructure networks [17], transport networks [18, 19], and so forth. Although still gathering momentum, the previous studies mainly focus on theoretical exploration [20]. However, in the face of external shocks, especially against the backdrop of the enormous negative impact of the COVID-19 on the global economy, the assessment of the resilience of the urban network structure is of great and far-reaching significance to enhance urban abilities to prevent system failures and to make contingency plans and promote the high quality and sustainable development of the region.

As the main carrier of flow elements, intercity population migration pushes the rapid flow and optimization of production factors within a region. Therefore, urban population migration networks based on geospatial "flow data" have attracted increasing attention of scholars in the recent years. Some earlier studies mainly used static data, such as census data and population sampling survey data, which cannot reveal the increasingly complex interaction between cities from the perspective of space of flows [21]. As mobile Internet has found its way into people's everyday life, the global positioning system (GPS), location-based services (LBSs), and location sharing services are increasingly used to discover the geographic locations of users and their individual preferences, travel routes, activity trails, and social networks, thereby elucidating their daily spatiotemporal behavior [3]. LBSs can provide accurate information such as starting points, ending points, and even individual spatial trajectories. The association between regional entities can be visualized as a variety of practical measurable association flows. Accordingly, related studies based on various big data are in full swings, such as microblog check-in [22], Tencent migration [3], Baidu migration data [23], and Facebook [24]. From the point of view of the research time-period, many studies took the public holidays as examples, such as China's Spring Festival, the National Day, and the Mid-Autumn Festival travel rush, the daily time-period was also involved [3, 22, 25, 26]. In addition, the research contents have included population flow intensity, scale, spatial distribution, travel pattern, etc. The spatial scale of related researches has involved from global to local cities. Overall, multiperspective and multiscale research on population flow networks is becoming an important way to understand the complex and dynamic spatial relationships between cities based on LBS data. However, the existing literature on population flow networks is mostly focused on a certain year, lacking vertical comparison from the perspective of evolution. Second, although directed weighted population flow networks were constructed, the directionality of population flow has been neglected in the analysis process, thereby leading to the underestimation or overestimation of the asymmetric relationship between urban nodes. The above illustrates that in the era of big data with rapid urban developments, the spatial patterns and scale characteristics of urban networks have already changed from static to complex spatiotemporal dynamics. Therefore, it is necessary to study the dynamic pattern characteristics of urban networks and their resilience based on spatiotemporal big data. In addition, Autonavi Map (AMap) application records population flow routes using its LBS device in real-time, dynamically, completely, and systematically, producing Autonavi migration big data. The available data period is from June 2018 to present, which makes it possible to characterize the population flow during the SFTR in different periods.

From the national strategy to promote the Yangtze River Economic Belt, urban agglomeration in the middle reaches of the Yangtze River plays a vital role in China's regional development pattern because it is an important support to promote the coordinated development of the Yangtze River Economic Belt. Moreover, Wuhan, a core city of the urban agglomeration, served as the high-risk area for the outbreaks of COVID-19 during the SFTR in 2020. The large-scale population flow among regional cities, especially within the urban agglomeration, further exacerbated the spread of COVID-19. It is, therefore, necessary to clearly understand the population flow between cities within the region, determine its spatiotemporal characteristics, and measure its network structure resilience. Taking 31 cities of the urban agglomeration in the middle reaches of Yangtze River as the research subjects, through gathering the Autonavi migration big data during the SFTR in 2019, 2020, and 2021, the spatiotemporal complex characteristics of the urban network of urban agglomeration in the middle reaches of the Yangtze River are revealed herein from the perspective of population flow. Then, the structural resilience of the regional urban networks is estimated when perturbations or major events occurred. Furthermore, it is expected to put forward specific strategies and suggestions in combination with the research conclusions to promote the region for further development. This study not only provides a new research perspective and methodological reference for research on urban networks, but also contributes to further regional coordination and sustainable development.

## 2. Study Area

The urban agglomeration in the middle reaches of the Yangtze River, also known as the "Triangle of Central China", is a super large national urban agglomeration located in the middle reaches of the Yangtze River, mainly formed by the Wuhan metropolitan area, Changsha-Zhuzhou-Xiangtan city group, and Poyang Lake city group. In this study, 31 cities defined in the development plan of urban agglomeration in the middle reaches of the Yangtze River issued by the National Development and Reform Commission in April 2015 were selected as the research objects (Figure 1). Specifically, the study area includes 13 cities in the Hubei province (Wuhan, Huangshi, Ezhou, Huanggang, Xiaogan, Xianning, Xiantao, Qianjiang, Tianmen, Xiangyang, Yichang, Jingzhou, and Jingmen), 8 cities in the Hunan province (Changsha, Zhuzhou, Xiangtan, Yueyang, Yiyang, Changde, Hengyang, and Loudi), and 10 cities in the Jiangxi province (Nanchang, Jiujiang, Jingdezhen, Yingtan, Xinyu, Yichun, Pingxiang, Shangrao, Fuzhou, and Ji'an), which covers a total area of approximately $31.7 \times 10^4\,\text{km}^2$. Among them, Wuhan, Changsha, and Nanchang, on the one hand, as provincial capitals, lead the social and economic development of the Wuhan metropolitan area, Changsha-Zhuzhou-Xiangtan city group, and Poyang Lake city group, respectively. On the other hand, the three core cities of the urban agglomeration form a "tripartite confrontation" spatial structure and play important dispersion and aggregation functions in the region. At the end of 2019, the regional GDP was 9.3833 trillion yuan, and the total resident population was 130.6491 million. As an important part of the Yangtze River Economic Belt, the urban agglomeration in the middle reaches of the Yangtze River not only connect the eastern and western regions, connect the South and the North, but also are the central area to implement the strategy of promoting the rise of the central region, comprehensively deepening the reform and opening up and promoting the new urbanization. What's more, the urban agglomeration is positioned as a new growth pole for China's economy, a leading area for new urbanization in the central and western regions, a demonstration area for inland open cooperation, and a leading area for "two oriented" social construction. Therefore, it plays an important role in China's regional development pattern.

## 3. Data and Methods

*3.1. Data Sources.* Taking 31 cities of the urban agglomeration in the middle reaches of the Yangtze River as examples,

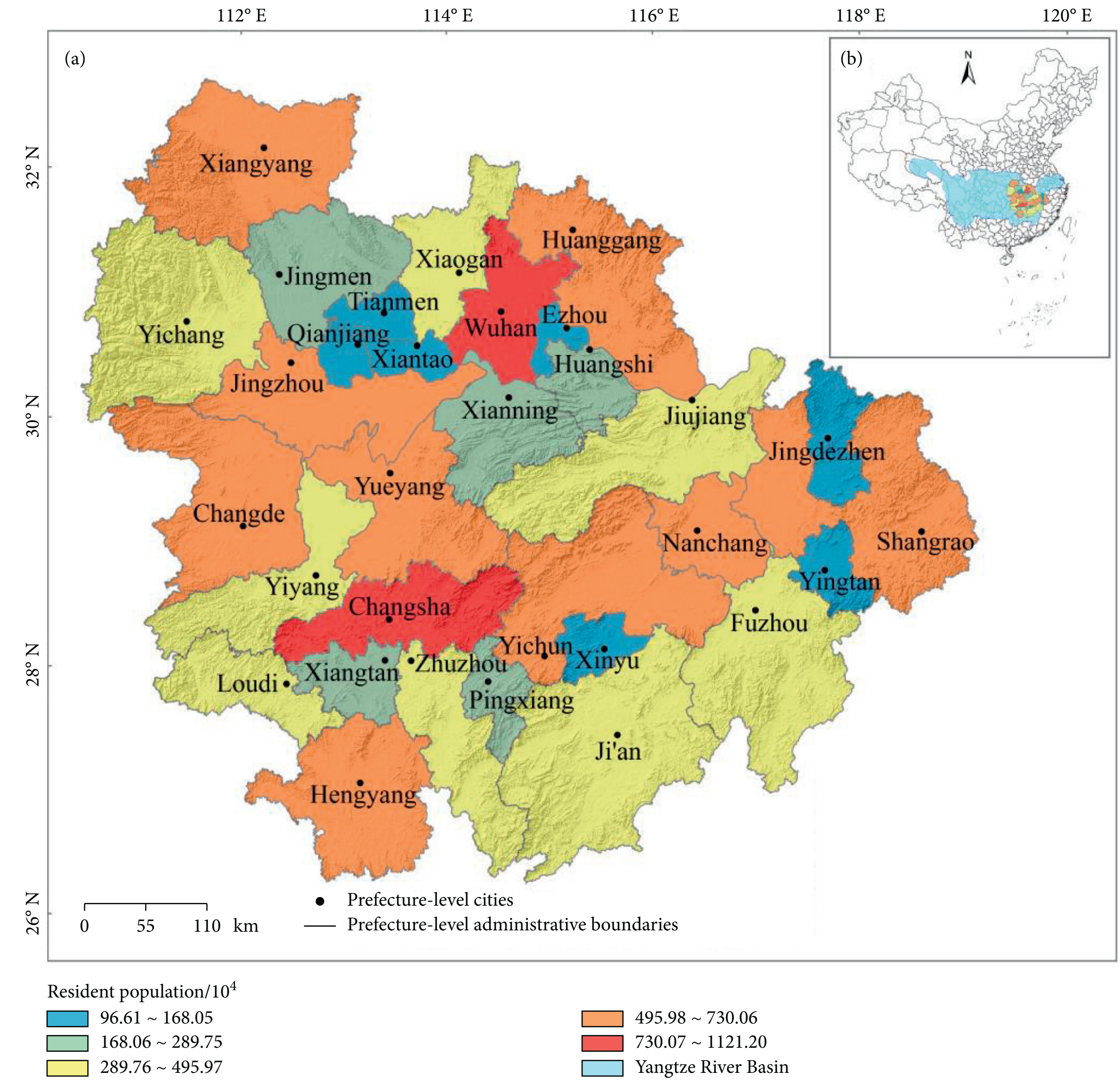



Figure 1: Study area: (a) spatial distribution of resident population; (b) the position of the study area in the Yangtze River Basin and China.

through gathering the Autonavi migration index during the Spring Festival travel rush in 2019, 2020, and 2021, the intercity population migration networks within the urban agglomeration are constructed to analyze spatiotemporal characteristics and resilience changes. The data analyzed in this study, Autonavi migration index, were derived from traffic big data provided by the Autonavi Company, one of the solution providers of digital map, navigation, and location service in China. According to statistics from the 2021 China Mobile Internet Development Report, as of the end of 2020, the number of monthly active users of Amap reached 555.86 million, while the Baidu map ranked second with 434.73 million users (https://www.questmobile.com.cn/research/report-new/143). That is, the Amap service ranked first among the travel service apps (Baidu Map, Didi Chuxing, and Tencent Map). Autonavi migration data, similar to Tencent migration and Baidu migration data, represents the index of population migration between cities, which is obtained by the number of migrants in each city and the total number of migrants in China. Each record of Autonavi migration data is mainly composed of four fields: source city, target city, actual migration index, and migration willingness index. The actual migration index is the record after completing the travel purpose from one city to another via Autonavi navigation service in Amap, which, to a large extent, represents the population travel on the highway. The migration willingness index refers to the behavior of searching target places and then planning driving routes via Autonavi navigation, that is, people's travel desire, rather than the real migration behavior. Through screening and cleaning the data for 120 days during the SFTR in 2019 (from January 21 to March 1, 2019), 2020 (from January 10 to February 18, 2020), and 2021 (from January 28 to March 8, 2021), the available data of three periods are obtained.

### 3.2. Methods

*3.2.1. Construction of Population Flow Network.* According to the complex network theory, regarding cities and connection between cities as nodes and edges, respectively, the population flow network among 31 cities was described. A directional weighted matrix $L = (L_{ij})$ was used to characterize the flow of the population during one day, where the $L_{ij}$ is the population flow intensity from city $i$ to city $j$. Then, $40 \times 3$ directed weighted asymmetric matrices of $31 \times 31$ were constructed based on Autonavi migration data. The matric can be expressed as follows [3]:

$$L = \begin{array}{c} \\ i_1 \\ i_2 \\ \cdot \\ \cdot \\ \cdot \\ i_{(n-1)} \\ i_n \end{array} \begin{array}{c} \begin{array}{cccccc} j_1 & j_2 & \cdot\cdot\cdot & j_{(n-1)} & j_n \end{array} \\ \begin{bmatrix} 0 & L_{12} & \cdot\cdot\cdot & L_{1(n-1)} & L_{1n} \\ L_{21} & 0 & \cdot\cdot\cdot & L_{2(n-1)} & L_{2n} \\ \cdot & \cdot & \cdot\cdot\cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot\cdot\cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot\cdot\cdot & \cdot & \cdot \\ L_{1(n-1)} & L_{2(n-1)} & \cdot\cdot\cdot & 0 & L_{n(n-1)} \\ L_{1n} & L_{2n} & \cdot\cdot\cdot & L_{n(n-1)} & 0 \end{bmatrix} \end{array} \quad (1)$$

*3.2.2. Social Network Analysis Methods*

(1) *Weighted degree.* The weighted degree represents the sum of the weights of the arcs connected with a city node, the larger the value, the stronger the connection between cities. In a directional weighted network, the weighted degree of a node is the sum of the weighted out-degree and the weighted in-degree of the one. Weighted out-degree measures the total population flow intensity from city node $i$, while weighted in-degree measures the total population flow intensity to node $i$. The formula is as follows:

$$W_i = W_i^{\text{in}} + W_i^{\text{out}}, \quad (2)$$

where $W_i$ is the weighted degree of city $i$, $W_i^{\text{in}}$ is the weighted in-degree of city $i$, and $W_i^{\text{out}}$ is the weighted out-degree of city $i$.

(2) *Node symmetry.* Node symmetry is used to describe the difference reflected by the in-degree and out-degree of every node [27]. However, the number relationship of intercity migration cannot be well reflected only by the in-degree and out-degree. Therefore, this study introduced the population flow intensity of the directed weighted network to calculate the node symmetry index (NSI) and then judge whether the city node is the receiver or the sender. The formula of NSI of the city node $i$ is as follows:

$$\text{NSI}_i = \frac{W_i^{\text{in}} - W_i^{\text{out}}}{W_i^{\text{in}} + W_i^{\text{out}}}. \quad (3)$$

(3) *Clustering coefficient.* The clustering coefficient is used to describe the interconnection level of nodes [28]. When certain nodes are closely connected, they can form a network cluster. The formula of clustering coefficient is as follows [25]:

$$C_i = \frac{2B_i}{m_i(m_i - 1)}, \quad (4)$$

where $C_i$ is the clustering coefficient and $B_i$ is the number of paths between the node and the neighboring nodes of $m_i$.

(4) *PageRank algorithm.* PageRank algorithm is originally used to evaluate the rank of a web page through the complicated hyperlink relationships in a network [29], revealing the importance of a particular web page relative to other web pages in the network. It is used herein to measure the criticality of city nodes in the population flow network. The calculation equation of the PR value of node $i$ is as follows [30]:

$$\text{PR}(p_i) = \frac{1-d}{N} + d \sum_{p_j \in M(i)} \frac{\text{PR}(p_j)}{L(j)}, \quad (5)$$

where $N$ is the number of all nodes, $d$ is the damping coefficient (a value between zero and one), $M(i)$ is the set of nodes connected to node $i$ and out of the chain (in-degree), and $L(j)$ is the number of nodes connected by node $j$ to the outside (out-degree).

*3.2.3. Assessment of Network Structure Resilience.* Network structure resilience refers to the ability of the network system to respond to sudden external shocks and disturbances. For the urban network system, the interference to a certain urban node may lead to the domino effect, thereby affecting the partial cascading of the network [31]. The resilience of the network structure herein is measured by the response of network after node failure and node attack. Node failure mainly considers the impact of natural disasters on different urban nodes, such as mud-rock flow, snowstorm, haze, etc. The resilience of the urban network structure in the face of interference is observed via removing the city nodes in turn. Node attack mainly considers the impact of man-made damages, such as terrorist attacks, military conflicts, or other artificial forces, remove nodes in the descending rank of node criticality (PR value), simulating the nodes fail from the most "important" ones [17]. Both strategies assume that the node will fail immediately after removal, and all edges are directly connected to the node will also be removed.

Network efficiency is a physical quantity used to describe the diffusion ability of elements among the networks, which is generally evaluated by path-related indicators of the network. In a network with high network efficiency, the propagation and exchange of elements between nodes can be realized more quickly, which is conducive to promoting the learning, innovation, and communication among nodes, and can enhance the resistance of the network to external shocks so that the network has a higher resilience. The network efficiency ($E$) has been widely regarded as an important metric to measure the network response [32, 33]; the larger the value, the better the performance of the network. The changes to its value indicate changes in global efficiency due to disturbances. For the directional weighted network, the formula is as follows:

$$E = \frac{1}{N(N-1)} \sum_{i \neq j \in G} \frac{1}{D_{ij}} \times AW_{ij}, \tag{6}$$

where $D_{ij}$ is the shortest path length from urban node $i$ to $j$, $N$ is the number of urban nodes in the network, and $G$ is the set of urban nodes in the network after removing a certain node, $AW_{ij}$ is the average weighted degree of the edges in the shortest path from node $i$ to node $j$.

## 4. Results

*4.1. Characteristics of Population Migration during SFTR.* To analyze the characteristics of population migration of urban agglomeration in the middle reaches of Yangtze River during SFTR, a time-sequential distribution chart of the daily summary (average) migration scale during the SFTR in 2019, 2020, and 2021 was produced (Figure 2). Figure 2 shows the following characteristics of the population flow. First, in a year, the changing trend of migration willingness and actual migration scale is basically consistent during the study period, and the migration willingness scale is always higher than the actual migration scale. Second, the migration willingness and actual migration index before the Spring Festival (the 16th day) shows a trend of first increase and then decrease. Due to the influence of the prevention and control policies of COVID-19, the people migration scale in 2020 and 2019 is higher than that in 2021. In addition, compared with 2019, the peak of population migration before the Spring Festival in 2020 and 2021 is slightly delayed. Third, due to the particularity of COVID-19, to prevent COVID-19 from rapidly spreading with population movements, the Wuhan Municipal Government completely closed the city at 10 : 00 a.m. on January 23, 2020 (the 14th day). Nevertheless, approximately 5 million people still left Wuhan [34]. Therefore, the migration willingness and actual migration index after the Spring Festival (the 16th day) in 2020 show a downward trend, which is far lower than the migration level in the same period in 2019 and 2021. With the effective prevention and control of COVID-19, the population migration after the Spring Festival in 2021 shows a rapid increasing trend compared to that before the year. On the last day of the Spring Festival Holiday (the 21st day), the population migration scale of the study area in 2021 exceeds that in 2019, and reaches the largest scale of people migration during SFTR. As the last traditional festival during SFTR, the Lantern Festival also ushered in a small peak of population flow. Since then, the population transportation of SFTR has gradually come to an end.

### *4.2. Spatiotemporal Characteristics of Population Flow Networks*

*4.2.1. Spatiotemporal Patterns of Popular Flow Networks.* With the help of ArcGIS software, the population flow networks were visualized based on the summarized directional weighted matrix of population flow intensity. The overall spatiotemporal pattern of population flow networks of urban agglomeration during the SFTR in the middle reaches of Yangtze River is presented in Figure 3. Meanwhile, the population migration intensity is divided into four levels, the higher the level, the stronger the population migration scale between cities.

Figure 3(a) shows the spatial pattern of the population flow network in 2019. From the results of the grading map, population flow between cities in the first level (81–177) mainly revolves around 12 routes around Wuhan, Changsha, and Nanchang, showing a characteristic of large-scale and basically symmetrical population migration between neighboring cities. The spatial pattern of the second level (41–80) still presents a central radial structure, which to some extent is a complement of the surrounding cities of the first level. Only the Changsha–Zhuzhou–Xiangtan city group shows the population flow across city space. In addition, it is found that there is a relatively strong connection between Xiangtan and Zhuzhou, and no major core cities were involved. The third level (10–40) mainly reflects the interaction among cities around the core cities, characterized by more in the northwest and less in the southeast. The fourth level (<10) accounts for about 84% of all routes, indicating that the intensity of population flow across region space in most cities is still weak.

The spatial pattern of the population flow network in 2020 (Figure 3(b)) indicates that the population flow routes of the first three levels are extremely sparse. The number of routes is 3, 12, and 68, respectively. The routes in the first level are only Wuhan–Xiaogan, Xiangtan–Changsha, and Changsha–Yiyang. As the core city of the Poyang Lake city group, Nanchang only appears in the second level, and the connection with Yiyang is relatively weak. The rest of the population flow connections also correspond to the first level in 2019. It indicates that the outbreak of COVID-19 in Wuhan had a serious impact not only on Wuhan and its surrounding cities but also on the closely linked cities within the urban agglomeration. Although there are population flow routes across region space in the third level, the number is few. For instance, Huanggang–Jiujiang, Jingzhou–Yueyang, and Changde–Jingzhou. The population flow routes in the fourth level have a relatively large scale, accounting for about 90% of all routes. This illustrates that during the epidemic in Wuhan, the closure policy had a greater impact on all cities within the urban agglomeration, causing an attenuation on the Spring Festival travel rush in China.

The population flow network during the SFTR in 2021 is shown in Figure 3(c) with levels 1 to 3 including 16, 23, and 103 routes, respectively. Compared with that of 2019, its population migration intensity has increased; on the other hand, both Wuhan and Nanchang are joined by new urban

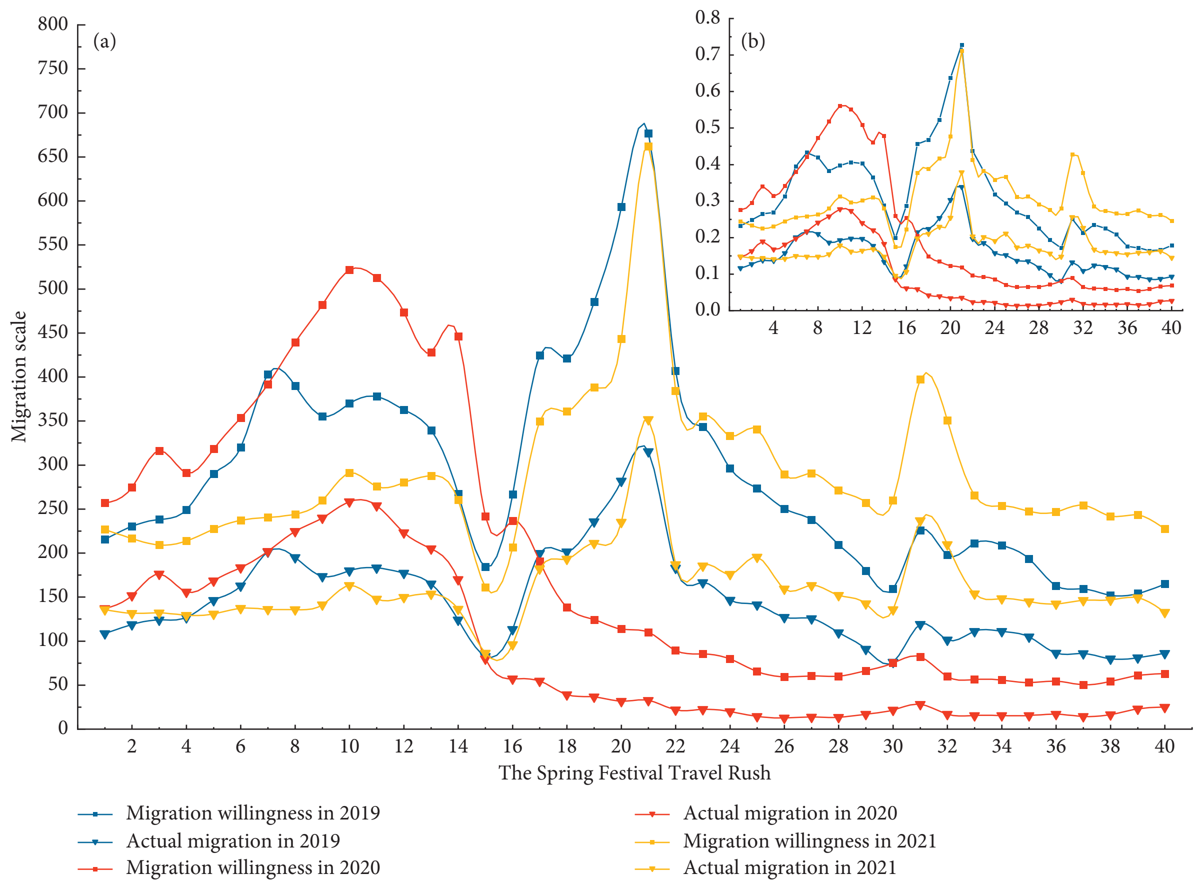



Figure 2: The population migration scale of the study area during the SFTR: (a) daily summary migration scale; (b) daily average migration scale. Blue, red, and yellow segments represent 2019, 2020, and 2021, respectively. The box and inverted triangle represent the migration willingness and the actual migration scale, respectively.

nodes at the first level, such as Ezhou and Jiujiang. In the second level, the population flow routes around Poyang Lake city group have increased significantly, which are reflected in Yingtan and Fuzhou. There are no significant differences in the third level of the population flow networks in 2019 and 2021. Surprisingly, the number of the routes in the fourth level is almost the same as in 2019 and 2021.

From the perspective of the urban agglomeration in the middle reaches of the Yangtze River, there are 930 routes in each population flow network during the three periods, that is, the population flow networks are all strongly connected graphs. However, the intensity of population flow during the SFTR in 2019, 2020, and 2021 is quite different. The maximum migration scale of the population flow network during the SFTR from 2019 to 2021 is 177, 87, and 213, respectively. Correspondingly, the average migration scale of each route is 6.18, 3.68, and 6.90, respectively. In addition, the spatial distribution of population flow during the SFTR in the urban agglomeration also shows temporal and spatial differences. In general, the networks present a spatial distribution pattern of "dense on the northwestern side and sparse on the southeastern side". Polycentricity is a prominent feature of the region. Specifically, the population flow between core cities (Wuhan, Changsha, and Nanchang) is relatively weak. In the first and second levels, no corridors connecting the three suburban groups are formed. The population migration intensity between Wuhan and Changsha during the SFTR in 2019 and 2021 is only at the third level. What's worse, its connection is weaker in 2020, only at the fourth level. Second, the population flow connections are mainly gathered on the northwestern side, between the Wuhan metropolitan area and the Changsha–Zhuzhou–Xiangtan city group, which is the main driving force for the development of the region. Only in 2020, due to the quarantine of Wuhan, the connection between them tends to weaken. In comparison, the interaction between the Poyang Lake city group on the southeast side and other regions is weaker, especially in 2020. Nevertheless, the results of their urban cluster structure reveal that the population flow networks of three periods surprisingly have the same urban cluster structure. That is, 13 cities in the Hubei province form cluster 0, 8 cities in the Hunan province form cluster 1, and 10 cities in the Jiangxi province form cluster 2. The cluster structure in different periods is completely consistent with the provincial boundary, indicating that there is an obvious provincial boundary segmentation effect on the urban agglomeration, which makes the population flow elements within the urban agglomeration focus only on the interior of provinces, and the integration level of the region is relatively low.

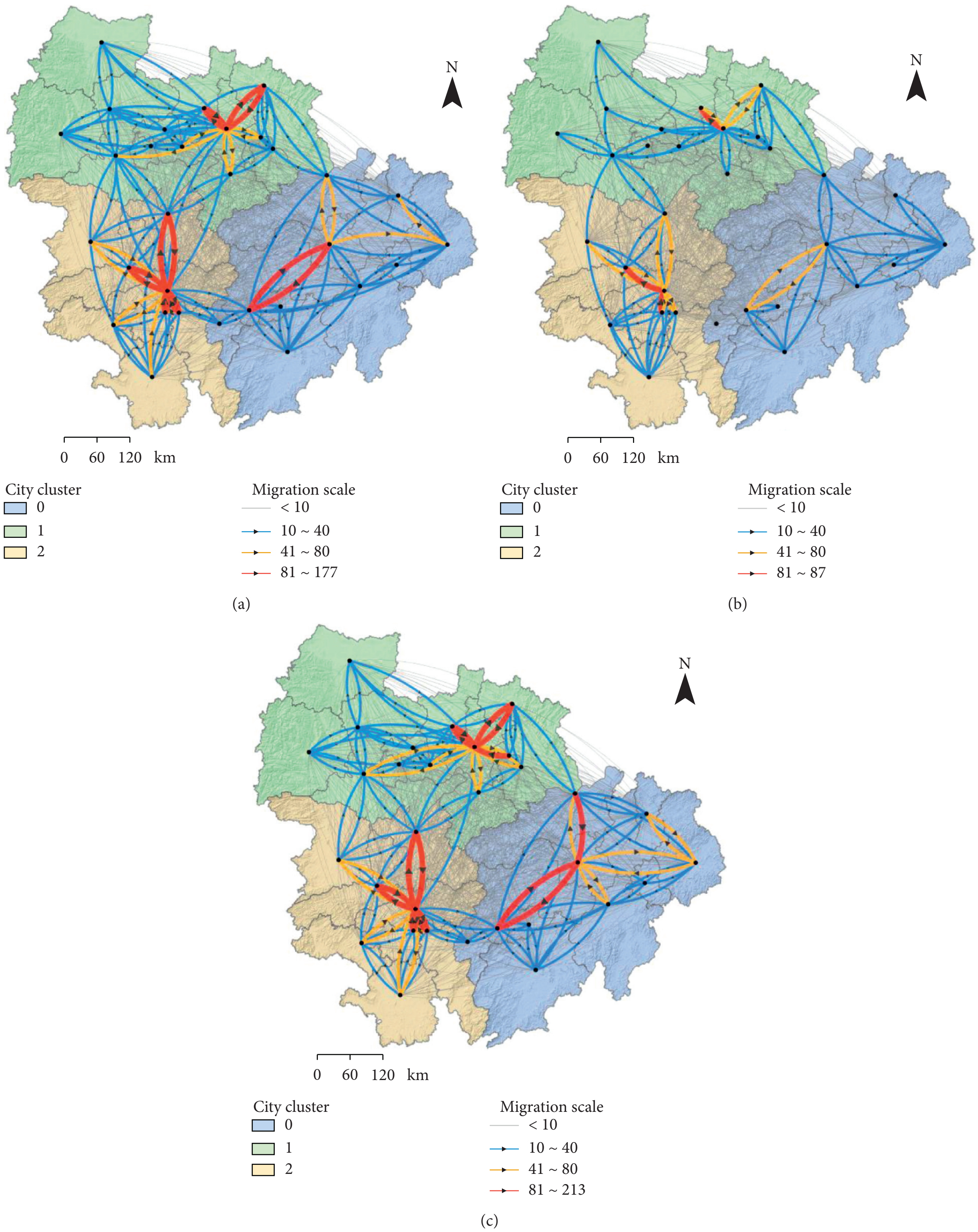



Figure 3: Population flow networks of the middle reaches of Yangtze river: (a) 2019; (b) 2020; and (c) 2021. The color of the route represents the intensity of population flow, and the change from gray to red represents an increase in the intensity. The color of the polygon represents city cluster.

*4.2.2. Popular Migration Routes of the Study Period.* By drawing the actual migration index and the migration willingness index, the popular migration routes (Figure 4) during the three periods of SFTR were obtained. The popular migration routes in 2019 are Xiaogan–Wuhan, Wuhan–Xiaogan, Xiangtan–Changsha, Huanggang–Wuhan, Wuhan–Huanggang, Yiyang–Changsha, and Changsha–Xiangtan. In 2020, the routes are Xiangtan–Changsha, Wuhan–Xiaogan, Changsha–Yiyang, Changsha–Yueyang, Xiaogan–Wuhan, Changsha–Xiangtan, and Wuhan–Huanggang. In 2021, the routes are Xiaogan–Wuhan, Xiangtan–Changsha, Wuhan–Xiaogan, Yiyang–Changsha, Wuhan–Huanggang, Huanggang–Wuhan, and Changsha–Xiangtan.

Combined with the popular flow networks (Figure 3), we can see clearly that the most popular migration routes in 2019 mainly focus on Wuhan, involving the nearest cities Xiaogan and Huanggang. Then, the routes are Xiangtan and Yiyang to Changsha. They reflect the migration effect from surrounding cities to core cities. During the SFTR in 2020, under the influence of COVID-19, the popular migration routes in the study area have changed greatly compared with that in 2019. As a subcore city, Changsha has become the dominant city of population migration. In addition, it is characterized by the popular migration from the core city to the surrounding cities. Popular migration routes in 2021 are similar to those in 2019, mainly centered around the two core cities of Wuhan and Changsha, and the difference is mainly reflected that the population migration intensity in 2021 is higher than that in 2019.

*4.2.3. Dominant Flow and Node Symmetry.* By extracting the dominant flow (maximum inflow or outflow) data of each city in the urban agglomeration, the population flow paths of the dominant flow during the period of SFTR were drawn. The results are expressed in the form of curves, whose colors represent different city clusters (Figure 5). In addition, in order to analyze the characteristics of urban population inflow and outflow more concretely and intuitively, we overlay the results of the node symmetry index with the dominant flow. According to the results of the node symmetry index, the city nodes were divided into four types, such as strong outflow, weak outflow, weak inflow, and strong inflow nodes.

From the dominant flow and city clusters, the number of city clusters of population dominant flow in the SFTR periods of 2019, 2020, and 2021 in the study area is 4, 5, and 3, respectively. There are different structures of city clusters corresponding to each period, among which 2019 and 2021 are more similar, showing a stronger clustering degree of 2021 than that of 2019. Specifically reflected as follows: for cluster 0, 13 cities in Hubei province not only have the same clustering characteristics but also have the same direction of population dominant flow in 2019 and 2021. However, the cluster was divided into three smaller city clusters (clusters 0, 3, and 4) in 2020 (Figure 5(b)). In addition, Jingmen, Jingzhou, Xiangyang, and Tianmen, which are used to taking Wuhan as the first inflow city, changed their inflow targets to the adjacent marginal cities. For cluster 1, the structure of population dominant flow in 2019 and 2021 is the same. Pingxiang, which belongs to cluster 2 in 2019 and 2021, had been attracted to cluster 1 with Changsha as the core in 2020. Besides, the dominant flow city in Changsha changed from Xiangtan to Yiyang. For cluster 2, although the three periods have different characteristics, on the whole, the changing trend is from weak to strong. In 2019, cluster 2 has not yet been merged with cluster 3 on the edge of the urban agglomeration. However, as the dominant flow city in Shangrao changed from Jingdezhen to Nanchang and the dominant flow city in Pingxiang changed from Changsha to Yichun, 10 cities in the Jiangxi Province (city groups around Poyang Lake) within the urban agglomeration formed a larger city cluster. Therefore, there were only three urban clusters in 2021, and they were relatively complete boundaries divided by provincial regions.

Node symmetry describes the difference between the population inflow and outflow interactions for every city node. The outflow nodes are the nodes with a negative node symmetry index. They are more important as senders in the population flow network. According to the negative value, they were divided into two types: strong outflow and weak outflow nodes. Correspondingly, the inflow nodes, including strong inflow and weak inflow nodes whose node symmetry index is positive, are primarily receivers in the network.

From the perspective of the node types (Figure 5), the strong outflow nodes mainly distribute around the core cities (Wuhan, Changsha, and Nanchang). The spatial distribution of the three periods is slightly different. As far as Wuhan metropolitan area is concerned, there are only two strong outflow cities (Xiantao and Ezhou) in 2019, all weak outflow and inflow cities in 2020, and three strong outflow cities (Xiantao, Ezhou, and Xiaogan) in 2021. Consequently, Wuhan has changed from a weak inflow node in 2019 to a strong inflow node in 2021. As a strong inflow city in 2019 and 2021, Changsha changed its direction and became a weak inflow node in 2020. Although Nanchang is known for its weak inflow, it turned around in 2020 and became a weak outflow node. Some weak outflow cities in 2019, such as Jingmen, Xianning, Yueyang, and Loudi, became weak inflow nodes in 2020, but quickly returned to the state of 2019 in 2021. In addition, there are some from weak to strong outflow cities such as Xiaogan, Qianjiang, and Yiyang, as well as inflow cities such as Wuhan and Pingxiang.

To sum up, the population flow of most urban nodes presents asymmetric characteristics during the SFTR in the study area. The core cities gradually show the characteristics of inflow from weak to strong, while its neighboring cities show the phenomenon of a strong outflow. This indicates that during the period of SFTR, the provincial capital cities still play a powerful central collection role. They are not only the travel targets of the surrounding cities, but also the destinations of the marginal cities.

*4.3. Node Criticality.* To estimate the importance of city nodes, on the premise of considering the intensity of population flow, the PageRank algorithm was applied to

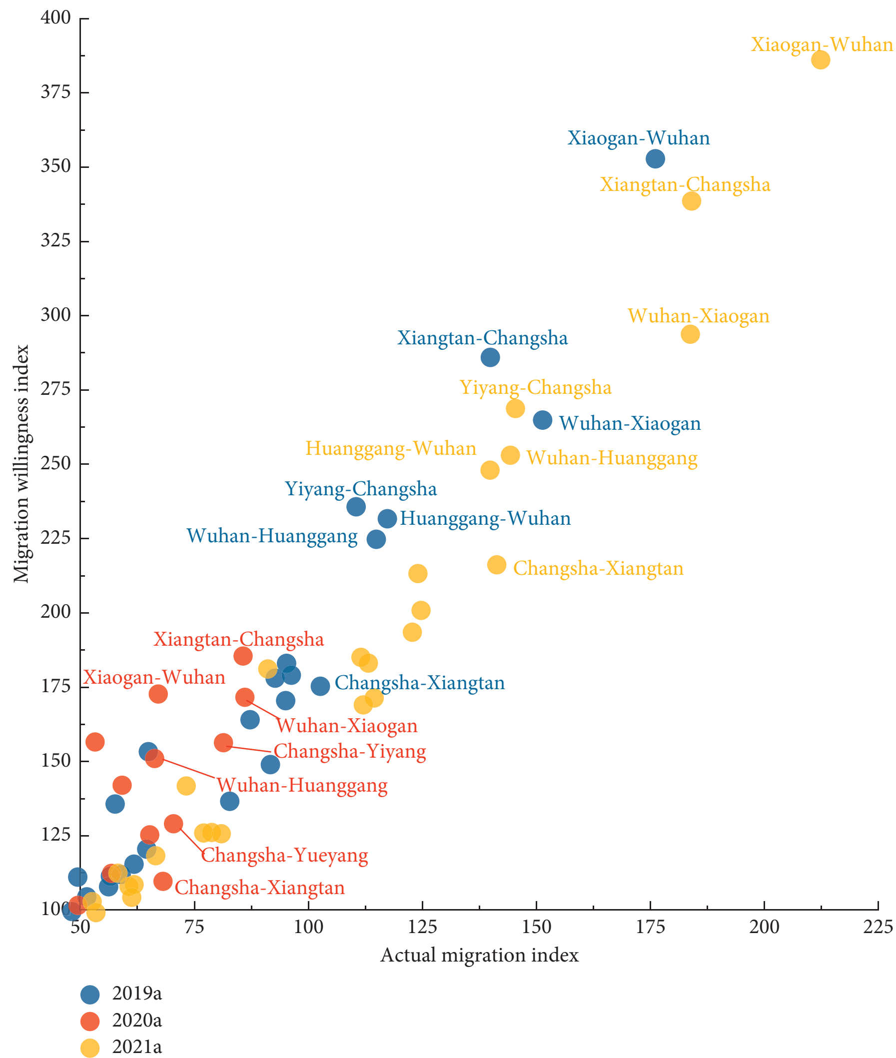



Figure 4: Popular migration routes in 2019, 2020, and 2021. The blue, red, and yellow represent the routes in 2019, 2020, and 2021, respectively.

calculate the PR values of cities in the population flow networks, with the damping coefficient set to 0.85. The higher the PR of a city, the higher status in the network and the more critical the city node is, accordingly. The PR values of city nodes during the SFTR in the study area were divided into five different levels via Natural Breaks, which can be shown in Figure 6. The nodes with higher PR values are mainly concentrated in the provincial capital cities. Although Nanchang jumped to the first level in 2021 because of its higher PR value (0.63), it still lags far behind Wuhan (0.11) and Changsha (0.10), which reveals that Nanchang, the core leading city of the Poyang Lake urban group, is emerging and is expected to form a tripartite confrontation with Wuhan and Changsha in the urban agglomeration. Jingzhou, as a city with a strong inflow of population in 2019 and 2020, has its status improved significantly in 2020. Six cities in its surroundings also rose in statuses, such as the three cities Changde, Yiyang, and Yueyang, from the third to the second level, while Xiangyang, Jingmen, and Yichang rose from the fourth to the third level. In the normal years of 2019 and 2021, the cities with a second-level and third-level PR value are mainly distributed around and between capital cities, which are bridges for the communication of population elements between capital cities. The cities with a low PR value are some marginal cities, such as Fuzhou, Ji'an, Hengyang, Loudi, Changde, Yichang, etc. For the cities with a lower PR value, such as Pingxiang, Xinyu, Yingtan, Tianmen, Qianjiang, etc., due to the small number of permanent residents, they have no advantages in the population flow network, thereby having the lowest status of cities.

### 4.4. Resilience Analysis of Urban Network Structure

*4.4.1. Network Response to Node Failure.* As a kind of network disruption, the failure of an urban node may be the loss of city caused by disaster; for example, Wuhan was quarantined from January 23 to April 8, 2020 because of the person-to-person transmission characteristic of COVID-19. To reflect the impact of disasters that different urban nodes

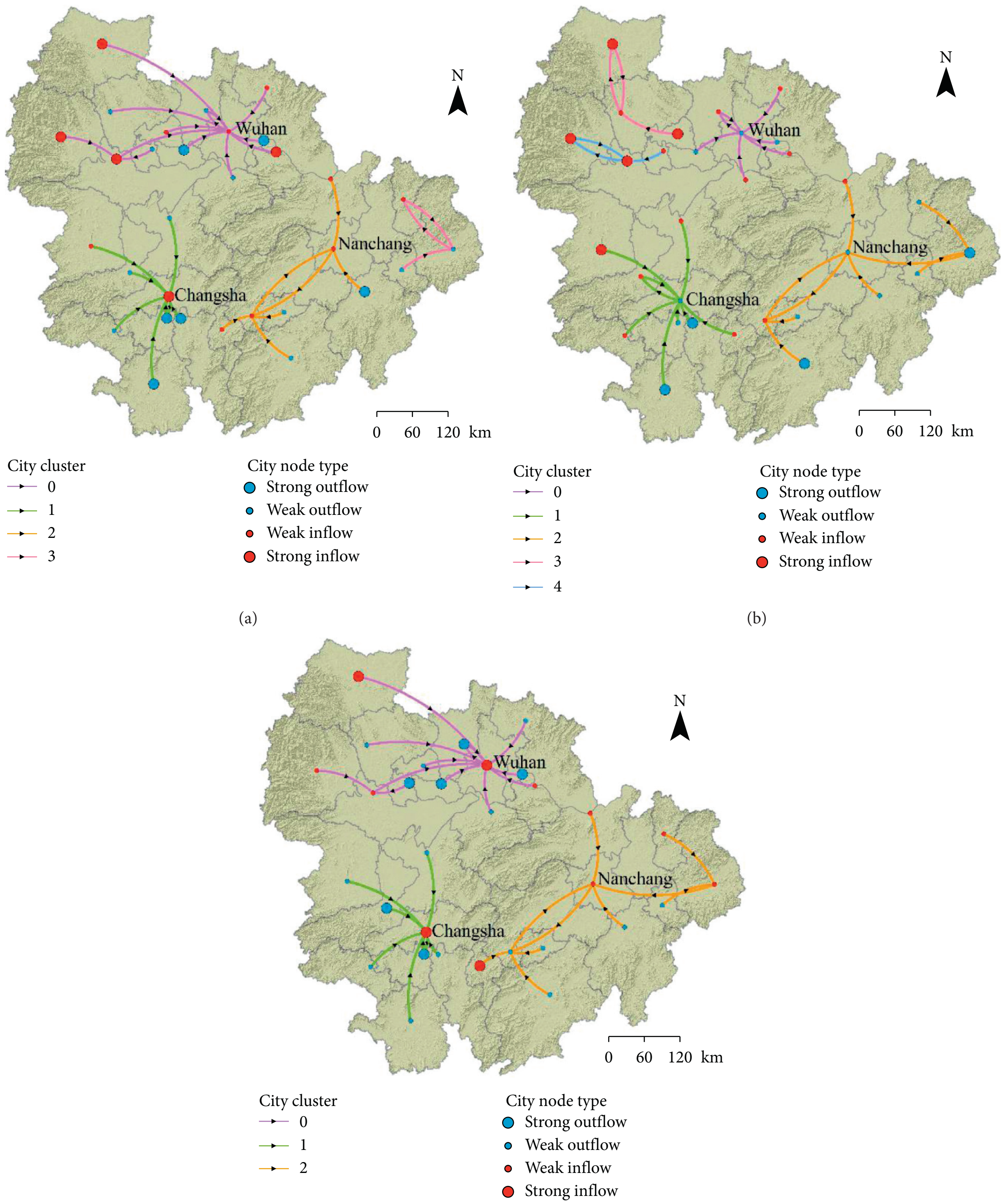
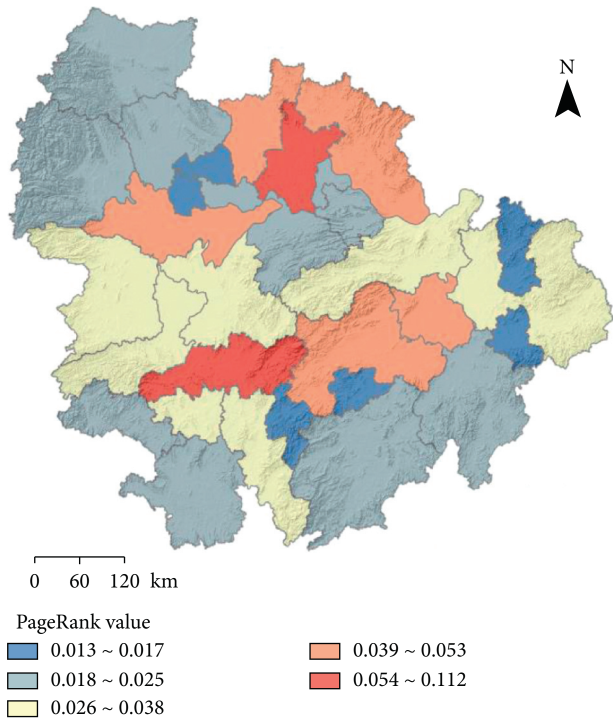
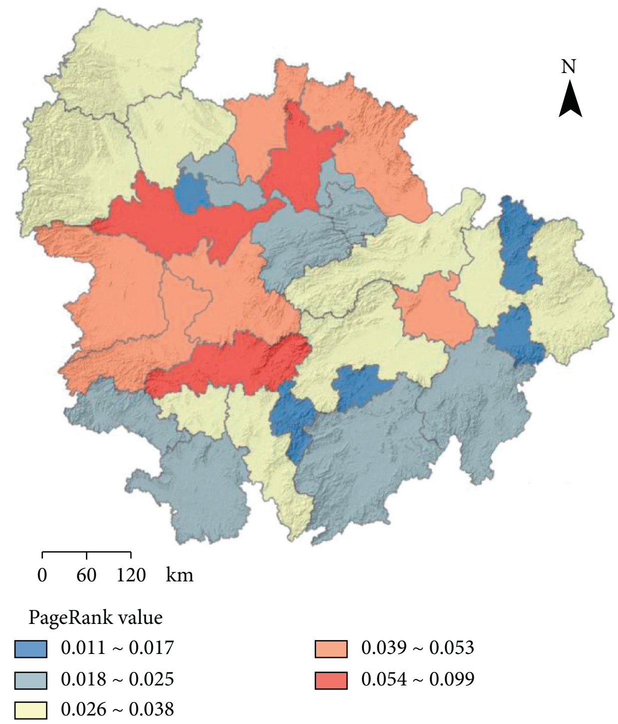



Figure 5: Dominant flow and node type distributions: (a) 2019; (b) 2020, and (c) 2021. The colors of lines represent the different city clusters. The blue and red nodes express population outflow and inflow, respectively. The size of nodes shows the degree of node symmetry: large (strong) and small (weak).

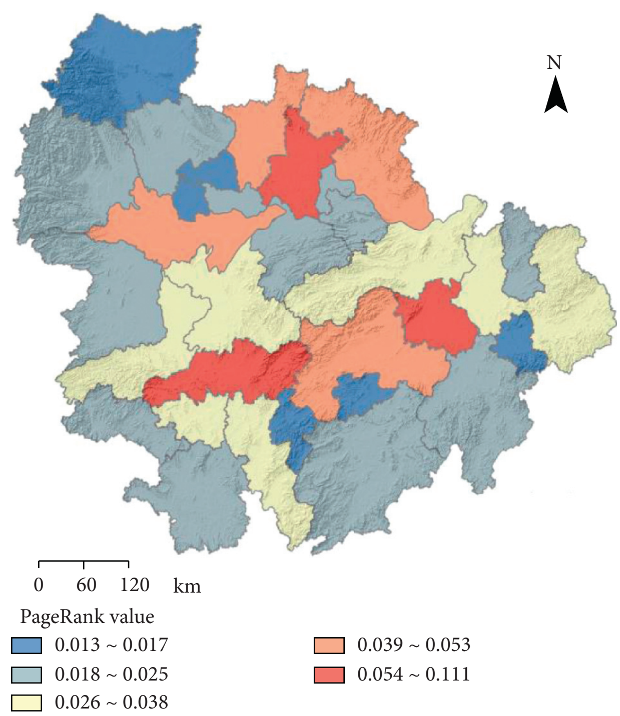
interrupted on regional networks, we conducted, respectively, simulation failures on 31 urban nodes in the networks. Network response is expressed by the network efficiency, the higher the value, the better the transmission performance of the network. The changes in network efficiency are shown in Figure 7. Obviously, there are significant differences in the




(a)




(b)




(c)

FIGURE 6: Spatial distribution of the PageRank value: (a) 2019; (b) 2020, and (c) 2021. The color from blue to red represents the different levels from low to high of PageRank.

structural resilience of the population flow networks during the SFTR in the study area. From the overall network efficiency after urban node failures, the result of 2021 (11.05–14.50) is slightly higher than that in 2019 (10.27–13.00), far higher than that in 2020 (6.28–8.24), revealing that after a severe epidemic, the propagation efficiency of the population flow networks in 2021 has been improved, and has exceeded that in 2019. However, in 2020,

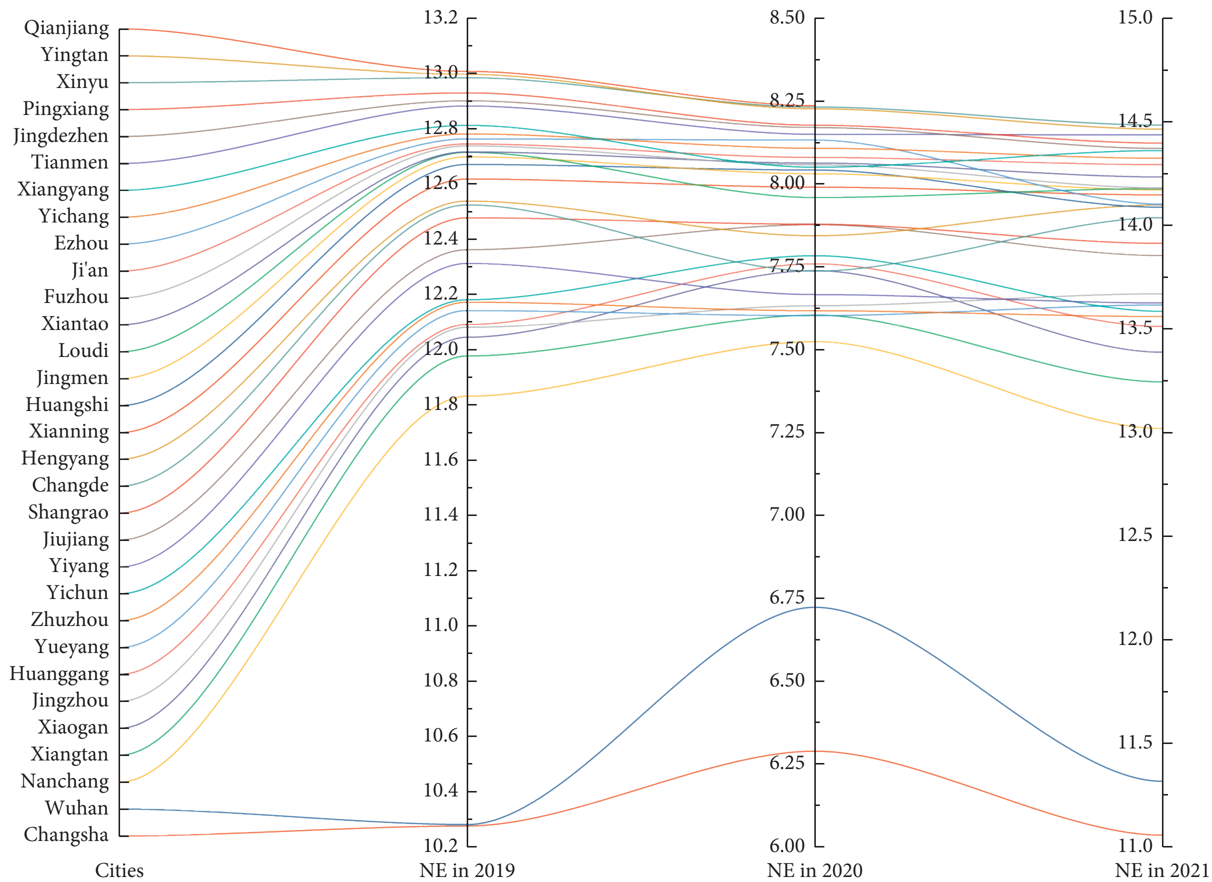



Figure 7: The changes in the network efficiency (NE) after node failure in different cities in 2019, 2020, and 2021.

the transmission efficiency of the population flow network is at a low point, which reflects the fragility of the network structure.

It is worth noting that after the failures in Changsha or Wuhan in 2019, the response of population transportation networks is almost the same, illustrating that the two primary core hub cities in the study area have relatively equal contributions to the Spring Festival travel rush. Obviously, because of the disturbance of Wuhan and travel restrictions in 2020, the overall network efficiency has reduced. Changsha has become the leading city of the urban agglomeration and has taken on the responsibility of population transportation. In addition, the response of the network to Xiaogan or Huanggang failures is insensitive because of urgent travel constraints caused by the rising number of patients with infection, while Zhuzhou, Yueyang, Yiyang, Changde, Hengyang, and Loudi within the Changsha–Zhuzhou–Xiangtan city group have a significant increase in the impact on the resilience of the network structure, because they are far away from the areas seriously affected by COVID-19. Nanchang, however, as the third core city of the study area, as well as the third city that affects the resilience of network structure in each period, still has not been prominent in the overall network. In other words, during the SFTR in 2020, the Wuhan metropolitan area has a serious disturbance on the population flow, which made the population flow barycenter in the urban agglomeration more inclined to the Changsha–Zhuzhou–Xiangtan city group, thereby increasing the pressure of epidemic prevention and control in the Hunan province. However, because of its lower level in the population flow network, the Poyang Lake city group does not play an obvious role in population diversion, thus the pressure of epidemic prevention and control is relatively small.

The first five cities in 2021 that affect the structural resilience of the population flow network are consistent with those in 2019, namely, Changsha, Wuhan, Nanchang, Xiangtan, and Xiaogan. Moreover, the impact of failures of the four cities, Zhuzhou, Yichun, Huangshi, and Ezhou, on the resilience of the network structure has been significantly enhanced due to their strong interaction with the core cities in the subregion. On the contrary, there are some cities whose population collection and distribution capacity is significantly reduced, such as Jingzhou, Xianning, and Xiantao, which might be subjected to homogeneous radiation from the surrounding cities. In addition, there are always some marginal cities, such as Qianjiang, Yingtan, Xinyu, Pingxiang, Jingdezhen, Tianmen, etc. While their risk of failure may not cause irreparable damage to the urban agglomeration, alienated relationships with other cities in the region may lead to their inability to obtain rapid resources. Therefore, attention should be paid to the improvement of the marginal nodes, which will not only help to enhance regional resilience but will also be of great benefit to the ability of the city itself to resist risks and disasters.

*4.4.2. Network Response after Node Attack.* As another kind of network disruption, node attack usually attacks the node with the maximum load, such as maximum degree, weighted degree, or centrality. For regional urban systems, malicious attacks may be terrorist attacks, military conflict, or other human forces, which are unpredictable and uncontrollable. Therefore, an attack against a certain city may lead to a wide range of network topology failures, thereby leading to the loss of urban network functions. According to the descending order of node criticality, the attacks on the city nodes were simulated. It is assumed that all links connected to a node will be removed after it was attacked, and the routes among other nodes will be redistributed. Figure 8 is the urban network response to node attacks during the three periods of SFTR. The $x$-axis is the percentage of nodes attacked, and the $y$-axis is the corresponding network efficiency. We can see that as the percentage of failed nodes increases, the resilience of the network structure decreases continuously. Obviously, although there is the highest population flow intensity of urban agglomeration in the middle reaches of the Yangtze River during the SFTR in 2021, its descent rate of the network efficiency is the fastest. Especially when attacking Nanchang and Jingzhou, the cities ranked third and seventh, and the network performance will be lower than that of 2019. Secondly, when the percentage of urban nodes that failed reaches 22%, the response of the population flow network in 2019 is almost the same as that in 2021, which is about 5.80. After that, with the interruption of urban nodes, the network efficiency in 2021 still declined at the fastest speed, and the resilience of urban network structure is always lower than that that in 2019. When the attack rate of urban nodes is slightly higher than 93%, the network efficiency almost drops to 0, which is similar to that in 2020. However, when the urban network efficiency drops to 0 in 2019, the attack rate of nodes exceeds 96%. This indicates that the urban population flow network of the study area during the SFTR in 2019 is more resilient. Although the vitality of population flow between cities is relatively low in 2020, its overall network response is relatively stable. Conversely, the urban population flow network in 2021 exposed regional vulnerability and resource disequilibrium under the background of large-scale population mobility.

Generally speaking, in the population flow network during the SFTR, the enhancement of population mobility intensity promotes the communication and cooperation between cities, as well as accumulates the potential pressure and uncertainty. Furthermore, Jingzhou, Huanggang, Yichun, Xiaogan, etc., which are closer to the center of the regional network, have undertaken the reconfiguration of most urban elements after the core cities were attacked. They perform the potential functions of population collection and distribution and maintaining the stability of the network

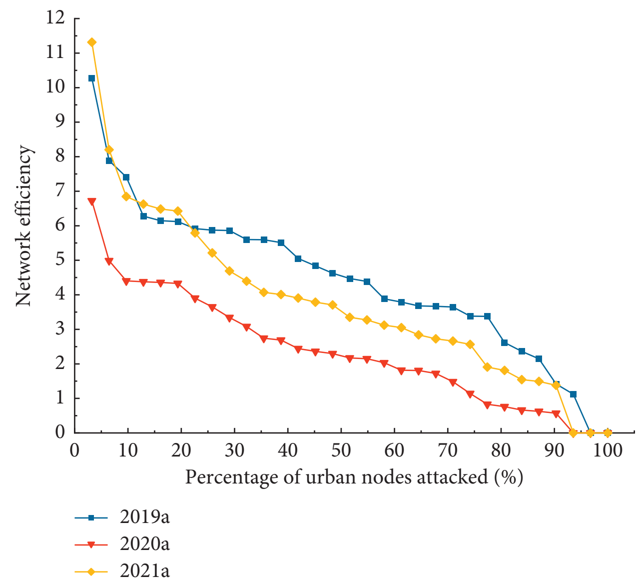



Figure 8: The changes in the network efficiency after attacking urban nodes in 2019, 2020, and 2021.

structure. Therefore, while emphasizing on the development and construction of core cities, we should pay more attention to the construction and support of such in-between cities in the transportation hubs, resource allocation, and information sharing to improve the resilience of regional urban systems.

## 5. Discussion

Increasingly, interweaving physical infrastructure and virtual networks present a more abundant connotation. Research on regional resilience and dynamic process of urban systems has been paid more and more attention by the scholars of geography and other fields. However, due to limitations in data acquisition, most of the related studies just stay in static expresses of characteristics. To some extent, the characteristics of the dynamic and complex population mobility networks have been neglected. What's worse, in the process of economic and technological globalization, the long-term chronic stress and short-term effects faced by regions and cities have become more prominent. Cities as multiagent systems play an indispensable role in functional integration and organizational management in regions [35]. Population floating, as the main carriers of regional intercity flow elements, the quantification and analysis of the dynamic properties of which is a powerful basis for understanding regional development. Therefore, it is urgent to measure and analyze the regional urban networks from the perspective of spatiotemporal evolution and resilience.

In this study, from the characteristics of population flow of the urban agglomeration in the middle reaches of the Yangtze River, it is revealed that the travel trends of residents during SFTR in 2019, 2020, and 2021 are consistent, that is, it increased with the approaching and ending of the Spring Festival Holiday and showed a lower trend on the day of the Spring Festival. The result is similar to the research-related results in other regions. For example, Lai and Pan [30] analyzed the characteristics and spatial pattern of population flow among cities in China in 2018 based on Tencent migration data and found that the population travel during SFTR showed certain regularity. In addition, from the spatiotemporal patterns and node criticality of the population flow networks, even if Wuhan was quarantined on January 23, 2020, the population flow barycenter was only more inclined to Changsha and the corridor cities between them. In other words, Nanchang, the third largest city in the urban agglomeration, still had not played a significant role. Among the popular migration routes of the urban agglomeration, Nanchang did not take a place. With the rise of its urban functional status, the situation of tripartite confrontation in the region has slightly appeared. On the one hand, the urban agglomeration in the middle reaches of the Yangtze River is the largest cross-region urban agglomeration in China. The distance between the core cities is more than 250 km, which is much higher than 100–200 km between the core cities of the top three urban agglomerations (Beijing–Tianjin–Hebei, Yangtze River Delta, and Pearl River Delta). Among the 31 node cities in the study area, Xiangyang and Shangrao are the cities with the farthest distance, which is about 700 km. On the other hand, the economic capacities of Wuhan, Changsha, and Nanchang has a certain gap and are far lower than the core cities of large urban agglomerations, such as Beijing, Shanghai, Guangzhou, Shenzhen, and so forth. Therefore, the integrative development of the urban agglomeration poses a severe test to both the economic power and radiation capacity of core cities. This is also the reason for the great differences in the development of the provinces within the urban agglomeration. What's more, the results of regional resilience via interruption simulations reveal that large-scale population floating may aggravate the instability of the regional urban network. Especially for the core cities and their corridor cities, a reasonable planning and development of their urban transportation infrastructure systems is the prerequisite to ensure regional stability.

Although an important planning area in the national development strategy, the urban agglomeration in the middle reaches of the Yangtze River, for which the government has given great support in development policies and funds, its inherent urban contact still restricts the integrated and coordinated development of the region. As we have shown in the results of this study, it is difficult for the population migration within the urban agglomeration based on Autonavi migration data to cross the provincial boundaries and connect the core cities. Therefore, in order to promote and guarantee the highly efficient flow of various elements within the urban agglomeration, especially the flow of materials and resources, we try to put forward strategies and suggestions from the perspective of urban connections for the reference of government managers and decision-makers. First, the government should focus on optimizing the spatial patterns among the core cities of the urban agglomeration. A stable "triangle" top-level network spatial structure with Wuhan, Changsha, and Nanchang as the vertices could be formed in the region, which can ensure the cross-regional cooperative transportation capacity of material resources to a certain extent. In recent years, regional uncertainty risks (such as floods, mud-rock flows, infectious diseases, etc.) have occurred increasingly. Once a core city is in a serious disaster area, its abilities in personnel coordination and resource dispersion may be slightly insufficient. Then, as the nearest neighbor and the most powerful helper, the core cities in the adjacent regions play a leading role in resource allocation and time compression. Therefore, the strong connection between core cities can not only soften the inherent relationships within the region but also enhance regional resilience. Second, the radiation driving role of core cities should be further stimulated. Wuhan, Changsha, and Nanchang, as regions with significant concentrations of various elements at the present stage, although having the highest centrality and power within the provincial scope, their radiation driving abilities to the surrounding subtier cities are still inadequate. We should make more full use of their own advantages to spread its innovative resources such as technology, talents, and information to surrounding cities. By strengthening relationships between adjacent regions and promoting differentiation constructions, it may be possible to preferentially improve the local interaction abilities to transform "weak connections" into "strong connections", and thus drive the development of the overall urban agglomeration. Third, an increasingly elaborate regional cooperation mechanism, a flat network development mechanism with complementary functions and cross-regional cooperation, should be constructed to guide the transformation of the jurisdiction based on the hierarchical system to the network system. Province level connectivity routes across regions could be enriched by breaking administrative barriers. In addition, it should be oriented by the urban agglomeration as a whole, opening up a new prospect for high-quality synergetic development of the region. Meanwhile, urban hierarchies within the urban agglomeration should also be valued to avoid the occurrence of phenomena such as "valuing the core, ignoring the edges". Finally, timely revision and update of emergency plans and disaster mitigation measures are extremely important to safeguard regional resilience. For instance, when a large-scale or sudden population migration occurs in a city, there should be correspondingly adequate prediction and response capacities to ensure the safe and sustainable working of the region. On the one hand, core cities should minimize the possibility of node failures. It can protect urban node security by strengthening emergency system construction and risk preparedness mechanisms. On the other hand, edge cities should aim at boosting the node's resilience to the risks, facilitating the circulation of elements and elevating node centrality.

Inevitably, there are some limitations in data acquisition, such as lag in socioeconomic statistics, confidentiality strategies of related sectors, etc. We extracted the scale of regional population floating using Autonavi migration data; however, it is still difficult to obtain the accurate number of passengers. The population flow networks constructed herein greatly presented the population floating patterns on the highway, which has obvious short-distance transportation characteristics, even during the Spring Festival travel rush [25]. However, as a part of population mobility, the overall and comprehensive relationship of population mobility between cities still needs more data to characterize. In addition, the specific reasons resulting in the abnormal population flow of urban agglomeration in the middle reaches of the Yangtze River were not further analyzed because they involve many aspects and are extremely complex. What's worse, although some other scholars have used multisource big data to estimate the migration population [36], the number of floating populations in a region itself is unstable, so there is also a larger deviation. In future relative research, the evolution and resilience of population flow networks with multisource big data based on long time series could be a promising challenge, especially for combining local and global interactions. Furthermore, the factors driving the abnormal flow of population and the influence mechanism of network structure resilience needs to be further revealed and discussed.

## 6. Conclusion

In this study, the spatiotemporal characteristics and resilience of population flow networks of urban agglomeration in the middle reaches of the Yangtze River during the Spring Festival travel rush in 2019, 2020, and 2021 were analyzed and evaluated. In the first place, Autonavi migration data (a total of 120 days) were obtained from the Autonavi big data platform based on the LBS technology to construct the regional urban networks. Second, drawing on complex network theory and methods, we analyzed the characteristics of the urban network structure during the SFTR in each period at the population flow distribution levels, the popular population floating routes, the integration scale of population floating, and the criticality of urban nodes. What's more, the resilience of the urban network structure was estimated based on node failures and attacks. In practice, this study can provide a branch for guiding rational population flow and planning of the regional urban system, thereby enhancing the further resilient development of the region. The main conclusions are as follows:

(1) During the 40-day Spring Festival travel rush, the changing trends of population flow in the three periods are consistent, showing travel peaks as the festival draws near and ends, with the highest travel peak being the last day of the Spring Festival Holiday. The population floating scale in 2020 and 2021 shows significant fluctuations before and after the Spring Festival. After the Spring Festival in 2020, the epidemic prevention and control policies restricted the travel of a large number of people; therefore, the population migration showed a downward trend. The SFTR in 2021 was a critical period for epidemic prevention and control. To curb the spread of the coronavirus, under the appeal of the China government in advocating off-peak travel and stay put during the Spring Festival, the population migration before the Spring Festival was relatively conservative, and its migration willingness was only comparable to the actual migration scale in 2019, and also lower than that in the same period of other years. However, as prevention and control work in an orderly way, the population migration after the Spring Festival immediately rebounded to the level of 2019, and surpassed that in 2019 after the Spring Festival Holiday.

(2) The intercity population flow network based on the Autonavi actual migration scale of the three periods during the SFTR has obvious differences in connection intensity and hierarchical distribution. It is reflected in the overall attenuation after the impact of the epidemic and the continuous improvement under effective prevention and control. Moreover, there is an obvious clustering phenomenon within each province of urban agglomeration, that is, under abnormal disturbance and influence, the urban cluster structure is still highly consistent. The popular migration routes in 2019 and 2021 mainly focused on Wuhan, which was characterized by the agglomeration of surrounding cities to core cities. In 2020, Changsha became the first popular distribution node. In addition, it is characterized by the migration trend from the core provincial capital cities to the surrounding small cities.

(3) The population floating intensity of most urban nodes presents asymmetric characteristics during the SFTR from the perspective of the node symmetry index. On the whole, the siphon effect of the core cities is gradually increasing, while the neighboring cities are the primary objects. According to the criticality of nodes, Nanchang ranked third after Wuhan and Changsha, but its core role has not been highlighted, indicating that a tripartite confrontation in the urban agglomeration gradually forms. Besides, Jingzhou, Xiaogan, Huanggang, Yichun, and other cities should not be underestimated in their status in the regional urban network.

(4) The disturbances of a city node may lead to a decrease in the overall transmission efficiency of the regional network. The higher the core position of a node is, the more significant the decline in the network efficiency. In addition, although the intensity of population flow during the SFTR in 2021 is higher than that in 2019 and 2020, the instability of the network structure is higher, and the potential risks and pressures are even greater. Therefore, attention should also be paid to the construction and development of the peripheral cities in the aspects of resource allocation and infrastructure arrangement to improve the regional resilience and ability to respond to disturbances.

## Data Availability

Autonavi migration data can be obtained from the Amap traffic big data released by Autonavi Company (https://trp.autonavi.com/migrate/page.do).

## Conflicts of Interest

The authors declare that there are no conflicts of interest regarding the publication of this paper.

## Acknowledgments

This study was funded by the National Natural Science Foundation of China (Nos. 42071216 and 41661025).

## References

[1] M. Castells, "Grassrooting the space of flows," *Urban Geography*, vol. 20, no. 4, pp. 294–302, 1999.

[2] S. M. Li, "Population migration, regional economic growth and income determination: a comparative study of dongguan and meizhou, China," *Urban Studies*, vol. 34, no. 7, pp. 999–1026, 1997.

[3] J. Pan and J. Lai, "Spatial pattern of population mobility among cities in China: case study of the National Day plus Mid-Autumn Festival based on Tencent migration data," *Cities*, vol. 94, pp. 55–69, 2019.

[4] S. Chen, J. Yang, W. Yang, C. Wang, and T. Bärnighausen, "COVID-19 control in China during mass population movements at New Year," *The Lancet*, vol. 395, no. 10226, pp. 764–766, 2020.

[5] H. De Haas, "Migration and development: a theoretical perspective," *International Migration Review*, vol. 44, no. 1, pp. 227–264, 2010.

[6] J. Liu, J. Y. Hao, Y. Y. Sun, and Z. W. Shi, "Network analysis of population flow among major cities and its influence on COVID-19 transmission in China," *Cities*, vol. 112, Article ID 103138, 2021.

[7] H. Bherwani, S. Anjum, S. Kumar et al., "Understanding COVID-19 transmission through Bayesian probabilistic modeling and GIS-based Voronoi approach: a policy perspective," *Environment, Development and Sustainability*, vol. 23, no. 4, pp. 5846–5864, 2021.

[8] Z. Sun, H. Zhang, Y. Yang, H. Wan, and Y. Wang, "Impacts of geographic factors and population density on the COVID-19 spreading under the lockdown policies of China," *The Science of the Total Environment*, vol. 746, Article ID 141347, 2020.

[9] C. Zhao, N. Li, and D. P. Fang, "A conceptual framework for modeling critical infrastructure interdependency: using a multilayer directed network model and targeted attack-based resilience analysis," *Computing in Civil Engineering*, pp. 347–354, 2015.

[10] Y. J. Shi, G. F. Zhai, L. H. Xu et al., "Assessment methods of urban system resilience: from the perspective of complex adaptive system theory," *Cities*, vol. 112, Article ID 103141, 2021.

[11] R. Boschma, "Towards an evolutionary perspective on regional resilience," *Regional Studies*, vol. 49, no. 5, pp. 733–751, 2015.

[12] C. S. Holling, "Resilience and stability of ecological systems," *Annual Review of Ecology and Systematics*, vol. 4, no. 1, pp. 1–23, 1973.

[13] M. Bruneau, S. E. Chang, R. T. Eguchi et al., "A framework to quantitatively assess and enhance the seismic resilience of communities," *Earthquake Spectra*, vol. 19, no. 4, pp. 733–752, 2003.

[14] Y. Wang, J. Zhan, X. Xu, L. Li, P. Chen, and M. Hansen, "Measuring the resilience of an airport network," *Chinese Journal of Aeronautics*, vol. 32, no. 12, pp. 2694–2705, 2019.

[15] M. Janić, "Reprint of "Modelling the resilience, friability and costs of an air transport network affected by a large-scale disruptive event"," *Transportation Research Part A Policy and Practice*, vol. 81, pp. 77–92, 2015.

[16] K. Anand, P. Gai, S. Kapadia, S. Brennan, and M. Willison, "A network model of financial system resilience," *Journal of Economic Behavior and Organization*, vol. 85, pp. 219–235, 2013.

[17] Y. T. Ding, M. Zhang, S. Chen, and R. Nie, "Assessing the resilience of China's natural gas importation under network disruptions," *Energy*, vol. 211, Article ID 118459, 2020.

[18] M. Li, H. Wang, and H. Wang, "Resilience assessment and optimization for urban rail transit networks: a Case Study of Beijing Subway Network," *IEEE Access*, vol. 7, pp. 71221–71234, 2019.

[19] N. Mou, S. Sun, T. Yang et al., "Assessment of the resilience of a complex network for crude oil transportation on the Maritime Silk Road," *IEEE Access*, vol. 8, pp. 181311–181325, 2020.

[20] X. H. Feng, C. L. Xiu, L. M. Bai, Y. X. Zhong, and Y. Wei, "Comprehensive evaluation of urban resilience based on the perspective of landscape pattern: a case study of Shenyang city," *Cities*, vol. 104, Article ID 102722, 2020.

[21] X. Liu, Z. Neal, and B. Derudder, "Featured graphic. City networks in the United States: a comparison of four models," *Environment and Planning A*, vol. 44, no. 2, pp. 255-256, 2012.

[22] A. Li, N. Mou, L. Zhang, T. Yang, W. Liu, and F. Liu, "Tourism flow between major cities during China's national day holiday: a social network analysis using weibo check-in data," *IEEE Access*, vol. 8, pp. 225675–225691, 2020.

[23] S. Wei and L. Wang, "Examining the population flow network in China and its implications for epidemic control based on Baidu migration data," *Palgrave Communications*, vol. 7, Article ID 145, 2020.

[24] X. Wei and X. A. Yao, "Constructing and analyzing spatial-social networks from location-based social media data," *Cartography and Geographic Information Science*, vol. 48, no. 3, pp. 258–274, 2021.

[25] R. Zhang, J. H. Pan, and J. B. Lai, "Network structure of intercity trips by Chinese residents under different travel modes: a case study of the Spring Festival travel rush," *Complexity*, vol. 2021, Article ID 1283012, 2021.

[26] Q. S. Gu, H. P. Zhang, M. Chen, and Y. Xie, "Hierarchical network structures and regional differentiations of tourist source destinations of Nanjing based on cellular signaling data," *Scientia Geographica Sinica*, vol. 39, no. 11, pp. 1739–1748, 2019, (in Chinese).

[27] N. Limtanakool, M. Dijst, and T. Schwanen, "A theoretical framework and methodology for characterising national urban systems on the basis of flows of people: empirical evidence for France and Germany," *Urban Studies*, vol. 44, no. 11, pp. 2123–2145, 2007.

[28] B. J. L. Berry, "Cities as systems within systems of cities," *Papers of the Regional Science Association*, vol. 13, no. 1, pp. 146–163, 1964.

[29] H. Ishii, R. Tempo, and E.-W. Bai, "A web aggregation approach for distributed randomized PageRank algorithms," *IEEE Transactions on Automatic Control*, vol. 57, no. 11, pp. 2703–2717, 2012.

[30] J. B. Lai and J. H. Pan, "China's city network structural characteristics based on population flow during Spring Festival travel rush: empirical analysis of "Tencent Migration" big data," *Journal of Urban Planning and Development*, vol. 146, no. 4, Article ID 04020018, 2020.

[31] S. S. Chopra and V. Khanna, "Understanding resilience in industrial symbiosis networks: insights from network analysis," *Journal of Environmental Management*, vol. 141, pp. 86–94, 2014.

[32] X. Li and R. Xiao, "Analyzing network topological characteristics of eco-industrial parks from the perspective of resilience: a case study," *Ecological Indicators*, vol. 74, pp. 403–413, 2017.

[33] S. Latora and A. Marchiori, "Attack vulnerability of scale-free networks due to cascading breakdown," *Physical Review*, vol. 70, no. 3, pp. 5101–5114, 2001.

[34] C. J. Fan, T. M. Cai, Z. Y. Gai, and Y. R. Wu, "The relationship between the migrant population's migration network and the risk of COVID-19 transmission in China—empirical analysis and prediction in prefecture-level cities," *International Journal of Environmental Research and Public Health*, vol. 17, no. 8, Article ID 2630, 2020.

[35] Y. Wei, J. E. Wang, W. Song, C. L. Xiu, L. Ma, and T. Pei, "Spread of COVID-19 in China: analysis from a city-based epidemic and mobility model," *Cities*, vol. 110, Article ID 103010, 2020.

[36] Z. Liu, J. L. Qian, Y. Y. Du, J. W. Yi, and Y. Sun, "Multi-level spatial distribution estimation model of the inter-regional migrant population using multi-source spatio-temporal big data: a case study of migrants from Wuhan during the spread of COVID-19," *Journal of Geo-information Science*, vol. 22, no. 2, pp. 147–160, 2020, (in Chinese).

Hindawi
Complexity
Volume 2021, Article ID 1935557, 16 pages
https://doi.org/10.1155/2021/1935557

*Research Article*

# An Assessment Approach to Urban Economic Resilience of the Rust Belt in China

**Shan Man,[1,2] Xiangli Wu ,[2] Yongchun Yang,[1,3] and Qingmin Meng[4]**

[1]*College of Earth and Environmental Sciences, Lanzhou University, Lanzhou 730000, China*
[2]*College of Geographical Science, Harbin Normal University, Harbin 150025, China*
[3]*Key Laboratory of Western China's Environmental Systems, Ministry of Education of the People's Republic of China, Lanzhou University, Lanzhou 730000, China*
[4]*Department of Geosciences, Mississippi State University, Mississippi State, Starkville, MS 39762, USA*

Correspondence should be addressed to Xiangli Wu; jndxwxl@163.com

Received 13 May 2021; Revised 28 July 2021; Accepted 15 August 2021; Published 10 September 2021

Academic Editor: Xiuliang Yuan



Urban economic resilience provides a novel perspective on the sustainable development of urban and regional economy. Selecting 37 prefecture-level cities in the Northeast China that also known as the rust belt in China as a regional study sample that has experienced significant economic decline and out-migration in the last 20 years compared to many other regions in China, this study aims to construct an index system using the data collected in 2005, 2010, and 2016. This study evaluates urban economic resilience including five socioeconomic aspects: diversity, capabilities related to revenue and expenditure, innovation environment, trend of development, and openness. We analyze the spatial and temporal evolution characteristics of urban economic resilience, explore the key factors contributing to urban resilience and then provide decision-making suggestions to enhance it. We find the following: (1) urban economic resilience in the Northeast China has gradually increased over time, but spatial heterogeneity of resilience was prominent. Specifically, coastal cities were more economically resilient than inland cities. (2) Urban economic resilience in the Northeast China is significantly contributed by the diversity of an economic system and the trend of development, which contribute to resilience with weights of 0.214 and 0.216, respectively. The dominant factors contributing to urban economic resilience are different among diverse urban economic types and size. (3) To enhance urban economic resilience, comprehensive economic cities need to focus on increasing the diversity of economic structures. Resource-based and old industrial cities should focus on raising the innovation environment. Coastal cities should focus on increasing the diversity of their economic structures and creating positive trends of economic and social development. Agricultural cities should focus on creating positive trends of economic and social development.

## 1. Introduction

Cities across the world are facing chronic pressure and acute shock nowadays from many perspectives, such as the ecological, social, economic, and environmental challenges. Scientists and urban managers have carried out relevant research and emergency work, and the construction of resilient cities is one of the important benchmarks [1]. Typically, there are four main backgrounds for the rise of resilient cities: first, at the ecological level, climate, geological disasters, and ecological deterioration have led to increasing urban emergencies. Extreme weather such as global warming, frequent ice disasters, and heavy rain seriously threatens the development of cities. Second, at the economic level, since the 2008 Western financial crisis, the global economy has been sluggish, urban unemployment has increased significantly, and factory closures and large-scale strikes have become frequent, which all call for the urban economy to shift from relying on only service industries to a diversified economy. Third, at the social level, the uncertainty of urban development has become a common problem in all countries in the world. The uncertainty faced by most cities is exacerbated by the instability of the natural, societal, economic, and political environment. Fourth, at the planning level, a goal of urban planning is to deal with uncertainty and regulate the healthy development of the city [2].

With the acceleration of globalization, the economic systems of different countries are often subject to various shocks, such as economic cycle fluctuations, policy adjustments, financial crises, and technological innovations. The development of urban economic system becomes more difficult to predict due to these social, economic, and environmental fluctuations. Facing the increasing regional and global markets' threats and opportunities, the resilience assessments that cities should take to maintain the stable and healthy development of economic systems become a grand challenge, which is more challenging for declining urban areas such as the Northeast China, which is the typical rust belt in China [3]. Urban economic resilience has gradually become a hot issue for scholars nationally and internationally.

Northeast China is the cradle of China's industry, including the three provinces of Liaoning, Jilin, and Heilongjiang and the eastern part of Inner Mongolia (Figure 1). Before the reform and opening up, the state-owned economy in Northeast China occupied a dominant position, played an important role in Chinese economic development progress, and was the foundation of China's industrial economic development. Northeast China has provided strong support for the rapid development of its economy. After the Reform and Opening Up policy in 1978, resource-exhausted cities appeared owing to the excessive exploitation and use of resources in resource-based cities, and the developmental momentum of cities in Northeast China gradually weakened. In general, cities in the Northeast China have low economic resilience and are easily affected by the economic crisis and other shocks. Moreover, as China's traditional old industrial basis, the economic recession and revival of Northeast China is of great significance to the development of the country and region. Urban economic resilience provides a new perspective for analyzing the economic revitalization of old industrial bases. Studying the characteristics and main contributing factors of urban economic resilience of the old industrial bases in Northeast China is of great significance for exploring the path of improving urban economic resilience of old industrial bases and realizing a fresh round of economic revitalization. Therefore, it is necessary to study the urban economic resilience of Northeast China [4].

However, since the implementation of the market economy system, a series of problems have emerged in the economic development of Northeast China. Although the implementation of China's Northeast revitalization strategy has improved the economic strength and social development level of cities in Northeast China, the economic development of Northeast China still encountered such problems as a lack of development power, weak urban agglomeration, and difficulties in transforming resource-based cities. The economic growth rate in Northeast China is gradually lagging behind the national average. In general, urban economic resilience in Northeast China still has the characteristics of fragility [5]. Therefore, under the background of the prominent urban economic problems in Northeast China, this study established the index system of urban economic resilience, which is conducive to improving the research on the index system of urban economic resilience in China. Moreover, this study clarified the main contributing factors of urban economic resilience, which can provide a basis for the coordinated development of cities in Northeast China.

At the same time, the index system established in this study can also be used as a reference for scholars in other developing countries to study urban economic resilience and help them find the factors that contribute the most to urban economic resilience. In addition, the policy conclusions of this study can provide reference for the economic development of the old industrial areas in other countries.

## 2. Literature Review

The concept of "resilience" was first applied to ecology [6] and subsequently to geography [7], economics [8], and sociology [9, 10]. Research on urban resilience began once the concept of resilience was extended from social-ecological systems to urban systems. The concept also has been widely used in the agricultural and biological environmental sciences [11–13], psychology [14], and energy engineering [15, 16]. In addition to the research on the concept of urban resilience, scholars have also carried out abundant research on its characteristic elements [17], research content framework [18], and evaluation system [19]. The research scope has been extended to different levels, such as individuals and communities [20]. Among them, urban economic resilience is an important perspective of urban resilience research, and it is the deepening of urban resilience research in the economic aspect. Based on prevalent research [21–25], it is defined as the ability of the urban economic system to maintain and promote its long-term economic development, responding to cyclical economic crises, increased competitors, unexpected company closures, and technological changes. Generally, four manifestations of urban economic resilience under external shocks and disturbances have been identified [26]: (1) a restoration to the stable state before the disturbance; (2) the maintaining of some functions and structures, and moving to a substable state; (3) the failure to maintain the stability of structures and functions, and gradually declining; and (4) the readjusting of its structure and function according to the situation after the disturbance, and then moving to another developmental state. Currently, research on the evaluation of urban economic resilience includes means of measuring resilience [27–29] and discovering factors influencing it [30, 31].

Research on resilience in China started later than that in the West. With the accelerating process of urbanization, the number of uncertain factors and unknown risks facing the country's complex urban system is increasing, and intensive research on urban resilience is thus underway. Different fields have different emphasis on resilience research, but the consensus among them is to use it to study the response of an object to external changes or interference [32]. The global financial crisis of 2008 led to "economic resilience" emerging


Figure 1: The location of Northeast in China.

as the focus of research on urban resilience. Scholars thus began exploring methods to measure it, especially quantitatively. Current research on urban economic resilience is mainly related to four aspects: (1) using the concept of resilience to guide urban planning and design [33, 34], (2) the exploration of concepts and theories related to urban resilience [35, 36], (3) the design of index systems for the evaluation of urban economic resilience [37–39] and its factors influencing it [40, 41], and (4) studies of the economic resilience of special types of cities [42, 43].

Although research has contributed significantly to the theory and practice of urban economic resilience, critical issues still need to be explored and revealed. First, the concept of urban economic resilience has penetrated into urban and regional development and management, but there is not a standard theoretical system available for regional urban systems. Second, although many scholars have attempted to build measurement systems for urban economic resilience, no commonly practicable index systems are available for regional assessment. Further research on evaluating urban economic resilience is thus needed. Third, the spatial and temporal scales of urban economic resilience are poor, and the spatial and temporal scales' exploration will provide deep reveals of it. At last, regional resilience assessment in rust belt regional has not been studied, which is an important social and economic issue in both developed and developing countries. To the best of our knowledge, this study is the first research to regionally and temporally assess urban economic resilience in a developing country. How to measure urban economic resilience is an important issue in current urban development research. Many studies have comprehensively measured the spatial and temporal characteristics of urban resilience by constructing a multidimensional and multistage evaluation system [27, 28, 37, 38], but few focus on the economic perspective to measure urban economic resilience. Moreover, the existing studies on urban economic resilience do not classify the types of cities.

In view of the above, this study enriches and complements the research in this field, and we constructed an evaluation index system of urban economic resilience from five socioeconomic aspects: diversity, capabilities related to revenue and expenditure, innovation environment, trend of development, and openness, and used multi-indicator comprehensive evaluation method to measure the spatiotemporal evolution characteristics of urban economic resilience in 37 cities of the rust belt of China. Then, we classified the cities in Northeast China into comprehensive economic cities, resource-based and old industrial cities, coastal cities, and agricultural cities, analyzed the factors contributing to the economic resilience of these four types of cities, and obtained the factors contributing to the economic resilience of different types of cities. Therefore, this study not only enriches the research on the index system of urban economic resilience in developing countries but also diversifies the types of cities studied. It also provides strategical decisions for sustainable development of rust belt cities in China.

## 3. Data and Methods

*3.1. Study Area and Data.* The research object of this paper was the prefecture-level cities in Northeast China (i.e., the rust belt in China). We studied 37 prefecture-level cities, excluding the Yanbian Korean Autonomous Prefecture, Da Hinggan Ling Prefecture, Hinggan League, and Xilingol League because of unavailable data. The data were drawn mainly from the *China City Statistical Yearbook*, *China City Construction Statistical Yearbook*, *Regional Economic Statistical Yearbook*, the statistical yearbooks of the relevant provinces and cities, economic census yearbooks, and statistical bulletins in 2005, 2010, and 2016. If the data of a factor are missed, we use the average value of the related records in the previous and the following years to interpret correspondingly.

*3.2. The Resilient Assessment Index System.* The evaluation index of urban economic resilience is a conceptual expression of the causes and mechanisms of urban economic resilience. Each index system or conceptual model is trying to find the root cause of resilience. It mainly involves sustainable development theory, system synergy theory, vulnerability theory, and so on. Sustainable development theory focuses on guiding the sustainable and healthy development of the urban economy. System synergy theory emphasizes the coordinated development of the various subsystems of the urban economy, thereby promoting overall urban development. High resilience is related to low vulnerability, so resilience can be interpreted in depth according to vulnerability theory.

Currently, the representative urban resilience evaluation system comes from the resilient city framework index system proposed by the Rockefeller Foundation, which is based on four types of index systems: health and well-being, economy and society, infrastructure and environment, leadership, and strategy [44]. And some scholars have proposed a framework for evaluating community resilience based on the Rockefeller Foundation's index system. This framework consists of four parts: external resources, assets, capabilities, and qualifications [45]. In addition, the Regional Research Institute of the State University of New York released an evaluation system for resilience to respond to rapid population growth, economic recession, and natural disasters in metropolitan areas, which is mainly composed of 12 indicators in three dimensions: regional economic capacity, society-population, and community connectivity. Some scholars also constructed a 4R evaluation index system based on the three aspects of society-politics-technology, economy, and nature, with the four characteristics of robustness, redundancy, resourcefulness, and rapidity [27]. Moreover, some scholars also use the interview method, questionnaire survey method, and other methods to construct evaluation models for evaluating urban resilience [28, 46].

Compared with the existing research, the evaluation index system in this paper makes the original research on urban resilience more pertinent. The evaluation index system established in this paper is based on the urban resilience evaluation system and conducts in-depth research on its economic resilience. It divides the economic resilience index system into five subsystems: diversity, capabilities related to revenue and expenditure, innovation environment, trend of development, and openness. On the one hand, it can make the evaluation results of economic resilience more accurate, thereby helping to make targeted suggestions for the optimization of urban economic resilience in Northeast China; on the other hand, it can improve the economic resilience evaluation index system and improve the scientificity of the measurement system.

By referring to the available index systems for the evaluation of urban economic resilience [24, 25, 34, 35], we adopted principles based on five aspects of an economic system including diversity, capabilities related to revenue and expenditure, innovation environment, trend of development, and openness, in combination with physical and geographical conditions, historical bases, and policy-related advantages of Northeast China to determine initial indices. We then consulted ten experts in human geography, regional economics, urban geography, and ecology to screen and initialize indices and then determine the final evaluation indices. The experts are scholars in the field of resilience and vulnerability research. We used the importance and availability of indicators as selection principles, and the retention and deletion of the index is based on 1/3 of the number of experts. When no less than 1/3 of the experts think the index needs to be adjusted, the index will be retained or deleted. Finally, the retained indicators are summarized into Table 1, which forms a hierarchical indexing assessment system including five system levels, 11 criterion levels, and 30 factor levels.

For the diversity system: the diversified development of an urban economy can enhance the ability of the urban economy to resist cyclical economic crisis. If one aspect of the economy goes wrong, other aspects can continue to support the development of the city. Therefore, the improvement of urban economic diversity can enhance urban economic resilience. The concept of economic diversity is usually equated with industrial diversity when studying urban economic diversity. Therefore, we selected the structural diversity index for primary, secondary, and tertiary industry and the structural diversity index for 19 subsectors in primary, secondary, and tertiary industry to represent the industrial diversity, and we selected the proportion of the tertiary industry as percentage of GDP to represent the ability of industrial transformation and upgrading, so as to measure the diversity of the urban economy from the aspects of industrial diversity and transformation and upgrading ability.

For the capabilities related to revenue and expenditure system: revenue and expenditure capacity is an important index to measure the living standard of residents. Urban revenue and expenditure capacity includes the revenue and expenditure capacity of individual residents and government public finance. The revenue and expenditure of public finance have the function of adjustment. It can relieve the unreasonable distribution phenomenon and realize social equity. Therefore, revenue and expenditure capacity in this

Table 1: An index system to assess urban economic resilience in the rust belt of China.

| System level | Criterion level | Specific factor level (formula) |
|---|---|---|
| A1 diversity | B1 industrial diversity | C1 structural diversity index for primary, secondary, and tertiary industry (the specific algorithm is detailed in Section 3.3.1) |
| | | C2 structural diversity index for 19 subsectors in primary, secondary, and tertiary industry (the specific algorithm is detailed in Section 3.3.1) |
| | B2 industrial transformation and upgrade | C3 tertiary industry as percentage of GDP (tertiary industry output/ gross domestic product) |
| A2 capabilities related to revenue and expenditure | B3 personal revenue- and expenditure-related ability | C4 per capita GDP (gross domestic product/population) |
| | | C5 per capita disposable income ((total household income-income tax paid-social security expenditure paid by individual-accounting subsidy)/family population) |
| | | C6 per capita household saving deposits rate (household saving deposits at year-end/gross domestic product/population) |
| | B4 governmental revenue and expenditure ability | C7 public fiscal revenue as percentage of GDP (public finance income/ gross domestic product) |
| | | C8 public fiscal expenditure as percentage of GDP (public finance expenditure/gross domestic product) |
| A3 innovation environment | B5 technological innovation environment | C9 science and technology expenditure as percentage of public fiscal expenditure (expenditure for science and technology/public finance expenditure) |
| | | C10 per 10,000 people subscribers of Internet services (number of subscribers of Internet services/population/10000) |
| | | C11 per 10,000 people patent applications (number of patent applications/population/10000) |
| | | C12 per 10,000 people scientific papers (number of scientific papers/ population/10000) |
| | B6 ecological development environment | C13 per capita public recreational green area (public recreational green area/population) |
| | | C14 green coverage of built district (green coverage area/built area) |
| | | C15 ratio of industrial solid wastes comprehensive utilized (number of utilized industrial solid wastes/total number of industrial solid wastes) |
| | B7 basic social environment | C16 per 10,000 people full-time teachers in regular institutions of higher education (number of full-time teachers in regular institutions of higher education/population/10000) |
| | | C17 per 10,000 people doctors (number of doctors/population/10000) |
| | | C18 proportion of education expenditure in fiscal expenditure (expenditure for education/public finance expenditure) |
| A4 trend of development | B8 economic development trend | C19 GDP growth rate ((gross domestic product of the next year-gross domestic product of the previous year)/gross domestic product of the previous year $*$ 100%) |
| | | C20 investment in focused assets growth rate ((investment in focused assets of the next year-investment in focused assets of the previous year)/investment in focused assets of the previous year $*$ 100%) |
| | | C21 total value of social retail goods growth rate ((total value of social retail goods of the next year-total value of social retail goods of the previous year)/total value of social retail goods of the previous year $*$ 100%) |
| | B9 social development trend | C22 natural population growth rate ((population of the next year-population of the previous year)/population of the previous year $*$ 100) |
| | | C23 employment growth rate ((employment of the next year-employment of the previous year)/employment of the previous year $*$ 100) |
| | | C24 social security and employment expenditure as percentage of fiscal expenditure (expenditure for social security and employment/public finance expenditure) |

Table 1: Continued.

| System level | Criterion level | Specific factor level (formula) |
|---|---|---|
| A5 openness | B10 foreign trade openness | C25 total value of imports and exports of goods as percentage of GDP (total value of imports and exports/gross domestic product) |
| | | C26 amount of foreign capital actually utilized as percentage of GDP (amount of foreign capital actually utilized/gross domestic product) |
| | | C27 proportion of output value from investment by foreign enterprises, and those of enterprises in Hong Kong, Macao, and Taiwan as percentage of gross industrial output value (output value from investment by foreign enterprises, and those of enterprises in Hong Kong, Macao, and Taiwan/gross industrial output value) |
| | B11 domestic trade openness | C28 output value of domestic enterprises as percentage of gross industrial output value (output value from domestic enterprises/gross industrial output value) |
| | | C29 GDP average highway freight traffic (highway freight traffic/gross domestic product) |
| | | C30 per capita highway passenger traffic (highway passenger traffic/population) |

study mainly consists of two aspects. One is personal revenue and expenditure ability, including per capita GDP, per capita disposable income, and per capita household savings deposits rate. The other is governmental revenue and expenditure ability, including public fiscal revenue as percentage of GDP and public fiscal expenditure as percentage of GDP.

For the innovation environment system, innovation environment influences the transformation of the regional economic growth model. The improvement of the innovation environment is conducive to the progress of science and technology, thus promoting the growth of the regional economy and improving the sustainable development ability of urban economy. Therefore, the improvement of the innovation environment is conducive to enhancing urban economic resilience. We selected three important indicators of innovation environment: one is technological innovation environment, including science and technology expenditure as percentage of public fiscal expenditure, per 10000 people subscribers of Internet services, per 10000 people patent applications, and per 10000 people scientific papers. Two is ecological development environment, including per capita public recreational green area, green coverage of built district, and ratio of industrial solid wastes comprehensive utilized. Three is basic social environment, including per 10000 people full-time teachers in regular institutions of higher education, per 10,000 people doctors, and the proportion of education expenditure in fiscal expenditure.

For the development trend system, the development trend system mainly reflects the law, speed and trend of urban development, and can also predict the future development direction of a city. Only by understanding the direction of future development can we better withstand the unknown risks. The development trend of this study mainly includes two aspects. One is economic development trend, including GDP growth rate, investment in focused assets growth rate, and total value of social retail goods growth rate. The other is social development trend, including natural population growth rate, employment growth rate, and social security and employment expenditure as percentage of fiscal expenditure.

For the openness system, increasing the openness of the economic system will increase the flow of resources. The more open the regional economy is, the higher the level of economic development will be, and the stronger the economic resilience will be. Under the development trend of economic globalization, enhancing the openness of the economic system has become the main development direction of the world economy. We selected two aspects of openness. One is foreign trade openness, including total value of imports and exports of goods as percentage of GDP, amount of foreign capital actually utilized as percentage of GDP and proportion of output value from investment by foreign enterprises, and those of enterprises in Hong Kong, Macao, and Taiwan as percentage of gross industrial output value. The other is domestic trade openness, including output value of domestic enterprises as percentage of gross industrial output value, GDP average highway freight traffic, and per capita highway passenger traffic.

### 3.3. Assessment Approaches

*3.3.1. Shannon Index.* We used the Shannon index to measure the value of C1 and C2 (Table 1). Suppose the total number of employees in a city is $A$; $A$ is then divided into "$n$" types of industries according to the economic activities that they engage in. The employees in various industries are $A_i(1,2,3,\ldots,n)$; then, $\sum_{i=1}^{n} A_i = A$. According to the formula for Shannon's entropy, the informational entropy of the industrial structure is defined as

$$H = -\sum_{i=1}^{n} P_i \ln P_i = -\sum_{i=1}^{n} \left(\frac{A_i}{\sum_{i=1}^{n} A}\right) \ln\left(\frac{A_i}{\sum_{i=1}^{n} A_i}\right), \qquad (1)$$

where $H$ is the value of the industry's structural diversity index and $P_i$ is the proportion of employees in various industries.

*3.3.2. Entropy Method.* We used the entropy method to give weights to the indicators of the specific index level and the system level. It is mainly a measure of uncertainty in the state of the system. The higher the entropy is, the more balanced is the system's structure, and the smaller the difference in indices. The lower the entropy is, the more unbalanced the system structure is, and the larger the difference in indices:

$$e_j = -k\sum_{i=1}^{m} y_{ij}\ln\left(y_{ij}\right), k = \frac{1}{\ln m}, 0 \le e_j \le 1 \tag{2}$$

$$\omega_j = \frac{\left(1 - e_j\right)}{\sum_{i=1}^{n}\left(1 - e_j\right)}, \tag{3}$$

where $e_j$ is the informational entropy of the index, $y_{ij}$ is the standardized value of the index, $\omega_j$ is its weight, $m$ is the number of samples, and $n$ is the number of indices in the specific index level.

*3.3.3. Multi-Index Comprehensive Evaluation Method.* We use the multi-index comprehensive evaluation method to measure urban economic resilience by weighting resilience and summing it.

First, the value of each subsystem's resilience is calculated by

$$t = \sum_{i=1}^{m} a_i \cdot \omega_i, \quad i = 1\ldots, l, \tag{4}$$

where $t$ is the value of the resilience of each subsystem, $a_i$ is the standardized value of index $i$ at the specific index level, $\omega_i$ is the weight of index $i$ at the specific index level of each system, and $l$ is the number of specific indices of each system level.

Then, we calculate the value of urban economic resilience:

$$T = \sum_{j=1}^{n} t_j \cdot W_j, \quad j = 1\ldots, z, \tag{5}$$

where $T$ is the value of urban economic resilience, $t_j$ is the value of the resilience of subsystem $j$ at the system level, $W_j$ is its weight, and $z$ is the number of indices in the system level.

## 4. Temporal and Spatial Characteristics

Applying the above methods, we obtained the urban economic resilience of the 37 prefecture-level cities in Northeast China. The results are summarized in Table 2.

*4.1. Temporal Evolution Characteristics.* On the whole, urban economic resilience in Northeast China gradually increased from 2005 to 2016 (Table 2). Its average value was 0.300 in 2005, 0.318 in 2010, and 0.340 in 2016. Growth from 2005 to 2010 was smaller than that from 2010 to 2016. This is mainly because the total economic output of cities in Northeast China has achieved continuous growth since implementing the strategy of revitalizing the old industrial base in Northeast China. However, due to the different levels of development at each stage, it shows stage characteristics. The period from 2005 to 2007 was in the early stage of implementation of the strategy, and the economic growth rate of cities in Northeast China was slow. In 2007, Northeast China's economy only accounted for 8.4% of the country's total. In addition, from 2007 to 2009, the global financial crisis struck, and some major natural disasters occurred, such as the snow disaster during the Spring Festival transport in Northeast China. Due to the impact of natural disasters and other related factors, the average growth rate of urban economic resilience in Northeast China from 2005 to 2010 was relatively small. With the assistance of national policies and self-adjustment, the economic growth of Northeast China has accelerated in 2010. By 2012, Northeast China's economy accounted for 8.8% of the country's total, reaching the highest proportion since the implementation of the strategy. Since 2013, China has paid more attention to regional economic development and improved its national development strategy based on local conditions. Therefore, the average growth rate of urban economic resilience in Northeast China from 2010 to 2016 was relatively large.

The temporal evolution of urban economic resilience (Table 2) shows that Dalian was the most economically resilient and Suihua was the least in 2005, 2010, and 2016. As a coastal city, Dalian is a main economic, industrial, and tourism hub in the northeastern and eastern China. It has prominent geographical advantages, a good living environment, high economic diversity, and a high degree of internationalization. Since the Reform and Opening Up, it has developed rapidly and has been less affected by the sluggish economic development of Northeast China. Therefore, Dalian had the highest urban economic resilience among cities in the region. Suihua, located in the hinterland of the Songnen Plain, is an important agricultural city with a large-scale base for commodity grain production. The proportion of agriculture here was much higher than the average levels of the country and those of Heilongjiang Province. It has few leading enterprises in industrial development, and the developmental system of its service industry is incomplete. This has led to low per capita GDP. Thus, Suihua was the least economically resilient of the cities in Northeast China.

Cities with declining resilience include Hulun Buir, Siping, Liaoyuan, Baishan, Jixi, Hegang, Yichun, and Mudanjiang. The largest decline in urban economic resilience was Mudanjiang, mainly due to its eight forestry bureaus. In the past decade, the transformation of forestry has been difficult, and the tourism industry has been severely restricted by seasons, which has led to a large loss of population, so the urban economic resilience has declined the most.

Cities with increasing resilience include Chifeng, Tongliao, Shenyang, Dalian, Anshan, Fushun, Benxi, Dandong, Jinzhou, Yingkou, Fuxin, Liaoyang, Tieling, Huludao, Changchun, and so on. The largest increase in urban economic resilience was Changchun, because Changchun has been relying on industries such as automobiles, agricultural

Table 2: Values of urban economic resilience in Northeast China in 2005, 2010, and 2016.

| City (abbreviation) | 2005 | 2010 | 2016 | City (abbreviation) | 2005 | 2010 | 2016 |
|---|---|---|---|---|---|---|---|
| Chifeng (CF) | 0.233 | 0.231 | 0.297 | Siping (SP) | 0.221 | 0.202 | 0.218 |
| Tongliao (TL) | 0.285 | 0.288 | 0.291 | Liaoyuan (LYu) | 0.255 | 0.229 | 0.222 |
| Hulun Buir (HB) | 0.401 | 0.328 | 0.363 | Tonghua (TH) | 0.277 | 0.253 | 0.330 |
| Shenyang (SY) | 0.508 | 0.512 | 0.637 | Baishan (BS) | 0.278 | 0.289 | 0.275 |
| Dalian (DL) | 0.564 | 0.642 | 0.644 | Songyuan (SYu) | 0.245 | 0.222 | 0.262 |
| Anshan (AS) | 0.293 | 0.334 | 0.371 | Baicheng (BC) | 0.232 | 0.213 | 0.241 |
| Fushun (FS) | 0.281 | 0.303 | 0.302 | Harbin (HBi) | 0.398 | 0.432 | 0.548 |
| Benxi (BX) | 0.273 | 0.342 | 0.351 | Qiqihar (QQH) | 0.332 | 0.315 | 0.332 |
| Dandong (DD) | 0.296 | 0.329 | 0.362 | Jixi (JX) | 0.312 | 0.326 | 0.282 |
| Jinzhou (JZ) | 0.319 | 0.345 | 0.366 | Hegang (HG) | 0.280 | 0.293 | 0.253 |
| Yingkou (YK) | 0.294 | 0.350 | 0.358 | Shuangyashan (SY) | 0.257 | 0.303 | 0.288 |
| Fuxin (FX) | 0.287 | 0.298 | 0.288 | Daqing (DQ) | 0.338 | 0.335 | 0.405 |
| Liaoyang (LY) | 0.302 | 0.339 | 0.362 | Yichun (YC) | 0.249 | 0.283 | 0.244 |
| Panjin (PJ) | 0.355 | 0.439 | 0.398 | Jiamusi (JMS) | 0.307 | 0.355 | 0.339 |
| Tieling (TLi) | 0.259 | 0.297 | 0.268 | Qitaihe (QTH) | 0.207 | 0.235 | 0.288 |
| Chaoyang (CY) | 0.226 | 0.230 | 0.276 | Mudanjiang (MDJ) | 0.396 | 0.391 | 0.345 |
| Huludao (HLD) | 0.239 | 0.295 | 0.332 | Heihe (HH) | 0.246 | 0.330 | 0.272 |
| Changchun (CC) | 0.359 | 0.336 | 0.519 | Suihua (SH) | 0.177 | 0.195 | 0.200 |
| Jilin (JL) | 0.321 | 0.334 | 0.455 | Average value | 0.300 | 0.318 | 0.340 |

products processing, and equipment manufacturing for stable and rapid development. In recent years, it has been actively developing emerging industries and attracting investment around the pillar industries and emerging industries mentioned above, so the urban economic resilience has increased the most.

*4.2. Characteristics of Spatial Distribution.* The spatial distribution of urban economic resilience in the rust belt of China had a gradient distribution of "high in the east and low in the west." By 2016, low urban economic resilience areas were mainly concentrated in small cities in the middle of Heilongjiang and southwestern Jilin, which have a lower economic ranking and slower development speed. High urban economic resilience areas were mainly concentrated in large cities with high economic rankings and faster development speed.

In terms of the urban economic resilience index in the Northeast China, the trend of "high in the east and low in the west" reflects that the overall urban economic resilience in the eastern coastal areas was higher than that in the western inland areas. In 2005, the overall urban economic resilience of Northeast China was relatively low, and that of Suihua, Qitaihe, Siping, Chaoyang Baicheng, and Chifeng was very weak. By 2010, after the financial crisis and natural disasters, with the assistance of national policies and self-adjustment, the economic growth in the Northeast region has accelerated. Although the spatial differentiation of urban economic resilience still exists, the number of cities with low economic resilience has significantly decreased, especially the economic resilience of cities in the eastern part of the Northeast China has grown rapidly. Since then, countries and regions have paid more and more attention to regional economic development. Therefore, urban economic resilience in Northeast China grew rapidly from 2010 to 2016, and low economic resilience areas were mainly concentrated in central Heilongjiang and southwestern Jilin.

We used ArcGIS software to classify urban economic resilience in Northeast China. The region was divided into four levels according to resilience: worse resilience, low resilience, medium resilience, and high resilience (Figure 2). In 2005 (Figure 2(a)), 27 cities in Northeast China had worse and low resilience levels, which were distributed widely in each province. Ten cities had medium and high resilience levels. In 2010 (Figure 2(b)), 19 cities had worse and low economic resilience, mainly distributed in northwestern Heilongjiang Province, western and southern Jilin Province, and northwestern Liaoning Province. 18 cities had medium and high resilience levels and were mainly distributed in eastern Heilongjiang, central Jilin, and southeastern Liaoning. In 2016 (Figure 2(c)), 18 cities had worse and low economic resilience in Northeast China, mainly distributed in central Heilongjiang, southwest Jilin, western and northern Liaoning, and eastern Inner Mongolia. 19 cities had medium and high resilience levels, mainly distributed in southeast Heilongjiang, central Jilin, and southeast Liaoning.

From 2005 to 2016, the number of cities with worse and low economic resilience gradually decreased and that of cities with medium and high economic resilience gradually increased in the region. Economic resilience of cities at low levels of economic development, such as Liaoyuan, Siping, Baicheng, Suihua, and Yichun, remained low while that of cities at medium levels of economic development, such as Dalian and Shenyang, remained medium. The overall distribution was "high in the east and low in the west."

## 5. Contribution of Each Subsystem and Promote Suggestions

The assessment of urban economic resilience is a complex system, and the contribution of each subsystem, including its contribution to urban economic resilience, is different in different cities. The greater the contribution of the subsystem is, the more likely it is to become the main factor that affects

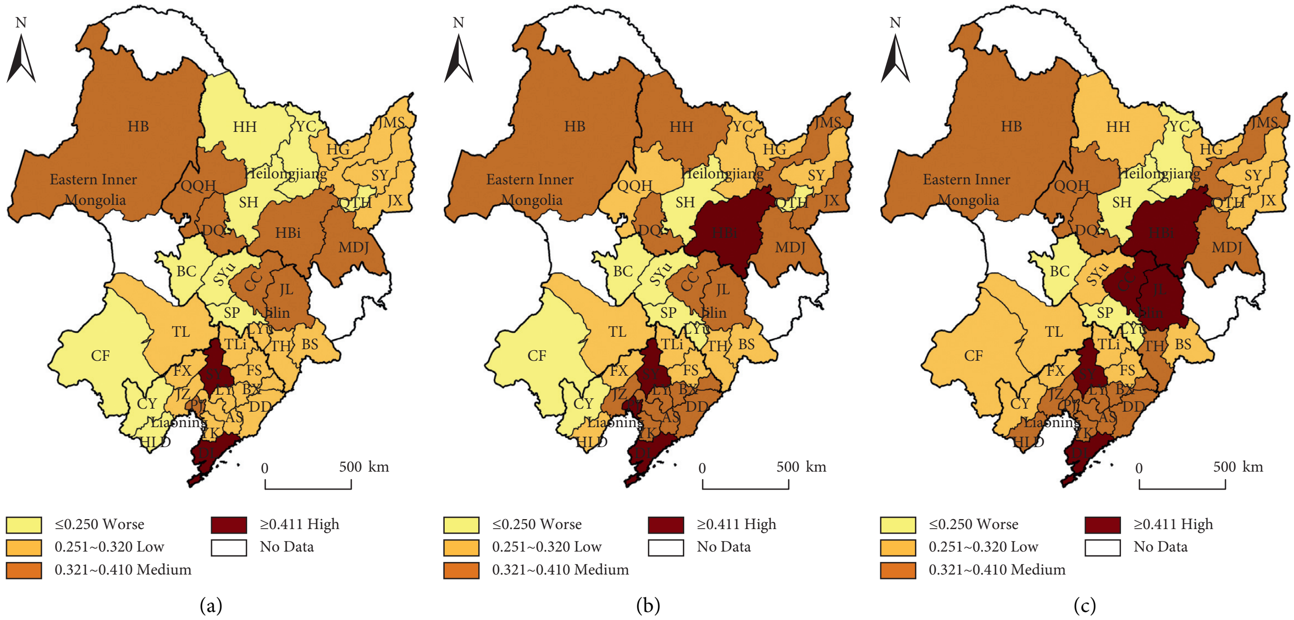



Figure 2: Map of distributions of urban economic resilience in Northeast China in 2005, 2010, and 2016.

urban economic resilience. The contribution of each subsystem in this paper can be determined according to its proportion in the urban economic resilience. The specific algorithm that we used is the entropy method (Section 3.3.2).

*5.1. Overall Characteristics of Subsystems' Contributions.* According to the average values of subsystem weight (Table 3), it is clear that urban economic resilience in the rust belt of China was significantly contributed by the subsystems of diversity and trends of development, with weights of 0.214 and 0.216, respectively. Among the five subsystems, the system of diversity made the largest contribution to Harbin's economic resilience, with a weight of 0.336, and the smallest contribution to Yichun's economic resilience, with a weight of 0.143. The subsystem of capabilities related to revenue and expenditure made the largest contribution to Benxi's economic resilience, with a weight of 0.273, and the smallest contribution to Dandong's economic resilience, with a weight of 0.124. The innovation environment subsystem made the largest contribution to Suihua's economic resilience, with a weight of 0.370, and the smallest contribution to Harbin, with a weight of 0.130. The subsystem of trends in development made the largest contribution to Chifeng's economic resilience, with a weight of 0.357, and the smallest contribution to Shenyang, with a weight of 0.135. The openness subsystem made the largest contribution to Heihe's economic resilience, with a weight of 0.306, and the smallest contribution to Anshan's economic resilience (a weight of 0.124).

*5.2. Spatial Mapping of Subsystem's Contribution.* We applied ArcGIS to the classification of contributions of each subsystem. Four levels were grouped: low, medium, relatively high, and high. The diversity subsystem made the relatively high and high contribution to urban economic resilience in 18 cities, and these cities were concentrated in Liaoning Province (Figure 3(a)). The capability subsystem related to revenue and expenditure made the relatively high and high contribution to urban economic resilience in five cities, but these cities were scattered in space (Figure 3(b)). The innovation environment system made the relatively high and high contribution to urban economic resilience in 14 cities, and these cities were concentrated in central and eastern Heilongjiang Province (Figure 3(c)). The trend of development system made the relatively high and high contribution to urban economic resilience in 18 cities, and these cities were mainly in western Heilongjiang Province, southern Liaoning Province, and eastern Inner Mongolia Autonomous Region (Figure 3(d)). The openness system made the relatively high and high contribution to urban economic resilience in 11 cities, and these cities were concentrated in northern Heilongjiang Province, northern Jilin Province, and central Liaoning Province (Figure 3(e)).

*5.3. Major Factors Contributing Economic Resilience of Different Types of Cities.* Cities in the rust belt in China can be divided into four types according to development: comprehensive economic cities, resource-based and old industrial cities, coastal cities, and agricultural cities. The cities can be divided into large, medium-sized, and small according to their urban population (Table 4). Considering the weights of the subsystems at the relatively high and high levels as the factor that makes a greater contribution to urban economic resilience, the results are summarized as follows: comprehensive economic cities were mainly large cities and were significantly contributed by the subsystem of diversity. Resource-based and old industrial cities as a whole were significantly contributed by the subsystem of innovation environment (seven cities). In terms of scale, large cities were contributed by the diversity subsystem, medium-sized

Table 3: Weights of the contribution of each subsystem on the economic resilience of cities in Northeast China.

| City | Diversity system | Capabilities related to revenue and expenditure system | Innovation environment system | Trend of development system | Openness system | Total |
|---|---|---|---|---|---|---|
| Chifeng | 0.148 | 0.167 | 0.178 | 0.357 | 0.150 | 1 |
| Tongliao | 0.150 | 0.150 | 0.233 | 0.315 | 0.151 | 1 |
| Hulun Buir | 0.186 | 0.227 | 0.187 | 0.208 | 0.192 | 1 |
| Shenyang | 0.280 | 0.148 | 0.138 | 0.135 | 0.299 | 1 |
| Dalian | 0.309 | 0.140 | 0.137 | 0.262 | 0.152 | 1 |
| Anshan | 0.264 | 0.171 | 0.301 | 0.140 | 0.124 | 1 |
| Fushun | 0.207 | 0.135 | 0.151 | 0.346 | 0.161 | 1 |
| Benxi | 0.243 | 0.273 | 0.148 | 0.148 | 0.188 | 1 |
| Dandong | 0.292 | 0.124 | 0.152 | 0.267 | 0.164 | 1 |
| Jinzhou | 0.212 | 0.147 | 0.144 | 0.353 | 0.144 | 1 |
| Yingkou | 0.324 | 0.126 | 0.156 | 0.237 | 0.158 | 1 |
| Fuxin | 0.254 | 0.133 | 0.207 | 0.138 | 0.267 | 1 |
| Liaoyang | 0.285 | 0.174 | 0.171 | 0.172 | 0.198 | 1 |
| Panjin | 0.191 | 0.198 | 0.225 | 0.192 | 0.195 | 1 |
| Tieling | 0.203 | 0.178 | 0.256 | 0.192 | 0.170 | 1 |
| Chaoyang | 0.301 | 0.172 | 0.178 | 0.185 | 0.164 | 1 |
| Huludao | 0.227 | 0.181 | 0.195 | 0.220 | 0.177 | 1 |
| Changchun | 0.212 | 0.181 | 0.158 | 0.246 | 0.204 | 1 |
| Jilin | 0.251 | 0.140 | 0.145 | 0.200 | 0.264 | 1 |
| Siping | 0.151 | 0.151 | 0.318 | 0.182 | 0.199 | 1 |
| Liaoyuan | 0.176 | 0.176 | 0.177 | 0.199 | 0.272 | 1 |
| Tonghua | 0.192 | 0.198 | 0.220 | 0.192 | 0.198 | 1 |
| Baishan | 0.184 | 0.184 | 0.187 | 0.261 | 0.183 | 1 |
| Songyuan | 0.191 | 0.199 | 0.195 | 0.225 | 0.191 | 1 |
| Baicheng | 0.156 | 0.156 | 0.165 | 0.353 | 0.171 | 1 |
| Harbin | 0.336 | 0.170 | 0.130 | 0.188 | 0.176 | 1 |
| Qiqihar | 0.190 | 0.198 | 0.194 | 0.221 | 0.197 | 1 |
| Jixi | 0.166 | 0.198 | 0.304 | 0.166 | 0.166 | 1 |
| Hegang | 0.172 | 0.201 | 0.252 | 0.202 | 0.173 | 1 |
| Shuangyashan | 0.211 | 0.173 | 0.170 | 0.170 | 0.275 | 1 |
| Daqing | 0.189 | 0.208 | 0.168 | 0.221 | 0.215 | 1 |
| Yichun | 0.143 | 0.152 | 0.268 | 0.139 | 0.299 | 1 |
| Jiamusi | 0.168 | 0.173 | 0.266 | 0.175 | 0.218 | 1 |
| Qitaihe | 0.261 | 0.128 | 0.248 | 0.135 | 0.228 | 1 |
| Mudanjiang | 0.179 | 0.201 | 0.230 | 0.218 | 0.172 | 1 |
| Heihe | 0.153 | 0.140 | 0.132 | 0.269 | 0.306 | 1 |
| Suihua | 0.153 | 0.157 | 0.370 | 0.159 | 0.161 | 1 |
| Average value | 0.214 | 0.171 | 0.201 | 0.216 | 0.198 | 1 |

Note: to make the data clearer, values in the table were all rounded to three decimal places.

cities were contributed by the innovation environment subsystem, and small cities were contributed by the developmental trends system. Coastal cities as a whole were greatly contributed by the diversity subsystem (five cities) and subsystem of developmental trends (five cities), no matter whether large cities, medium-sized cities, or small cities. Agricultural cities were mainly small and medium-sized cities and were greatly contributed by the subsystem of developmental trends (four cities). In terms of scale, medium-sized cities were contributed by the innovation environment subsystem, and small cities were contributed by the developmental trends system.

### 5.4. Promote Suggestions

*5.4.1. Promote Diversity.* The diversity subsystem can be improved mainly by enhancing industrial diversity and the ability of industries to transform and upgrade. Industrial diversity can be improved by extending the industry chain and promoting the diversity of regional enterprises. The extension of the industry chain can reinforce the connection between upstream and downstream enterprises, promote industry gathering in the region, and provide a basis for resource sharing between industries. The diversity of regional enterprises can allow different types of enterprises and talents to coexist, alleviate the problems posed by the large proportion of state-owned enterprises in Northeast China, and provide more employment opportunities and economic income.

The system of diversity made the largest contribution to the urban economic resilience of comprehensive cities (Dalian, Changchun, Shenyang, and Harbin) and some coastal cities (Dandong, Jinzhou, Yingkou, and Huludao). These cities can take the following specific measures to improve their economic resilience. In terms of extending the industrial chain, it can be achieved by introducing upstream and downstream industries. Introduced industries need to

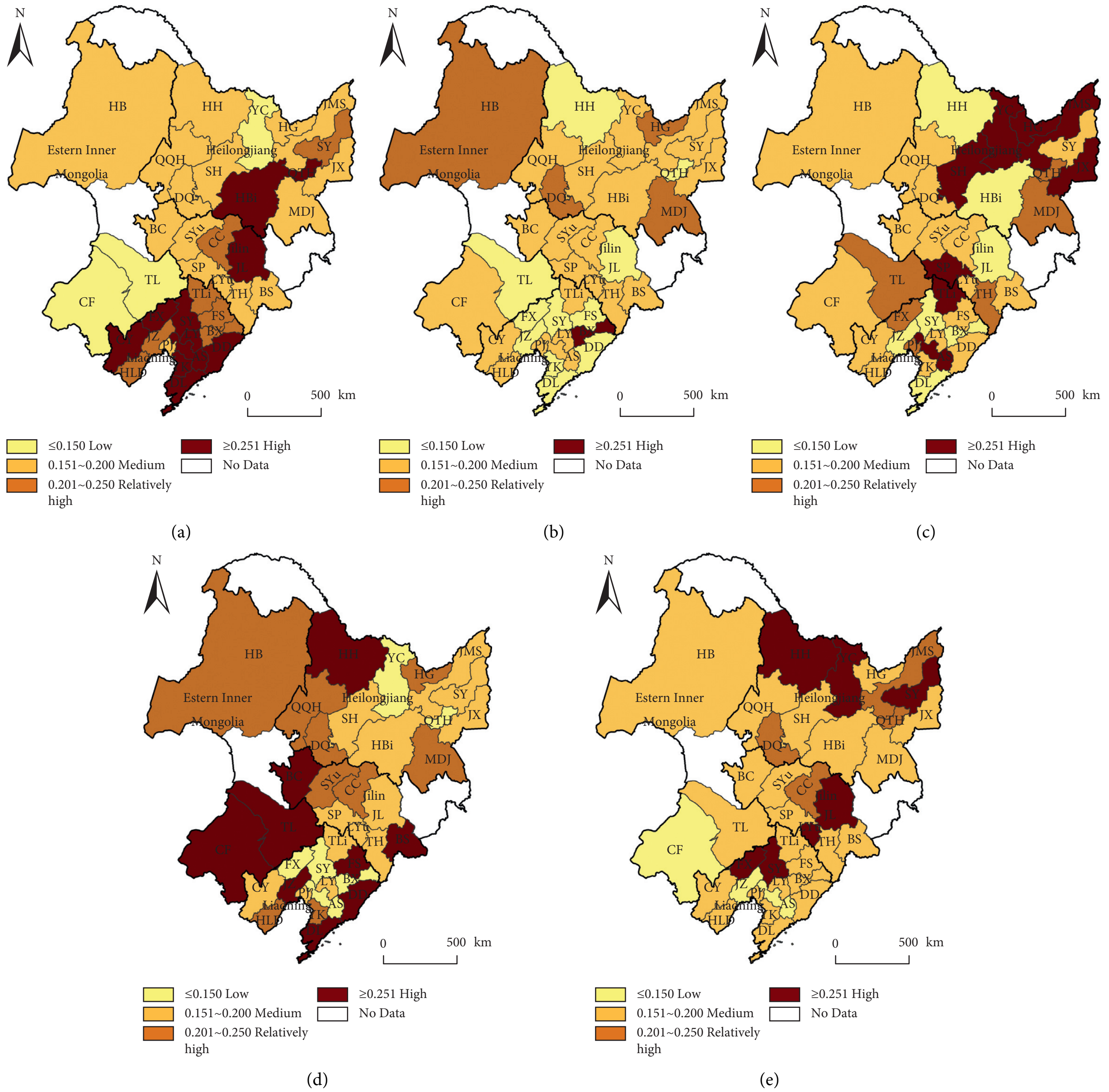



Figure 3: Spatial distribution of the degree of contribution of subsystems on urban economic resilience. (a) Contribution degree of the diversity system. (b) Contribution degree of capabilities related to revenue and expenditure system. (c) Contribution degree of the innovation environment system. (d) Contribution degree of the trend development system. (e) Contribution degree of the openness system.

Table 4: Weight ratings of subsystems contributing economic resilience for different types and scales of cities.

| Type | Scale | City | Diversity system level | Capabilities related to revenue and expenditure system level | Innovation environment system level | Trend of development system level | Openness system level |
|---|---|---|---|---|---|---|---|
| Comprehensive economic city | Large city | Dalian | High | Low | Low | High | Medium |
| | | Changchun | R-high | Medium | Medium | R-high | R-high |
| | | Shenyang | High | Low | Low | Low | High |
| | | Harbin | High | Medium | Low | Medium | Medium |

Table 4: Continued.

| Type | Scale | City | Diversity system level | Capabilities related to revenue and expenditure system level | Innovation environment system level | Trend of development system level | Openness system level |
|---|---|---|---|---|---|---|---|
| Resource-based and old industrial city | Large city | Anshan | High | Medium | High | Low | Low |
| | | Fushun | R-high | Low | Medium | High | Medium |
| | | Jilin | High | Low | Low | Medium | High |
| | | Daqing | Medium | R-high | Medium | R-high | R-high |
| | Medium-sized city | Chifeng | Low | Medium | Medium | High | Low |
| | | Benxi | R-high | High | Low | Low | Medium |
| | | Fuxin | High | Low | R-high | Low | High |
| | | Panjin | Medium | Medium | R-high | Medium | Medium |
| | | Jixi | Medium | Medium | High | Medium | Medium |
| | | Yichun | Low | Medium | High | Low | High |
| | | Mudanjiang | Medium | R-high | R-high | R-high | Medium |
| | Small city | Huludao | R-high | Medium | Medium | R-high | Medium |
| | | Liaoyuan | Medium | Medium | Medium | Medium | High |
| | | Tonghua | Medium | Medium | R-high | Medium | Medium |
| | | Baishan | Medium | Medium | Medium | High | Medium |
| Coastal cities | Large city | Dalian | High | Low | Low | High | Medium |
| | Medium-sized city | Panjin | Medium | Medium | R-high | Medium | Medium |
| | | Dandong | High | Low | Medium | High | Medium |
| | | Jinzhou | R-high | Low | Low | High | Low |
| | | Yingkou | High | Low | Medium | R-high | Medium |
| | Small city | Huludao | R-high | Medium | Medium | R-high | Medium |
| Agricultural city | Medium-sized city | Panjin | Medium | Medium | R-high | Medium | Medium |
| | | Siping | Medium | Medium | High | Medium | Medium |
| | Small city | Hulun Buir | Medium | R-high | Medium | R-high | Medium |
| | | Songyuan | Medium | Medium | Medium | R-high | Medium |
| | | Baicheng | Medium | Medium | Medium | High | Medium |
| | | Heihe | Medium | Low | Low | High | High |

*Note.* ① According to the *Notice on Adjusting the Standard of City Scale Division*, the standard for the scale for cities was based on the population of their urban areas. Small cities (fewer than 500,000people), medium-sized cities (500,000–1 million), large cities (1 million–5 million), megacities (5 million–10 million), and supercities (over 10 million) ("more than" includes the number itself, and "less than" does not include it). ② R-high in the table refers to relatively high.

be related to the existing industries. On the one hand, this way can provide more diverse employment opportunities and increase economic income. On the other hand, this way can promote the agglomeration of enterprises within the city and make the connection between upstream and downstream enterprises and existing enterprises closer, such as face-to-face contact with key personnel and timely information acquisition. In terms of promoting the transformation and upgrading of industries, it can be carried out in stages: first, cities can take the lead in introducing enterprises engaged in basic industries, including technology, transportation, energy, raw materials, and other industrial types of enterprises. This can reduce the unit production costs of other enterprises and improve economic benefits. Next, cities need to introduce emerging enterprises. Cities should first understand the supply and demand of labor in existing industries and then selectively introduce new enterprises according to the existing employment structure, thereby promoting the stable development of the urban economy.

*5.4.2. Promote Capabilities Related to Revenue and Expenditure.* The system of capabilities related to revenue and expenditure mainly includes micro- and macrolevels. And it made the largest contribution to the urban economic resilience of Benxi, Hulun Buir, Daqing, Mudanjiang, and Hegang. These cities can take the following specific measures to improve their economic resilience. At the microlevel, the abovementioned cities can increase the income of urban residents by increasing the wages of employees and improving the wage protection system. For example, by increasing the adjustment of the increase in wages of employees, improving and implementing the minimum wage system, and promoting enterprises to establish a collective wage bargaining system to improve the income and expenditure capacity of the urban economy. At the same time, cities should improve measures conducive to expanding consumption, for example, by implementing low fees for public welfare places such as museums, memorials, exhibition halls, and other public welfare places to broaden the scope of consumption. At the macrolevel, first, cities should play the role of tax macrocontrol. This is mainly through two ways: "tax increase" and "tax reduce." For example, when the economy is booming, the national income level and consumption capacity will increase, and tax increase can prevent the effect of excessive economic expansion; while the economy is in recession, the national income level and consumption power will decrease, and tax reduce can attract investment. Second, cities can increase fiscal macrocontrol, mainly by adjusting the structure of fiscal expenditures, such as increasing fiscal spending on science and education, culture, public health, and other social undertakings, thus promoting scientific and technological progress, developing the environmental protection industry, and promoting the transformation of socioeconomic development.

*5.4.3. Promote Innovation Environment.* The innovation environment system can provide an impetus for innovative activities. And it made the largest contribution to the urban economic resilience of resource-based and old industrial cities (Anshan, Fuxin, Panjin, Jixi, Yichun, Mudanjiang, and Tonghua). These cities can take the following specific measures to improve their economic resilience. First, local governments should ensure equitable access to innovation resources. Specifically, it can be achieved by cultivating and introducing high-tech enterprises, high-level universities, and scientific research institutions with high innovation ability. Second, full play should be given to the leading role of entrepreneurs and stimulate their enthusiasm for innovation. At the same time, enterprises must establish a fair reward mechanism that can help distribute dividends to talents who made important contributions to innovation. Specifically, enterprises can implement equity incentive and dividend incentive for innovative talents, so that entrepreneurs can get more benefits and spiritual incentives for their innovative activities. Third, we must focus on cultivating and attracting all types of talent and formulate policies that promote the mobility and introduction of talent, for example, settling down for talents and solving problems such as children's enrollment and spouse employment. Fourth, we need to strengthen domestic and international collaborative innovation and make full and effective use of resources and markets for foreign innovation, such as strengthening trade cooperation with neighboring countries, in order to broaden urban innovative development path.

*5.4.4. Promote Developmental Trends.* The system of developmental trends includes two aspects: economic and social development. A good urban economic development trend is a necessary condition for improving the ability of the urban economic system to resist uncertain risks and promoting the stable development of the urban economy. The developmental trend system made the largest contribution to the urban economic resilience of agricultural cities (Hulun Buir, Songyuan, Baicheng, and Heihe) and some coastal cities (Dalian, Panjin, Dandong, Jinzhou, Yingkou, and Huludao). The following measures can be taken to maintain a good trend of urban economic development for the above cities: the first is to adjust the economic structure to promote growth. This is mainly through adjusting the industrial structure to bring the urban economic development of the Northeast region to a new level. Northeast China has a large number of old industrial base cities and resource-dependent cities. Thus, adjusting the industrial structure can promote rapid urban economic growth. Second, we need to pay attention to population-related issues, adjust social distribution, and solve employment problems. Population aging is a problem in the economic development of cities in Northeast China. It can be alleviated through such measures as adjusting population policies, improving the population structure, increasing the working age range of the labor force, encouraging older people to continue working, and improving the social security system. Third, we need to protect the resources and the environment to improve the living standards of residents. The available resource reserves need to be protected. At the same time, more attention should be given to environmental issues. The protection of resources and the environment can guarantee an improvement in the quality of life. For example, cities can encourage the development of new pollution-free enterprises while adjusting the industrial structure and encourage the use of clean production materials and technologies for existing industrial enterprises.

*5.4.5. Promote Openness.* The openness system includes two aspects: openness to foreign and domestic trade. And it made the largest contribution to the urban economic resilience of Heihe, Yichun, Shenyang, Shuangyashan, Liaoyuan, and other cities. The following strategies can be adopted in enhancing the openness of foreign trade for above cities: first of all, the business environment should be improved to enhance the intensity of investment attraction, because the quality of business environment directly affects the intensity of investment attraction. In recent years, one of the reasons for the slow economic development of cities in Northeast China is the lack of a good business environment, which affects the enthusiasm of foreign investment. Therefore, the government of each city should create a stable institutional environment to ensure the legitimate rights of investors. For example, the government can actively promote the openness of government affairs to establish its reputation, curb the phenomenon of arbitrary fees to reduce the burden of foreign enterprises, and standardize and simplify all kinds of examination and approval, licensing, and registration procedures to create a good business environment. Secondly, cities can further develop the northeast Asian market by attracting the investment of multinational corporations, so as to improve the level of urban import and export trade. At present, the proportion of investment by multinational companies in cities in the Northeast is relatively small. Therefore, cities should highlight their characteristics and focus on attracting the investment of large foreign transnational corporations, to improve the openness of domestic trade. Due to the cost and convenience of transportation directly affect the possibility of trade between cities, they can start by improving infrastructure and promoting the construction of a comprehensive regional transportation network. Specific measures include speeding up the construction of dead-end roads in each province, and speeding up the construction of high-speed rail between cities.

## 6. Conclusion and Discussion

### 6.1. Conclusions

*6.1.1. Forming a Comprehensive Urban Economic Resilience Measurement Model.* Based on the results of this research, the urban economic resilience system is mainly contributed by the five subsystems of urban economic diversity, capabilities related to revenue and expenditure, innovation environment, trends of development, and openness. We constructed an index system for urban economic resilience in Northeast China based on certain principles and index selection methods.

*6.1.2. Urban Economic Resilience in Northeast China Has Gradually Increased over Time.* In terms of time, the urban economic resilience of cities in Northeast China gradually increased from 2005 to 2016. The growth was small in 2005–2010 and large in 2010–2016. From 2005 to 2016, Dalian was the most resilient city and Suihua was the least, and their maximum and minimum values increased year by year. In terms of spatial distribution, urban economic resilience in Northeast China showed a gradient differentiation characteristic of "high in the east and low in the west," and the economic resilience of coastal cities was higher than that of inland cities.

*6.1.3. The Most Contributing Factors Are Different for Different Types of Cities.* Judging from the average values of the weights of the various subsystems, urban economic resilience in Northeast China is greatly contributed by the systems of diversity and developmental trends, with weights of 0.214 and 0.216, respectively. The economic resilience of comprehensive cities was mainly contributed by the diversity system, and that of resource-based and old industrial cities was mainly contributed by the systems of innovation environment systems. The economic resilience of coastal cities was mainly contributed by the systems of diversity and developmental trends, and that of agricultural cities was mainly contributed by the developmental trends system. Different types of cities should enhance their economic resilience based on the factors contributing to their economic systems.

*6.2. Discussion.* The results of this paper have guiding significance for the sustainable development of urban economy in Northeast China. Compared with the urban resilience index system established in the existing studies [27, 28, 37, 38], the index system in this paper is specifically applied to the field of urban economic resilience, and it is more detailed. And taking into account the actual situation of the development of Northeast China, we have taken the level of industrial diversification and its transformation and upgrading ability into account when constructing the index system. Therefore, the index system is in line with the actual economic development in Northeast China and can be used as a reference for future research on urban economic resilience. Moreover, different from the existing economic resilience research [40, 47], we subdivided the urban types in the study area. Therefore, this study not only enriches the research on the index system of urban economic resilience in developing countries but also diversifies the types of cities studied.

Although we have made some contributions to the study of urban economic resilience, there are still some areas that need to be improved: first, few studies have empirically considered the urban economic resilience of cases in China, and theoretical research is relatively scarce as well. Therefore, some factors were not taken into account here when constructing the index system of urban economic resilience in Northeast China, especially subjective factors such as economic policies. In future research, we intend to improve the index system for the evaluation of urban economic resilience from different perspectives. Second, this paper studied economic resilience at the city scale. No comparative analysis of economic resilience was provided between scales. The scale of research can be changed for further research. Finally, we did not consider the interaction between contributive factors in the context of urban economic resilience in Northeast China. A correlation test of these factors will be an important subject for research in the future.

## Data Availability

The data were drawn mainly from the China City Statistical Yearbook, China City Construction Statistical Yearbook, Regional Economic Statistical Yearbook, the statistical yearbooks of the relevant provinces and cities, economic census yearbooks, and statistical bulletins.

## Conflicts of Interest

The authors declare that there are no conflicts of interest regarding the publication of this paper.

## Acknowledgments

This research was supported by the Philosophy and Social Science Foundation of China, no. 16BJY039.

## References

[1] T. Li, "New progress in study on resilient cities," *Urban Planning International*, vol. 32, no. 5, pp. 15–25, 2017, in Chinese.
[2] L. Chen, X. Zhu, and J. Sun, "The basic concept, mechanism and planning ideas of resilient cities," *Modern Urban Research*, vol. 9, pp. 18–24, 2017, in Chinese.
[3] X. Hu and C. Yang, "Building a role model for rust belt cities? Fuxin's economic revitalization in question," *Cities*, vol. 72, pp. 245–251, 2018.
[4] L. Zhao, W. Wang, Y. Zhang, R. Li, and D. Wu, "The spatio-temporal evolution about urban vulnerability on the Northeast region since Northeast revitalization," *Economic Geography*, vol. 34, no. 12, pp. 69–77, 2014, in Chinese.
[5] J. Sun and X. Sun, "Research progress of regional economic resilience and exploration of its application in China," *Economic Geography*, vol. 37, pp. 1–9, 2014, in Chinese.

[6] C. S. Holling, "Resilience and stability of ecological systems," *Annual Review of Ecology and Systematics*, vol. 4, no. 4, pp. 1–23, 1973.

[7] W. N. Adger, "Social and ecological resilience: are they related?" *Progress in Human Geography*, vol. 24, no. 3, pp. 347–364, 2000.

[8] A. Rose, "Economic resilience to natural and man-made disasters: multidisciplinary origins and contextual dimensions," *Environmental Hazards*, vol. 7, no. 4, pp. 383–398, 2007.

[9] N. J. Wood, C. G. Burton, and S. L. Cutter, "Community variations in social vulnerability to Cascadia-related tsunamis in the U.S. Pacific Northwest," *Natural Hazards*, vol. 52, no. 2, pp. 369–389, 2010.

[10] F. Landry, J. Dupras, and C. Messier, "Convergence of urban forest and socio-economic indicators of resilience: a study of environmental inequality in four major cities in eastern Canada," *Landscape and Urban Planning*, vol. 202, 2020.

[11] S. T. A. Pickett, M. L. Cadenasso, and J. M. Grove, "Resilient cities: meaning, models, and metaphor for integrating the ecological, socio-economic, and planning realms," *Landscape and Urban Planning*, vol. 69, no. 4, pp. 369–384, 2004.

[12] R. Leichenko, "Climate change and urban resilience," *Current Opinion in Environmental Sustainability*, vol. 3, no. 3, pp. 164–168, 2011.

[13] J. Ahern, "From fail-safe to safe-to-fail: sustainability and resilience in the new urban world," *Landscape and Urban Planning*, vol. 100, no. 4, pp. 341–343, 2011.

[14] P. Lu and D. Stead, "Understanding the notion of resilience in spatial planning: a case study of Rotterdam, The Netherlands," *Cities*, vol. 35, no. 4, pp. 200–212, 2013.

[15] M. Thornbush, O. Golubchikov, and S. Bouzarovski, "Sustainable cities targeted by combined mitigation-adaptation efforts for future-proofing," *Sustainable Cities and Society*, vol. 9, pp. 1–9, 2013.

[16] S. Meerow and J. P. Newell, "Spatial planning for multifunctional green infrastructure: growing resilience in Detroit," *Landscape and Urban Planning*, vol. 159, pp. 62–75, 2017.

[17] S. Tyler and M. Moench, "A framework for urban climate resilience," *Climate & Development*, vol. 4, no. 4, pp. 311–326, 2012.

[18] K. C. Desouza and T. H. Flanery, "Designing, planning, and managing resilient cities: a conceptual framework," *Cities*, vol. 35, pp. 89–99, 2013.

[19] M. Suárez, E. Gómez-Baggethun, J. Benayas, and D. Tilbury, "Towards an urban resilience index: a case study in 50 Spanish cities," *Sustainability*, vol. 8, no. 8, p. 774, 2016.

[20] D. Zhao, B. Shi, Y. Chen, C. Chen, and P. Lei, "Effect of collaborative management on enhancement mechanism of urban resilience," *Management Review*, vol. 8, pp. 207–214, 2016.

[21] R. Hassink, "Regional resilience: a promising concept to explain differences in regional economic adaptability?" *Cambridge Journal of Regions, Economy and Society*, vol. 3, no. 1, pp. 45–58, 2010.

[22] R. Martin, "Regional economic resilience, hysteresis and recessionary shocks," *Journal of Economic Geography*, vol. 12, no. 12, pp. 1–32, 2012.

[23] R. Wink, "Regional economic resilience: policy experiences and issues in Europe," *Raumforschung und Raumordnung*, vol. 72, no. 2, pp. 83-84, 2014.

[24] S. Meerow, J. P. Newell, and M. Stults, "Defining urban resilience: a review," *Landscape and Urban Planning*, vol. 147, pp. 38–49, 2016.

[25] M. Spaans and B. Waterhout, "Building up resilience in cities worldwide - rotterdam as participant in the 100 resilient cities programme," *Cities*, vol. 61, pp. 109–116, 2017.

[26] J. Simmie and R. Martin, "The economic resilience of regions: towards an evolutionary approach," *Cambridge Journal of Regions, Economy and Society*, vol. 3, no. 1, pp. 27–43, 2010.

[27] L. Oxborrow and C. Brindley, "Regional resilience in recessionary times: a case study of the East Midlands," *International Journal of Retail & Distribution Management*, vol. 40, no. 11, pp. 882–899, 2012.

[28] B. Fingleton, H. Garretsen, and R. Martin, "Recessionary shocks and regional employment: evidence on the resilience of U.K. regions," *Journal of Regional Science*, vol. 52, no. 1, pp. 109–133, 2012.

[29] A. Sharifi, "Resilient urban forms: a macro-scale analysis," *Cities*, vol. 85, pp. 1–14, 2019.

[30] J. R. Holm and C. R. Østergaard, "Regional employment growth, shocks and regional industrial resilience: a quantitative analysis of the Danish ICT sector," *Regional Studies*, vol. 49, no. 1, pp. 95–112, 2015.

[31] R. Capello, A. Caragliu, and U. Fratesi, "Spatial heterogeneity in the costs of the economic crisis in Europe: are cities sources of regional resilience?" *Journal of Economic Geography*, vol. 15, no. 5, pp. 951–972, 2015.

[32] Q. Xie, "Enlightenment of resilient city construction policy in developed country," *Scientific Decision Making*, vol. 24, no. 4, pp. 60–75, 2017, in Chinese.

[33] K. Liu, "Emergence and theory construction of resilient planning in the background of socialist market economy: based on urban planning practice in Shenzhen (1979-2011)," *City Planning Review*, vol. 38, no. 11, pp. 59–64, 2014, (in Chinese).

[34] C. Xiu, Y. Wei, and Q. Wang, ""Evaluation of urban resilience of Dalian city based on the perspective of "Size-Density-Morphology"" *Acta Geographica Sinica*, vol. 73, no. 12, pp. 2315–2328, 2018, in Chinese.

[35] Y. Chen and K. Li, "Overseas studies progress and practice exploration on resilient city: based on the perspective of public administration," *Chinese Public Administration*, vol. 33, no. 1, pp. 137–143, 2017, in Chinese.

[36] X. Zang and Q. Wang, "Case studies on resilience," *Resilience and Urban Disasters*, vol. 37, no. 22, p. 94, 2019, in Chinese.

[37] J. Liu and Z. Zeng, "The construction and empirical study of the evaluation index system of resilient city," *E-Government*, vol. 3, pp. 82–88, 2014, in Chinese.

[38] L. Zhou, "Resilient Cities: risk Management and indicator construction--Also discuss the international case," *Journal of Beijing Administration Institute*, vol. 13–20, 2016 in Chinese.

[39] X. Qi, J. Zhang, and W. Xu, "A study on the evaluation of the development of county economic resilience in zhejiang province," *Zhejiang Social Sciences*, vol. 5, pp. 40–46+156, 2019, in Chinese.

[40] Y. Xu and C. Wang, "Influencing factors of regional economic resilience in the 2008 financial crisis: a case study of Zhejiang and Jiangsu provinces," *Progress in Geography*, vol. 36, no. 8, pp. 986–994, 2017, in Chinese.

[41] X. Feng, C. Xiu, L. Bai, Y. Zhong, and Y. Wei, "Comprehensive evaluation of urban resilience based on the perspective of landscape pattern: a case study of Shenyang city," *Cities*, vol. 104, 2020.

[42] H. Guan, P. Zhang, W. Liu, and J. Li, ““A comparative analysis of the economic transition process of China’s old industrial cities based on evolutionary resilience theory,” *Acta Geographica Sinica*, vol. 73, no. 4, pp. 771–783, 2018, in Chinese.
[43] L. Li, P. Zhang, J. Tan, and H. Guan, “A regional economic resilience approach to the economic revitalization process in Liaoning Old Industrial Base, China,” *Scientia Geographica Sinica*, vol. 39, no. 1, pp. 116–124, 2019, in Chinese.
[44] M. Quan and Y. Chen, “Analysis of evaluation system & Mode of overseas resilient city,” *Housing Science*, vol. 39, pp. 1–6, 2019, in Chinese.
[45] S. Woolf, J. Twigg, P. Parikh, A. Karaoglou, and T. Cheaib, “Towards measurable resilience: a novel framework tool for the assessment of resilience levels in slums,” *International Journal of Disaster Risk Reduction*, vol. 19, pp. 280–302, 2016.
[46] S. Slocum and C. Kline, “Regional resilience: opportunities, challenges and policy messages from Western North Carolina,” *Anatolia*, vol. 25, no. 3, pp. 403–416, 2014.
[47] T. Zhang, “A study of temporal and regional characteristics of the economic resilience of the cities in jiangsu province,” *Journal of Wuxi Vocational Institute of Commerce*, vol. 18, pp. 1–8, 2018, in Chinese.

Hindawi
Complexity
Volume 2021, Article ID 6265751, 14 pages
https://doi.org/10.1155/2021/6265751

*Research Article*

# Geospatial Analysis and Research on Social and Spatial Inequality of Compulsory Education: A Case Study of Hangzhou, China

**Ge He** [1] **and Qinshi Huang** [1,2]

[1]School of Architecture and Urban Planning, Nanjing University, Nanjing 210093, China
[2]School of Civil Engineering and Architecture, Zhejiang University of Science and Technology, Hangzhou 310023, China

Correspondence should be addressed to Qinshi Huang; huangqinshi@zust.edu.cn

Received 10 May 2021; Accepted 23 August 2021; Published 31 August 2021

Academic Editor: Jing-Hu Pan



Equal compulsory education is an important way to realize social and spatial equality, while the uneven allocation of educational resources in different regions and groups results in inequality of opportunity and solidification of social strata. Traditional research conducted on the basis of fixed search range ignores the special institutional background of Chinese school district system. In this paper, an improved Gaussian two-step floating catchment area model is developed taking into consideration the school district system, while the bivariate local spatial analysis method and geographically weighted regression model are employed to study the social and spatial differentiation of compulsory education accessibility and its capitalization effects in Hangzhou. Results show that (1) the improved Gaussian two-step floating catchment area model is more in line with the national condition of China's "nearby schooling" policy; (2) the accessibility of compulsory schools in Hangzhou shows an obvious core-periphery typology, and the aggregation effect of primary school accessibility is more significant than that of secondary schools; (3) compared to groups with high socioeconomic status, vulnerable groups are highly disadvantaged in terms of access to educational services; (4) spatial heterogeneity exists in education capitalization, and the areas where education accessibility has the strongest impact on housing prices are in the central city with rich high-quality educational resources; (5) high-quality educational resources, high-priced communities, clusters of high socioeconomic status groups, and communities enjoying high-level education accessibility are highly consistent in all spaces, which is the spatial expression of educational inequality. The research on Hangzhou, a regional central city, provides a theoretical basis and technical support for the humanistic shift in the allocation of educational resources.

## 1. Introduction

As economic globalization and neoliberalism advance, the gap between rich and poor has widened drastically. Given that differences in individual family background and ability endowments naturally exist and are difficult to eliminate, it is a top priority for the country's education efforts to rationalize the allocation of educational resources through the process of spatial redistribution and to alleviate educational inequality caused by unequal social and economic status. Education is an important manifestation of social governance and the degree of justice as it forms the foundation of people's livelihood. Hence, educational equality has a direct impact on the quality of people's lives. However, issues with imbalance and inadequacy in China's education remain prominent, which may lead to significant social problems such as class consolidation. Because of the scarcity of educational resources and differentiation of local government's economic strength, there are extensive regional, urban-rural, and interschool differences in the allocation of educational resources [1, 2]. Since educational resources directly determine students' performance and play a decisive role in their future career choices [3–5], differences in the allocation of educational resources highlight the key role of education in the production and reproduction of identity by social groups [6, 7]. Groups with high socioeconomic status (SES) are able to use endogenous attributes such as economy, culture, and social capital to achieve monopoly on quality educational resources [8]. Subsequently, future generations who enjoy quality educational services can further consolidate their social status. This kind of intraclass identity transmission, brought about by the differentiation of educational resource allocation, ensures the continuity of upper-class social status and cuts off the possibility of upward mobility of the bottom masses, intensifying social differentiation and stratification.

Accurately assessing the service capacity of educational resources is a prerequisite for exploring educational inequality. Since Hansen proposed the concept of accessibility [9], scholars use various spatial and temporal methods to quantitatively analyze the differences in spatial accessibility of educational facilities [10–18]. These models, without exception, determine whether residents are located within the service area by artificially delineating service thresholds for educational facilities [12, 17, 19]. However, unlike the compulsory education system in Western countries, China's "nearby schooling" policy has set an aspatial threshold for access to educational services [11, 20], and the service capacity of educational resources cannot be measured simply by how easy it is to reach the nearest school from a given location. In fact, under the school district system, social attributes are attached to educational facilities. These are essentially geographically distributed public goods, making the measurement of accessibility and equity more complex. The "nearby schooling" policy does not refer to the nearest school in terms of space but the right given to all school-age teenagers to attend schools within their registered household region. Therefore, ignoring the role of school district when calculating accessibility of education will enlarge the error [14, 21]. Accordingly, combining factors such as school location, service quality, and school district constraints, we define education accessibility in the Chinese context as the ability of a population to obtain services from a school which he or she is allowed to attend.

A correct understanding of the relationship between the difference among groups and educational accessibility is the key to grasping the connotation of educational equality and promoting it. Traditionally, educational equality has been about ensuring all students enjoy equal access to education. However, educational equality cannot simply be equated to equal education opportunities. Value judgments need to be made on the distribution process and relative results about whether they are reasonable and good, since the equity of distribution results or status does not necessarily mean equality. In fact, the difficulty in accessing public goods and services varies significantly among groups with different backgrounds, economic capabilities, and social status. Existing research points out that low-income groups [22], migrants [23], racial/ethnic minorities [24, 25], and other vulnerable groups are usually disadvantaged in terms of access to public services [26–28]. When investigating educational equality, we should not only concentrate on "place prosperity" but also pay attention to "people prosperity" [29, 30]. Educational equality is to achieve equal educational opportunities based on individual differences. When it comes to the index of accessibility, it should not only include the geographical barriers that people overcome to access public services in the physical world but also take into consideration social factors such as individuals' class, status, and reputation [31, 32]. Most prior studies in China equate educational equity with education equality and treat the equal allocation of educational resources as a criterion, whose focus is geographical justice rather than demographically targeted equality [21, 33]. With increasing awareness of social problems, a few researchers focused on exposing social inequality by exploring the differences in the ability of various social groups to access quality educational services [34–36]. Few scholars have assessed the quality of educational services through quantitative methods from the perspective of social equality [37]. Only by matching education accessibility with different social groups and accurately identifying the status of individuals enjoying high-quality education services, we can allocate educational resources in a targeted manner and realize "people prosperity," which will lead us to the ultimate goal of equal education.

As the basic unit of access to educational services, housing, which has both the geographical attribute of immovability and the social attribute of providing people with living space and corresponding eligibility for admission, provides us with another perspective to connect geographic territory and society [38]. The nature of housing as a commodity enables residents to pursue high-level educational resources by housing purchase [39, 40]. Because of the scarcity of school district housing, the value of the public attribute of education accessibility is inevitably capitalized through the surrounding housing [41, 42], which leads to a filtering mechanism for groups that cannot afford the premium of high-quality educational resources [43, 44] and exacerbating social stratification and socio-spatial inequality. Accurately identifying the degree of capitalization of education accessibility is critical to understanding the mechanism by which inequitable allocation of educational resources causes social problems. This will aid governments in deciding the structure of educational supply as well as the direction of educational policy reform.

As such, this research breaks through the limitations of traditional accessibility measurements and integrates the complex characteristics of spatial accessibility, service comprehensiveness, and access threshold of educational resources. It measures the equity of educational resource allocation from various aspects such as spatial distribution and social group matching, while exploring the capitalization effect of education accessibility with a view to gain insights into education, housing, and urban development policy making.

## 2. Study Area and Data

*2.1. Study Area.* Hangzhou, the capital of Zhejiang Province, is one of the most populated cities situated in the Yangtze River Delta. It is also one of the political, economic, and cultural centers of southeast coastal China. The chosen study area is the area around the city highway which is regarded as the main city, while public primary and secondary schools as well as residential communities are selected as subjects for analysis. The study mainly covers administrative districts including Shangcheng, Xiacheng, and Binjiang, most of Xihu, Gongshu, and Jianggan, and parts of Xiaoshan and Yuhang, covering a total area of about 1,040.41 $km^2$ (Figure 1).

Although Hangzhou has always been at the forefront of basic education across the whole country, the distribution of its educational resources is still differentiated where key primary and secondary schools are generally located in areas closer to the city center. Compared to regular primary and secondary schools, the educational resources of the schools in the inner city are more abundant. Consequently, the educational equality in the periphery of the city is greatly reduced, and construction and allocation of educational resources in related areas are still in progress. Given that private schools do not recruit students based on school districts, their impact on surrounding communities over a short time is negligible. Therefore, public primary and secondary schools in the main city are taken as research subjects.

*2.2. Data.* Data of school districts came from the List of School District Scope of Public Primary and Secondary Schools published by Hangzhou Municipal Education Bureau in 2020. There are 303 public schools within the study area, including 210 primary schools and 93 secondary schools.

As primary schools do not hold entrance examinations into secondary schools in Hangzhou, there are no official data on the ranking of primary schools. Operationally, we formed the hierarchy of all primary schools based on ranking information provided by local real estate agency websites such as Fangtianxia.com and 51souxue.com and random interviews with second-hand housing transaction platforms. Among 210 primary schools, prestigious schools generally recognized by the public with good reputation are assigned a value of 4, prestigious schools with a certain influence are assigned a value of 3, schools with limited influence are assigned a value of 2, and other schools are assigned a value of 1.

Unlike Western countries, although the examinations to enter high schools can reflect the quality of secondary schools in China, the data published by various schools have different calibers and some schools do not publicize relevant data [45]. Consequently, this article uses the enrolment rate of top eight senior high schools (also known as “best eight”) as the basis, combined with the evaluation given by local real estate agency, to classify 90 secondary schools. Several other Chinese scholars adopted similar methods and treated enrolment data combined with social subjective opinions as an indicator of school quality [46]. Also, some Western studies utilized similar methods and obtained robust estimation results [47, 48]. Similar to the assignment mechanism of primary school, secondary schools with better reputation are given higher scores.

The housing price data are collected from the list offered by Hangzhou Real Estate Market Comprehensive Management Service Center and supplemented with information from China Housing Price Quotes Platform (http://creprice.cn) in 2020. The demographic and social characteristics of the streets come from the results of the sixth census in Hangzhou, in 2010. The longitude and latitude of primary and secondary schools, school districts, and residential communities are generated by Google Earth. Geographic elements such as study area, traffic road networks, water systems, and mountains are vectorized using ArcGIS to establish a spatial database in the study area.

*2.3. Method.* As an improved model of two-step floating catchment area (2SFCA) method, the Gaussian two-step floating catchment area (G2SFCA) method is widely used to analyze spatiotemporal accessibility since it emphasizes the fact that residents’ demand for public goods will decrease with distance. Among most impedance functions, Gaussian decay function, rather than exponential function or power function, can best describe the supply-demand relationship between residents and the park that is normally distributed with distance [49, 50]. This allows G2SFCA to more realistically calculate accessibility under the influence of distance.

As the name suggests, the classic G2SFCA method is composed of two steps as follows:

$$R_j = \frac{S_j}{\sum_{k\in\{d_{kj}\le d_0\}} G(d_{kj}, d_0) N_k},$$
$$A_i = \sum_{i\in\{d_{ij}\le d_0\}} G(d_{kj}, d_0) R_j, G(d_{kj}, d_0) = \begin{cases} \dfrac{e^{-(1/2)\times(d_{kj}/d_0)^2} - e^{-(1/2)}}{1 - e^{-(1/2)}}, & d_{kj}\le d_0, \\ 0, & d_{kj}\le d_0, \end{cases} \tag{1}$$

where $R_j$ is the supply-demand ratio of the facility $j$, $S_j$ means the supply capacity of facility, $N_k$ indicates the sum of the population of each demand point $k$ within a threshold distance ($d_0$) from $j$, $A_i$ is the accessibility of each demand point $i$, and $G(d_{ij}, d_0)$ means the Gaussian distance between $i$ and $j$ while $d_{ij}$ is the geographic distance between $i$ and $j$.

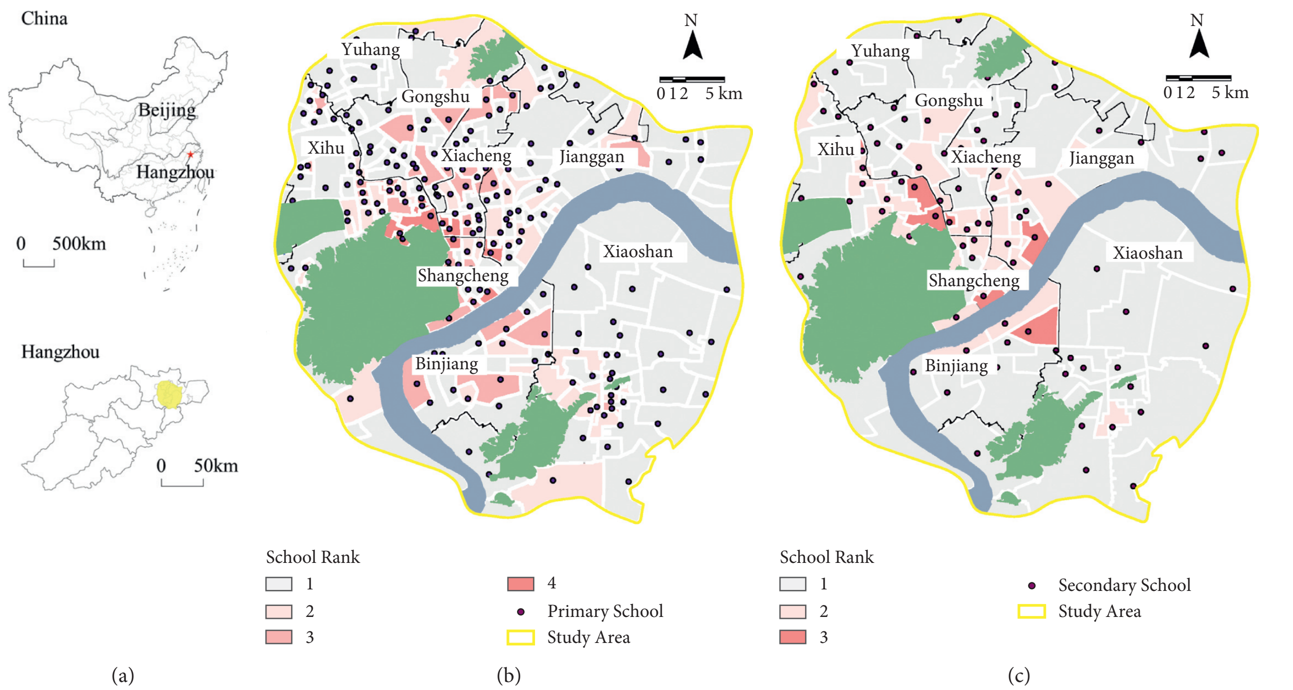



Figure 1: Study area and subject. (a) Location of the study area. (b) Distribution of primary schools and school districts. (c) Distribution of secondary schools and school districts.

However, there are several aspects that can be improved. Facility capacity is represented by a single index in most studies, which is not accurate enough. On the other hand, under school district policy, it is not the distance between the supply point and demand point that decides which gets the right to enjoy the service. To improve the accuracy of the model, we modified the correlation coefficient of the supply and demand relationship of public primary and secondary school.

(1) We enriched the indicator of the supply capacity (i.e., school service quality) to make sure the school's hardware and software facilities and residents' needs will not be missed out. According to current standards [51], there are several indexes that should be combined to inspect service capacity. We selected school level $l$, area $a$, number of teachers $t$, and number of students enrolled in 2020 $se$ to measure the performance of each school. (2) Before delineating the service area according to the distance threshold of the supply point and selecting the serviceable demand point, it is necessary to establish a collection of communities that can be served by particular school according to the delimited school district. Since every community in a certain school's district can enjoy educational service no matter how far it is to the school in spatial aspect, $d_0$ is decided by the distance from the farthest community. The advanced G2SFCA formula is listed as follows:

$$R'_j = \frac{S'_j}{\sum_{k\in T_j} G\left(d_{kj}, d_{\max}\right) N_k}, \tag{2}$$

$$A'_i = \sum_{i\in T_j} G\left(d_{kj}, d_{\max}\right) R'_j, \tag{3}$$

$$S'_j = l \times \left(w_1 \times (a \times a_0) + w_2 \times (t \times t_0) + w_3 \times se\right), \tag{4}$$

where $T_j$ refers to the collection of all communities in the school $j$'s district. $d_{\max}$ is the distance between the school $j$ and the farthest community in its district. In equation (4), $a_0$ is the official standard for area per student while $t_0$ means the official standard for teachers. $w_1$, $w_2$, and $w_3$ are the weights of school area, number of teachers, and number of students. According to relevant studies, teacher quality shows a stronger relationship than school facilities to pupil achievement [52], which enables a school with more teachers to provide better service. Besides, the number of students enrolled directly determines the school's ability to absorb and serve school-age children. Therefore, it was decided that $w_1 = 0.2$, $w_2 = 0.4$, and $w_3 = 0.4$.

In order to march towards "people prosperity," a spatial correlation analysis of the education accessibility among different groups characterized by socioeconomic attributes was added based on census data. In this level of research, the study area was divided into 500 m × 500 m grid adding up to 3,793 grids in total. We assigned accessibility data and census data to our grids based on their location. The Bivariate Local Indicator of Spatial Association (biLISA) method [53, 54] is employed to identify the spatial agglomeration characteristics of the supply capacity of public primary and secondary schools and the demand levels of different social groups, followed by exploring the matching degree between the two. The calculation formula of biLISA is

$$I^m_{xy} = z^m_x \sum_n W_{mn} z^n_y, \tag{5}$$

where $z_x^m$ represents the standardized value of education accessibility at grid $m$ and $z_y^n$ stands for the value at grid $n$. $W_{mn}$ is the spatial weight matrix between $m$ and $n$. $I_{xy}^m$ refers to the linear correlation between the variable value $z_x^m$ at grid $m$ and the mean value $z_y^n$ of another variable at the surrounding research unit $n$.

To explore the average effect of education accessibility on housing prices, we employed geographically weighted regression (GWR) to investigate the spatial heterogeneity and detect the spatial pattern of the impact of educational resource quality on housing prices. The model is as follows:

$$P_b = \beta_0 (u_b, v_b) + \sum_{c=1}^{q} \beta_c (u_b, v_b) f_{bc} + \varepsilon_b, \tag{6}$$

where $P_b$ is the housing price of residential community $b$, $(u_b, v_b)$ is the spatial location coordinates of $b$, $\beta_c (u_b, v_b)$ represents the $c$-th regression parameter at $b$, $f_{bc}$ represents the coefficient value of the $c$-th variable at observation point $b$, and $\varepsilon_b$ is the random error.

Housing prices are affected by various factors, most of which are measured by hedonic attributes. According to existing studies, hedonic determinants include architecture attributes (basic functions of a house such as age, decoration, and greening rate), location attributes (distance to the city center, metro station or landscape etc.), and neighborhood attributes (such as schools, hospitals, and parks in the surrounding area). Taking into account the availability of relevant data, we separate accessibility attributes from neighborhood attributes so as to analyze the mechanism of the impact school district has on housing price.

The socioeconomic data involved and their descriptive statistics are shown in Tables 1 and 2. Comparing the rank of education accessibility with the rank of the house price or population, we attempt to identify school deserts for different populations.

## 3. Results and Discussion

*3.1. Spatial Distribution of Public Educational Facilities.* In recent years, the supply of public education services in Hangzhou has shown a rising trend, which is closely related to economic development and population growth. From 2009 to 2019, the total number of students in school increased by 42.5% and the number of schools increased by 26.3% [55]. At present, the level of educational resource supply in Hangzhou is at the forefront of the whole country. The number of public education degrees offered per 10,000 people is 606.7, far exceeding the national average. The student-to-teacher ratio of primary and secondary schools is 15 : 1 and 12 : 1, which exceeds the standard set by “Opinions on the Establishment of the Standards for the Establishment of Primary and Secondary School Staff” from the Central Editing Office which set them at 19 : 1 and 13.5 : 1. However, there are obvious shortcomings in the spatial allocation of educational resources. Several smallest streets own the most concentrated educational resources, which reflects the lack of equity in the allocation of educational resources.

*3.2. Accessibility Distribution of Educational Resources.* The accessibility of education calculated via improved G2SFCA is shown in Figure 2. On the whole, the distribution shows an obvious core-periphery typology where the accessibility was higher in the center and lower in the surrounding areas. High-value primary schools are mainly clustered in the central area including part of Xihu, Shangcheng, and Xiacheng (in the red circle), which is the inevitable result of high concentration of quality schools and dense road networks. Different from primary school accessibility, there is no obvious high-value cluster relating to secondary schools. Apart from the central group comprising Xihu, Shangcheng, and Xiacheng, there are two other groups showing the same pattern, namely, in Binjiang and Xiaoshan (in the green circle). Inside each group, the accessibility value shows a gradual decrease from the inside to the outside. The reason may be that the dense population had a relatively high demand in the inner city which narrows the gap among regions. In addition, the development stage of the city changed from the “Xihu Era” to the “Qiantang River Era” and the construction of infrastructure is comparable on each side.

When looking into the different characteristics of education accessibility, it was concluded that there are several factors that affected the access to educational services. School quality is the most decisive factor to education accessibility. Communities located in school districts with better reputation and wider recognition generally have higher accessibility, as shown in Figure 2. This relationship is evident and strongly proven in most areas including Xihu and Shangcheng. Quality schools are favored in the process of financial allocation to education and can provide higher level services consequently, which leads to the guaranteed accessibility over average level.

However, there is no absolute connection between the high quality of the school and the high accessibility of school district housing. First, the supply capacity of schools is determined by the quality of school, area, number of teachers, and number of students enrolled. Limited by land supply, campus area in the central city is significantly smaller than that in the suburbs. To control the quality of teaching, elite elementary schools often impose strict restrictions on the number of students enrolled, which weakens the superiority of schools in rankings. As a result, there are some schools topping the list while their demand points face high pressure for access to educational resources. Apart from that, since accessibility is decided by supply, demand, and the gap between them, the higher the demand for the school district, the lower the accessibility of each demand point. The places where residential communities are concentrated have a correspondingly greater demand for educational resources, making local schools incapable of meeting the oversaturated demand for educational services. With too many communities to serve, top schools failed to provide satisfactory amount of educational resources. On the contrary, less demand allows ordinary schools the opportunity to maximize the potential of educational resources.

Distance is not the key factor that affects ease of access to educational facilities since communities in the same district generally have the same level of accessibility. In fact, parents are willing to travel farther for better educational service.

Table 1: Socioeconomic indicators.

| Variable type | Variables | Explanation | Data source |
|---|---|---|---|
| Dependent variable location | Housing price | Data on housing prices | Listing data |
| | Central potential | The product of the distances to Wulin Square[1] and the nearest secondary commercial center | Google Earth |
| | Traffic potential | Distance to the nearest subway station | Google Earth |
| | Landscape potential | Distance to the nearest large-scale landscape resource | Google Earth |
| Architecture | House age | Year from the completion of house | Listing data |
| | Greening rate | The ratio of green area to the planned area of the community | Listing data |
| Accessibility | Accessibility to primary school | Quality of primary school per student | Calculated data |
| | Accessibility to secondary school | Quality of secondary school per student | Calculated data |
| Neighborhood | Lifestyle facilities | Distance to the nearest shopping mall or supermarket | Google Earth |
| | Medical facilities | Distance to the nearest hospital, clinic, or community healthcare center | Google Earth |
| | Financial facilities | Distance to the nearest bank branch (not including ATM) | Google Earth |
| | Leisure facilities | Distance to the nearest park, square, campus, or cultural and sports facilities | Google Earth |
| | F&B facilities | Distance to nearest restaurant | Google Earth |
| | Business facilities | Distance to the nearest commercial building | Google Earth |

[1]Wulin Square is recognized to be the city center.

Table 2: Descriptive statistics of variables.

| Variables | Unit | Maximum | Minimum | Mean | Median |
|---|---|---|---|---|---|
| Housing price | US dollar[2] | 21,033.67 | 1,432.46 | 5,300.65 | 5,100.51 |
| Central potential | km$^2$ | 1,377.68 | 0.34 | 39.84 | 23.48 |
| Traffic potential | m | 29,331.91 | 5.81 | 3,044.88 | 1,784.13 |
| Landscape potential | m | 8,626.91 | 22.31 | 2,188.66 | 1,814.43 |
| House age | Years old | 41.00 | 3.00 | 16.06 | 16.00 |
| Greening rate | % | 75.46 | 10.00 | 30.46 | 30.00 |
| Accessibility to primary school | — | 185.59 | 0.00 | 1.43 | 0.85 |
| Accessibility to secondary school | — | 6.80 | 0.00 | 0.72 | 0.54 |
| Lifestyle facilities | m | 2,145.22 | 3.11 | 278.97 | 218.45 |
| Medical facilities | m | 2,245.79 | 1.43 | 261.06 | 190.80 |
| Financial facilities | m | 3,150.78 | 7.66 | 359.87 | 244.95 |
| Housing price | m | 3,859.43 | 2.49 | 484.54 | 373.22 |
| Central potential | m | 1,372.29 | 1.24 | 134.41 | 107.80 |
| Traffic potential | m | 8,921.95 | 9.13 | 1,118.31 | 776.98 |

[2]According to IRS, yearly average exchange rates for converting Chinese yuan into US dollars in 2020 is RMB6.900 : US$1: https://www.irs.gov/individuals/international-taxpayers/yearly-average-currency-exchange-rates.

Figure 3 shows the mapping results and the geographic distribution of accessibility calculated by the classic G2SFCA method which does not coincide with reality. Zhuantang Subdistrict in Xihu District and the area along the river in Binjiang District (in the red circle) display the highest accessibility. However, these areas do not have abundant high-quality educational resources and the population density is relatively high, which means that the surrounding communities do not have the potential to enjoy access to high-quality education. The reason why education accessibility in the central city is underestimated is that the radius of school district in the central area is much smaller than 2,000 meters. When calculating with traditional methods, demand points that are out of the school district are also taken into consideration, which leads to nonexistent competition among irrelevant demand points. As a result, the corresponding education accessibility is unsatisfactory since the needs of residents are incorrectly identified. For the same reason, education accessibility in periphery areas is overestimated. Hence, it is important in the analysis of education accessibility to decide the number of demand points according to the scope of the school district. This will minimize the over- or underestimation of demand under current school district system in China.

*3.3. Spatial Heterogeneity of Education Accessibility.* In order to evaluate the spatial pattern between education accessibility and socioeconomic characteristics, we analyzed the mapped spatial distribution of five variables, namely, the presence of migrants, residents with rural *hukou* (China's *hukou* system is a family registration program that serves as a domestic passport, regulating population distribution and rural-to-urban migration;

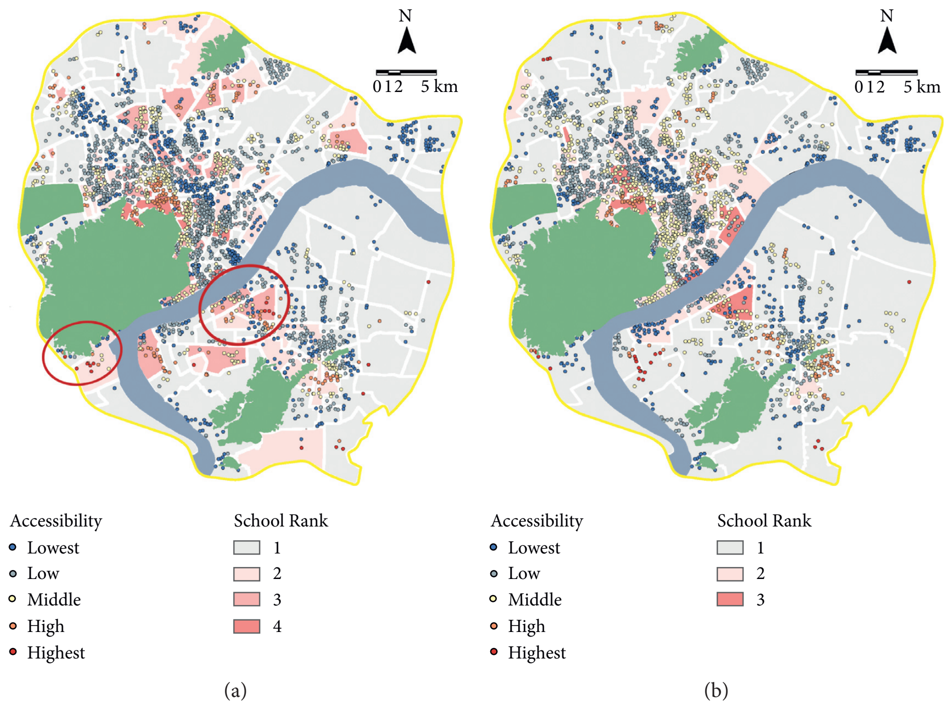



Figure 2: The distribution of education accessibility in Hangzhou with improved G2SFCA. (a) Primary school. (b) Secondary school.

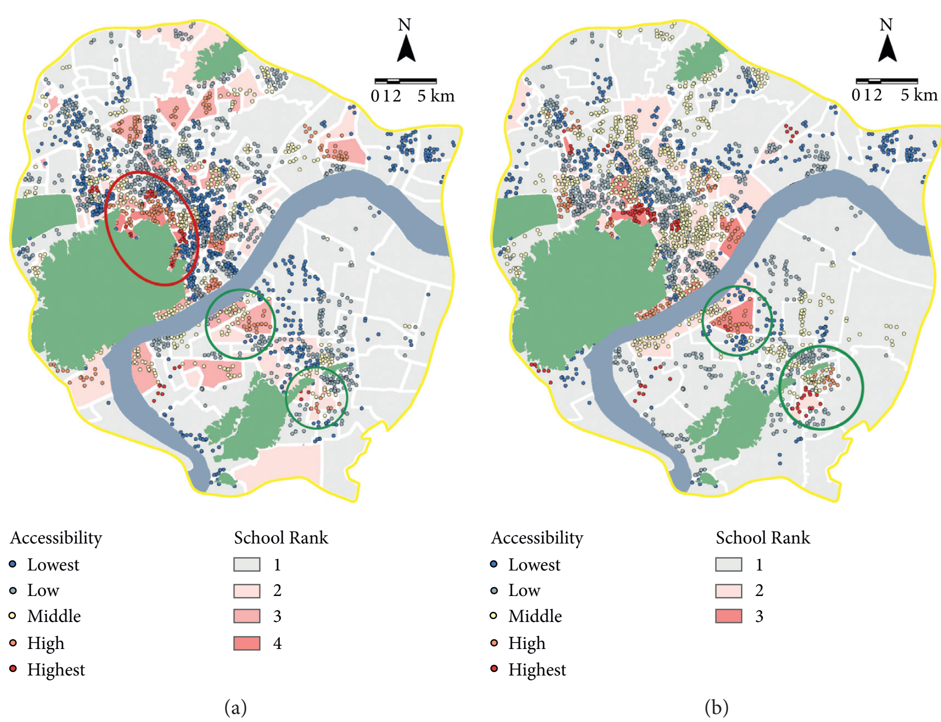



Figure 3: The distribution of education accessibility in Hangzhou with classic G2SFCA. (a) Primary school. (b) Secondary school.

residents with rural *hukou* are denied from the same rights and benefits enjoyed by urban residents), civil servants, professional technical personnel, and housing price (as an indicator for income). Figure 4 displays the spatial clustering of socioeconomic indicators and the distribution of education accessibility. According to the results of biLISA, the social spatial differentiation pattern of education accessibility is closely related to the characteristics of social groups. High percentages of different groups and high rates of education accessibility show a state of coupling or mismatch depending on the socioeconomic status of the group.

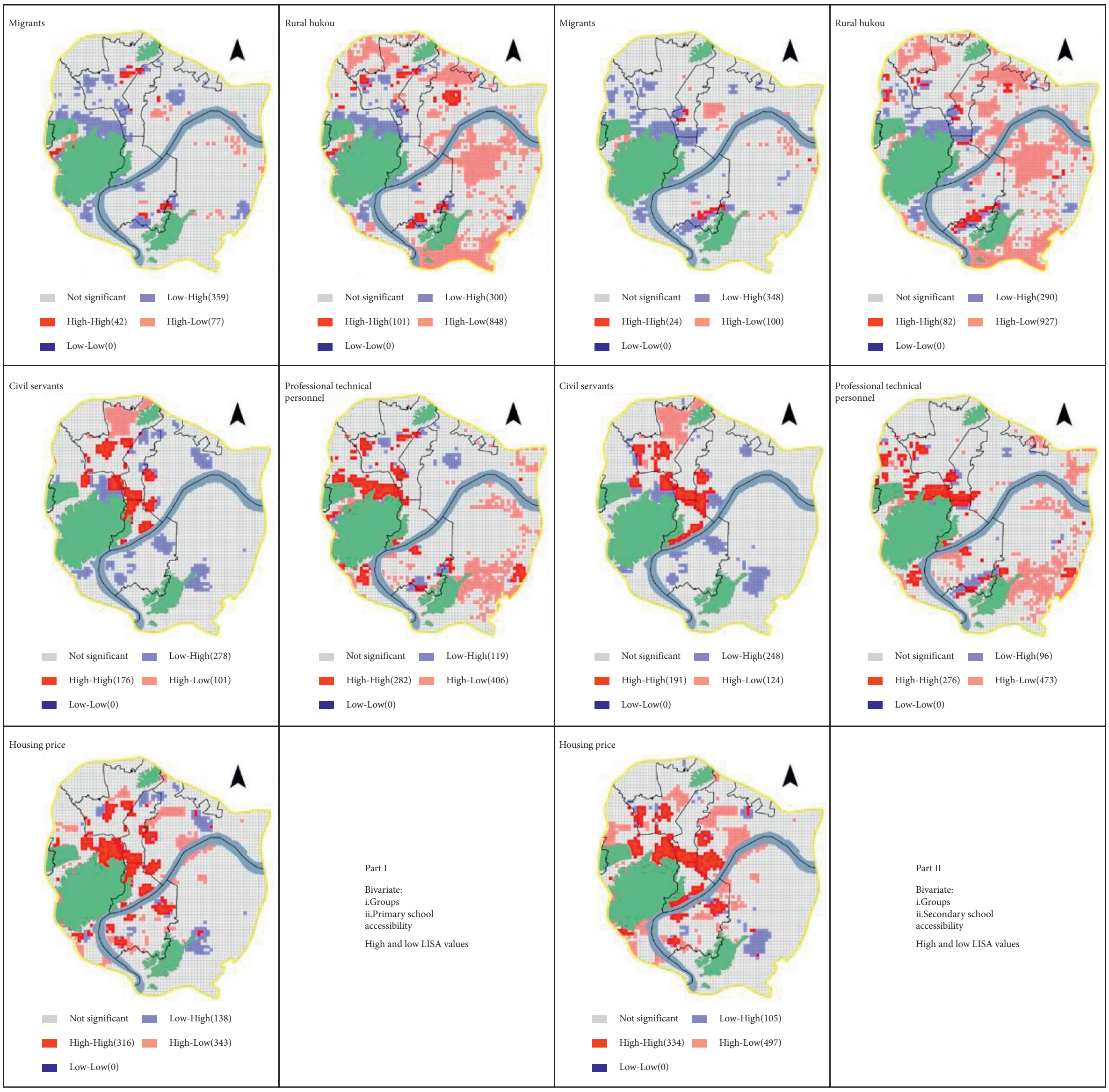



FIGURE 4: LISA high and low cluster of education accessibility among different groups.

Vulnerable groups such as migrants and residents with rural *hukou* lack sufficient political power to influence the allocation of primary and secondary schools or strong economic power to actively seek high-quality educational resources. Although the government has been carrying out construction of high-level teaching structures (such as the Three-Year Action Plan for Building a Better Education in Hangzhou), the beneficiaries are only limited to parents who can afford a house in the city center.

Out of consideration to reduce rental costs, migrants are more inclined to live in the suburban new city with convenient public transport [56]. Therefore, areas with a high presence of migrants are mainly located in Gongshu and the vicinity of the ring road. This phenomenon of spatial mismatch between educational resources and groups has resulted in rather few migrants who can meet their education needs, and the total number of migrants who cannot is far less than the migrants who can. Residents with rural *hukou* are more widely distributed in the periphery of the investigated area where education accessibility is lower. The concentration of rural *hukou* in the periphery of the city such as Gongshu and Xiaoshan is particularly obvious, where the area is vast and there are still a large number of rural and village committees except for the well-developed streets. However, compared to these places, city center is favoured for its high population density when the local government builds new educational facilities. As a result, migrant residents can hardly meet their needs for high-level education services.

An analysis done on education accessibility on socioeconomic status showed that the middle class has absolute upper hand in the competition for high-quality educational resources. This is specifically expressed by the high degree of spatial coupling among high SES groups, high-quality educational resources, and high-level education accessibility. Similar to the research results from studies conducted in

Sweden and France, groups with low socioeconomic status have a higher degree of inequality in educational opportunity [57].

In terms of occupation, areas with highly educated individuals, such as civil servants, teaching and research personnel, and professional technical personnel, are consistent with the range of high-quality school districts. A possible explanation for this outcome could be that there is still an invisible system that prioritizes quality educational resources for formerly respected and established occupations, namely, employees of government and institution, despite the dissolution of "unit system" under the influence of market economy. Under the situation where schools have limited supply capacity, the scope of school district was forced to shrink, and only nonemployees with strong financial capabilities are able to access these renowned schools.

As for income, housing price can partly represent income level since areas with high housing prices usually have higher concentration of high-income groups. According to the results from biLISA analysis, the distribution of education accessibility presents significant spatial heterogeneity among high- and low-income groups. Most areas with high housing prices and high levels of education accessibility are located in the central city with well-developed transportation systems and rich educational resources. The results suggest that high-income groups have completed the monopoly of high-quality educational resources through the filtering effect of the real estate market and housing purchases. Areas with less access to educational services and high housing price are scattered around the central area where landscape resources are rich, the potential explanation to which is that citizens who choose these areas are not attracted by teaching resources but other city services.

Unlike vulnerable groups, the middle class who enjoys high education accessibility has an overwhelming advantage over the disadvantaged middle class that does not. Under active parental choices, the middle class generally has good education accessibility and the stratification of middle class school districts has already occurred. High-quality educational resources guarantee that housing in the surrounding district enjoys above-average levels of education accessibility, which attracts high SES groups who are willing and capable to fight for better educational services for their children. The high SES groups will gather near renowned schools. This housing filtering mechanism forms education-driven residential spatial differentiation, which has a great impact on the social structure within the school district.

The distribution of access to secondary schools is heterogeneous and shows certain mismatch while education services failed to reach more groups. This can be explained by the changes in supply and demand during the process of calculation. Secondary school districts are larger, and the number of demand points in each district has doubled. Therefore, the pressure on each school to provide more quality education resources has increased, and the level of access enjoyed by each demand point has decreased significantly. At the same time, larger school district means families have to travel further to seek educational services, which leads to an increase in the average commuting distance from the community to the school, further weakening the school's service capacity. This downward trend is not obvious when accessibility of secondary schools is analyzed separately. However, when it is combined with the social characteristics of the group, the spatial differentiation is accurately captured. The matching attributes of biLISA analysis can not only describe the degree of spatial differentiation but also improve accuracy when identifying mismatches that are ignored by single-item analysis.

*3.4. Capitalization Effect of School Quality on Housing Price.* GWR is applied to explore the spatial heterogeneity of educational capitalization. Before GWR is conducted, the ordinary least squares (OLS) regression is needed to identify insignificant factors. Ultimately, there are eleven factors, i.e., accessibility to primary school, accessibility to secondary school, greening rate, central potential, traffic potential, landscape potential, lifestyle facilities, medical facilities, recreational facilities, F&B facilities, and business facilities. The adjusted $R^2$ of the model is 0.508, indicating that the socioeconomic factors chosen can partially explain housing prices. We focus on the impact of educational resources on housing prices and try to portray the capitalization effect of school quality on housing prices.

Figure 5 depicts the coefficient distribution of various urban services. The regression coefficients of primary school accessibility are mostly positive in the study area (Figure 5(a)), indicating that homebuyers are willing to pay a certain premium to live near a primary school, which is consist with Song's conclusion [58]. The coefficient distribution of accessibility to primary schools is evidently regular. The quality of primary schools has a strong and positive influence on housing prices. The regression coefficient manifested a radial distribution, with areas with better educational resources, namely, the junction of Shangcheng, Xiacheng, and Xihu (in red circle), displayed the highest value. Areas with poorer educational resources and lower housing prices, namely, center of Binjiang, Xiaoshan, and Jianggan and peripheral areas of Xihu and Gongshu, have become clusters with high negative values.

The average value of the regression coefficient of the accessibility to secondary school increased from 0.08 to 0.11, indicating that the impact of secondary school on housing prices is greater. The highest capitalization rate appears in the junction of Xihu and Gongshu (in red circle in Figure 5(b)) where there are high-quality secondary schools as well as many high-quality primary education resources, making it the center of high-quality "double school districts" in Hangzhou. Because of the scarcity of high-quality dual school districts, houses in dual school districts tend to be occupied for a long time and the turnover rate is low [21]. The quality of educational resources rather than community characteristics such as building age and community environment and its location becomes the key factor that influences housing prices. After the promulgation of the second-child policy, young parents are more inclined to choose dual-school-district housing under the consideration

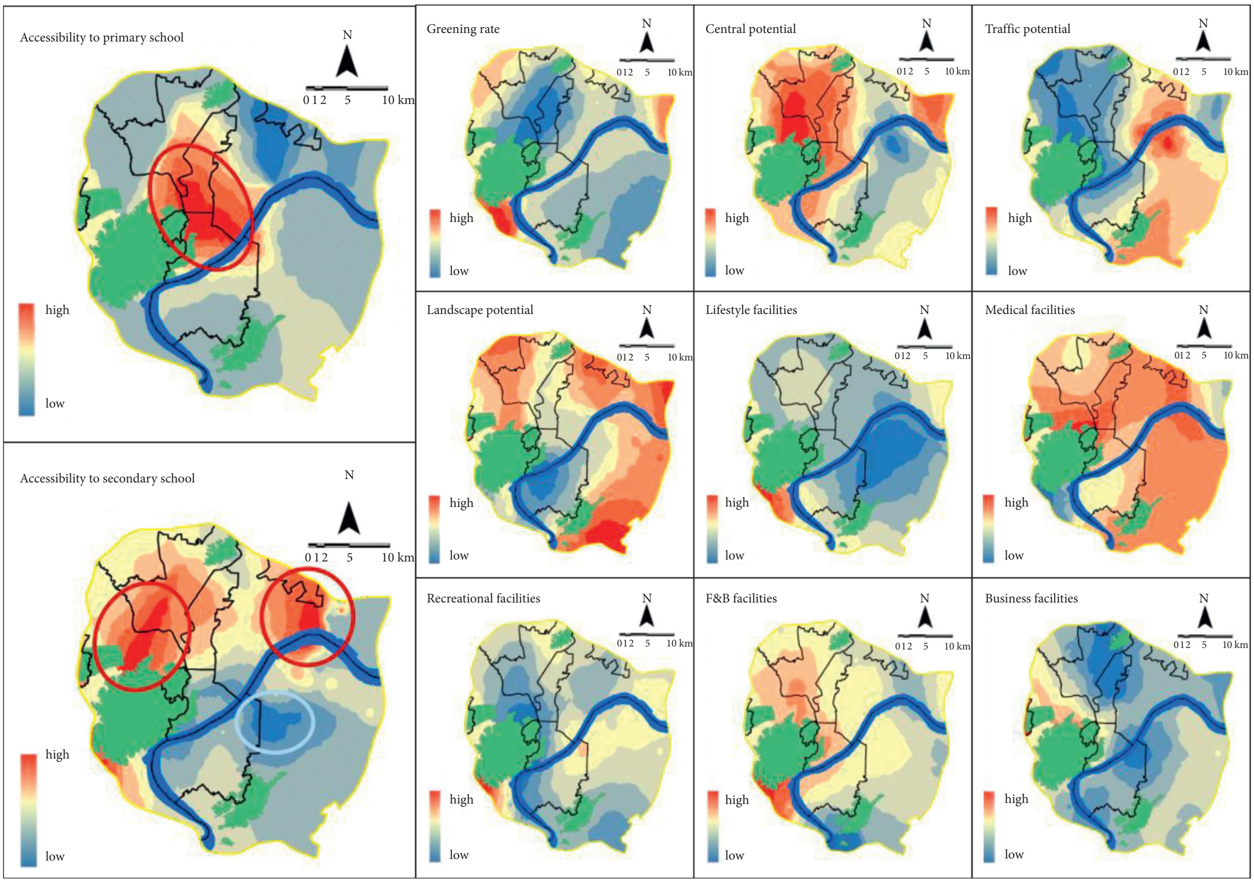



FIGURE 5: Spatial distribution of regression coefficient from the perspective of education accessibility based on the GWR model.

of the convenience in the primary-secondary examination, which leads to the increase of housing prices. Furthermore, the capitalization effect gradually decreased from the city center to the northwest and southeast regions, where relatively high-quality educational resources can be found. Areas with low regression coefficient are located along the Yangtze River and Xiaoshan District (in the blue circle). The housing prices in the relevant areas are slightly affected by secondary school accessibility and are more related to the housing conditions, the location of the community, and the abundance of surrounding supporting facilities, which indicates that the capitalization of educational resources is not thorough.

Given the differences in the natural attributes, social needs, and many other factors, compulsory primary and secondary education resources have significantly different impact on housing prices. Although secondary school has a stronger influence on housing prices in the investigated areas, the scope was reduced. In addition, the standard deviation of the regression coefficient of secondary school accessibility exceeds that of primary school accessibility, indicating that the difference in the impact of secondary school on housing prices in different regions has a trend of expanding. Unlike primary schools, high-quality secondary school access has a greater effect on the improvement of college entrance examination scores and is more likely to be sought after by parents. Therefore, the value-added effect of most ordinary secondary schools on housing prices is limited. However, the pursuit of housing with access to the secondary school superimposed with high-quality primary schools solidified the distribution pattern of the educational resources and expanded the degree of differentiation, which would result in the continuous increase of the positive impact of quality educational resources on housing prices.

### 3.5. Educational Inequality under the School District System.

The original intent of the "nearby schooling" policy is to eliminate the phenomenon of school selection by students' score, family background, and their social relationship under the circumstances that differences in educational resources exist objectively and are difficult to eliminate in the short term. Furthermore, the purpose is to weaken the impact of family background on the access to educational opportunities, thereby achieving educational equality. The approach of controlling school choice coincides with the studies on educational equality conducted in European and American countries [59].

However, the implementation of linking educational resources with market-oriented housing has enabled the school district system to create a bidding market for high-quality educational resources through "school district housing" transactions, indirectly raising the threshold of school choice. Powerful family background still guarantees

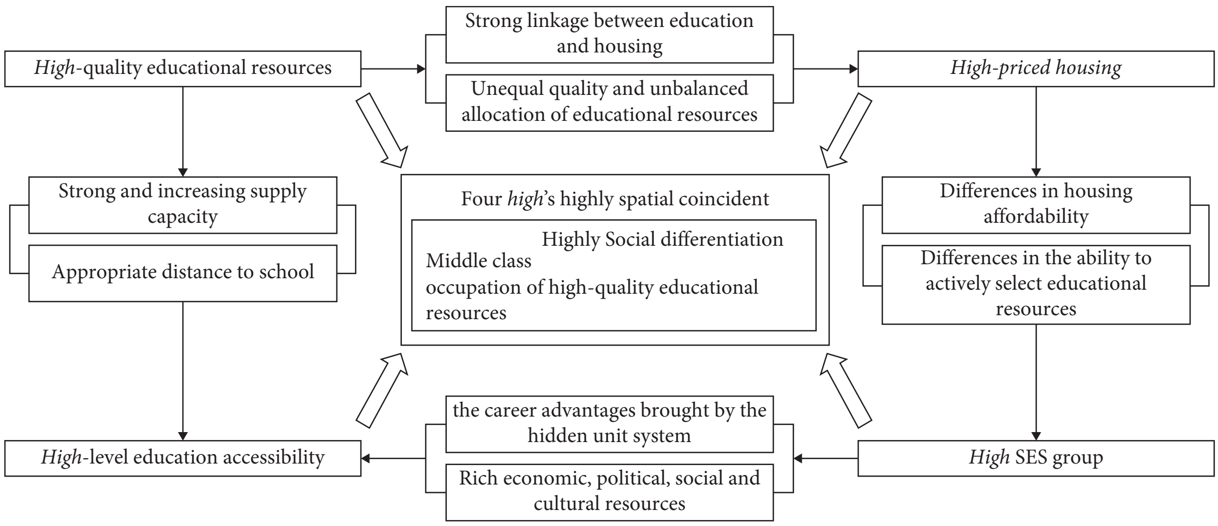



Figure 6: Social spatial differentiation and the mechanism of educational resource quality.

the residents an overwhelming advantage in obtaining quality educational resources with the existence of the giant gap among groups in their SES in China. In order to compete for scarce high-quality public educational resources for their children, the wealthy middle class is competing to purchase "school district housing" and thus forming clusters in high-quality school districts. This kind of gentrification triggered by the school-choosing behavior of the middle class as the core driving force guarantees the intergenerational transformation and transmission of economic capital, social capital, and cultural capital of wealthy families, but it may also cause social issues such as educational inequality and social spatial heterogeneity.

The mechanism of educational inequality and social mobility problems caused by differences in educational resources under the school district system is as follows (Figure 6): high-quality educational resources ensure the high-level education accessibility of school district housing and attract parents to obtain high-quality educational services by purchasing school district housing. However, because of uneven quality and unbalanced allocation of educational resources, parents' competition for school district housing directly raises housing prices near high-quality educational resources under the market economy system. Because of the filtering effect of the real estate market, groups with high SES can use endogenous attributes such as finances, culture, and social capital to pursue high-level education accessibility and high-quality educational resources, whereas low-income groups lose opportunities to enjoy quality education and gain subsequent ability to choose careers since they cannot afford high housing prices. The spatial expression of this screening mechanism is that high-quality educational resources, high-priced communities, clusters of high socioeconomic status groups, and communities enjoying high-level education accessibility are highly consistent. Societal performance is that the middle class has completed the occupation of high-quality educational resources, and the disadvantaged groups are confined to the margins and bottom of society, resulting in the increased differentiation of social classes.

School district housing under the "nearby schooling" policy is a direct result of the differences in urban educational resources. The competition for quality school district housing has evolved into a competition of political, economic, social, and cultural strength between families, which runs counter to the original policy intentions and the concept of educational equality. If left unchecked, scarce high-quality educational resources will continue to gather in urban centers and to the upper class. As a result, ordinary families will fall into the vicious cycle of being unable to obtain high-quality educational services and improve their ability to seek high-income jobs to improve their family's economic situation and obtain quality education services for the next generation, losing possible opportunities for upward mobility. In view of the deep effect of educational equality in alleviating social differentiation and promoting social mobility, it is necessary to rationally plan the scope of school districts, narrow the interschool gap, enhance the spatial and intergroup balance in the allocation of educational resources, and make every effort to achieve educational equality.

## 4. Conclusions

Based on the particularity of the compulsory education resource supply pattern under the school district system, this research proposes an improved Gaussian two-step floating catchment area method, which measures the difficulty of obtaining educational services and its social effects from both social and spatial aspects and draws the following conclusions:

(1) Different from the classic G2SFCA method, which uses the distance threshold $d_0$ as the standard for the service scope of the supply facilities, the improved G2SFCA method is more in line with China's "nearby schooling" policy which determines enrolment qualifications by the school district. Consequently, the improved G2SFCA method we propose

that incorporates school district effect into traditional models can provide a more realistic appraisal of accessibility.

(2) In developed coastal areas, the distribution of education accessibility shows an obvious core-periphery typology, and the aggregation effect of primary school accessibility is more significant than that of secondary schools. In terms of the factors that are decisive to education accessibility, school level plays the most important role in determining the level of regional accessibility. High-quality school districts allow the community to enjoy above-average educational services. However, given that high-quality schools are usually accompanied with limited number of student enrolment and competition subsequently gets more intense, there is no doubt that high quality does not guarantee high accessibility.

(3) In the process of rapid urbanization, compared with groups with high socioeconomic status, vulnerable groups in inflow areas are highly disadvantaged in terms of access to educational services. Lacking sufficient political power and strong economic capability prevents migrants and residents with rural *hukou* from enjoying quality education, whereas groups with high SES gather in the school district through housing purchases and complete their monopoly on quality educational resources, indicating that the phenomenon of stratification in school districts has occurred.

(4) The accessibility to primary schools and secondary schools has significant impacts on housing prices where the commercial housing market is mushrooming. Furthermore, spatial heterogeneity exists in educational capitalization. The highest regression coefficients appear in areas with rich high-quality educational resources. As for secondary school accessibility, areas with high values of the regression coefficients are located in the scarce dual school districts, suggesting that parents are willing to pay high housing prices for good schools and residents in different regions may have disparate preferences for quality education. Fully understanding the complex relationship between education resources and housing prices helps the government to tailor education policies to local conditions.

(5) High-quality educational resources, high-priced communities, clusters of high socioeconomic status groups, and communities enjoying high-level education accessibility are highly consistent spatially, which is the spatial expression of educational inequality. The crucial means to promote educational equality and social mobility relies on reducing the differences in educational quality between schools and promoting reform of the school district system.

The education system plays an important role in promoting social equality. A good education system can make up for disadvantages in individuals' family background and ability endowments. It will in turn allow the rational allocation of educational resources and alleviate educational inequality caused by differences in economic and social status [60]. It turns out that countries where socioeconomic inequality between families is similar may have varying degrees of social disparities in educational accessibility. Wide gaps in income inequality exist both in New Zealand and Japan, but there is much greater educational equality in the former than the latter [61]. The same is true for Belgium and Finland, where income inequality is moderate but educational inequality varies dramatically [62]. Therefore, designing and implementing a rational education policy is the key to promoting educational equity.

The "nearby schooling" policy is an active exploration of Chinese governments. It restricts school choice and weakens the impact of family economic status on the acquisition of educational opportunities. Still, we have to admit that although nonselective schools is the common ground shared by more equal countries (such as East Asian and Nordic countries) [62], the unignorable differences in educational resources and the rapid development of the commercial housing market in China have limited this policy and even had a negative impact. Because of the intensification of the educational resources differentiation and social spatial heterogeneity, the school district system failed to reach its goal in creating a spatial balance of educational resources and equality between groups. The research results supplement the mechanism of educational equity under special circumstances, namely, the school district system, and further demonstrate the importance of education policies that promote educational equality.

It is hardly possible to achieve the equal utilization of high-quality educational resources in cities overnight, and educational equality requires long-term efforts of the government, education departments, and academia. The next step in the research on the allocation of urban educational resources should focus on how to adopt feasible policy measures and implementation paths to promote the optimization, restoration, reform, and innovation of the school district system, continuously reduce the difference in the quality of education between regions and schools, and achieve spatial balance of education quality and the social equity of educational opportunities to the maximum. Measures such as group school-running, strong schools leading weak schools, grand school district system (multi-school planning), interschool mobility of excellent teachers, and tilting of education funds towards weaker schools should be actively explored to promote the full development of urban education so that the goal of social equity and spatial justice can be realized one day.

## Data Availability

Data for this work are available upon request from the corresponding author.

## Conflicts of Interest

The authors declare that there are no conflicts of interest regarding the publication of this paper.

## Acknowledgments

This work was supported by the National Natural Science Foundation of China (41771184).

## References

[1] J. Golley and S. T. Kong, "Inequality in intergenerational mobility of education in China," *China and World Economy*, vol. 21, no. 2, pp. 15–37, 2013.

[2] Y. Chen and S. Feng, "Access to public schools and the education of migrant children in China," *China Economic Review*, vol. 26, pp. 75–88, 2013.

[3] C. Harmon, H. Oosterbeek, and I. Walker, "The returns to education: microeconomics," *Journal of Economic Surveys*, vol. 17, no. 2, pp. 115–156, 2003.

[4] I. Gordon and V. Monastiriotis, "Education, location, education: a spatial analysis of English secondary school public examination results," *Urban Studies*, vol. 44, no. 7, pp. 1203–1228, 2007.

[5] Q. Wu, X. Zhang, and P. Waley, "When Neil Smith met Pierre Bourdieu in Nanjing, China: bringing cultural capital into rent gap theory," *Housing Studies*, vol. 32, no. 5, pp. 659–677, 2017.

[6] S. R. Lucas, "Effectively maintained inequality: education transitions, track mobility, and social background effects," *American Journal of Sociology*, vol. 106, no. 6, pp. 1642–1690, 2001.

[7] W. Boterman, S. Musterd, C. Pacchi, and C. Ranci, "School segregation in contemporary cities: socio-spatial dynamics, institutional context and urban outcomes," *Urban Studies*, vol. 56, no. 15, pp. 3055–3073, 2019.

[8] M. Oberti, "Social and school differentiation in urban space: inequalities and local configurations," *Environment and Planning A: Economy and Space*, vol. 39, no. 1, pp. 208–227, 2007.

[9] W. G. Hansen, "How accessibility shapes land use," *Journal of the American Institute of Planners*, vol. 25, no. 2, pp. 73–76, 1959.

[10] M. Oloko-oba, O. Akintunde, A. T. Alaga et al., "A geospatial approach to evaluation of accessibility to government primary schools in Ilorin west local government area, Kwara state, Nigeria," *European International Journal of Science and Technology*, vol. 4, pp. 96–107, 2015.

[11] S. Yin, X. Hu, Z. Ma, and W. Song, "Spatial and temporal effects of housing price in urban school districts based on education equity: a case study of public pimary school in the main urban area of nanjing," *Economic Geography*, vol. 39, no. 9, pp. 82–93, 2019.

[12] S. Williams and F. Wang, "Disparities in accessibility of public high schools, in metropolitan Baton Rouge, Louisiana 1990–2010," *Urban Geography*, vol. 35, no. 7, pp. 1066–1083, 2014.

[13] N. Page, M. Langford, and G. Higgs, "Measuring spatial accessibility to services within indices of multiple deprivation: implications of applying an enhanced two-step floating catchment area (E2SFCA) approach," *Applied Spatial Analysis and Policy*, vol. 12, no. 2, pp. 321–348, 2019.

[14] T. Pengfei, L. Jing, and G. Chen, "Spatial accessibility analysis of primary schools at the county level based on the improved potential model: a case study of Xiantao city, Hubei province," *Progress in Geography*, vol. 36, no. 6, pp. 697–708, 2017.

[15] H.-M. Kim and M.-P. Kwan, "Space-time accessibility measures: a geocomputational algorithm with a focus on the feasible opportunity set and possible activity duration," *Journal of Geographical Systems*, vol. 5, no. 1, pp. 71–91, 2003.

[16] R. Wiśniewski, M. Stępniak, and B. Szejgiec-Kolenda, "Accessibility of public services in the age of ageing and shrinking population: are regions following trends," *Geografiska Annaler: Series B, Human Geography*, vol. 103, no. 1, pp. 55–74, 2021.

[17] E. Talen, "School, community, and spatial equity: an empirical investigation of access to elementary schools in West Virginia," *Annals of the Association of American Geographers*, vol. 91, no. 3, pp. 465–486, 2001.

[18] H. Wen, Y. Xiao, E. C. M. Hui, and L. Zhang, "Education quality, accessibility, and housing price: does spatial heterogeneity exist in education capitalization?" *Habitat International*, vol. 78, pp. 68–82, 2018.

[19] A. I. Moreno-Monroy, R. Lovelace, and F. R. Ramos, "Public transport and school location impacts on educational inequalities: insights from São Paulo," *Journal of Transport Geography*, vol. 67, pp. 110–118, 2018.

[20] Y. Jiang, B. Li, and R. Zhao, "An allocation scheme of students based on IPM under the nearby enrollment policy," in *Proceedings of the 2019 10th International Conference on Information Technology in Medicine and Education (ITME)*, pp. 260–264, Qingdao, China, 2019.

[21] Y. Xu, W. Song, and C. Liu, "Social-spatial accessibility to urban educational resources under the school district system: a case study of public primary schools in nanjing, China," *Sustainability*, vol. 10, no. 7, p. 2305, 2018.

[22] L. Stiefel, R. Berne, P. Iatarola, and N. Fruchter, "High school size: effects on budgets and performance in New York city," *Educational Evaluation and Policy Analysis*, vol. 22, no. 1, pp. 27–39, 2000.

[23] J. Hennebry, J. McLaughlin, and K. Preibisch, "Out of the loop: (In) access to health care for migrant workers in Canada," *Journal of International Migration and Integration*, vol. 17, 2015.

[24] K. R. Ihlanfeldt and D. L. Sjoquist, "Job accessibility and racial differences in youth employment rates," *The American Economic Review*, vol. 80, no. 1, pp. 267–276, 1990.

[25] S. Farber, M. Z. Morang, and M. J. Widener, "Temporal variability in transit-based accessibility to supermarkets," *Applied Geography*, vol. 53, pp. 149–159, 2014.

[26] J. Merrell, F. Kinsella, F. Murphy, S. Philpin, and A. Ali, "Accessibility and equity of health and social care services: exploring the views and experiences of Bangladeshi carers in South Wales, UK," *Health and Social Care in the Community*, vol. 14, no. 3, pp. 197–205, 2006.

[27] J. Wolch, J. Wilson, and J. Fehrenbach, "Parks and park funding in Los Angeles: an equity-mapping analysis," *Urban Geography*, vol. 26, no. 1, pp. 4–35, 2013.

[28] R. M. Fernandez, "Race, space, and job accessibility: evidence from a plant relocation," *Economic Geography*, vol. 70, no. 4, pp. 390–416, 1994.

[29] S. Pinch, *Cities and Services*, Routledge & Kegan Paul, London, UK, 1985.

[30] L. K. Snow, "Economic development breaks the mold: community-building, place-targeting, and empowerment zones," *Economic Development Quarterly*, vol. 9, no. 2, pp. 185–198, 1995.

[31] J. Kuyvenhoven and W. R. Boterman, "Neighbourhood and school effects on educational inequalities in the transition from primary to secondary education in Amsterdam," *Urban Studies*, Article ID 0042098020959011, 2020.

[32] A. L. Mabogunje, *The Development Process: A Spatial Perspective*, Holmes & Meier, London, UK, 1980.
[33] W. Hu and R. Wang, "Segregation in urban education: evidence from public schools in Shanghai, China," *Cities*, vol. 87, pp. 106–113, 2019.
[34] J. R. Logan, E. Minca, and S. Adar, "The geography of inequality," *Sociology of Education*, vol. 85, no. 3, pp. 287–301, 2012.
[35] D. Satz, "Unequal chances: race, class and schooling," *Theory and Research in Education*, vol. 10, no. 2, pp. 155–170, 2012.
[36] Q. Wu, X. Zhang, and P. Waley, "Jiaoyufication: when gentrification goes to school in the Chinese inner city," *Urban Studies*, vol. 53, no. 16, pp. 3510–3526, 2016.
[37] J. Petrie, J. Jerry, and H. Craig, *Consolidation of Schools and Districts: What the Research Says and What it Means*, National Education Policy Center, Boulder, CO, USA, 2011.
[38] S. Yin, Z. Ma, W. Song, and C. Liu, "Spatial justice of a Chinese metropolis: a perspective on housing price-to-income ratios in nanjing, China," *Sustainability*, vol. 11, no. 6, p. 1808, 2019.
[39] C. M. Tiebout, "A pure theory of local expenditures," *Journal of Political Economy*, vol. 64, no. 5, pp. 416–424, 1956.
[40] C. Liu and W. Song, "Perspectives of socio-spatial differentiation from soaring housing prices: a case study in nanjing, China," *Sustainability*, vol. 11, no. 9, p. 2627, 2019.
[41] D. Leech and E. Campos, "Is comprehensive education really free?: a case-study of the effects of secondary school admissions policies on house prices in one local area," *Journal of the Royal Statistical Society: Series A*, vol. 166, no. 1, pp. 135–154, 2003.
[42] K. Burger, "The socio-spatial dimension of educational inequality: a comparative European analysis," *Studies In Educational Evaluation*, vol. 62, pp. 171–186, 2019.
[43] Y. S. Lee, "School districting and the origins of residential land price inequality," *Journal of Housing Economics*, vol. 28, pp. 1–17, 2015.
[44] S. Machin and K. G. Salvanes, "Valuing school quality via a school choice reform," *The Scandinavian Journal of Economics*, vol. 118, no. 1, pp. 3–24, 2016.
[45] S. Zheng, W. Hu, and R. Wang, "How much is a good school worth in Beijing? Identifying price premium with paired resale and rental data," *The Journal of Real Estate Finance and Economics*, vol. 53, no. 2, pp. 184–199, 2016.
[46] H. Wen, X. Bu, and Z. Qin, "Spatial effect of lake landscape on housing price: a case study of the West Lake in Hangzhou, China," *Habitat International*, vol. 44, pp. 31–40, 2014.
[47] D. M. Brasington and D. Hite, "School choice and perceived school quality," *Economics Letters*, vol. 116, no. 3, pp. 451–453, 2012.
[48] R. Jacobsen, J. W. Snyder, and A. Saultz, "Informing or shaping public opinion? The influence of school accountability data format on public perceptions of school quality," *American Journal of Education*, vol. 121, no. 1, pp. 1–27, 2014.
[49] M.-P. Kwan, "Space-time and integral measures of individual accessibility: a comparative analysis using a point-based framework," *Geographical Analysis*, vol. 30, no. 3, pp. 191–216, 1998.
[50] X. Zhao, Q. Zheng, X. Liu, and M. Jin, "Study on the planning and allocation of urban park green space based on 2SFCA improved model-A case study of Futian district, Shenzhen," *Chinese Landscape Architecture*, vol. 34, no. S2, pp. 95–99, 2018.
[51] Notice on Carrying out the Supervision and Evaluation Work of Modern Schools in Zhejiang Province, http://jyt.zj.gov.cn/art/2020/4/14/art_1532973_42574733.html.
[52] J. S. Coleman, "Equality of educational opportunity," *Equity & Excellence in Education*, vol. 6, no. 5, pp. 19–28, 1968.
[53] L. Anselin, "Local indicators of spatial association—LISA," *Geographical Analysis*, vol. 27, no. 2, pp. 93–115, 1995.
[54] Y. Xiao, Z. Wang, Z. Li, and Z. Tang, "An assessment of urban park access in Shanghai-implications for the social equity in urban China," *Landscape and Urban Planning*, vol. 157, pp. 383–393, 2017.
[55] S. NBo, *China City Statistical Yearbook*, China Statistics Press, Beijing, China, 2019.
[56] H. Qinshi, W. Song, L. Liu et al., "Exploring residential heterogeneity through multiscalar lens: a case study of Hangzhou, China," *Complexity*, vol. 2021, Article ID 3798183, 12 pages, 2021.
[57] K. Yang Hansen and J.-E. Gustafsson, "Identifying the key source of deteriorating educational equity in Sweden between 1998 and 2014," *International Journal of Educational Research*, vol. 93, pp. 79–90, 2019.
[58] W. Song, N. Mao, P. Chen, Y. Yuan, and Y. Wang, "Coupling mechanism and spatial-temporal pattern of residential differentiation from the perspective of housing prices: a case study of Nanjing," *Dili Xuebao/Acta Geographica Sinica*, vol. 72, pp. 589–602, 2017.
[59] J. Östh, E. Andersson, and B. Malmberg, "School choice and increasing performance difference: a counterfactual approach," *Urban Studies*, vol. 50, no. 2, pp. 407–425, 2013.
[60] M. Liao and S. Hua, "Educational inequality analysis: international comparison," *International Journal of Business and Social Science*, vol. 2, no. 16, pp. 88–93, 2011.
[61] R. Nash, "Inequality/difference in New Zealand education: social reproduction and the cognitive habitus," *International Studies in Sociology of Education*, vol. 13, no. 2, pp. 171–194, 2003.
[62] M. Duru-Bellat, "Social inequality and schooling," in *International Encyclopedia of the Social & Behavioral Sciences*, J. D. Wright, Ed., pp. 325–330, Elsevier, Oxford, UK, 2nd edition, 2015.

Hindawi
Complexity
Volume 2021, Article ID 7199044, 12 pages
https://doi.org/10.1155/2021/7199044

*Research Article*

# The Spatial Agglomeration and Industrial Network of Strategic Emerging Industries and Their Impact on Urban Growth in Mainland China

**Jiaming Li, Peiyuan Zhang, Jialin He, and Dongqi Sun**

*Institute of Geographic Sciences and Natural Resources Research (CAS), 11A, Datun Road, Chaoyang District, Beijing 100101, China*

Correspondence should be addressed to Dongqi Sun; sundq@igsnrr.ac.cn

Received 3 May 2021; Accepted 8 August 2021; Published 31 August 2021

Academic Editor: Jing-Hu Pan



Strategic emerging industries (SEIs) are an important industrial policy to promote innovation, develop advanced manufacturing, and upgrade the economy in China. The research explores the impact of SEIs on urban economic growth in mainland China, from agglomeration externalities and network externalities. The results show that SEIs have a significant impact on growth, and network externalities are generally more important than agglomeration externalities. Although the agglomeration of large-scale listed enterprises in the city promotes growth, too many large enterprises in the cluster, that is, the lack of small and medium-sized supporting enterprises, will harm urban growth. Meanwhile, a city with a higher degree of centrality has a larger GDP, but the strength has a negative impact.

## 1. Introduction

Strategic emerging industries (SEIs) are those key industries that the Chinese government considers to have an important impact on global economic competition. Those, specifically, include 9 industries: new generation of information technology industry, high-end equipment manufacturing industry, new material industry, bioindustry, new energy automobile industry, new energy industry, energy conservation and environmental protection industry, digital creative industry, and related service industry. Since they were proposed in 2010, SEIs have received great attention from the government. "The Fourteenth Five-Year Plan for National Economic and Social Development of the People's Republic of China and the Outline of Long-Term Goals for 2035," a key document for China's development in the next five years, regards strategic emerging industries as an important part of China's modern industrial system and high-quality economic development.

In addition to the impact on the future development of China, more importantly, the dual characteristics of these industries are worthy of attention. On the one hand, 7 out of 9 industries belong to manufacturing, so the agglomeration economy, that promotes the rapid growth of manufacturing, has an impact on the outlay and development of SEIs. On the other hand, the importance of R&D and information to the development of the industry makes these industries show a certain degree of characteristics of the network. The industrial network is usually a characteristic of the high-end service industry. For SEIs, are agglomeration economies and network externalities equally important? Or which one is more important? Actually, many advanced manufacturing industries are important; they are also affected by both the agglomeration economy and network externalities. This research has reference significance for thinking about the development of advanced manufacturing industries.

*1.1. Literature Review.* Agglomeration economy and network externalities are from two different theoretical frameworks. The former focuses on the mutual influence between enterprises within the region, while the latter pays more attention to the influence of the outside on the region.

The benefits of local development from agglomeration economies mainly come from three aspects: close upstream and downstream industrial connections, knowledge and information spillover, and large-scale labor market pool [1–3]. To a large extent, Marshall believed that agglomeration economies were formed by specialized division of labor by different enterprises in the same industry. But in reality, the proximity of plants across industries also often occurs. Namely, those plants from different industries also benefit from being close to each other. Jacobs [4] argued the spatial agglomeration of diverse industries and people helps the emergence of new ideas, thereby promoting the development of the local economy. Ellison et al. [5] also developed coagglomeration indices and found some highly coagglomerated industry pairs in the United States. Studies have shown that coagglomeration of multiple industries is conducive to the development of emerging industries and high-tech industries, while for traditional industries, it is mostly the agglomeration of enterprises in the same industry [6, 7]. Industry clusters, formed by the agglomeration of related industries, are important for regional growth [8]. Not only are the industry clusters collections of enterprises and even public sectors, but they also emphasize the cooperation network formed between enterprises as an important source of competitive advantage.

Compared with the cooperation network based on spatial proximity in industrial clusters, network externality emphasizes the benefits from the information or financial connections outside the region or city. Therefore, network externality highlights the impact of nonlocalization. Rosenthal and Strange [9] also argued the effect of economic activities outside a certain area on the economic activities within that area should not be ignored. Certainly, the study of industrial agglomeration is not completely ignored by external influences. For example, the concept of market potential is introduced to measure the influence of other cities on one city. However, the calculation of market potential is still based on the distance between cities. So it is still not beyond spatial proximity and distance decay.

With the rise of globalization and flow space, interconnected agglomerations and cities have become the engine of economic growth [10, 11]. Based on the Sassen [12] and Castells [13] elaboration of the global cities system, Taylor and Derudder [14] developed the interlocking network model (INM) to measure network externality. Urban network externalities are mostly from two aspects: function complementarities between cities and borrowed size. The former elaborates that functional networks between cities could create synergies and complementarities [15, 16]. The latter showed the benefits that small and medium-sized cities get from large cities. When Alonso [17] first introduced the concept of borrowed size, it only explained that the agglomeration benefits that small and medium-sized cities acquired due to their proximity to large cities. Meijers and Burger [18] recently argued that the concept can be easily generalized to network connectivity between large and small cities. It means small and medium cities can gain benefits by being well-positioned in both geographic space and urban networks.

The researches of two strands of agglomeration externalities and network externalities are separated and there are few comparatively empirical studies. So it is unknown about the relative importance of agglomeration externalities vis-à-vis network externalities. The same is true for SEIs and other industries in China. Some studies have shown that strategic emerging industries are mainly concentrated in the economically developed eastern part of China, and their distribution centers are shifting southward [19, 20]. Zhao [21] showed the producer network in China. One of the purposes of the research is to answer which externalities are more important for ESIs and urban growth.

## 2. Data and Methodology

*2.1. Data.* The dataset of this study includes 1069 listed enterprises from Shanghai Stock Exchange (SHSE) and Shenzhen Stock Exchange (SZSE) in strategic emerging industries and their associated 18713 enterprises. China Securities Index Co., Ltd., and Shanghai Stock Exchange released 1117 sample enterprises included in Strategic Emerging Composite Index in 2017. We collected 1069 in the 1117 sample enterprises. Furthermore, we obtained the branches of these listed enterprises and the affiliated enterprises they invest in, by cooperating with Shanghai HeHe Information Technology Co., Ltd. Listed enterprises generally have strong strengths and can be regarded as core enterprises, leading the development of strategic emerging industries in China, while the branches and the affiliated enterprises can be taken as supporting ones. There are 1,900 sample enterprises in the research. Besides, we also obtained data such as GDP, employees, and fixed asset investment from the statistical yearbook.

*2.2. Methodology.* In general, this study did not use complex analysis methods. The analysis of industrial agglomeration is mainly based on the number of urban enterprises, while for network externality, three basic indicators, degree centrality, strength, and average degree of the neighbors, are measured [22].

*2.2.1. Degree Centrality*

$$K_i = \sum_j a_{ij}, \quad j \in n,\ j \neq i, \tag{1}$$

where the sum runs over the set $n$ of neighbors of $i$. If there are links between city $i$ and city $j$, then $a_{ij} = 1$; otherwise, $a_{ij} = 0$.

*2.2.2. Strength.* The strength of a node integrates the information about both its connectivity and the importance of the weights of its links.

$$S_i = \sum_j a_{ij} * \omega_{ij}, \quad j \in n,\ j \neq i, \tag{2}$$

where $a_{ij}$ is the same as above and $\omega_{ij}$ is the weight, which is the number of enterprise connections between city $i$ and city $j$.

*2.2.3. The Average Degree of the Neighbors.* To identify the degree correlation of industry network, we need to measure the average degree of the neighbors of each city. The average degree of the neighbors is identified:

$$N_i = \frac{1}{k_i} \sum_j a_{ij} * k_j, \quad j \in n,\ j \neq i,$$
$$N_i^{\omega} = \frac{1}{k_i} \sum_j a_{ij} * \omega_{ij} * k_j, \quad j \in n,\ j \neq i. \tag{3}$$

If high-degree cities tend to link with those cities with high-degree centrality, this tendency is referred to as an assortative network. Otherwise, tendency of high-degree/low-degree cities and that of low-degree/high-degree neighbors to connect are referred to as disassortative ones [23].

## 3. The Spatial Agglomeration Characteristics and Cluster Construction of China's SEIs

Nearly 20,000 core and supporting enterprises are located in 329 prefecture-level and above cities and 24 counties directly administrated by the province in China. Among China's 333 prefecture-level administrative units, only 9 cities in a few provinces such as Tibet, Qinghai, and Xinjiang have no supporting enterprises. Specifically, 1069 core enterprises are distributed in 141 prefecture-level and above cities. Except for the 4 municipalities in China, more than 40% of the cities have core enterprises in SEIs; 18,843 supporting enterprises are distributed in 352 cities. Only 19 of the 293 prefecture-level cities have no supporting enterprises. Overall, SEIs are widely distributed in Chinese cities.

*3.1. Agglomeration and Distribution of Core Enterprises in SEIs.* Macroscopically, most cities in the eastern coastal region have core enterprises of SEIs. In the central and northeastern regions, they are mainly distributed in central cities and surrounding areas, while in the northwest, they are scattered in some cities such as Lanzhou and Urumqi. The spatial distribution of SEIs' core enterprises in the eastern, central, and western regions is uneven.

From the perspective of regional distribution and agglomeration, core enterprises are mainly concentrated in the Yangtze River Delta, Beijing-Tianjin-Hebei, Pearl River Delta, and Chengdu-Chongqing regions (Figure 1). Among them, core enterprises have the highest degree of agglomeration in the Yangtze River Delta. The number of core enterprises in the Yangtze River Delta has reached 308, accounting for 28.81% of the 1,069 core enterprises. In this region, there are 8 cities with more than 10 core enterprises. The number of core enterprises in the Beijing-Tianjin-Hebei region and the Pearl River Delta region is about 200 (209 and 193, respectively). However, the characteristics of spatial distribution are significantly different. In the Beijing-Tianjin-Hebei region, the distribution of core enterprises is very concentrated. More than 86% of core enterprises (180) are located in Beijing, while the number of core enterprises in Tianjin is 15, and in other cities, there are only 1–2. The distribution of core enterprises in the Pearl River Delta is relatively even. There are more than 5 core enterprises in 7 cities, and 4 of them have more than 10 core enterprises. There are 48 core enterprises in the Chengdu-Chongqing region, mainly located in Chengdu and Chongqing (30 and 11, respectively).

Although core enterprises in SEIs are widely distributed in China's 142 cities, core enterprises are highly concentrated in a few cities and their municipal districts. More than 40% of the 142 cities have only one core enterprise (59 cities, accounting for 41.55%). And more than 50% of the core enterprises are concentrated in 8 cities including Beijing and Shenzhen (Table 1). Even in these big cities, the majority of core enterprises are concentrated in a few municipal districts. The agglomeration degree of core enterprises in SEIs at the county (district) level is significantly higher than that at the city level. Among the 2,846 county-level administrative units, only 387 have core enterprises, accounting for only 13.57%. And there is only one in 223 counties (districts and county-level cities). The top five districts with the total number of core enterprises are Haidian District in Beijing, Nanshan District in Shenzhen, Bao'an District in Shenzhen, Chaoyang District in Beijing, and Pudong New District in Shanghai (the core enterprises are 85, 42, 34, 33, and 22, respectively). These top five municipal districts have gathered close to one-fifth of the country's core enterprises in SEIs (19.30%).

*3.2. Agglomeration and Distribution of Supporting Enterprises in SEIs.* Beijing-Tianjin-Hebei region, Yangtze River Delta region, Pearl River Delta region, and Chengdu-Chongqing region are still the regions with the largest number of supporting enterprises (Figure 2). In the Yangtze River Delta and the Pearl River Delta, the number of supporting enterprises in most cities is relatively large, while the distribution of core enterprises shows the characteristics of the core-periphery in the Beijing-Tianjin-Hebei and Chengdu-Chongqing regions. In the latter two regions, most supporting enterprises are concentrated in regional central cities, while the number of small and medium-sized cities is significantly smaller. In most of the central and western provinces, supporting enterprises of SEIs are also mainly concentrated in provincial capitals.

The cities with more than 1,000 supporting enterprises are still Beijing, Shanghai, and Shenzhen. The total number of supporting enterprises in the three cities is close to 5,000 (4711), accounting for 1/4 of all supporting enterprises. There are 37 cities with more than 100 supporting enterprises. The total number of supporting enterprises in these cities reached 13,067, accounting for nearly 70%. It is worth noting that only 11 of these 37 cities are located in the north of China. In terms of SEIs, there is a big gap between northern cities and southern cities in China.

The scope of enterprise agglomeration is further measured. The results show the agglomeration space of core enterprises and supporting enterprises is the same, and the distances between every two enterprises are gathered in

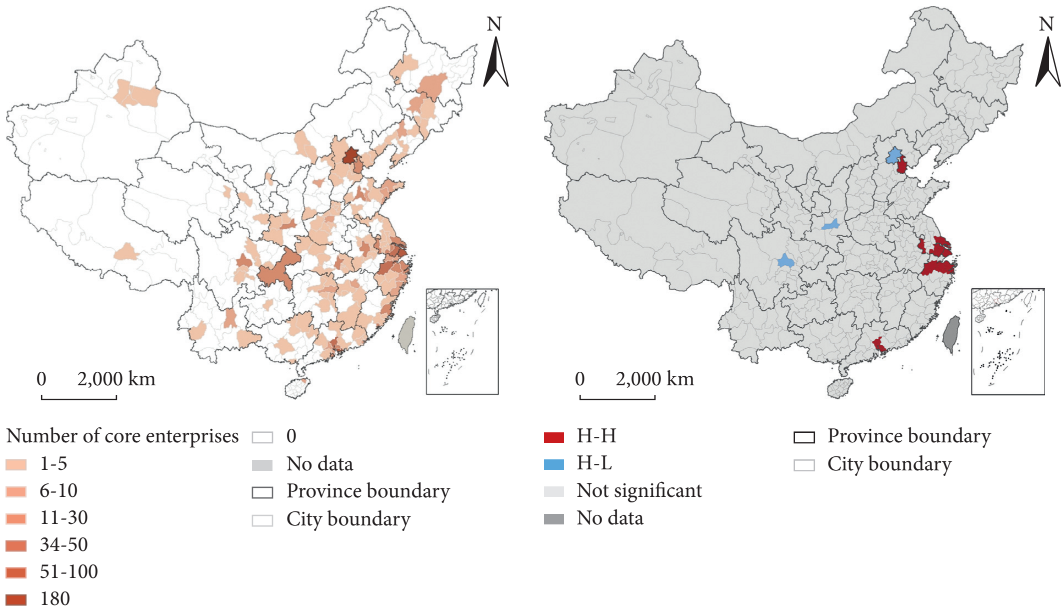



Figure 1: Spatial distribution and agglomeration of core enterprises in SEIs.

Table 1: Cities with 10 or more core enterprises in SEIs and their numbers.

| City | Number of core enterprises | City | Number of core enterprises |
|---|---|---|---|
| Beijing | 180 | Ningbo | 15 |
| Shenzhen | 98 | Zhuhai | 15 |
| Shanghai | 91 | Jinan | 15 |
| Hangzhou | 47 | Xiamen | 14 |
| Suzhou | 39 | Hefei | 13 |
| Guangzhou | 34 | Fuzhou | 13 |
| Chengdu | 30 | Dongguan | 13 |
| Wuhan | 27 | Nantong | 12 |
| Nanjing | 21 | Shaoxing | 12 |
| Xi'an | 18 | Chongqing | 11 |
| Wuxi | 16 | Shenyang | 10 |
| Tianjin | 15 | | |
| Percentage | Primate city<br>16.84 | Top three cities<br>34.52 | Top ten cities<br>56.22 |

90 km, 1100 km, and 1900 km (Figure 3). (1) 90 km is roughly equal to the distance from a large city to a small city in its market hinterland but is significantly less than the distance between cities of 800–1300 km. It means SEIs are mainly centered on regional central cities and are clustered in the surrounding 100 km. The SEI industrial cluster, composed of large core enterprises and small and medium supporting enterprises, is mainly concentrated in the regional central city and its surrounding area about 100 kilometers. (2) The second peak of the distances between every two enterprises is 1100 km, which is just within the range of the distance between cities (800 km–1300 km). Furthermore, the distance between Beijing-Tianjin-Hebei or the Pearl River Delta and the Yangtze River Delta and Chengdu-Chongqing area is about 1,000 km. The industrial clusters are located in these urban agglomeration areas. So the 1100 km is the distance between industrial clusters of SEIs. (3) 1900 km, the third peak, is roughly equivalent to the distance from the Beijing-Tianjin-Hebei to the Pearl River Delta. It also shows the distance between the distant industrial clusters in China. In general, large core enterprises and small and medium-sized supporting enterprises are concentrated on major urban agglomeration areas centered on large cities and develop several regional industrial clusters of SEIs in China.

*3.3. The Structure and Distribution of Strategic Emerging Industrial Clusters.* In addition to the number of enterprises, the structural characteristics presented by the proportions of different types of enterprises also have an important impact on the development of industrial clusters. Studies have shown there are Marshall-type industrial clusters dominated by small and medium-sized enterprises and wheel-type industrial clusters dominated by large, medium, and small enterprises.

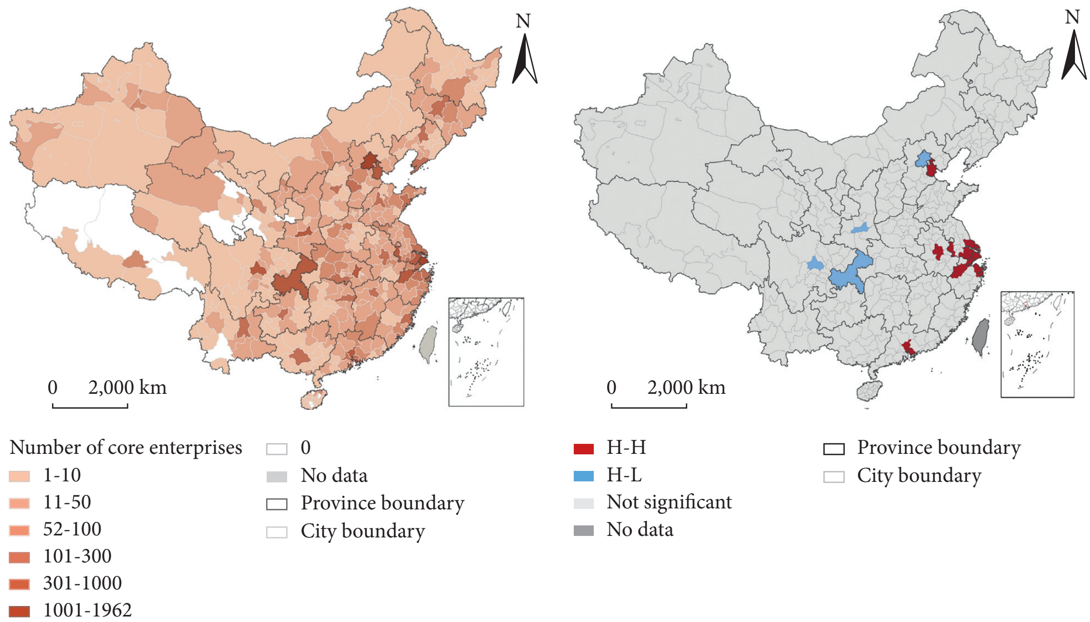



Figure 2: Spatial distribution and agglomeration of supporting enterprises in SEIs.

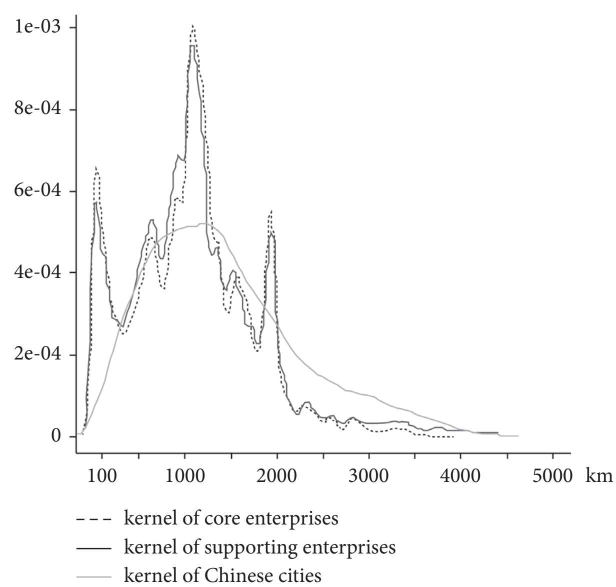



Figure 3: The spatial agglomeration characteristics of core enterprises and supporting enterprises.

The research will further analyze the cluster structure and characteristics of SEIs in different cities based on the proportions between large core enterprises and small and medium-sized supporting enterprises. The SME support ratio of clusters in Chinese cities (the ratio of the number of small and medium-sized supporting enterprises to the number of large core enterprises in the city) has a median number of 17, and the lower quartile and upper quartile of the SME support ratio of each city are 12.17 and 27, respectively. According to this, cities can be divided into nonsupported type (cities with leading enterprises but no supporting enterprises), support-deficient type (cities with SME support ratio less than 12.17), balanced type (cities with SME support ratio between 12.17 and 27), core-deficient type (cities with SME support ratio more than 27), and noncore type. Those cities with less than 5 enterprises related to SEIs are considered to be non-industrial cluster cities.

In terms of types, since core enterprises are mainly concentrated in central cities such as Beijing, Shanghai, and Shenzhen, nearly one-half of the cities are noncore or core-deficient types (Figure 4). There are 76 balanced cities, accounting for about 1/5. Support-deficient and nonsupported cities account for less than 10%. In addition, there are slightly more than one-fifth of cities that have not formed strategic emerging industrial clusters. It is worth noting that, contrary to what most people think, most large cities with

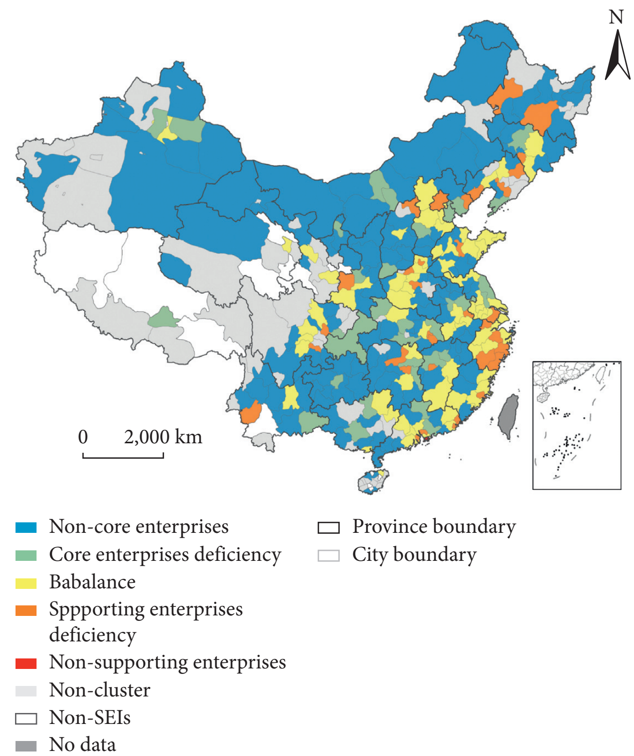



Figure 4: Spatial distribution of strategic emerging industrial clusters in China.

more core enterprises are balanced. There are not many cases of insufficient supporting enterprises due to the development of core enterprises. Actually, among the 23 cities with more than 10 core enterprises, only Beijing, Suzhou, Dongguan, and Shaoxing have a relatively lower SME support ratio. Tianjin and Chongqing even have a relatively lower proportion of core enterprises. The core enterprises and supporting enterprises in most cities are relatively balanced. Most the support-deficient cities are due to undeveloped industrial clusters. Half of the cities of this type have less than 20 supporting enterprises, and 8 cities have 10 or fewer supporting enterprises.

In terms of the distribution of various cities, the relatively developed urban agglomeration areas along the eastern coast such as the Beijing-Tianjin-Hebei, Yangtze River Delta, Pearl River Delta, Shandong Peninsula, and the west coast of the Straits have a relatively good ratio of core and supporting enterprises, mainly belonging to balanced and support-deficient types. Only in the fringe areas far from regional central cities, there are a small number of core-deficient and noncore cities and a certain degree of core-periphery distribution.

In the near western (Chongqing, Sichuan, and Shaanxi) and central provinces, different types of cities are mixed. There are fewer balanced cities, and most cities are core-deficient and noncore. In particular, there are 14 core-deficient cities in the 8 provinces in these regions, accounting for more than 40% (43.75%) of this type of city. From the perspective of the distribution of various types of cities, Anhui and Jiangxi, which are close to the coastal areas, have a trend of coordinated development with urban agglomeration of the Yangtze River Delta and the west straits coast. In general, some strategic emerging industrial clusters have been formed in the near western and central regions, but these clusters still suffered from insufficient core enterprises.

For far western and northeastern provinces, western Sichuan and Hainan are dominated by noncore and noncluster cities. Only 5 of the 64 noncluster cities are not in these regions. Most of the noncore cities also have undeveloped clusters. There are nearly 70 noncore cities in these regions. More than 60% of cities have less than 10 supporting enterprises, and only three cities have more than 20 supporting enterprises. Among the 9 support-deficient cities, only Harbin and Anshan have more than 20 related enterprises. In other cities, the number of enterprises is relatively small. In the fringe areas of western and northeastern China, the development of strategic emerging industries is relatively slow.

## 4. Networking Characteristics of China's SEIs

Although dominated by manufacturing, SEIs are cutting-edge, innovative, and rapid iterative. Therefore, factors such as information, R&D, and capital are also crucial to the development of SEIs. Among the 18,843 supporting enterprises invested by 1069 core enterprises of SEIs, nearly 7,000 are related to information and R&D, accounting for nearly 40% (37.10%), and the number of financial and business service enterprises is 2,738, accounting for nearly 15% (14.53%), while manufacturing enterprises account for only 20% (20.16%). Different from the manufacturing process, the allocation of elements such as information, R&D, and capital is not only localized but has obvious network characteristics.

*4.1. Network Structure of SEIs.* From the perspective of the overall structure of the SEIs network, SEIs present a diamond pattern with Beijing, Shanghai, Guang-Shen (Guangzhou, Shenzhen), and Chengdu as the apex. In the diamond-type industrial connection network, the strength of connections between cities in the east of the Beijing-Guang-Shen line is significantly higher than that in the west. This result is consistent with the conclusions of some urban network research [24].

Degree centrality and strength, two key indicators of network analysis, represent different meanings in the study of industrial linkages. Degree centrality reflects the extensiveness of a city's connections across the country. For strength, it is the sum of the number of industrial connections between a city and other cities (represented by the number of enterprises connected between cities in this study) and reflects the total frequency of connections with other cities. More importantly, this indicator of a city is significantly affected by a few key cities that have large-scale industrial connections with the city.

The results show that Beijing's degree centrality reaches 297, indicating that Beijing has industrial connections with another 332 prefecture-level administrative units in China, followed by Shenzhen's degree of 194, and Shanghai's degree of 163 (Figure 5). The top ten cities in terms of extensive connections all have more than 100 degrees. Similar to the result of degree centrality, Beijing, Shenzhen, and Shanghai are still the three cities with the highest strength (3815, 1717, and 1580, respectively). But the divergence of the strength between cities is more obvious than that of degree centrality. For example, the values of the strength between Shenzhen and Shanghai is only 45% and 41% of Beijing's, respectively; and Hangzhou, the fourth-highest city, has a connection strength of only 922, which is also less than 60% of the third-highest city, Shanghai.

Generally speaking, Beijing is the most advantageous city in the industrial connection network. The values of both the degree and strength are much larger than in other cities including Shenzhen and Shanghai. Besides, those cities with a higher degree of centrality generally have higher strength, but there are also slight differences. For example, Dongguan, which ranks seventh in degree, ranks 18th in strength, while Tianjin, which ranks 12th in degree, ranks among the top 10 in strength.

The industrial connection network formed by degree centrality and the strength differ particularly in terms of matching. In the topological network formed by the SEIs, the degree centrality and its corresponding average degree of the neighbors show an obvious negative correlation, and the correlation coefficient is −0.848 ($P \leq 0.01$) (Figure 6). It means cities with high degrees tend to be linked to cities with low degrees. So, from the perspective of degree centrality, the entire network is disassortative. It is exactly the opposite of the result of the network weighted by contact frequency. The global weight matching coefficient is 0.559. So, from the perspective of strength, the weighted network is assortative (Figure 7). Cities connected with high-strength cities have a higher average strength than their neighbors. This result shows that when only considering the industrial links between cities, the SEIs network has a center-periphery structure. Small and medium-sized cities tend to have connections with large cities such as Beijing, Shenzhen, and Shanghai, while the connections between small and medium-sized cities are not close, showing the characteristics of vertical connections. When the number of connected enterprises is used as the weight, the large-scale and high-strength corporate connections between large cities have greatly increased the average strength of the neighbors, resulting in the weighted network showing the characteristics of assortativity. Therefore, the disassortativity of the topological network is mainly caused by the fact that small and medium-sized cities are less interconnected and mainly connected with large cities, while the assortativity of the weighted network is caused by the large-scale company connections between large cities. However, this does not change the overall structural characteristics of the hierarchical vertical connection industrial network. In general, the SEIs network presents multiple local networks with vertical connections centered on large cities, and the core cities in the local network are closely connected with high intensity.

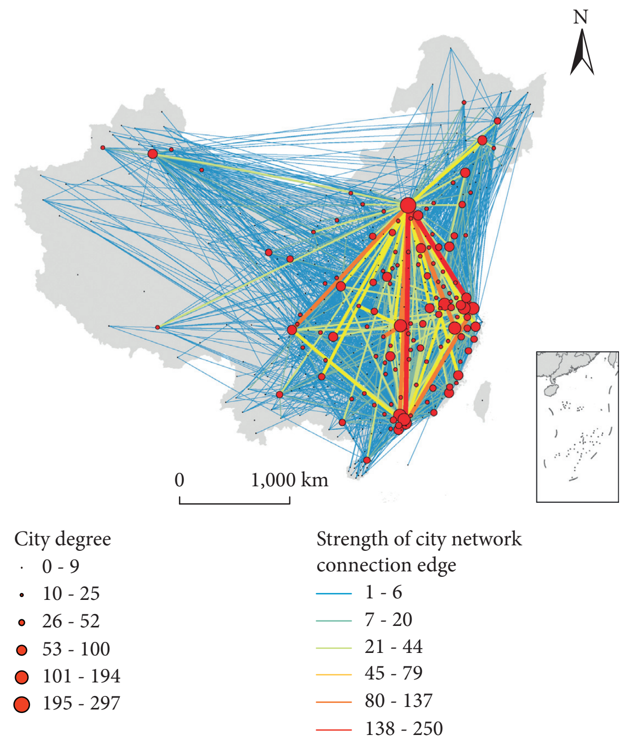



Figure 5: Network characteristics of China's SEIs.

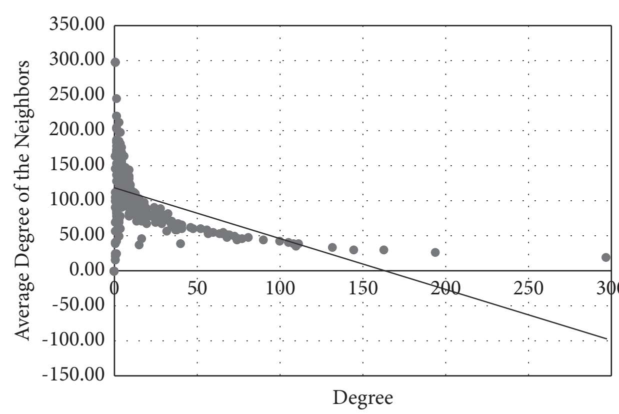



Figure 6: Degree correlation of the topological network of China's SEIs.

*4.2. Industrial Connection Network of Major Node Cities.* To further clarify the differentiated characteristics of different cities in the industrial connection network, a comparative analysis will be made on the industrial connection network of major cities. For Beijing, its investment in external regions is focused on central cities in the eastern and central regions (Figure 8). For example, Beijing has a high investment scale in Shanghai, Wuhan, Xi'an, Chengdu, and Shenzhen. However, Beijing has no obvious investment tendency in neighboring areas such as Hebei, Shandong, and Shanxi. Except for Tianjin that has invested in more than 100 related enterprises, cities such as Xi'an, Jinan, Shijiazhuang, and Shenyang and other cities have invested in less than 60

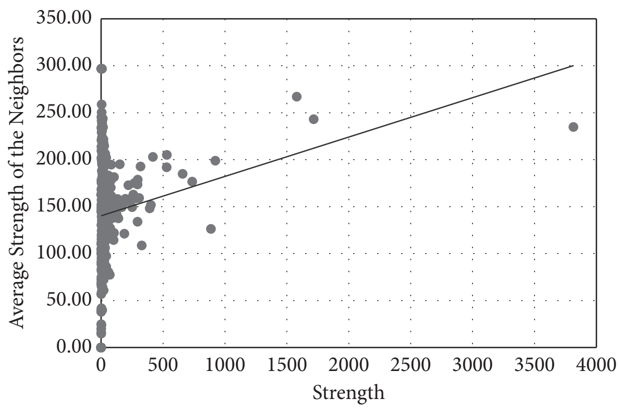



Figure 7: Strength correlation of the topological network of China's SEIs.

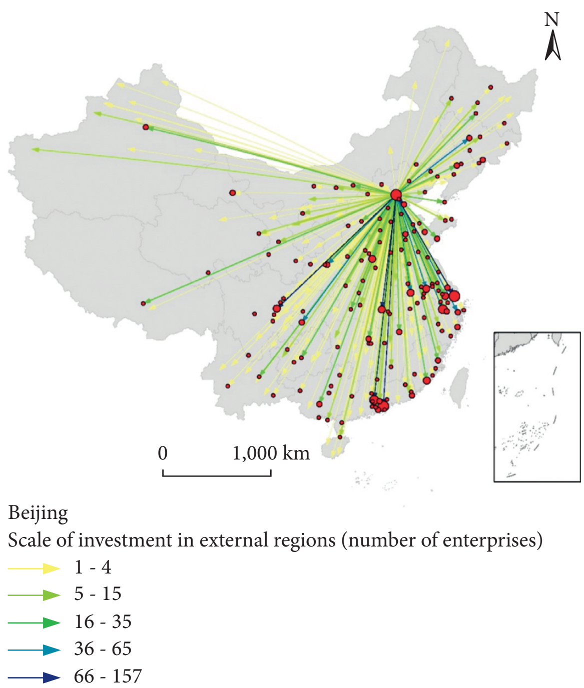



Figure 8: The scale and direction of Beijing's SEIs investment in external regions.

related enterprises. The largest number of investment enterprises in the northern central city is Xi'an, which has invested in 58 enterprises, ranking 9th in the number of enterprises investing in external regions, and lower than central and southwestern central cities such as Wuhan and Chengdu.

Shenzhen is the second-largest node city after Beijing in China's SEI system. Except for its investment enterprises ranked second, Shenzhen also has a greater degree of investment across the country than Shanghai (Figure 9). Only 16.03% of enterprises investing outside Shenzhen are concentrated in the Pearl River Delta region, and among 22 cities with more than 10 invested enterprises from Shenzhen, only 4 cities are located in the Pearl River Delta area. At the same time, cities with a large number of enterprises invested

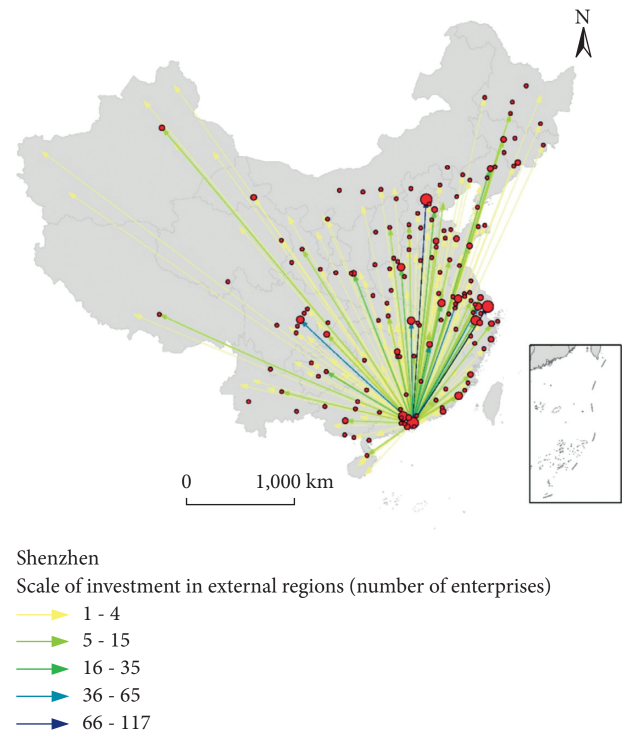



Figure 9: The scale and direction of Shenzhen's SEIs investment in external regions.

by Shenzhen include not only major urban agglomeration areas such as the Yangtze River Delta, the Pearl River Delta, Beijing-Tianjin-Hebei, and Chengdu-Chongqing, but also central cities in the northeast and southwest regions such as Shenyang and Kunming.

Apart from investing in large-scale enterprises in Beijing and Shenzhen, Shanghai's main investment cities are located in the Yangtze River Delta such as Jiangsu and Nanjing (Figure 10). There are 18 cities with more than 10 invested companies from Shanghai, and half is located in the Yangtze River Delta. Unlike Beijing's investment across all over the country, invested enterprises from Shanghai are mainly concentrated in the Yangtze River Delta region, accounting for 62.70%. Nearly one-third of the enterprises investing outside Shanghai is located in the Yangtze River Delta region (32.95%). There are fewer companies from Beijing investing in the Beijing-Tianjin-Hebei region. Less than 1/3 (32.75%) of invested companies from Beijing are located in the Beijing-Tianjin-Hebei region, and even less than 10% (8.14%) in Tianjin and Hebei province. In addition, Wuhan in the middle reaches of the Yangtze River and Chengdu and Chongqing in the upper reaches are also cities where Shanghai tended to invest.

Except for Beijing, Shenzhen, and Shanghai, other cities take their provinces as the main contact areas, so these cities are more inclined to be provincial-level central nodes. For example, four of the top seven cities that have the largest number of enterprises that Guangzhou invests in external regions, are the Pearl River Delta and surrounding areas (Foshan, Shenzhen, Changsha, and Shaoguan). Among the top 5 cities that have the largest number of enterprises that

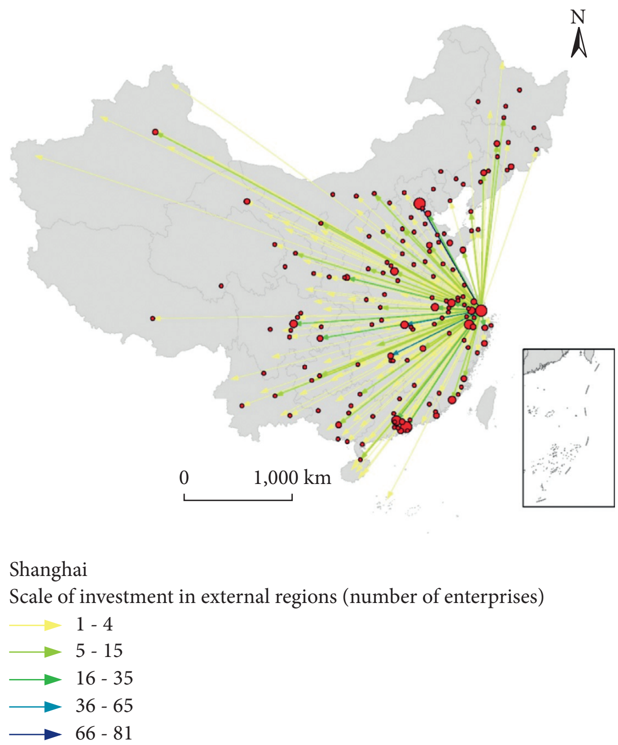



FIGURE 10: The scale and direction of Shanghai's SEIs investment in external regions.

Hangzhou invests in, only Beijing is not a city in Zhejiang Province. The number of enterprises investing in Hangzhou and other cities in Zhejiang Province accounts for 55.42% of the total of invested enterprises from Hangzhou. There are 13 cities with more than 10 enterprises invested in Nanjing (Figure 11). Among them, Beijing, which ranks second, and fifth Shenzhen, are not cities in Jiangsu Province.

Beijing is not only an export destination for investment by enterprises related to SEIs across the country but also an absorption destination for related investments. Beijing is the city with the largest number of enterprises invested by Shanghai, Shenzhen, and Nanjing, and the second-largest number of enterprises invested by Guangzhou and Nanjing. Therefore, Beijing plays a very important role in the construction of China's SEIs system and the process of industrial development, and it is also an important national node city in the industrial system. Shenzhen is second only to Beijing in the degree of industrial connections in China and is a quasi-national node city. Shanghai has the characteristics of a certain regional node city because its industrial connections are mainly concentrated in the Yangtze River Delta. The industrial connection network of other cities takes the province where it is located as the main spatial scope.

## 5. The Impact of Spatial Agglomeration and Industrial Networks on Industrial Development

To further clarify the influence of industrial agglomeration and industrial network on urban development, a multiple regression model is built based on the Cobb Douglas production function to analyze the influence of four variables, including cluster size, cluster structure, degree, and strength on urban economic growth. The model is as follows:

$$\ln Y = \ln K + \ln L + \ln \mathrm{Scale\_large} + \ln \mathrm{Structure} + \ln \mathrm{Degree} + \ln \mathrm{Strength}, \quad (4)$$

where $Y$ is GDP in 2018, $K$ is capital stock in 2017, $L$ is the number of employees in 2017, Scale_large is the number of core enterprises in SEIs, and Structure is the cluster structure of SEIs (specifically, the number of core enterprises/the number of supporting enterprises), Degree is the degree centrality of the city, and Strength is the strength of the city (Table 2).

The results show that the agglomeration and network of SEIs have a significant impact on urban economic growth. Among them, the number of core enterprises and degree of centrality have obvious positive effects on urban economic growth, while the cluster structure and the strength have negative effects (Table 3). From the perspective of industrial agglomeration and clusters, core enterprises have a significant positive impact, indicating that the more the core enterprises, the stronger the promotion of urban growth. However, the higher the proportion of core enterprises is, the less conducive they are to urban economic growth. The result shows the important role of supporting enterprises in SEIs clusters. At the same time, when the variable of the total number of enterprises in SEIs is added to the model, we find that the total number of enterprises has no significant effect on urban economic growth. In other words, the scale of SEIs clusters has little effect on urban growth.

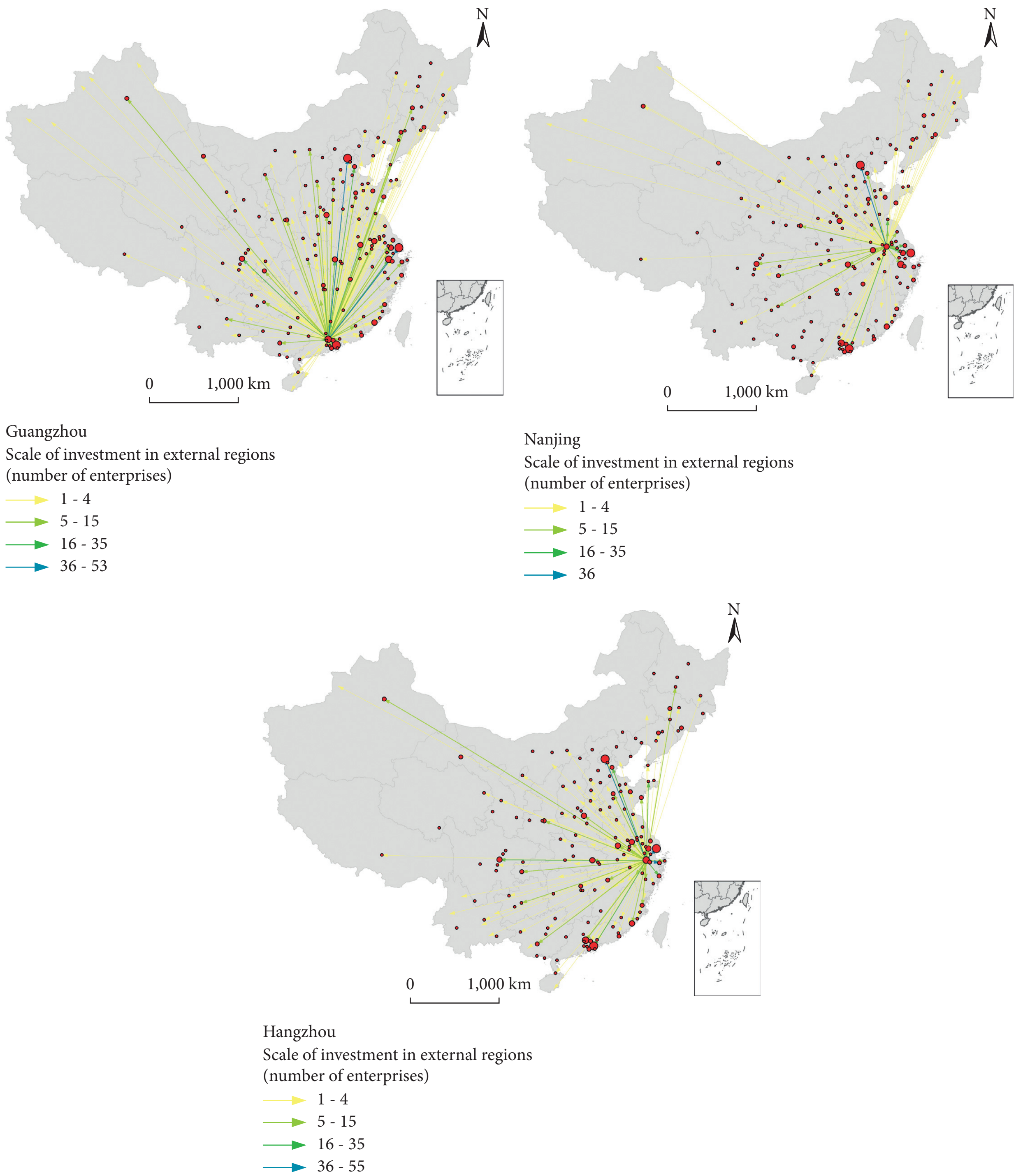



FIGURE 11: The scale and direction of SEIs investment of Guangzhou, Nanjing, and Hangzhou in external regions.

TABLE 2: Descriptive statistics of cities.

| | GDP in 2018 (billions) | Capital stock in 2017 (billions) | Number of employees in 2017 (thousands) | Number of large enterprises in cities ($n$) | Cluster structure of strategic emerging industry (/) | Degree ($n$) | Node strength ($n$) |
|---|---|---|---|---|---|---|---|
| Mean | 465.30 | 901.57 | 891.19 | 7.05 | 0.05 | 31.84 | 150.17 |
| Std. dev | 544.07 | 819.95 | 1142.99 | 19.23 | 0.04 | 40.02 | 387.16 |
| Max | 3267.99 | 5254.88 | 8128.59 | 180.00 | 0.23 | 297.00 | 3815.00 |
| Min | 27.42 | 66.70 | 83.19 | 0.00 | 0.00 | 1.00 | 1.00 |

Table 3: Regression results of industrial agglomeration, cluster structure, and linkage network.

| | Model 1 | Model 2 | Model 3 | Model 4 |
|---|---|---|---|---|
| $K$ | 0.5770*** (0.000) | 0.5092*** (0.000) | 0.5523*** (0.000) | 0.5138*** (0.000) |
| $L$ | 0.5316*** (0.000) | 0.5268*** (0.000) | 0.5269*** (0.000) | 0.5240*** (0.000) |
| Scale_large | | 0.0741** (0.047) | | 0.1211** (0.013) |
| Structure | | −0.5208 (0.327) | | −1.0768* (0.056) |
| Degree | | | 0.1369** (0.023) | 0.1565** (0.012) |
| Strength | | | −0.0804* (0.095) | −0.1470*** (0.007) |
| $R^2$ | 0.9497 | 0.9505 | 0.9509 | 0.9525 |

From the perspective of the industrial network, the degree centrality has a positive effect, and the influence of degree is the most stable among the four variables. It means that the extensiveness of the city's connections in the industrial network has a significant role in promoting urban economic growth. But the strength has a negative effect.

However, compared with capital stock and employees, the impact of strategic emerging industries on urban economic growth is still relatively small. The $R$ square of the two variables relative to GDP reaches 0.9497, while the $R$ square only increases by 0.0028 after adding the four variables. To some extent, SEIs is still not a core driving force for urban economic growth.

## 6. Concluding Remarks

SEIs are a relatively special type of industry. On the one hand, the SEIs are dominated by manufacturing industries such as information equipment, new energy, and medicine, which are characterized by a relative agglomeration of manufacturing industries in nearby regions. On the other hand, because these industries require a large amount of capital, information, and R&D investment, they also show the networked characteristics of high-end service industries. Based on the relevant data of China's SEIs, this paper deeply analyzes the characteristics of industrial cluster agglomeration and networking and finds that both the spatial agglomeration and networking of SEIs have a significant impact on the development of Chinese cities.

Enterprises related to SEIs are widely distributed in Chinese cities, but they are also highly concentrated. More than 40% of cities have core enterprises in SEIs, and more than 93% of prefecture-level cities have at least one supporting company. However, most core and supporting enterprises are concentrated in a few cities, such as Beijing, Shanghai, and Shenzhen, and a few regions such as the Yangtze River Delta, Beijing-Tianjin-Hebei, Pearl River Delta, and Chengdu-Chongqing region. The spatial distribution of both core enterprises and supporting enterprises presents a center-periphery model centered on large cities, and they are mainly concentrated within a range of 90 km around large cities. At the same time, China has formed several large industrial clusters, based on major urban agglomeration areas.

From the perspective of industrial clusters, the proportion between core enterprises and supporting enterprises in the eastern region is more balanced. The central region generally shows the characteristics of insufficient core enterprises, while the western and northeastern regions have no core or support company due to undeveloped industrial clusters. At the same time, noncluster cities with less than 5 related enterprises are also mainly distributed in the northeast, western ethnic minority areas, and Hainan.

From the perspective of the network, the diamond network pattern of SEIs is the same as that of the whole industry, but the development level of the western Chengdu-Chongqing region is relatively low. Beijing, Shenzhen, and Shanghai are the first-tier cities in the industrial network. Among them, Beijing occupies a core position in China's strategic emerging industrial network. Nearly 300 cities have industrial connections with Beijing. Shenzhen also has the potential to become a national center. Relatively speaking, Shanghai is more inclined to be the center city of the Yangtze River Delta and the Yangtze River Basin, and more than 60% of its external investment enterprises are located in the Yangtze River Delta region. Other cities, such as Guangzhou, Hangzhou, Chengdu, etc., are more regarded as provincial central cities, and their industrial connections are mostly within their provinces.

In general, both industrial agglomeration and the network of SEIs have a significant impact on the growth of cities. In particular, the extensiveness of the industrial connection network plays an important role in promoting the growth of cities, followed by the agglomeration of core enterprises in cities. However, it does not dwarf the importance of small and medium-sized supporting enterprises. An excessively high proportion of large core enterprises in industrial clusters harms urban economic growth.

## Data Availability

The enterprises data used to support the findings of this study may be released upon application to the Shanghai HeHe Information Technology Co., Ltd., which can be contacted at the URL: https://www.qixin.com/?from=baidusemBrand1.

## Conflicts of Interest

The authors declare that they have no conflicts of interest.

## References


[1] G. Ellison and E. L. Glaeser, "Geographic concentration in U.S. Manufacturing industries: a dartboard approach," *Journal of Political Economy*, vol. 105, no. 5, pp. 889–927, 1997.

[2] P. Krugman, "Increasing returns and economic geography," *Journal of Political Economy*, vol. 99, no. 3, pp. 483–499, 1991.

[3] A. Marshall, *Principles of Economics*, Macmillan, London, UK, 1920.

[4] J. Jacobs, *The Death and Life of Great American Cities*, Vintage Books, New York, NY, USA, 2012.

[5] G. Ellison, E. L. Glaeser, and W. R. Kerr, "What causes industry agglomeration? Evidence from coagglomeration patterns," *The American Economic Review*, vol. 100, no. 3, pp. 1195–1213, 2010.

[6] G. Duranton and D. Puga, "Nursery cities: urban diversity, process innovation, and the life cycle of products," *American Economic Review*, vol. 91, 2001.

[7] V. Henderson, A. Kuncoro, and M. Turner, "Industrial development in cities," *Journal of Political Economy*, vol. 103, 1995.

[8] E. J. Feser, *Industry Clusters and Economic Development: A Learning Resource*, University of Illinois, Illinois, IL, USA, 2004.

[9] S. S. Rosenthal and W. C. Strange, "Evidence on the nature and sources of agglomeration economies," *Handbook of Regional and Urban Economics*, vol. 4, pp. 2119–2171, 2004.

[10] R. Florida, *Who's Your City? How the Creative Economy in Making where to Live the Most Important Decision of Your Life*, Basic Books, New York, NY, USA, 2008.

[11] P. McCann and Z. J. Acs, "Globalization: countries, cities and multinationals," *Regional Studies*, vol. 45, pp. 17–32, 2011.

[12] S. Sassen, *The Global City*, Princeton University Press, Princeton, NJ, USA, 1991.

[13] M. Castells, *The Rise of the Network Society*, Blackwell, Oxford, UK, 1996.

[14] P. J. Taylor and B. Derudder, *World-city Network: A Global Urban Analysis*, Routledge, London, UK, 2nd edition, 2016.

[15] R. Boix and J. Trullén, "Knowledge, networks of cities and growth in regional urban systems," *Papers in Regional Science*, vol. 86, no. 4, pp. 551–574, 2007.

[16] R. Capello, "The city network paradigm: measuring urban network externalities," *Urban Studies*, vol. 37, no. 11, pp. 1925–1945, 2000.

[17] W. Alonso, "Urban zero population growth," *Daedalus*, vol. 104, 1973.

[18] E. J. Meijers and M. J. Burger, "Stretching the concept of "borrowed size"," *Urban Studies*, vol. 54, no. 1, pp. 269–291, 2017.

[19] J. Hu and Y. L. Zhao, "Research on agglomeration degree and its changing tendency of China's strategic emerging industries-empirical evidence from listed enterprises," *Reform of Economic System*, no. 6, pp. 102–106, 2015.

[20] C. A. Tang, L. J. Zhang, and Q. Yin, "Evolution and optimization of spatial layout of strategic emerging industry in China," *Economic Geography*, vol. 5, pp. 101–107, 2018.

[21] M. Zhao, X. Liu, B. Derudder, Y. Zhong, and W. Shen, "Mapping producer services networks in mainland Chinese cities," *Urban Studies*, vol. 52, no. 16, pp. 3018–3034, 2015.

[22] S. Boccaletti, V. Latora, Y. Moreno et al., "Complex networks: structure and dynamics," *Complex Systems and Complexity Science*, vol. 424, 2006.

[23] M. Newman, "The structure and function of complex networks," *Proceedings of the National Academy of Sciences*, vol. 101, no. 1, pp. 5200–5205, 2004.

[24] K. Wu, C. L. Fang, and M. X. Zhao, "The spatial organization and structure complexity of Chinese intercity networks," *Geographical Research*, no. 4, pp. 711–728, 2015.

Hindawi
Complexity
Volume 2021, Article ID 3280777, 19 pages
https://doi.org/10.1155/2021/3280777

*Research Article*

# Modeling the Public Transport Networks: A Study of Their Efficiency

**Mary Luz Mouronte-López** [1,2]

[1]*Higher Polytechnic School, Universidad Francisco de Vitoria, Madrid, Spain*
[2]*Telefónica Chair, Universidad Francisco de Vitoria, Madrid, Spain*

Correspondence should be addressed to Mary Luz Mouronte-López; maryluz.mouronte@ufv.es

Received 19 May 2021; Revised 28 July 2021; Accepted 3 August 2021; Published 14 August 2021

Academic Editor: Jing-Hu Pan



The public transportation network (PTN) provides mobility and access to community resources, employment, medical care, infrastructures, and other resources in the city. This research studies the process of the formation of links among nodes in different real-world PTNs. We have found that this process may be appropriately explained by a generalized linear model (GLM) using local, global, and quasilocal similarity indexes as explanatory variables. In modeling, the response variable was described by a binomial probability density function, and the logit function was used as a link function. In the crossvalidation process, utilising a downsampling approach, both average accuracy and area under the receiver operating characteristic curve (AUC) metrics presented higher values than 0.99. The kappa parameter had magnitudes larger than 0.93 for most of the PTNs. In the final validation stage, recall and specificity metrics took the value 1. Accuracy and precision parameters were larger than 0.99 and 0.87, respectively, for the majority of PTNs. Only one of the PTNs required utilising a smoothed bootstrap approach in order to achieve better results. The similarity measures with the greatest influence on the model were determined. We also assessed the impact of link removal on the global efficiency of PTNs, considering several similarity indexes. Additionally, we find that most of the networks show low local and global efficiencies ($\leq$0.20), as well as travel times with a relevant variability, exhibiting standard deviations larger than 790 seconds. Significant similarities exist between the cumulative probability distributions of the local efficiency in all PTNs. With respect to the centrality measures, the eigenvector centrality presented a strong correlation with the hub/authority centralities (>0.80), while the pagerank showed a moderate, high, or very high correlation with the degree in all PTNs, >0.50.

## 1. Introduction

Link prediction methods have been the subject of research [1–4], which suggests several mechanisms to detect hidden connections. These mechanisms take into account the path information between pairs of nodes in order to estimate their common neighbors. They also consider a mutual information perspective in order to evaluate the similarity index between pairs of nodes. The conditional probability for the existence of a link is calculated, given the common neighbor of two nodes, as described in [5]. Finally, the weight of the links are considered, developing the mechanisms described in [6], which are based on the common neighbor, resource allocation (RA) [7], and adamic adar (AA) [8] indexes. The above is combined with the weighted mutual information (WMI) [9] score estimated between node pairs. Reference [10] suggests a new local information-based link prediction method, tie connection strength index (TCS), concerning the efficient paths between the target node-pair and their common neighbor. An adaptable parameter is presented in order to estimate the impact of the TCS and the topology of the network on the similarity of pairs of nodes. Reference [11] establishes a new type of triangle structure, which consists of one seed node, one common neighbor, and another node. Based on this, a new similarity index, named TRA index by the authors, is proposed for link prediction. The authors integrate the new triangle structure and the idea of RA [7] index [7]. Reference [12] proposed a new similarity measure based on the AA score, information related to communities generated from the topological structure of the network and the degree centrality. The link prediction algorithms use two open implementations of a bulk synchronous parallel programming model [13]. They are Apache Giraph and Apache Graphx. Reference [14] demonstrates that similarities with respect to structural features (eigenvectors) optimize the link prediction task in multiplex networks. This is done using a layer reconstruction method (LRM), which considers the unconnected node pairs in the target layer as similar, provided that they are not only analogous from the point of view of the target layer but also from the perspective of other layers. Tests on real multiplex networks show that LRM takes advantage of existing information redundancy in different layers.

The application of link prediction methods in real contexts has also been analyzed. A great deal of research is done on the analysis of social networks. Reference [15] carries out a comprehensive review and discusses some link prediction applications in social networks such as recommender systems, community detection, anomaly detection, and influence analysis. Because social networks are highly dynamic with the come-and-go of nodes and links, some research considers temporal aspects. Reference [16] characterizes the likelihood of a link between two nodes from both existing connectivity topology and the popularity of both nodes. Several datasets are considered in order to test and calculate the performance of algorithms. Reference [17] builds a linear model for integrating neighborhood similarity measures and node specific information and uses an evolutionary algorithm to locate the coefficients, which optimizes the prediction of links. The authors assign different weights to each index using the Covariance Matrix Adaptation Evolution Strategy (CMAES) [18, 19]). In addition, the protein-protein interaction (PPI) networks (PPI) have been examined using link prediction methods. Reference [20] utilises the support vector machine learning method for protein-protein interaction (PPI) prediction. Features, often used in social networks, like some similarity index, have been progressively put into practice to make predictions in PPI [21, 22].

This paper studies the link formation process in several PTNs using various similarity measures, which have been applied in a link prediction theoretical framework. The most influential indexes in the pattern followed by link formation between pairs of nodes are determined.

PTNs have been examined from different points of view. Thus, models have been implemented to analyze travel behaviours. Reference [23] forecasts, based on surveys, some characteristics related to the passenger flow. Reference [24] implements a Bayesian network to detect the relationships between travel happiness and several parameters that affect travel behavior. Reference [24] checks pretravel information-seeking behaviours of the passengers using data collected during an extensive public transport on-board survey. For this purpose, the authors implement a multivariate binomial logistic regression model. The model takes into account factors related to sociodemographics, aspects of the travelers, characteristics of the trip, and devices used for information consultation.

The main novelty of our research is that it shows that the link formation pattern in PTNs can be appropriately explained by means of a generalized linear model (GLM), which has local, quasilocal, and global similarity measures between nodes as explanatory variables. The response variable, which establishes whether or not a link exists between pairs of nodes, is described by a binomial probability density function. The link function used is the logit function.

Studies exist that analyze topological parameters in PTNs (degree distributions, path length distribution, and betweenness), as well as growth models. However there are no analysis that we know of, which does this demonstration on PTNs. Research exists, which has developed growth models for PTNs, based on other considerations. Reference [25] replicates some statistical features of PTNs, describing their evolution in terms of adding routes in *P-space*. The authors use a self-avoiding walk (SAW) as a route model. In the aforementioned *P-Space* [26], one node symbolizes one stop, and one link joins a pair of stops, if at least one route exists that supports a direct service between them. Reference [27] developed an area-based model of highway growth. Specifically, a binary logit model in order to estimate the new route growth probability of divided highways and secondary highways using high-quality geographic information system (GIS) data of land-use, population distribution, and highway network for the Twin Cities Metropolitan Area from 1958 to 1990 was obtained in [28]. A growth model that iteratively invested in constructing new links or incrementing the capacity of those existing was implemented. The objective of the research was to establish the impact the demand distributions and operational costs have on the evolution of a PTN. The model considered parameters related to grid geometry, demand characteristics, operating mode parameters (operational speed per mode, cost per km, and capacity). On the contrary, the model described in this paper explains the appearance of links in PTNs based on exclusively topological parameters.

The PTNs also been studied as complex systems [29, 30] describes a geospatial layout for distributing stops and uses a maximum allowable walking distance in order to link the routes. The PTNs are optimized, considering aspects as efficiency and robustness. Reference [31] studies common problems that have been found when a complex system scheme is used for the analysis of the topology of a transportation system (such as mechanisms for the evaluation of the scale-freeness, metrics for the analysis of the network structure, and examination of the vulnerability of the networks using methods with an unacceptable computational time). The vulnerability of the PTNs has also been analyzed in depth [26, 32].

This paper studies the impact that the removal of links, with certain similarity characteristics, has on the global efficiency of PTNs. The relationships between similarity characteristics and the local efficiency of nodes are also checked. Other research has analysed the effect that the node elimination has on the global efficiency of PTNs [33], and the robustness of PTNs has been examined from other points of view, such as the evolution of the giant component when several nodes are deleted [26, 33]. The fault propagation

[20, 26] from nodes with certain topological characteristics (highest betweenness, degree, eigenvector centralities, and pagerank) has also been analyzed. However, a detailed study of the effect on the global efficiency in PTNs when certain links are removed according to similarity indexes analysed in this research has not been found.

This paper also examines the correlation between some centrality measures and relates them to other traffic flow characteristics. Some research exists [34–37] that analyze the correlations between centrality measures in networks of different types. However, we focus on the study of centralities in PTNs and relate them to the flow of vehicles. These characteristics, that we know, have not been previously studied specifically in the PTNs presented here. Moreover, the networks analyzed here are of very different sizes and nationalities, which suggests that they can also operate differently, bringing generality to the analysis. The correlation between centrality measures can explain some of the patterns found in PTN, when a target attack or a fault propagation is suffered by them [26].

The same applies to the study of travel times. It has been shown that, in general, the size, complexity, and variability of available routes in PTNs produce trip times that are highly different between routes. We also study the local efficiency, demonstrating that there are commonalities between PTNs with respect to this feature.

The PTNs studied are AVL, CFL, RGTR, and TICE in Luxembourg, which has 1,372 nodes and 340,684 links; Island Transit in USA, which has 358 nodes and 5,946 links; Lanta in USA, which consists of 2,150 nodes and 91, 583 links; Linja-Karjala Oy in Kuopio, Finland, which has 551 nodes and 63,339 links; Metlink in New Zealand, which has 3007 nodes and 355621 links; Prague Public Transit Company (PPTC), Regional Organiser of Prague Integrated Transport (ROPIT) in Prague, which consists of 5,152 stops and 1,602,778 links; STAR in France, which consists of 1,415 stops and 9,477,213 links; Thunder Bay Transit in Ontario, Canada, which consists of 825 nodes and 78,247 links; TransAntofagasta in Chile, which has 650 nodes and 58 724,362 links; and finally, Sage in California, which has 31 stops and 66 links. It can be observed that the networks are of small, medium, and large sizes.

The vulnerability of AVL, CFL, RGTR, TICE; Linja-Karjala Oy, STAR; Thunder Bay Transit; and Trans-Antofagasta networks was analyzed in [26].

The objectives of this research were as follows:

(1) To analyze whether a GLM, which has as input variables certain measures of similarity between nodes, can correctly explain the formation of links. To establish which of the measures have greater significance in this process.

(2) To detect the influence that the links can have on the global efficiency of the network, according to their similarity characteristics.

(3) To find common features in the networks that allow to characterize their efficiency and trip times).

(4) To determine the relationships that may exist between some centrality measures (eigen vector, pagerank, betweenness, hub, and authority), as well as with other traffic flow characteristics.

## 2. Materials and Methods

*2.1. Overview of Used Resources.* Information related to the stops and routes based on the studied networks, which is available on the websites, was utilised. Several programs in R [38] and Python [39] were specifically implemented to carry out this research, using the R.3.6.0 and 3.8.3 version, respectively. The networks and igraph packages were used. In addition, the proxfun, caret, nortest, stats, vip, and rose packages in R were utilised.

The programmes specifically developed to perform this research allowed:

Processing of information related to the PTNs to be able to work with it (routes, stops, stop times, trips, and calendars) (in Python, ProcessPTNInf.py).

Construction and simplification of the graphs that describe a PTN. Obtaining the similarity measures between nodes (in R, ConstGraphCalcSim.R).

Estimation of centralities (in R and python, CalcCentralities.py and CalcCentralities.R).

Building of a binary classification model, evaluating their results (in R, ModelingPTN.R).

Obtaining frequency and cumulative probability distributions related to efficiency and trip times (in R, CalcDistr.R).

Get graphs showing the results (in R, DrawGraphs.R).

These programs followed the typical development life cycle with phases of specification, detailed design, coding, and testing.

*2.2. Overview of Used Methods*

*2.2.1. Generalized Linear Models.* This is the generalized linear model (GLM) we have used for the simulation of link formation in PTNs.

Consider the response $Y_i$ and the set of independent variables $X_i = (x_{i1}, x_{ip})$ for $i = 1, \ldots, n$. A GLM consists of both a random and a systematic component, as well as a link function.

Regarding the random component, it is assumed that $Y_i$, $1 \leq i \leq n$, are independent random variables described by a probability density function from the exponential family:

$$f(y;\theta,\phi) = \exp\left[\frac{y\theta - b(\theta)}{a(\phi)} + c(y,\phi)\right], \qquad (1)$$

where $a, b, c$ are known functions, and $\theta, \phi$ are parameters, called natural and dispersion parameters, respectively.

The systematic component relates some vector $(\eta_1, \ldots, \eta_n)$ to the $p$ features.

$$\eta_i(\beta) = x_i^t, \quad \beta = \beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2} + \cdots + \beta_p x_{ip}, \quad (2)$$

where $\beta = (\beta_0, \beta_1, \ldots, \beta_p)$ are called regression parameters.

The link function $g(\mu_i) = \eta_i = x_i^t\beta$ relates the linear predictor to the mean $\mu_i$ of $y_i$. If $\eta = \theta$, that is, if $\theta_i = \eta_i$, $\forall i$ holds. The link function is called the canonical link function.

The exponential family contains commonly used distributions such as gamma, normal, inverse Gaussian, Bernoulli, binomial, Poisson, geometric, negative binomial, and exponential.

In particular, a probability density function $f(y;\theta,\phi)$, characterized as a binomial distribution, where $n$ is the number of trials, can be defined as

$$\begin{aligned} f(y;\theta,\phi) &= \binom{n}{y}\mu^y(1-\mu)^{n-y} \\ &= \exp\left[y\ln(\mu) + (n-y)\ln(1-\mu) + \ln\binom{n}{y}\right] \\ &= \exp\left[y\ln\left(\frac{\mu}{1-\mu}\right) + n\ln(1-\mu) + \ln\binom{n}{y}\right]. \end{aligned} \quad (3)$$

Therefore,

$$\begin{aligned} \theta &= \ln\left(\frac{\boldsymbol{\mu}}{1-\boldsymbol{\mu}}\right), \\ \mathbf{b}(\theta) &= -\mathbf{n}\,\mathbf{ln}(1-\boldsymbol{\mu}) = \mathbf{n}\,\mathbf{ln}(1+\mathbf{exp}\,\boldsymbol{\theta}), \\ \mathbf{c}(\mathbf{y},\phi) &= \ln\binom{\mathbf{n}}{\mathbf{y}}. \end{aligned} \quad (4)$$

To evaluate the parameters of an exponential family, GLM maximum likelihood can be applied,

$$L(\theta) = \prod_{i=1}^{n} f(y_i;\theta,\phi). \quad (5)$$

Therefore, log-likelihood for the sample $y_1, \ldots, y_n$ is

$$l(\theta) = \sum_{i=1}^{n}\frac{y_i\theta_i - b(\theta_i)}{a(\phi)} + \sum_{i=1}^{n} c(y_i,\phi). \quad (6)$$

We use as link function $g$, a logit function. It returns values between 0 and 1 for any input,

$$g(\mu_i) = \ln\left(\frac{\mu_i}{1-\mu_i}\right). \quad (7)$$

In order to maximize $l(\theta)$ over all choices of coefficients $\beta \in R^p$, it is necessary to consider that each natural parameter $\theta_i$ may be expressed using the mean $\mu_i$ of the exponential family distribution. Taking it into account, and recalling that a link function exists, such as

$$g(\mu_i) = \eta_i, \quad (8)$$

which joins the mean $\mu_i$ to the parameter $\eta_i = x_i^t\beta$. It is possible to compute $\beta$ as in $\widehat{\beta}$ and then use these estimates to state that $g(\widehat{\mu}_i) = x_i^t, \widehat{\beta} i = 1, \ldots, n$; $\widehat{\mu}_i = g^{-1}\ (x_i^t\widehat{\beta}), i = 1, \ldots, n$.

Therefore, it is possible to establish

$$l(\beta) = \sum_{i=1}^{n} y_i\theta_i - b(\theta_i), \quad (9)$$

where the terms that do not depend on $\theta_i, i = 1, 2, \ldots, n$, have been removed.

If the canonical link function $g$ (8) considers

$$\theta_i = \eta_i = x_i^t\beta, \quad i = 1, \ldots, n, \quad (10)$$

$l(\beta)$ to maximize over $\beta$ is

$$l(\beta) = \sum_{i=1}^{n} y_i x_i^t\beta - b x_i^t\beta. \quad (11)$$

In order to maximize $l(\beta)$ to form $\widehat{\beta}$, it is possible to carry out iteratively reweighted least squares regressions (IRLS) [40, 41]. Finally, the coefficients $\widehat{\beta}$ can be managed as a result of a single weighted least squares regression, the last one in the IRLS succession.

Specifically in this research, it is shown that the pattern of link formation in various PTNs can be well explained through a GLM. In this case, the response $Y_i$ takes a categorical value, whether or not a link exists between two stops. The independent variables, $X_i = (x_{i1}, x_{ip})$, correspond to several indexes describing the similarity between stops. The probability density function $f(y;\theta,\phi)$ is characterized as a binomial distribution. The similarity indexes utilised as predictors are described in the labeled link building process in PTNs and the Supplementary materials section.

In order to check the importance of predictors using the $t$-test, it is required to examine if $\widehat{\beta}_{\mathbf{j}}$ $\forall j$ is normally distributed. This is checked by applying the Anderson–Darling test [30] with a significance level $\alpha = 0.05$. The considered hypotheses are as follows:

(i) Null hypothesis $H_0$: "$\widehat{\beta}_j$ is normally distributed"

(ii) Alternative hypothesis $H_a$: "$\widehat{\beta}_j$ is not normally distributed"

If $p -$ value $<\alpha$, $H_0$ is rejected, $H_a$ is accepted. Else $H_0$ is taken.

The R package nortest was utilised for the calculation of the Anderson–Darling test.

Once it has been verified that $\widehat{\beta}_j$ $\forall j$ is normally distributed, $t$ -tests [42] were carried out with a level of significance $\alpha$. This allows us to know the contribution of each individual explanatory variable, $X_{ij}$, to the model. The possible hypotheses are as follows:

(i) Null hypothesis $H_0$: "explanatory variable $X_{ij}$ has a slope that is equal to zero, that is, $X_{ij}$ is not useful to predict $Y_i$, $\widehat{\beta}_j = 0$"

(ii) Alternative hypothesis $H_a$: "explanatory variable $X_{ij}$ has a slope that is different from zero, that is, $X_{ij}$ contribute to predict $Y_i$, $\widehat{\beta}_j \neq 0$

The results obtained in the test can be:

(i) If $p-$ value $<\alpha$, $H_0$ is rejected, $H_a$ is taken

(ii) Else $H_0$ is accepted, $H_a$ is rejected

Next, the importance of the predictors is determined using a $t$ statistic estimator, which is defined as the ratio of the estimated parameter $\widehat{\beta}_j$ to the standard error SE $(\widehat{\beta}_j)$ of the estimation,

$$t - \text{statistic } \widehat{\beta}_j = \frac{\widehat{\beta}_j}{\mathrm{SE}(\widehat{\beta}_j)}. \tag{12}$$

For a given SE, the higher the value of the estimator, the higher value of the $t-$ statistic.

If the null hypothesis is accepted, a high estimator produce evidence against it, similar to when the $t-$ statistic is very far from the hypothesized value.

In order to implement the GLM model and to evaluate the importance of the predictors, the caret and vip packages in $R$ are used.

*2.2.2. Topological Representation of PTNs.* A PTN can be represented in a topological space named $L$-Space in which a network is mapped as a graph $G = (N; L)$, where $N$ is the set of nodes symbolizing the stops and $L$ is the set of links established between them. In the $L$-Space, one node represents one stop, and one link means a union between two consecutive stops. This tells us that there is a link between two stops, if one stop is the successor of the other on a route.

*2.2.3. Link Building Process in PTNs.* In each network, it was analyzed whether a GLM could adequately describe the link formation process. As was explained in Section 2.2.1, the caret package in $R$ was used in order to carry out the stages of training and validation of the model. The process was as follows.

The $L$-Space was constructed. All the loops and multiple links from the graph were deleted, obtaining a graph $G\prime$, where the maximal connected components were obtained. Then, with the largest cluster, the giant component (CG), the following operations were performed:

The number of pairs of connected and unconnected nodes were estimated, and several similarity measures were calculated for each one of them. Local, quasilocal, and global methods were applied.

The local similarity indexes used were: Adamic-Adar (dsimaa) [43], common neighbours (dsimcn), cosine (dsimcos) [44], cosine similarity on L+ (dsimcos_l) [45], hub promoted (dsimhpi) [46], jaccard (dsimjaccard) [47], hub depressed (dsimhdi) [3, 7], Leicht–Holme–Newman (dsimlhn_local) [48], preferential attachment (dsimpa) [49], and Sørensen (dsimsor) [50]. The global similarity measures used were: average commute time (dsimact) [37], normalized average commute time (dsimact_n) [51], Katz (dsimkatz) [52], L+ directly (dsiml) [45], matrix forest (dsimmf) [53], and random walk with restart (dsimrwr) [54]. Finally, the quasi-local measures of the similarity utilised were graph distance (dsimdis) and local path (dsimlp) [6, 55]. These indexes are described in detail in the Supplementary materials section.

The model has the values that describe the different similarities between pairs of nodes as input variables (features) and the indication of whether or not there is a link between them as output variable. In order to build the model, supervised learning is used. In this technique, the relations among the input variables (features) and outgoing ones (target) are learnt. That is, from some labeled examples (in each the correct input and output are known), the algorithm that is able to predict the value of the output for new cases not utilised in the learning (training process). For each PTN, a set of data is provided with different features, and the outcome or target (label) is known for each case (pair of nodes). The goal is to predict the label of new cases (pairs of nodes) with the minimum possible error. Since the outcome variable is a categorical value, whether or not a link exists, the prediction corresponds to a binary classification problem.

Crossvalidation is used as a procedure to estimate the model. Instead of splitting the dataset into a training and a test subset, in the crossvalidation mechanism, $k$ equal partitions of the dataset are made. The model is trained $k$ times: each time one of the partitions is taken as a test set, and the model is trained with the rest of the data (with the remaining $k-1$ folds). Each fold is used once as a test set. Finally, several predictions exist about the whole dataset. This process results in $k$ estimates of a parameter related to the effectiveness of the model. An average of an estimated parameter (EP) can be made,

$$\langle \mathrm{EP} \rangle = \frac{1}{k}\sum_{i=1}^{k} \mathrm{EP}. \tag{13}$$

EP can be accuracy (14), area under the curve (AUC) [56], and kappa [57].

These parameters are described as follows:

TP: truth positives, TN: truth negatives, FP: false positives, and FN: false negatives.

$$\text{accuracy} = \frac{\mathrm{TP}+\mathrm{TN}}{\mathrm{TP}+\mathrm{TN}+\mathrm{FP}+\mathrm{FN}}. \tag{14}$$

AUC: AUC represents the probability that a classifier ranks a randomly selected positive instance higher than a randomly chosen negative instance. This EP can be defined, in general terms, as follows, given a binary classification task that has $m$ positive and $n$ negative instances, respectively. The outputs of a binary classifier can be considered as a rigorously ordered list for these instances, which can be appropriately represented by $l_x$, which is an indicator function of a set $X$. Therefore, $c$ is a fixed classifier, where $y_{p1}, \ldots, y_{pm}$ are its outputs on the positive instances and $y_{n1}, \ldots, y_{nn}$ are its outputs on the negative instances. The AUC related to $c$ is described [58] as

$$\text{AUC} = \frac{\sum_{i=1}^{i=m} \sum_{j=1}^{n} l_{y_{pi} > y_{nj}}}{mn}, \quad (15)$$

which is the value of the Wilcoxon–Mann–Whitney statistic [59].

Kappa: this EP is defined as

$$\text{Kappa} = \frac{p_0 - p_c}{1 - p_c}, \quad (16)$$

where $p_0$ = Accuracy and

$$p_c = \frac{\text{TP} + \text{FN}}{\text{TP} + \text{FP} + \text{TN} + \text{FN}} + \frac{\text{FP} + \text{TN}}{\text{TP} + \text{FP} + \text{TN} + \text{FN}} + \frac{\text{TP} + \text{FP}}{\text{TP} + \text{FP} + \text{TN} + \text{FN}} + \frac{\text{TN} + \text{FN}}{\text{TP} + \text{FP} + \text{TN} + \text{FN}}. \quad (17)$$

Finally, an independent end estimation of the accuracy, recall, precision, and specificity of the model can be obtained using the validation set. The last three parameters are

$$\begin{aligned} \text{recall} &= \frac{\text{TP}}{\text{TP} + \text{TN}}, \\ \text{specificity} &= \frac{\text{TN}}{\text{TN} + \text{FP}}, \\ \text{precision} &= \frac{\text{TP}}{\text{TP} + \text{FP}}. \end{aligned} \quad (18)$$

In addition, the confusion matrix as an estimation of the provided solution was obtained in the end validation for each PTN. Table 1 describes the confusion matrix general concept for a binary classification problem.

The final validation was performed on 20% of the total samples.

The selection of the similarity measures to be used as input variables to the model required checking the existing correlation between them. To determine whether this correlation should be estimated using Spearman's or Pearson's method, we checked whether the variables were normally distributed. The Anderson–Darling test [60] was applied with a significance level equal to 0.05. The following hypotheses were used:

(i) $H_0$: "the sample comes from a normal distribution"

(ii) $H_a$: "the sample does not come from a normal distribution"

If $p-$value <0.05, $H_0$ is rejected; otherwise, $H_0$ is accepted.

The *R* package nortest was utilised for the calculation of the Anderson–Darling test.

*2.2.4. Study of the Efficiency.* In a graph, $G$, the distance between the two nodes ($i$ and $j$), $d(i, j)$, is the number of links that form the shortest path between them. If there is no link between $i$ and $j$, then $d(i, j) = \infty$. The efficiency between $i$ and $j$ [60] can be defined as

$$\text{Eff}_{ij} = \frac{1}{d(i, j)}, \quad \forall i \neq j. \quad (19)$$

Table 1: Confusion matrix for a binary classification problem.

| | Actual value (AV) | |
|---|---|---|
| Predicted value (PV) | AV 0 (no link exists) | AV 1 (A link exists) |
| PV 0 (no link exists) | Number of TN | Number of FP |
| PV 1 (A link exists) | Number of FN | Number of TP |

Since $\text{Eff}_{ij}$ is estimated based on the shortest path length between node pairs, an increase in $d(i, j)$ would result in a decrease in the local efficiency between $i$ and $j$.

In addition, the global efficiency of $G$ can be described as

$$\text{GlobEff}(G) = \frac{1}{N(N-1)} \sum_{i \neq j} \text{Eff}_{ij}. \quad (20)$$

This parameter is the average of the efficiencies calculated over all pairs of nodes in $G$. For a given number of nodes $N$, GlobEff $(G)$ increases with the addition of links. According to the previous definition $0 \leq \text{GlobEff}\ (G) \leq 1$, being the value 1 reached for a complete graph [61].

GlobEff $(G)$ has been estimated in several PTNs as one of its features [62, 63]. This research analyses the impact that the elimination of links between pairs of nodes, with certain similarity characteristics, has on the GlobEff of the GC in $G\prime$. The result could help to achieve better network planning, since, depending on which links are removed or built, higher or lower GlobEff can be obtained. Common characteristics regarding efficiency in PTNs are also identified.

The relationship between GlobEff and network density is also analyzed. This last characteristic for undirected graphs such as PTNs can be defined as

$$\text{density} = \frac{2 * \text{number of links in}\ G}{\text{number of nodes} * (\text{number of nodes} - 1)}. \quad (21)$$

*2.2.5. Correlations between Topological Measurements.* Certain investigations have been performed focusing on the study of centrality measures [35] in a PTN. In [36], the authors study some centralities in 58 existing social networks. Further studies examine the correlation between centrality metrics: using Pearson, Spearman, and Kendall methods [37]. The authors use the degree as the base to approximate three other metrics: closeness, betweenness, and eigenvector. They check the correlation between centrality metrics in several real networks, categorized as social, technological, and biological networks. Authors find that the betweenness occupies the highest coefficient, closeness is at the middle level, while eigenvector fluctuates dramatically between networks. They also put forward the idea that rank correlation performs better than the Pearson one in scale-free networks. In [40], several different real-world network graphs, representing several contexts (social club network, birds' social network, word adjacency network, airports network, games network, and related book network) with the number of nodes ranging from 34 to 332, were used. The authors classify the main centrality metrics into two categories: degree-based (degree and eigenvector centralities) and shortest path-based (betweenness, closeness, distance, and eccentricity centralities). They analyze the correlation between the aforementioned centrality metrics, showing that two degree-based centrality metrics (degree and eigenvector centrality) are highly correlated across all the studied networks. There is predominantly a moderate level of correlation between any two of the shortest path-based centrality metrics (betweenness, closeness, distance, and eccentricity). The authors explain that a poor correlation exists between a degree-based centrality metric and a shortest path-based centrality metric for regular random networks. As the variation in the degree distribution of the nodes increases, the correlation coefficient between the two classes of centrality metrics increases. Reference [34] uses a regression model to show a correlative relationship between passenger flow distribution and the conventional network properties (in/out degree, betweenness, and closeness) for the train system in Hague and Amsterdam cities.

Due to the classification, social, technological, and biological networks can encompass networks of very different types, and our investigation focuses on the study of centralities in PTNs. These correlations are studied in $G'$. Specifically, the following centralities are calculated:

(i) The degree of a node $i$, $k(i)$, for an undirected graph, $G$, such as a PTN, is [26, 64]

$$k(i) = \sum_{j=1}^{N} A_{ij}, \tag{22}$$

where

$A_{ij}$ is the element $ij$ of the adjacency matrix, $A$, such as $A_{ij} = 1$, if the node $i$ is linked to node $j$ and 0, otherwise.

(ii) The minimum distance between two nodes $i$, $j$ in $G$, $l$, is the length of the shortest path between them.

(iii) The betweenness centrality of a node $i$ in $G$, $\mathrm{BC}(i)$,is [26, 65]

$$\mathrm{BC}(i) = \sum_{u \neq i \neq w} \frac{\gamma_{u,w}(i)}{\gamma_{u,w}}, \tag{23}$$

where $\gamma_{u,w}$ is the total number of shortest paths from node $u$ to node $w$, and $\gamma_{u,w}(i)$ is the number of those paths that pass through $i$.

(iv) Regarding the eigenvector centrality of a node $i$ in $G$, $\mathrm{EC}(i)$ [26, 65, 66]: $\lambda_1, \lambda_2, \lambda_3, \ldots, \lambda_N$ are the eigenvalues of the adjacency matrix $A = \{A_{ij}\}$ of $G$. Then, the largest eigenvalue of matrix $A$ is $\lambda_{\max}$ with an eigenvector $e = [e_1, e_2, \ldots, e_N]^T$ such that $\lambda_{\max} * e_i = \sum_{j=1}^{N} A_{ij} * e_j$. The eigenvector centrality for node $i$ represented as $\mathrm{EC}(i)$ can be defined as

$$\mathrm{EC}(i) = \frac{1}{\lambda_{\max}} \sum_{j=1}^{N} A_{ij} * e_j. \tag{24}$$

(v) Pagerank, PR, of a node $i$ in $G$, is [26, 66–68]

$$\mathrm{PR}(i) = \frac{q}{N} + (1-q) \sum_{j: j \longrightarrow i} \frac{\mathrm{PR}(j)}{k_{\mathrm{out}}(j)}, \quad i = 1, 2, 3, \ldots, N, \tag{25}$$

where [26]

$N$ is the number of nodes in $G$, $\mathrm{PR}(j)$ is the pagerank of a node $j$, and $k_{\mathrm{out}}(j)$ is the outdegree of node $j$, being the sum of $(\mathrm{PR}(j)/k_{\mathrm{out}}(j))$ executed over the nodes pointing towards $i$. In the case of the PTNs, it is considered that $G$ is an undirected graph; therefore, $k_{\mathrm{out}}(j) = k(j)$.

$q$ is the damping parameter, $\in [0, 1]$.

(vi) A hub is a node that points to many relevant nodes, and an authority node is the one that is focused on by many important nodes. Both are based on the eigenvectors related to the highest eigenvalues of the matrices $AA^T$ and $A^TA$.

The hub centrality of the node $i$, denoted by HC $(i)$, is the $i$-th entry of the following vector $y$ satisfying equation:

$$AA^T y = \lambda y, \quad \text{where } \lambda \in R \text{ is the highest eigenvalue of } AA^T. \tag{26}$$

Similarly, the authority of a node $i$, symbolized by AC $(i)$, is the $i$-th entry of the following vector $x$ satisfying equation:

$$A^T A x = \lambda x, \quad \text{where } \lambda \in R \text{ is the highest eigenvalue of } A^T A. \tag{27}$$

For an undirected graph, such as a PTN, the adjacency matrix $A$ is symmetric. The two scores, $\mathrm{AC}(i)$ and $\mathrm{HC}(i)$, are identical.

## 3. Results and Discussion

*3.1. Link Building Process in PTNs.* As was previously displayed in 2.2.2, the network was represented in the *L-Space*. All loops and multiple links were eliminated, obtaining graph $G\prime$. This is where we calculate the existing maximum number of connected components. Table 2 contains information collected after the explained process, for all analysed networks, the number of links and existing nodes and clusters in $G\prime$. In addition, there are the number of nodes and links present in the largest cluster GC. As well as the fact some of them have several clusters, detection of clusters in cities over PTNs can also allow us to find urban groups, which are strongly connected through transportation. The comparison between PTN clusters and urban agglomerations can be used to estimate whether the PTNs are capable of supporting these human distributions [69]. Identifying under- and overserviced areas can also help in policy decisions, including infrastructure planning and local development [70].

Table 2: Number of nodes, links, clusters, and characteristics of the GC in $G'$ for all analyzed networks.

| Network | Number of nodes | Number of links | Number of clusters | GC | |
|---|---|---|---|---|---|
| | | | | Number of nodes | Number of links |
| AVL, CFL, RGTR, TICE | 1328 | 1924 | 3 | 1370 | 1921 |
| Island Transit | 358 | 420 | 2 | 271 | 313 |
| Lanta | 2150 | 2330 | 1 | 2150 | 2330 |
| Linja-Karjala Oy | 534 | 700 | 1 | 534 | 700 |
| Metlink | 3007 | 3583 | 3 | 2998 | 3574 |
| PPTC, ROPIT | 5152 | 6757 | 18 | 4985 | 6599 |
| Sage | | 36 | 1 | 31 | 36 |
| STAR | 1415 | 1993 | 2 | 1386 | 1965 |
| Thunder Bay Transit | 818 | 885 | 6 | 813 | 885 |
| TransAntofagasta | 645 | 962 | 1 | 645 | 962 |

As was explained in 2.2.1, we used the caret package in R for the building of the model. As described in 2.2.3, the model was trained $k$ times: each time one of the partitions was taken as a test set, and the model was trained with the rest of the data (with the remaining $k-1$ folds). Each fold was used once as a test set. Finally, several predictions exist about the whole dataset. This process results in $k$ estimates of the *accuracy*, *AUC*, and kappa parameters. Additionally, if two similarity measures had a correlation greater than 0.9, one of them was not considered in the prediction. Table 3 shows the similarity indexes that present a Spearman correlation higher than 0.9 with another.

In order to know the method to be used for the calculation of correlations, Pearson or Spearman, the Anderson–Darling test was applied with a significance level $\alpha = 0.05$. All networks showed a $p-$ value <0.05. Therefore, the null hypothesis, $H_0$ was rejected, inferring that the distributions did not follow a normal pattern. Spearman's method was used to calculate correlations.

The importance of each predictor in the model was estimated calculating the absolute value of the $t-$statistics [71], whose definition has been presented in 2.2.1 The importance of predictors is shown in Table 4.

Tables 5 and 6 show, in each PTN, the average of the estimators (accuracy, AUC ,and kappa) calculated over the $k$ times that the model was trained. Since the number of links between pairs of nodes was much lower than the number of unconnected pairs of nodes, the down-sampling approach was utilised, randomly removing the observations. In order to improve the results, artificial balanced samples were generated according to a smoothed bootstrap procedure [60] in the Thunder Bay Transit network. The rose package in R was used.

Table 7 shows, in each network, the confusion matrix [72] obtained in the final validation. In Table 8, accuracy, recall, precision, and specificity parameters are presented.

All networks showed good results applying down-sampling, according to the parameters chosen for the evaluation of the model. In the crossvalidation process, average accuracy and AUC values were higher than 0.99 and kappa larger than 0.93. In the validation stage, accuracy and recall showed values higher than 0.99, and specificity had a value equal to 1. The only exception was the Thunder Bay Transit network, where it was necessary to apply the rose method in order to achieve better kappa and precision values.

As a result, the process of building links was appropriately modeled using a GLM, which had some measures of similarities between nodes as input variables. The response variable, which establishes the existence or not of a link between pairs of nodes, is appropriately described by a binomial probability density function. The link function used is the logit function, as we explained in 2.2.1. The model has the novelties described in Section 1, with respect to other models that have already been developed for PTNs.

In most networks, the figure with the highest influence was dsimdis, followed by simact. In addition, the simcos_l and simlp showed high or moderate importance in some networks.

*3.2. Study of Trip Times.* The trip times are analyzed in order to estimate things in common between networks. Several statistical parameters are calculated (average, standard deviation, median, moda, maximum, and minimum values). The results and the frequency distribution are displayed in Table 9 and Figure 1, respectively.

The cumulative probability distributions are also checked. They are shown in Figure 2. The stats package in R was used. The similarity between two distributions is examined, applying the Kolmogorov–Smirnov test [73]. A significance level equal to 0.05 is taken, while the following hypotheses are considered:

(i) Null hypothesis ($H_0$): "the samples come from the same distribution."

(ii) Alternative hypothesis ($H_a$):"the samples come from different distributions."

If a $p-$value <0.05 is obtained in the test, the null hypothesis is rejected. Table S.1 shows the results obtained in the test.

It can be noted that similarities do not exist between the PTNs in relation to the trip times. All networks presented a high standard deviation. The lowest is 14.02 minutes (790.23523 seconds) and the highest is 11.12 hours (42,027.19610 seconds). This shows that the size, the complexity, and variability of available routes in the PTNs cause trip times to be highly inconsistent between routes. Trip times allow the evaluation of how travelers choose a service based on whether or not it is convenient. Trip times have

Table 3: Similarity measures that present a Spearman correlation higher than 0.9 with another.

| Network | Similarity measures |
| --- | --- |
| AVL, CFL, RGTR, TICE | dsimcn, dsimcos, dsimhdi, dsimhpi, dsimjaccard, dsimlhn_local, dsimsor, dsiml, dsimkatz, dsimmf, dsimrwr. |
| Island Transit | dsimcn, dsimcos, dsimhdi, dsimhpi, dsimjaccard, dsimlhn_local, dsimsor, dsimact_n, dsimdis, dsimkatz, dsimmf, dsimrwr, dsiml |
| Lanta | dsimcn,d simcos, dsimhdi, dsimhpi, dsimjaccard, dsimlhn_local, dsimsor, dsimact_n, dsiml, dsimkatz, dsimmf, dsimrwr |
| Linja-Karjala Oy | dsimcn, dsimcos, dsimhdi, dsimhpi, dsimjaccard, dsimlhn_local, dsimsor, dsiml, dsimkatz, dsimmf, dsimrwr |
| Metlink | dsimcn, dsimcos, dsimhdi, dsimhpi, dsimjaccard, dsimlhn_local, dsimsor, dsimact_n, dsiml, dsimkatz, dsimmf, dsimrwr |
| PPTC, ROPIT | dsimcn, dsimcos, dsimhdi, dsimhpi, dsimjaccard, dsimlhn_local, dsimsor, dsimact_n, dsiml, dsimkatz, dsimmf, dsimrwr |
| Sage | dsimcn, dsimcos, dsimhdi, dsimhpi, dsimjaccard, dsimlhn_local, dsimsor, dsimact_n, dsimdis, dsimkatz, dsimmf, dsimrwr Dsiml |
| STAR | dsimcn, dsimcos, dsimhdi, dsimhpi, dsimjaccard, dsimlhn_local, dsimsor, dsiml, dsimkatz, dsimmf, dsimrwr |
| Thunder Bay Transit | dsimcn, dsimcos, dsimhdi, dsimhpi, dsimjaccard, dsimlhn_local, dsimsor, dsimact_n, dsiml, dsimkatz, dsimmf, dsimrwr |
| TransAntofagasta | dsimcn, dsimcos, dsimhdi, dsimhpi, dsimjaccard,d simlhn_local, dsimsor, dsiml, dsimkatz, dsimmf, dsimrwr |

Table 4: Importance of predictors.

| Network | Similarity measures | Importance | Network | Similarity measures | Importance |
| --- | --- | --- | --- | --- | --- |
| AVL, CFL, RGTR, TICE | Dsimdis | 100 | Island Transit | Dsimact | 100 |
| | Dsimlp | 13.20776 | | dsimcos_l | 53.78080 |
| | Dsimpa | 11.61152 | | Dsimlp | 7.17716 |
| | Dsimaa | 10.37026 | | Dsimpa | 6.75717 |
| | dsimcos_l | 9.554070 | | Dsimaa | 0 |
| | dsimact_n | 4.423947 | | | |
| | Dsimact | 0 | | | |
| Lanta | dsimcos_l | 100 | Linja-Karjala Oy | Dsimdis | 100 |
| | Dsimdis | 94.94621 | | Dsimlp | 25.23217 |
| | Dsimpa | 10.94917 | | Dsimact | 23.95510 |
| | Dsimlp | 7.372249 | | Dsimaa | 17.12581 |
| | Dsimaa | 6.737758 | | Dsimpa | 12.22195 |
| | Dsimact | 0 | | dsimact_n | 4.047694 |
| | | | | dsimcos_l | 0 |
| Metlink | Dsimdis | 100 | PPTC, ROPIT | Dsimdis | 100 |
| | dsimcos_l | 48.71324 | | Dsimlp | 19.90804 |
| | Dsimaa | 10.70750 | | dsimcos_l | 14.18266 |
| | Dsimpa | 7.501585 | | Dsimact | 11.82702 |
| | Dsimact | 2.410628 | | Dsimaa | 3.308090 |
| | Dsimlp | 0 | | dsimact_n | 1.716437 |
| | | | | Dsimpa | 0 |
| Sage | Dsimact | 100 | STAR | Dsimdis | 100 |
| | dsimcos_l | 26.16491 | | Dsimlp | 9.89944 |
| | Dsimlp | 16.60197 | | Dsimact | 7.32344 |
| | Dsimpa | 11.17248 | | Dsimaa | 5.10874 |
| | Dsimaa | 4.439547 | | Dsimpa | 3.83377 |
| | Dsimhpi | 0 | | dsimcos_l | 0.10442 |
| | | | | dsimact_n | 0 |
| Thunder Bay Transit | Dsimdis | 100 | TransAntofagasta | Dsimdis | 100 |
| | dsimcos_l | 40.08267 | | Dsimact | 55.94523 |
| | Dsimaa | 38.62297 | | Dsimlp | 48.88970 |
| | Dsimlp | 34.96505 | | dsimcos_l | 41.40694 |
| | Dsimpa | 24.13330 | | dsimact_n | 28.06088 |
| | Dsimact | 0 | | Dsimaa | 27.07529 |
| | | | | Dsimpa | 0 |

Table 5: In AVL, CFL, RGTR, TICE, Island Transit, Lanta, Linja-Karjala Oy, Metlink networks, the average estimation of accuracy, AUC, and kappa calculated over the $k$ times in which the model was trained.

| Network | Training approach | Accuracy | AUC | Kappa |
|---|---|---|---|---|
| AVL, CFL, RGTR, TICE | Downsampling | 0.99996 | 1 | 0.98208 |
| Island Transit | Downsampling | 0.99986 | 1 | 0.98406 |
| Lanta | Downsampling | 1 | 1 | 1 |
| Linja-Karjala Oy | Downsampling | 1 | 1 | 1 |
| Metlink | Downsampling | 0.99994 | 1 | 0.93209 |

Table 6: In PPTC, ROPIT, Sage, STAR, Thunder Bay Transit, TransAntofagasta networks, the average estimation of accuracy, AUC, and kappa calculated over the $k$ times in which the model was trained.

| Network | Training approach | Accuracy | AUC | Kappa |
|---|---|---|---|---|
| PPTC, ROPIT | Downsampling | 1 | 1 | 0.99886 |
| Sage | | 1 | 1 | 1 |
| STAR | Downsampling | 1 | 1 | 1 |
| Thunder bay Transit | Downsampling | 0.99933 | 0.99960 | 0.79877 |
| | Smoothed bootstrap | 0.99999 | 1 | 0.99718 |
| TransAntofagasta | Downsampling | 1 | 1 | 1 |

Table 7: Final validation. Confusion matrix.

| Network | Training approach | Confusion matrix | | |
|---|---|---|---|---|
| | | | Actual value | |
| | | Predicted value | AV 0 | AV 1 |
| AVL, CFL, RGTR, TICE | Downsampling | PV 0 | 349548 | 0 |
| | | PV 1 | 14 | 384 |
| | | | Actual value | |
| | | Predicted value | AV 0 | AV 1 |
| Island Transit | Downsampling | PV 0 | 14506 | 0 |
| | | PV 1 | 2 | 62 |
| | | | Actual value | |
| | | Predicted value | AV 0 | AV 1 |
| Lanta | Downsampling | PV 0 | 923138 | 0 |
| | | PV 1 | 0 | 466 |
| | | | Actual value | |
| | | Predicted value | AV 0 | AV 1 |
| Linja-Karjala Oy | Downsampling | PV 0 | 56644 | 0 |
| | | PV 1 | 0 | 140 |
| | | | Actual value | |
| | | Predicted value | AV 0 | AV 1 |
| Metlink | Downsampling | PV 0 | 1795467 | 0 |
| | | PV 1 | 104 | 714 |
| | | | Actual value | |
| | | Predicted value | AV 0 | AV 1 |
| PPTC, ROPIT | Downsampling | PV 0 | 4966405 | 0 |
| | | PV 1 | 3 | 1319 |
| | | | Actual value | |
| | | Predicted value | AV 0 | AV 1 |
| STAR | Downsampling | PV 0 | 383136 | 0 |
| | | PV 1 | 0 | 393 |
| | | | Actual value | |
| | | Predicted value | AV 0 | AV 1 |
| Thunder Bay Transit | Downsampling | PV 0 | 131588 | 0 |
| | | PV 1 | 89 | 177 |
| | | Predicted value | AV 0 | AV 1 |
| | Smoothed bootstrap | PV 0 | 131676 | 0 |
| | | PV 1 | 1 | 177 |

Table 7: Continued.

| Network | Training approach | Confusion matrix | | |
|---|---|---|---|---|
| | | | Actual value | |
| | | Predicted value | AV 0 | AV 1 |
| TransAntofagasta | Downsampling | PV 0 | 82691 | 0 |
| | | PV 1 | 0 | 192 |
| | | | Actual value | |
| | | Predicted value | AV 0 | AV 1 |
| Sage | Downsampling | PV 0 | 171 | 0 |
| | | PV 1 | 0 | 7 |

Table 8: Final validation. Accuracy, recall, precision, and specificity parameters.

| Network | Training approach | Accuracy | Recall | Precision | Specificity |
|---|---|---|---|---|---|
| AVL, CFL, RGTR, TICE | Downsampling | 0.99999 | 1 | 0.96482 | 1 |
| Island Transit | Downsampling | 0.99986 | 1 | 0.96875 | 1 |
| Lanta | Downsampling | 1 | 1 | 1 | 1 |
| Linja-Karjala Oy | Downsampling | 1 | 1 | 1 | 1 |
| Metlink | Downsampling | 0.9999421 | 1 | 0.87286 | 1 |
| PPTC, ROPIT | Downsampling | 1 | 1 | 0.99773 | 1 |
| STAR | Downsampling | 1 | 1 | 1 | 1 |
| Thunder Bay Transit | Downsampling | 0.99933 | 1 | 0.66541 | 1 |
| | Smoothed bootstrap | 0.99999 | 1 | 0.99438 | 1 |
| TransAntofagasta | | 1 | 1 | 1 | 1 |
| Sage | | 1 | 1 | 1 | 1 |

Table 9: Trip time metrics (seconds).

| Network | Average | Standard deviation | Median | Max | Moda | Min |
|---|---|---|---|---|---|---|
| AVL, CFL, RGTR, TICE | 1,731.33538 | 790.23523 | 1800 | 4020 | 2280 | 60 |
| Island Transit | 5,385.63025 | 1,298.01961 | 2100 | 69000 | 2400 | 120 |
| Lanta | 3,532.24305 | 1,505.97579 | 3360 | 10140 | 1800 | 420 |
| Linja-Karjala Oy | 2,071.65844 | 841.08391 | 2040 | 4140 | 1260 | 60 |
| Metlink | 2,111.43848 | 1,293.83504 | 1980 | 43800 | 1500 | 240 |
| PPTC, ROPIT | 1,911.43488 | 1,087.87538 | 1820 | 30180 | 960 | 60 |
| Sage | 12,000 | 7,842.19357 | 11850 | 24000 | 1200 | 1200 |
| STAR | 43,533.94450 | 42,027.19610 | 83460 | 86340 | 84960 | 60 |
| Thunder bay Transit | 1,751.91089 | 892.014494 | 1380 | 4500 | 1200 | 300 |
| TransAntofagasta | 10,297.08240 | 1297.94464 | 10254 | 13244 | 9827 | 5999 |

been considered by some researchers to evaluate the performance of PTNs [74, 75].

### 3.3. Study of Efficiency

*3.3.1. Local Efficiency.* All networks showed a large majority of nodes with low local efficiency ≤0.20, as can be noted in Figures 3 and 4.

As was done with trip times, the similarity between local efficiency distributions is examined, applying the Kolmogorov–Smirnov test. A significance level equal to 0.05 is taken, resulting in the following hypotheses being considered:

(i) Null hypothesis ($H_0$): "the samples come from the same distribution."

(ii) Alternative hypothesis ($H_a$): "the samples come from different distributions."

If the $p-$value obtained in the test is <0.05, the null hypothesis is rejected.

The networks presented high analogies in the cumulative distributions of local efficiency. The test yielded a $p-$value >0.05 in all pairwise comparisons performed, as can be appreciated in Table S.2. Therefore, in general, if a stop is unavailable, the remaining connections between its neighbours are distinct from direct connections. This is revealed by the low value of the local efficiency [76].

*3.3.2. Global Efficiency.* The calculation of the GlobEff was carried out in the GC of $G\prime$, and it can be observed, according to the results depicted in Table 10, that the higher the density of $G\prime$, the higher the GlobEff.

Most of the analyzed networks presented a GlobEff of small value (<0.20). Some pieces of research use the GlobEff as a parameter to compare PTNs [77, 78], and others apply it to identify hubs [79, 80]. Consequently, the degree of a node

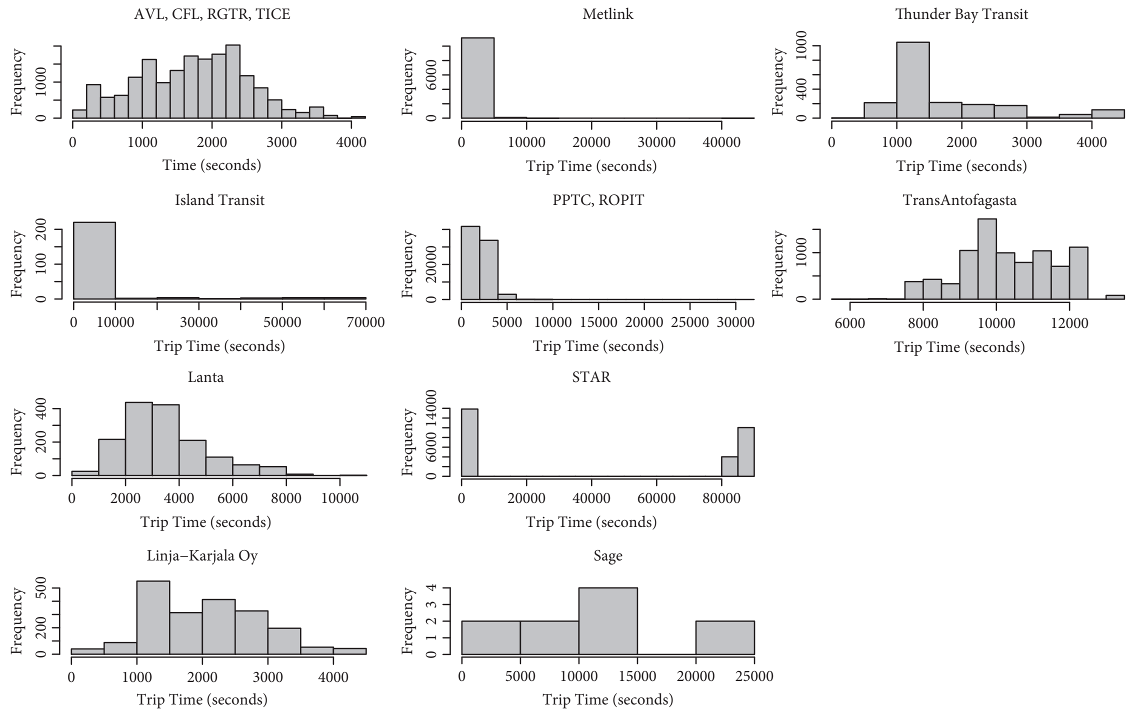



Figure 1: Histogram of trip time.

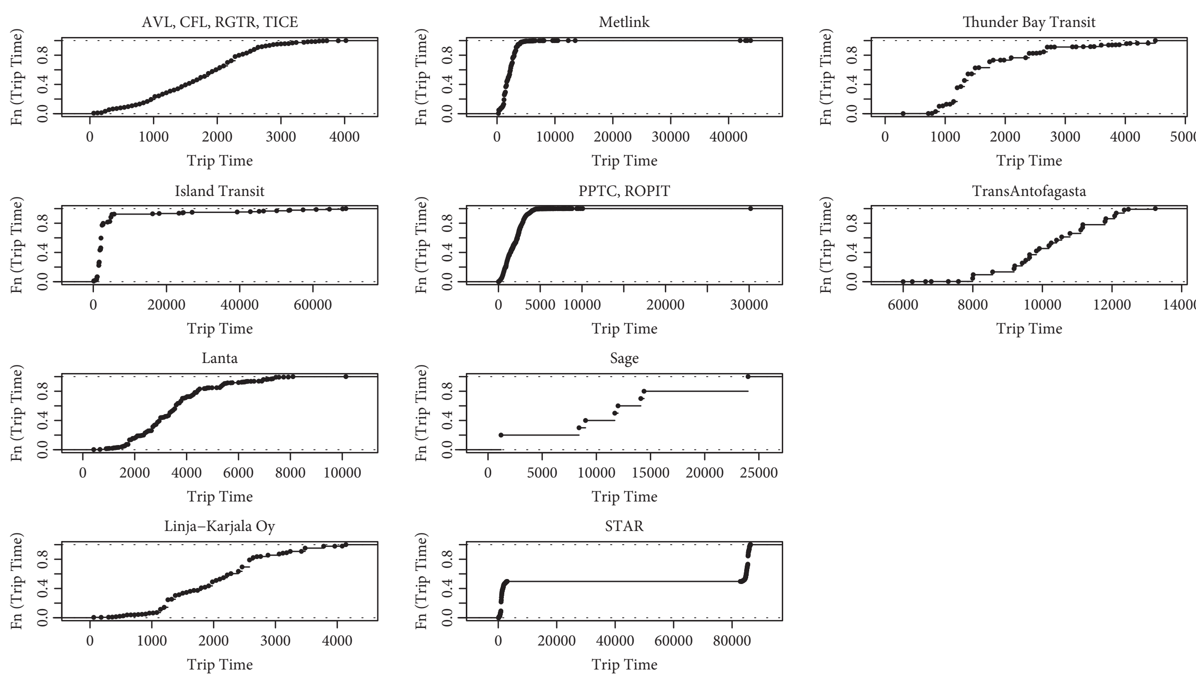



Figure 2: Cumulative probability distribution of trip times.

is ranked by comparing the changes in PTN efficiency after eliminating the node. In contrast, this research analyses the variation in GlobEff when links with certain similarity characteristics were removed. The results are shown in Table 11. Similarity measures with a correlation higher than 0.9 with another were not considered. It can be noted that in most of the networks, the link deletion in which a 75% reduction was reached most quickly was dsimpa and dsimlp, and the one that took the longest to reach was dsimcos_l. Figures 5–7 show the variation in GlobEff when certain links are removed.

Table 11 shows, for each similarity measure, the number of removed links that causes the reduction of GlobEff by 75%.

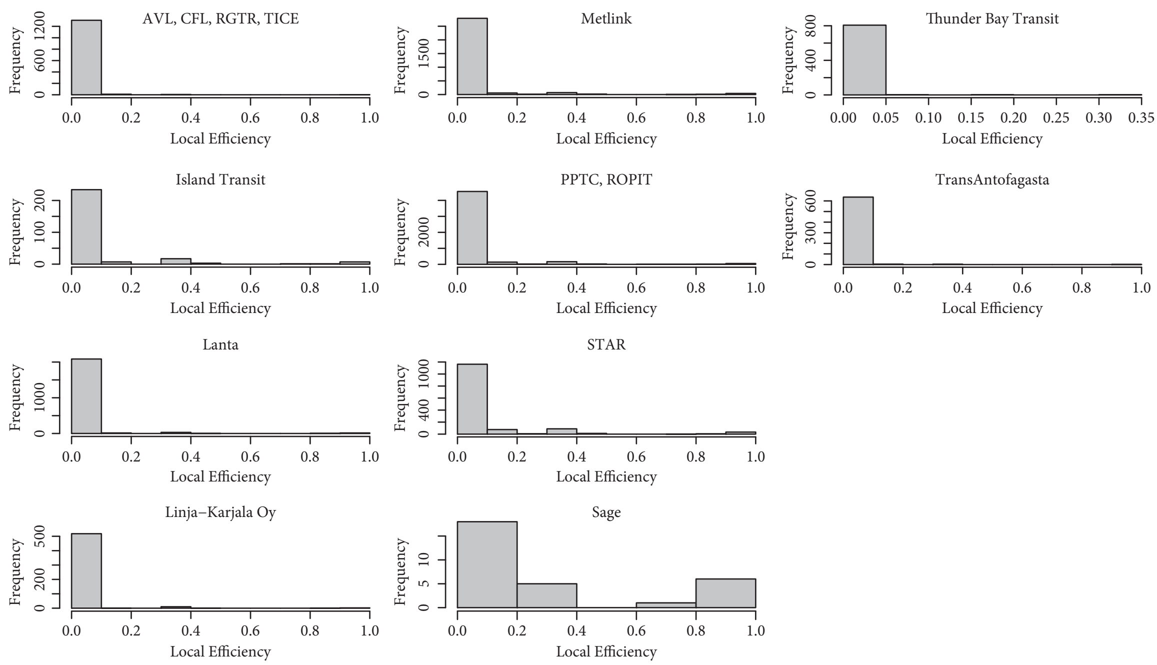



Figure 3: Histogram of local efficiency.

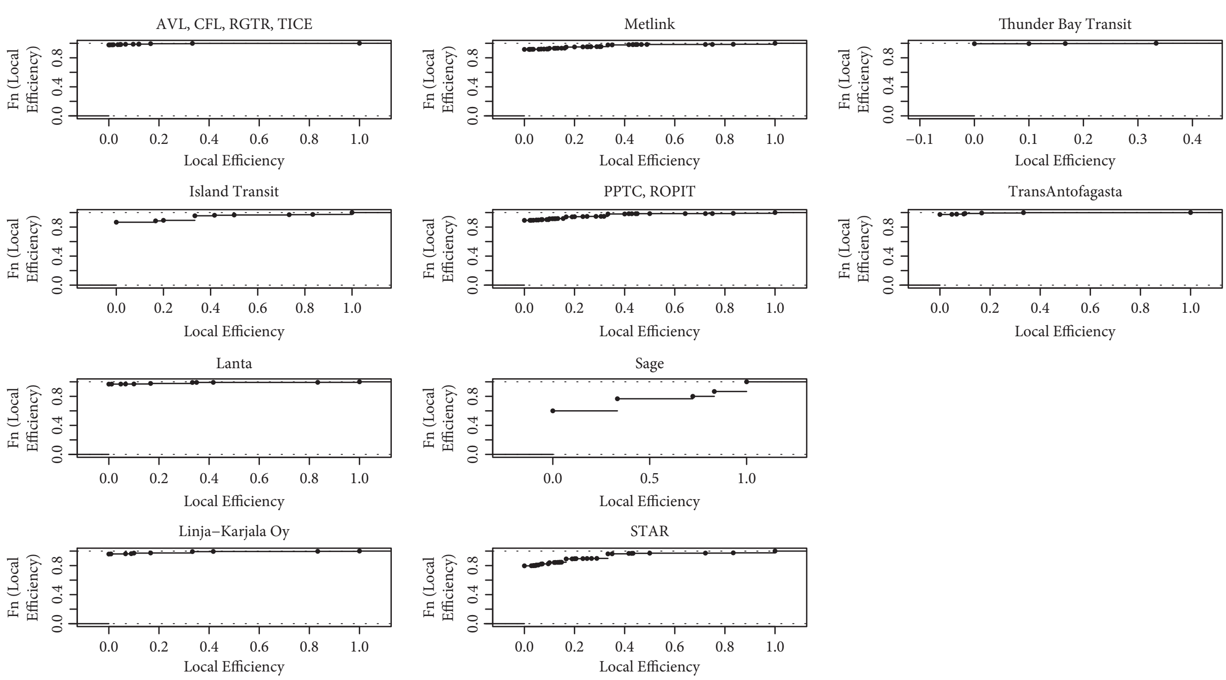



Figure 4: Cumulative probability distribution of local efficiency.

*3.3.3. Correlations between Topological Measurements.* The eigenvector, betweenness, pagerank, degree, hub, and authority centralities were calculated in $G\prime$, in order to study the correlation between them. The correlation of these variables with the amount of transport arriving and departing weekly from a stop were also estimated. Enabling us to know which method, Pearson or Spearman, should be used in the calculation, the Anderson–Darling test with a significance level $\alpha = 0.05$ was applied. In this way, it could be known whether or not the variables were normally distributed. The test yielded a $p$ – value <0.05 for all variables, so the null hypothesis $H_0$ was rejected, and the alternative hypothesis $H_a$ was accepted.

The correlations obtained by applying Spearman's method are shown in Tables S.3–S.12. In all networks, the eigenvector centrality presented a strong correlation with hub and authority centralities. Pagerank showed a moderate, high, or very high correlation with the degree. Therefore,

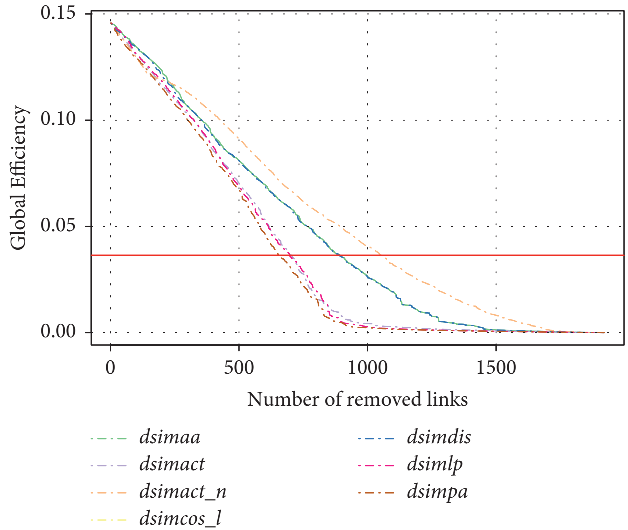
Table 10: GlobEff and density in GC in $G'$ for all analyzed networks.

| Network | GlobEff (GC in $G'$) | Density |
|---|---|---|
| AVL, CFL, RGTR, TICE | 0.14604 | 0.00219 |
| Island Transit | 0.07901 | 0.00856 |
| Lanta | 0.03248 | 0.00101 |
| Linja-Karjala Oy | 0.15180 | 0.00492 |
| Metlink | 0.05731 | 0.00080 |
| PPTC, ROPIT | 0.05291 | 0.00053 |
| Sage | 0.24383 | 0.07742 |
| STAR | 0.11220 | 0.00205 |
| Thunder Bay Transit | 0.07584 | 0.00268 |
| TransAntofagasta | 0.16649 | 0.00463 |

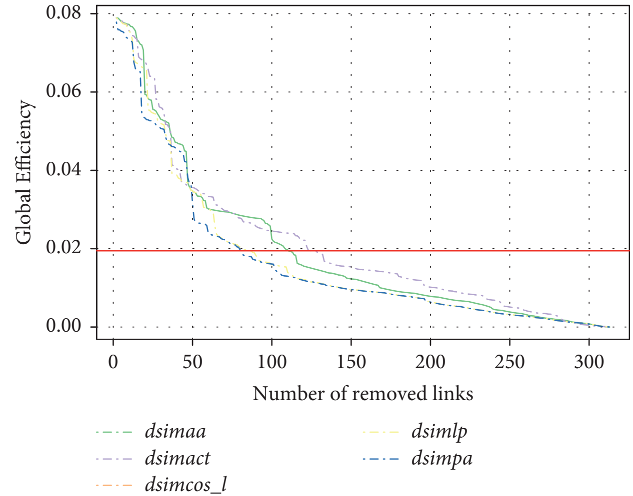
Table 11: Number of removed links that cause a 75% of reduction in the GlobEff.

| Network | Similarity measures | Number of removed links | Network | Similarity measures | Number of removed links |
|---|---|---|---|---|---|
| AVL, CFL, RGTR, TICE | Dsimpa | 655 | Island transit | dsimpa | 80 |
| | Dsimlp | 697 | | dsimlp | 83 |
| | Dsimact | 702 | | dsimaa | 110 |
| | Dsimdis | 890 | | dsimact | 123 |
| | Dsimaa | 892 | | dsimcos_l | 195 |
| | dsimact_n | 1052 | | | |
| | dsimcos_l | 1084 | | | |
| Lanta | Dsimpa | 177 | Linja-Karjala Oy | dsimpa | 181 |
| | Dsimlp | 203 | | dsimlp | 206 |
| | Dsimact | 358 | | dsimact | 221 |
| | Dsimaa | 549 | | dsimdis | 302 |
| | Dsimdis | 616 | | dsimaa | 303 |
| | dsimcos_l | 1450 | | dsimact_n | 395 |
| | | | | dsimcos_l | 399 |
| Metlink | Dsimpa | 1,193 | PPTC, ROPIT | dsimpa | 1500 |
| | Dsimlp | 1,217 | | dsimlp | 1730 |
| | Dsimact | 1,718 | | dsimact | 1850 |
| | dsimcos_l | 2,202 | | dsimaa | 2550 |
| | Dsimaa | 2,493 | | dsimdis | 2550 |
| | Dsimdis | 2,538 | | dsimcos_l | 3750 |
| Sage | Dsimlp | 20 | STAR | dsimact | 545 |
| | Dsimpa | 20 | | dsimpa | 640 |
| | Dsimaa | 21 | | dsimlp | 714 |
| | Dsimact | 21 | | dsimdis | 1,050 |
| | dsimcos_l | 24 | | dsimaa | 1,053 |
| | | | | dsimact_n | 1,172 |
| | | | | dsimcos_l | 1,198 |
| Thunder Bay Transit | Dsimpa | 89 | TransAntofagasta | dsimlp | 386 |
| | Dsimlp | 93 | | dsimact | 398 |
| | Dsimact | 118 | | dsimaa | 512 |
| | Dsimdis | 327 | | dsimact_n | 535 |
| | Dsimaa | 393 | | dsimdis | 550 |
| | dsimcos_l | 422 | | dsimcos_l | 553 |

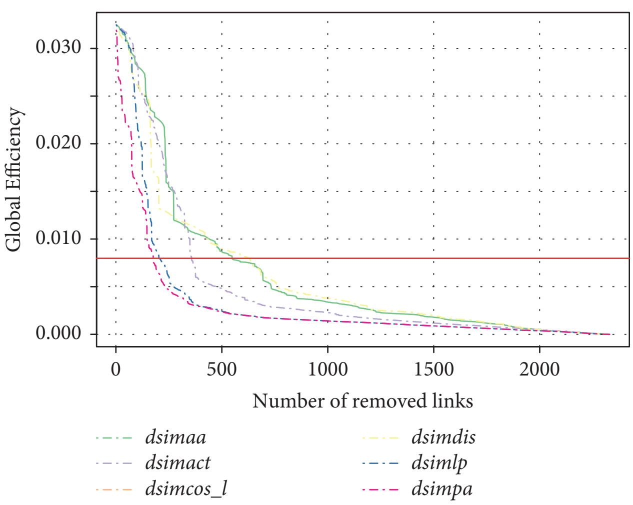
also in this network, a high degree usually has a significant influence. The pagerank and degree only presented a moderate or high correlation with betweenness in some networks, demonstrating that specifically in these few networks a node with a high degree also usually presents an important level of connectivity. Eigenvector and degree, in most networks, exhibited a low or very low correlation. Furthermore, the number of weekly buses arriving and departing from a bus stop showed no strong correlation with any of the centrality measures. Strong correlations between



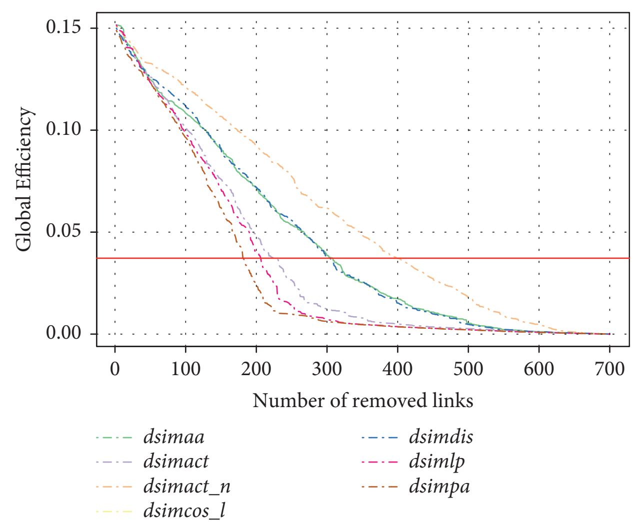

(a)




(b)

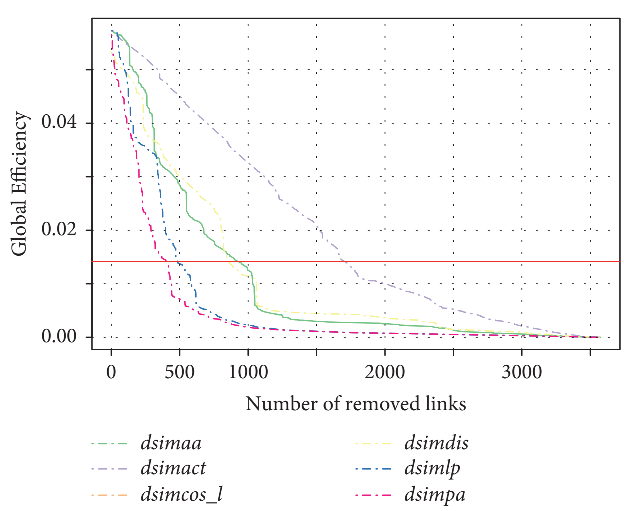



(c)

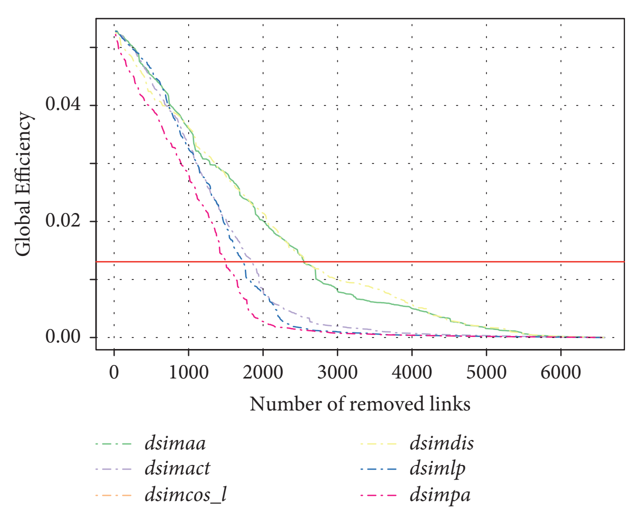



(d)

Figure 5: Variation in global efficiency when links with certain similarity characteristics are removed in AVL, CFL, RGTR, and TICE (a), Island Transit (b), Lanta (c), and Linja-Karjala Oy (d) networks.

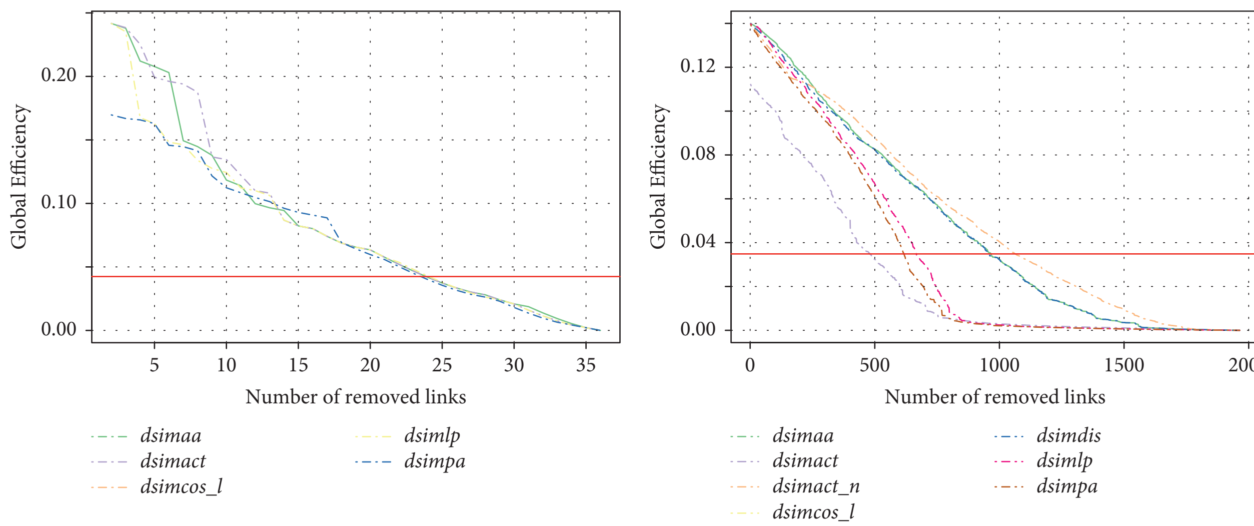



(a)



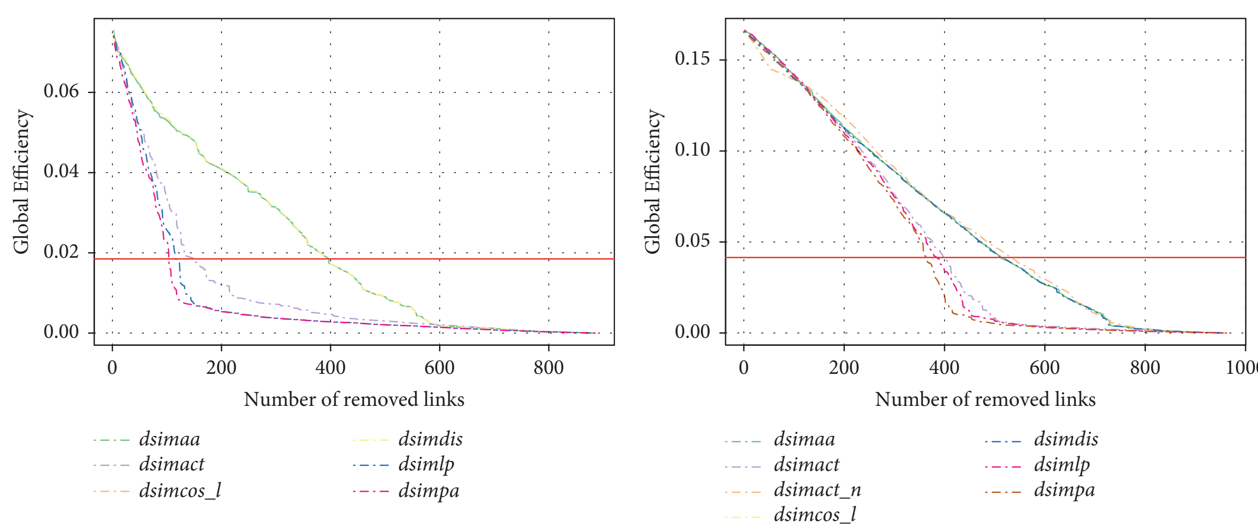

(b)

Figure 6: Continued.




(c) (d)

Figure 6: Variation in global efficiency when links with certain similarity characteristics are removed in Metlink (a), PPTC, ROPIT (b), Sage (c), and STAR (d) networks.




(a) (b)

Figure 7: Variation in global efficiency when links with certain similarity characteristics are removed in Thunder Bay Transit (a) and TransAntofagasta (b) networks.

degree and pagerank and degree and betweenness have also been found in some Chinese PTNs [78].

## 4. Conclusions

Regarding the model followed by the formation of links between stops, this research shows that it can be correctly explained through a generalized linear model, which has certain similarity measures as input variables. Although the similarity measures that explain the model are different among networks, in most of them, dsimdis has a higher significance. It has a value equal to 100. In addition, dsimcos_l and dsimlp presented relevant importance in some PTNs with values higher than 30. Additionally, dsimact and dsimpa showed values equal to 100 and larger than 10, respectively, in certain PTNs.

Regarding travel times, these showed a high variability between networks (with standard deviations greater than 790.23 seconds), as well as very different cumulative probability distributions ($p-$value $\geq 0.05$ in Kolmogorov–Smirnov test).

The study of local efficiency reveals that its cumulative distributions have strong analogies in all network distributions (Kolmogorov–Smirnov test showed $p-$values

<0.05). The local efficiency showed values ≤0.2 in the most of PTNs. Similarly, the overall efficiency exhibited reduced values (≤0.25). This seems to be a common feature of PTNs.

With respect to the centrality measures, they did not show correlation with the flow of vehicles, suggesting that traffic dynamics in the network may be strongly influenced by other different parameters as opposed to topological ones. In all networks, strong correlations of the eigenvector centrality with the hub and authority centralities were detected (with values higher than 0.80). The pagerank showed moderate, high, or very high correlation with the degree (it was larger than 0.5 in all networks). Therefore, these correlation characteristics seem to be a commonality in PTNs.

This research can be continued with a detailed study on the interactions between the different existing modes of transport modes in the cities. A multimodal transportation system, embodied as a multiplex network, can be considered in order to face the problem of urban mobility. In a multiplex network, a node symbolizes a specific origin/destination stop, which exists in each of the network layers. Nevertheless, the links are represented by a different layer of interaction determined by the type of transportation mode used for connecting two nodes.

## Data Availability

Information of stops, routes and trip times of AVL, CFL, RGTR, and TICE; Island Transit; Lanta; Linja-Karjala Oy; Metlink; PPTC, ROPIT; Sage; STAR; Thunder Bay Transit; and TransAntofagasta were retrieved from the operating companies' public web sites, the Deconet Public Transport Network Data, and GTFS Data Exchange repositories.

## Conflicts of Interest

The authors declare that there are no conflicts of interest regarding the publication of this paper.

## Acknowledgments

This work was partially funded by Telefonica Chair at Francisco de Vitoria University.

## Supplementary Materials

Supplementary Material includes (i) description of similarity measures (local, global, and quasilocal methods), (ii) tables related to the study of the trip times, (iii) tables regarding analysis of the local efficiency, and (iv) tables related to correlations between centrality measures. (*Supplementary Materials*)

## References

[1] I. Ahmad, M. U. Akhtar, S. Noor, and A. Shahnaz, "Missing link prediction using common neighbor and centrality based parameterized algorithm," *Scientific Reports*, vol. 10, no. 1, pp. 1–9, 2020.

[2] V. Martínez, F. Berzal, and J.-C. Cubero, "A survey of link prediction in complex networks," *ACM Computing Surveys*, vol. 49, no. 4, pp. 1–33, 2017.

[3] L. Lu and T. Zhou, "Link prediction in complex networks: a survey," *Physica Statistical Mechanics and Its Applications A*, vol. 390, no. 6, pp. 1150–1170, 2011.

[4] D. Malhotra and R. Goyal, "Link prediction in complex networks using information-theoretic measures," *Journal of Complex Networks*, vol. 8, no. 4, 2020.

[5] F. Tan, Y. Xia, and B. Zhu, "Link prediction in complex networks: a mutual information perspective," *PLoS One*, vol. 9, Article ID e107056, 2014.

[6] T. Zhou, L. Lü, and Y.-C. Zhang, "Predicting missing links via local information," *The European Physical Journal B*, vol. 71, no. 4, pp. 623–630, 2009.

[7] T. Zhou, L. Lü, and Y.-C. Zhang, "Predicting missing links via local information," *The European Physical Journal B*, vol. 71, no. 4, pp. 623–630, 2009.

[8] F. Gao, K. Musial, C. Cooper, and S. Tsoka, "Link prediction methods and their accuracy for different social networks and network metrics," *Scientific Programming*, vol. 2015, Article ID 172879, 13 pages, 2009.

[9] B. Zhu and Y. Xia, "Link prediction in weighted networks: a weighted mutual information model," *PLoS One*, vol. 11, pp. 1–13, 2016.

[10] Y. Yang, J. Zhang, X. Zhu, J. Ma, and X. Su, "Link prediction based on the tie connection strength of common neighbor," *International Journal of Modern Physics C*, vol. 30, no. 11, Article ID 1950089, 2019.

[11] S. Bai, L. Li, J. Cheng, S. Xu, and X. Chen, "Predicting missing links based on a new triangle structure," *Complexity*, vol. 2018, Article ID 7312603, 11 pages, 2018.

[12] A. Mohan, R. Venkatesan, and K. V. Pramod, "A scalable method for link prediction in large real world networks," *Journal of Parallel and Distributed Computing*, vol. 109, pp. 89–101, 2017.

[13] S. Gorlatch, U. Banerjee, R. De Nicola et al., "BSP (Bulk synchronous parallelism)," in *Encyclopedia of Parallel Computing*, D. Padua, Ed., Springer, Boston, MA, USA, pp. 192–199, 2011.

[14] A. M. Abdolhosseini-Qomi, S. H. Jafari, A. Taghizadeh, N. Yazdani, M. Asadpour, and M. Rahgozar, "Link prediction in real-world multiplex networks via layer reconstruction method," *Royal Society Open Science*, vol. 7, no. 7, Article ID 191928, 2020.

[15] N. N. Daud, S. H. Ab Hamid, M. Saadoon, F. Sahran, and N. B. Anuar, "Applications of link prediction in social networks: a review," *Journal of Network and Computer Applications*, vol. 166, Article ID 102716, 2020.

[16] T. Wang, X.-S. He, M.-Y. Zhou, and Z.-Q. Fu, "Link prediction in evolving networks based on popularity of nodes," *Scientific Reports*, vol. 7, no. 1, 2017.

[17] C. A. Bliss, M. R. Frank, C. M. Danforth, and P. S. Dodds, "An evolutionary algorithm approach to link prediction in dynamic social networks," *Journal of Computational Science*, vol. 5, no. 5, pp. 750–764, 2014.

[18] N. Hansen and A. Ostermeier, "Completely derandomized self-adaptation in evolution strategies," *Evolutionary Computation*, vol. 9, no. 2, pp. 159–195, 2001.

[19] N. Hansen, "The CMA evolution strategy: a comparing review," in *Towards a New Evolutionary Computation. Studies in Fuzziness and Soft Computing*, J. A. Lozano, P. Larrañaga, I. Inza, and E. Bengoetxea, Eds., vol. 192, pp. 75–102, Springer, Berlin, Germany, 2006.

[20] J. R. Bock and D. A. Gough, “Predicting protein-protein interactions from primary structure,” *Bioinformatics*, vol. 17, no. 5, pp. 455–460, 2001.

[21] L. Lü, C.-H. Jin, and T. Zhou, “Similarity index based on local paths for link prediction of complex networks,” *Physical Review E*, vol. 80, no. 4, Article ID 046122, 2009.

[22] C. Lei and J. Ruan, “A novel link prediction algorithm for reconstructing protein-protein interaction networks by topological similarity,” *Bioinformatics*, vol. 29, no. 3, pp. 355–364, 2013.

[23] X. Li, “Research on passenger flow forecast of urban rail transit,” *E3S Web of Conferences*, vol. 213, Article ID 03026, 2020.

[24] G. Yeboah, C. D. Cottrill, J. D. Nelson, D. Corsar, M. Markovic, and P. Edwards, “Understanding factors influencing public transport passengers’ pre-travel information-seeking behaviour,” *Public Transport*, vol. 11, no. 1, pp. 135–158, 2019.

[25] C. von Ferber, T. Holovatch, Y. Holovatch, and V. Palchykov, “Public transport networks: empirical analysis and modeling,” *The European Physical Journal B*, vol. 68, no. 2, pp. 261–275, 2009.

[26] M. L. Mouronte-López, “Analysing the vulnerability of public transport networks,” *Journal of Advanced Transportation*, vol. 2021, Article ID 5513311, 22 pages, 2021.

[27] D. Levinson and W. Chen, “Area-based models of highway growth,” *Journal of Urban Planning and Development*, vol. 133, no. 4, pp. 250–254, 2007.

[28] O. Cats, A. Vermeulen, M. Warnier, and H. van Lint, “Modelling growth principles of metropolitan public transport networks,” *Journal of Transport Geography*, vol. 82, Article ID 102567, 2020.

[29] M. L. Mouronte and R. M. Benito, “Structural properties of urban bus and subway networks of Madrid,” *Networks and Heterogeneous Media*, vol. 7, no. 3, pp. 415–428, 2012.

[30] B. P. V. Samson, G. A. T. Velez, J. R. Nobleza, D. Sanchez, and J. T. Milan, “Optimizing the efficiency, vulnerability and robustness of road-based para-transit networks using genetic algorithm,” *Lecture Notes in Computer Science*, vol. 2018, pp. 3–14, 2018.

[31] M. Zanin, X. Sun, and S. Wandelt, “Studying the topology of transportation systems through complex networks: handle with care,” *Journal of Advanced Transportation*, vol. 2018, Article ID 3156137, 17 pages, 2018.

[32] E. Rodríguez-Núñez and J. C. García-Palomares, “Measuring the vulnerability of public transport networks,” *Journal of Transport Geography*, vol. 35, pp. 50–63, 2014.

[33] B. Berche, C. von Ferber, T. Holovatch, and Y. Holovatch, “Resilience of public transport networks against attacks,” *The European Physical Journal B*, vol. 71, no. 1, pp. 125–137, 2009.

[34] D. Luo, O. Cats, and H. van Lint, “Can passenger flow distribution be estimated solely based on network properties in public transport systems?” *Transportation*, vol. 47, no. 6, pp. 2757–2776, 2020.

[35] M. L. Mouronte and R. M. Benito, “Structural analysis and traffic flow in the transport networks of Madrid,” *Networks and Heterogeneous Media*, vol. 10, no. 1, pp. 127–148, 2015.

[36] T. W. Valente, K. Coronges, C. Lakon, and E. Costenbader, “How correlated are network centrality measures?” *Connections*, vol. 28, no. 1, pp. 16–26, 2008.

[37] E. Costenbader and T. W. Valente, “The stability of centrality measures when networks are sampled,” *Social Networks*, vol. 25, no. 4, pp. 283–307, 2003.

[38] R Project for Statistical Computing, 2020, https://www.r-project.org/.

[39] Python, Python Lenguaje de Programación, 2020, https://www.python.org/.

[40] C. Shao, P. Cui, P. Xun, Y. Peng, and X. Jiang, “Rank correlation between centrality metrics in complex networks: an empirical study,” *Open Physics*, vol. 16, no. 1, pp. 1009–1023, 2018.

[41] P. J. Green, “Iteratively reweighted least squares for maximum likelihood estimation, and some robust and resistant alternatives,” *Journal of the Royal Statistical Society: Series B*, vol. 46, no. 2, pp. 149–170, 1984.

[42] P. Kragh and L. Theil, *Regression with Linear Predictors*, Springer, New York, NY, USA, 2010.

[43] L. A. Adamic and E. Adar, “Friends and neighbors on the web,” *Social Networks*, vol. 25, no. 3, pp. 211–230, 2003.

[44] A. Rodriguez, B. Kim, M. Turkoz et al., “New multi-stage similarity measure for calculation of pairwise patent similarity in a patent citation network,” *Scientometrics*, vol. 103, no. 2, pp. 565–581, 2015.

[45] F. Fouss, A. Pirotte, J.-M. Renders, and M. Saerens, “Random-walk computation of similarities between nodes of a graph with application to collaborative recommendation,” *IEEE Transactions on Knowledge and Data Engineering*, vol. 19, no. 3, pp. 355–369, 2007.

[46] E. Ravasz, A. L. Somera, D. A. Mongru, Z. N. Oltvai, and A.-L. Barabasi, “Hierarchical organization of modularity in metabolic networks,” *Science*, vol. 297, no. 5586, pp. 1551–1555, 2002.

[47] J. Paul, “Étude comparative de la distribution florale dans une portion des Alpes et des Jura,” *Bulletin del la Société Vaudoise des Sciences Naturelles*, vol. 37, pp. 547–579, 1901.

[48] E. A. Leicht, P. Holme, and M. E. J. Newman, “Vertex similarity in networks,” *Physical Review E*, vol. 73, no. 2, pp. 1–10, 2006.

[49] H. Chen, X. Li, and Z. Huang, “Link prediction approach to collaborative filtering,” in *Proceedings of the 5th ACM/IEEE-CS Joint Conference of Digital Library*, pp. 141-142, Denver, CO, USA, June 2005.

[50] T. J. Sørensen, “A method of establishing groups of equal amplitude in plant sociology based on similarity of species content and its application to analyses of the vegetation on Danish commons,” *Biologiske Skrifter, Kongelige Danske Videnskabernes Selskab*, København: I Kommission Hos E. Munksgaard, Denmark, 1948.

[51] D. J. Klein and M. Randić, “Resistance distance,” *Journal of Mathematical Chemistry*, vol. 12, no. 1, pp. 81–95, 1993.

[52] L. Katz, “A new status index derived from sociometric analysis,” *Psychometrika*, vol. 18, no. 1, pp. 39–43, 1953.

[53] P. Y. Chebotarev and E. V. Shamis, “The matrix-forest theorem and measuring relations in small social groups,” *Automation and Remote Control*, vol. 58, no. 9, pp. 1505–1514, 1997.

[54] J.-Y Pan, H.-J. Yang, C. Faloutsos, and P. Duygulu, “Automatic multimedia cross-modal correlation discovery”” in *Proceedings of the 10th ACM SIGKDD International Conference Knowledge Discovery & Data Mining KDD’04*, pp. 653–658, Seattle, WA, USA, August 2004.

[55] L. Lü, C. H. Jin, and T. Zhou, “Similarity index based on local paths for link prediction of complex networks,” *Physical Review E, Statistical, Nonlinear, and Soft Matter Physics*, vol. 80, no. 4, Article ID 046122, 2009.

[56] H. Aidos, P. Robert, W. Duin, and A. Fred, “The area under the ROC curve as a criterion for clustering evaluation”” in

*Proceedings of the 2nd International Conference on Pattern Recognition Applications and Methods (ICPRAM 2013)*, Barcelona, Spain, February 2013.

[57] J. Cohen, "A coefficient of agreement for nominal scales," *Educational and Psychological Measurement*, vol. 20, no. 1, pp. 37–46, 1960.

[58] C. Cortes and M. Mohri, "AUC optimization vs. error rate minimization," in *Proceedings of the 16th International Conference on Neural Information Processing Systems (NIPS '03)*, pp. 313–320, British Columbia, Canada, December 2003.

[59] M. Neuhauser, *Nonparametric Statistical Tests: A Computational Approach*, Chapman and Hall/CRC, New York, NY, USA, 2017.

[60] G. Menardi and N. Torelli, "Training and assessing classification rules with imbalanced data," *Data Mining and Knowledge Discovery*, vol. 28, no. 1, pp. 92–122, 2012.

[61] V. Latora and M. Marchiori, "Efficient behavior of small-world networks," *Physical Review Letters*, vol. 87, no. 19, pp. 198701–198705, 2001.

[62] T. B. Athul and G. Suresh Singh, "Total graph of regular graphs," *Advances in Mathematics: Scientific Journal*, vol. 9, no. 6, pp. 4213–4220, 2020.

[63] E. Bryan, C. VerSchneider, A. Darren, and Narayan, "Efficiency of star-like networks and the Atlanta subway network," *Physica A*, vol. 392, pp. 5481–5489, 2013.

[64] M. Natarajan, "Correlation coefficient analysis of centrality metrics for complex network graphs," in *Intelligent Systems in Cybernetics and Automation Theory. CSOC 2015. Advances in Intelligent Systems and Computing*, R. Silhavy and S. Roman, Eds., vol. 348Czech, Springer, 2015.

[65] M. E. J. Newman, "The structure and function of complex networks," *SIAM Review*, vol. 45, no. 2, pp. 167–256, 2003.

[66] M. Dehmer, *Structural Analysis of Complex Networks*, Birkhäuser, Vienna, Austria, 2011.

[67] E. Estrada, D. J. Higham, and N. Hatano, "Communicability betweenness in complex networks," *Physica A: Statistical Mechanics and Its Applications*, vol. 388, no. 5, pp. 764–774, 2009.

[68] C. Durón, "Heatmap centrality: a new measure to identify super-spreader nodes in scale-free networks," *PLoS One*, vol. 15, no. 7, Article ID e0235690, 2020.

[69] N. Perra and S. Fortunato, "Spectral centrality measures in complex networks," *Physical Review E*, vol. 78, no. 3, Article ID 036107, 2008.

[70] H. Yue, Q. Guan, Y. Pan, L. Chen, and J. Lv, "Detecting clusters over intercity transportation networks using K-shortest paths and hierarchical clustering: a case study of mainland China," *International Journal of Geographical Information Science*, vol. 2019, pp. 1–29, 2019.

[71] A. Bramson, M. Hori, B. Zha, and H. Inamoto, "Scoring and classifying regions via multimodal transportation networks," *Applied Network Science*, vol. 4, no. 1, p. 97, 2019.

[72] J. Xu, Y. Zhang, and D. Miao, "Three-way confusion matrix for classification: a measure driven view," *Information Sciences*, vol. 507, pp. 772–794, 2020.

[73] B. Llc, "Normality tests: Kolmogorov-Smirnov test, Pearson's chi-square test, Anderson-darling test, D'agostino's K-squared test, Jarque-bera Test," 2010.

[74] Y. Boujelbene and A. Derbel, "The performance analysis of public transport operators in Tunisia using AHP method," *Procedia Computer Science*, vol. 73, pp. 498–508, 2015.

[75] P. P. Kumar, M. Parida, and M. Swami, "Performance evaluation of multimodal transportation systems," *Procedia-Social and Behavioral Sciences*, vol. 104, pp. 795–804, 2013.

[76] L.-G. Mattsson and E. Jenelius, "Vulnerability and resilience of transport systems - a discussion of recent research," *Transportation Research Part A: Policy and Practice*, vol. 81, pp. 16–34, 2015.

[77] J. Lin and Y. Ban, "Complex network topology of transportation systems," *Transport Reviews*, vol. 33, no. 6, pp. 658–685, 2013.

[78] L. Zhang, J. Lu, B.-B. Fu, and S.-B. Li, "A review and prospect for the complexity and resilience of urban public Transit network based on complex network theory," *Complexity*, vol. 2018, Article ID 2156309, 36 pages, 2018.

[79] X. Zhang, W. Li, J. Deng, and T. Wang, "Research on Hub node identification of the public transport network of Guilin based on complex network theory," in *Safe, Smart, and Sustainable Multimodal Transportation Systems (CICTP 2014)*, J. Ma, D. Pan, H. Huang, and Y. Yin, Eds., American Society of Civil Engineers, Reston, VA, USA, pp. 302–1309, 2014.

[80] X. Yang, S. Lu, W. Zhao, and Z. Zhao, "Exploring the characteristics of an intra-urban bus service network: a case study of Shenzhen, China," *ISPRS International Journal of Geo-Information*, vol. 8, no. 11, pp. 486–517, 2019.

Hindawi
Complexity
Volume 2021, Article ID 2293524, 13 pages
https://doi.org/10.1155/2021/2293524

*Research Article*

# Research on the Hierarchical Spatial Structure of the Urban Agglomeration of the Yellow River Ji-Shaped Bend

**Fang Wang**,[1] **Mengyao Guo**,[1] **Xi Guo**,[1] **and Fangqu Niu**[2]

[1]*School of Public Administration, Inner Mongolia University, Hohhot 010070, China*
[2]*Institute of Geographic Sciences and Natural Resources Research, Chinese Academy of Sciences, Beijing 100101, China*

Correspondence should be addressed to Xi Guo; 13327118080@163.com and Fangqu Niu; niufq@lreis.ac.cn

Received 21 May 2021; Revised 16 July 2021; Accepted 28 July 2021; Published 13 August 2021

Academic Editor: Miao Zhang



Despite the rising interest in understanding the various uses of space of flows, few studies have combined the traditional static urban development level with dynamic space of flows concepts. In the context of the coordinated development of the urban agglomeration of the Yellow River Ji-shaped bend (UAYB), this study identifies the hierarchical spatial structure of the UAYB through a combination of Baidu migration big data and traditional data. The following conclusions can be drawn. (1) The cities with the strongest regional comprehensive power are Ordos, Taiyuan, Hohhot, Yinchuan, and Yulin, which cause the UAYB to present a significant "center-periphery" spatial pattern. (2) The biggest population flows mostly occur between cities in the same province, while interprovincial population flows mainly exist between cities with the strongest comprehensive power. (3) The hierarchical spatial structure of the UAYB forms a multitree structure, with Ordos as the core. (4) The attractiveness of the UAYB is very weak, being only slightly attractive to individual surrounding provinces, while the population outflow index to economically developed areas is high. Several policy implications are proposed, which can provide important insights for planning intercity connections among the UAYB, in order to achieve more coordinated regional development.

## 1. Introduction

With the strengthening of economic globalization, the comprehensive competitiveness of a country often depends on its urban agglomeration development [1, 2]. As early as the beginning of the twentieth century, the spatial structure of urban agglomeration has attracted the attention of scholars. Some classic theories have gradually formed, such as the central place theory, the law of order and scale of cities, and Zipf's law. Since the 1990s, globalization and informatization have profoundly affected the development and evolution of cities and regions around the world. Through various infrastructure and information networks, the connections between regions exceed the boundaries of central place theory, and the development of regions or cities is no longer carried out in a closed system [3, 4]. The flow of people, goods, technology, and information between cities within and outside urban agglomerations forms a dynamic and cooperative urban system, so urban agglomerations are no longer isolated systems [5]; however, the traditional theories have relatively static and isolated defects, and exploring the spatial structure of urban agglomeration no longer meets the needs of current development. Therefore, the perspective of "flow space," based on dynamic correlation, has gradually become a hotspot in research on the spatial structure of urban agglomeration. Castells first proposed that a "space of flows" can be used as a new perspective of the urban and regional structure and pointed out that the dynamic flow of elements can replace the static regional special structure [6]. In the context of space of flows, research on the network structure, function, and relationships between cities and regions at various scales (especially at the global scale) has attracted more attention from scholars.

To date, research on the spatial structure of urban agglomerations can basically be divided into two categories. On the one hand, only traditional static data are used; on the other hand, data of space of flows are used to focus purely on the flow pattern of elements, especially traffic flows. However, the urban spatial structure includes not only the flows of various elements but also the development level of each city in the region. To date, few studies have combined data of traditional static urban development level with data of dynamic element flows. Studying the spatial structure of urban agglomeration through the combination of traditional static data and population flow data can lead to results that are more accurate and closer to reality, which is of academic significance.

The Yellow River is the fifth largest river in the world, and as such, the Yellow River basin has a very important position in China's economic and social development and ecological security. In 2019, "ecological conservation and high-quality development of the Yellow River basin" was proposed as a major national strategy in China (YRCC, 2013; MOEE, 2020 [7, 8]). On January 3, 2020, the Chinese government officially proposed the concept of the urban agglomeration of Yellow River Ji-shaped bend (UAYB) and emphasized the promotion of the coordinated development of this urban agglomeration. "Ji" is a Chinese character, which is the pictographic character of the shape of the middle part of the Yellow River. Therefore, UAYB is an urban agglomeration in the middle reaches of the Yellow River, where it is Ji-shaped.

The UAYB straddles the central and western parts of China, and the new Eurasian Continental Bridge passes through the border. At the same time, it is also a rare resource-rich area in China and, even, in the world. Energy resources such as coal, natural gas, and rare Earth metals are extremely rich in this region, making it unique in China's development pattern. Therefore, with the in-depth economic development of the UAYB, its internal economic ties are also expected to become closer. As such, carrying out research on the socioeconomic connection of the UAYB has important reference value for exploring how to expand the radiation power of core cities in this urban agglomeration, accelerate the formation of a driving axis for regional development, and construct a coordinated development pattern.

In this paper, we take the UAYB as the research area, carry out regional multilevel spatial structure research through a new computational algorithm, and identify the hierarchical spatial structure of the UAYB. Our key contributions are: (i) we fill the gap in the literature related to comprehending the hierarchical spatial structure of urban agglomeration from the combination of traditional static data and population flow data and (ii) we provide a reference for the coordinated development of the UAYB. The rest of the paper is structured as follows: Section 2 summarizes the literature concerning the space of flows. Section 3 presents the study area, the data, and the methods used. In Section 4, we discuss the hierarchical spatial structure of the UAYB. Section 5 concludes this study and provides some further research possibilities.

## 2. Literature Review

In 1986, Friedmann proposed the importance of urban hierarchical network structure research from the perspective of urban agglomerations in the theory of world cities. He believed that cities are arranged "into a hierarchy of spatial articulations, roughly in accord with the economic power they command" [9]. Marshall pointed out, in 1989, that the spatial agglomeration and close contact of cities are necessary factors for the formation of urban agglomerations [10]. In China, Gu initially divided economic zones based on an analysis of the Chinese urban system [2]. Since then, a great deal of research methods for the spatial structure of urban agglomerations have emerged, mainly including system dynamics (SD), cellular automata (CA), pressure state response model (PSR), expansion index model (AGI), and other methods. These try to sum up the characteristics, connotations, and evolution of the spatial structure of urban agglomerations, by calculating the data for each city [11, 12].

With the development of globalization and informatization, scholars have paid more attention to the impact of the space of flows on urban agglomerations. Castells has been committed to the research of information networks and modern cities since the 1980s. He first proposed that space of flows can be used as a new perspective of the urban and regional structure and pointed out that the dynamic flow of elements will replace the traditional static regional special structure [6]. In the context of space of flows, research on the network structure, function, and relationships between cities and regions at various scales (especially at the global scale) has become a hot spot. In terms of research on the measurement methods of various flows between cities, POLYNET (European Multicenter Megacity Regional Sustainable Development Management Project) has demonstrated the superiority of the theory and method of space of flows by studying eight megacities in Europe [13]. Mitchelson and Wheeler used the US Postal Service as the basic data to assess information flow, functional connection, and hinterland range between cities, as identified from the perspective of space of flows [14]. However, the abovementioned research mostly relies on the background of the individual cities in the United States, which is unique. For this, Matsumoto used aviation flow data to reveal the characteristics of urban cyberspace, focusing on the spatial structure of international airport urban agglomerations [15].

In recent years, the research of space of flows based on big data has shown an increasing trend. Early studies mostly used the characteristics of intercity traffic flow data to identify the urban hierarchical system and network spatial structure, such as bus traffic flows [16, 17], highway traffic flows [18], flight flows [19], and freight volume flows [15]. For example, Ma et al. used passenger traffic flow data to study the multicenter structure of the Shandong coastal urban belt in China and found that it has obvious characteristics of scale benefits and internalization in the spatial structure [2, 20]. Cai et al. also used traffic flow data to study the spatial structure and multicenter characteristics of urban agglomerations in the Pearl River Delta in China and found that this urban agglomeration shows a balanced development trend [21]. However, with the development of information technology and the rise of urban network research, Internet resources (e.g., represented by the Baidu Index and Tencent big data) have gradually become new directions for urban globalization and integration research from the perspective of space of flows. For example, Qiu et al. used information flow and traffic flow data to identify and analyze the network structure characteristics of the Guangdong–Hong Kong–Macau Greater Bay Area from the two dimensions of internal and external connections, through measurement correction, spatial measurement, and social network analysis [22]. Zhou and Wang revised the relevant parameters of the gravity model, in order to measure the flows of Chinese interprovincial tourists [23].

Current research on China's regional spatial evolution is mostly concentrated in economically developed regions, such as the Yangtze River Delta, the Pearl River Delta, and the Beijing–Tianjin–Hebei region [24–28]. There has been very little research on the UAYB. In the context of the coordinated development of the UAYB, it is urgent to carry out relevant research to explore the scientific path of coordinated development.

## 3. Study Area and Research Methods

*3.1. Study Area.* The UAYB refers to the area located at the bend of the Yellow River, of nearly 557,000 square kilometers. A 3,000 kilometer section of the Yellow River flows through this area, west from the Baiyin city of Gansu province, through Ningxia Hui Autonomous Region, Inner Mongolia Autonomous Region, Shaanxi province, to Linfen city of Shanxi province. It includes the three capital cities of Taiyuan, Hohhot, and Yinchuan, as well as Wuzhong, Zhongwei, Wuhai, Bayannaoer, Baotou, Ordos, Shaanxi Yulin, Shanxi Shuozhou, Xinzhou, Lvliang, and so on, for a total of 21 cities (Figure 1).

*3.2. Data Source.* Static population data, such as official statistical yearbooks and traditional decennial census data referenced in the past, cannot reflect the complex interactions between cities in the context of rapid urbanization in China. Baidu Maps is one of the largest map and navigation service providers in China. The Baidu migration data provided by Baidu Maps is calculated by comparing the changes in user position and the number of intelligent terminal users whose positions have changed through all kinds of vehicles, such as railways, highways, and aviation. Therefore, Baidu migration data are able to aggregate anonymized location information and provide data on population outflow and inflow for different time periods and in different regions. The migration scale index, which indicates the daily population mobility intensity between different cities, has a uniform standard and is comparable in size. We used the "Baidu Migration" platform to obtain population migration data on Baidu Maps from January 1 to March 31, 2020 (http://qianxi.baidu.com/). With outflow cities as the ordinate and inflow cities as the abscissa, a $21 \times 21$ directed multivalued network matrix was obtained. In addition, from the *China Statistical Yearbook 2020* [29], *China City Statistical Yearbook* [30], and so on, basic data such as per capita GDP and the proportion of tertiary industries in 2019 were obtained.

*3.3. Research Methods*

*3.3.1. Social Network Analysis Methods.* Social network theory holds that society is a huge network composed of various relationships, where each actor is a node in the network. This work applied a node to represent each city in the network construction, with directed edges depicting population flows and edge weights measuring population flow. We analyzed the network characteristics of the UAYB by using social network analysis, from the two aspects of network density and centrality [31, 32].

① Network density

Network density refers to the degree of connection among cities in a social network. The calculation formula is as follows:

$$D = \frac{\sum_{i=1}^{N}\sum_{j=1}^{N} d\left(n_i, n_j\right)}{N(N-1)}, \quad (1)$$

where $D$ is the network density, $N$ is the number of nodes, and $d$ is the actual degree of connection between the two nodes.

② Degree centrality

Degree centrality indicates the number of other nodes directly connected to a node, where a node with a high degree of centrality maintains numerous contacts with other network nodes, which characterizes the importance of the node in the network [33, 34]. There are two types of measurements: in-degree centrality and out-degree centrality. In-degree centrality concerns the number of nodes connected internally to a primary node, whereas out-degree centrality refers to the number of nodes linked externally to this node. The expressions are as follows:

$$\mathrm{Ind}_i = \sum_{i=1}^{n} a_{ij}, \quad (2)$$

$$\mathrm{Outd}_i = \sum_{i=1}^{n} a_{ij}, \quad (3)$$

where $\mathrm{Ind}_i$ represents the in-degree centrality and $\mathrm{Outd}_i$ represents the out-degree centrality. If nodes $i$ and $j$ are connected, $a_{ij}$ is assigned a value of 1; otherwise, $a_{ij}$ takes on a value of 0.

*3.3.2. Evaluation Method of City Comprehensive Power.* The comprehensive power of a city is an important basis to evaluate its rank and core status. Referring to previous studies [2, 22], we selected nine first-level indicators and 12 second-level indicators, considering the three dimensions of attractiveness, economic level, and social, technological, and cultural development level, to construct a comprehensive power evaluation index system for cities (Table 1). In this paper, the three dimensions of attractiveness, economic


FIGURE 1: Study area.

TABLE 1: Evaluation index system and variable definition of city comprehensive power.

| The target layer | First level | Second level |
|---|---|---|
| Attractiveness | Centrality | Indegree centrality |
| Economic | Economic force | GDP per capita |
| | | Proportion of the tertiary industry |
| | Economic vitality | Number of commercial banks |
| | | Number of listed companies |
| Social science and technology culture | Living standard | Total retail sales of consumer goods |
| | Workforce | Total population |
| | Social vitality | Baidu Index (average day) |
| | Public service | Investment in fixed assets |
| | Science and technology culture | Number of universities per 10,000 people |
| | | Number of patents per 10,000 people |
| | | Number of museums per 10,000 people |

level, and social, technological, and cultural development were treated as equally important at first; then the entropy method was used to assign weights objectively to the indicators of each dimension. Finally, the comprehensive power score of each city was obtained, according to the weights of the indicators. The specific formula is as follows:

① The indicators of cities are made comparable through standardization treatment. The cities and 14 secondary indices are arranged, in order to form a matrix as follows:

$$X = \left(X_{ij}\right)_{m\times n}, \quad 1 \leq i \leq 21,\ 1 \leq j \leq 14. \tag{4}$$

② Calculate the entropy value of each index as follows:

$$e_j = -k\sum_{i=1}^{21}\left(p_{ij}\ln\, p_{ij}\right), \tag{5}$$

where $k = (1/\ln\, n)$ and $p_{ij} = (y_{ij}/\sum_{i=1}^{21} y_{ij})$.

③ Calculate the weight of indicators as follows:

$$w_j = \frac{1 - e_j}{\sum_{j=1}^{14} 1 - e_j}. \tag{6}$$

④ Calculate the scores of the comprehensive power score of each city as follows:

$$F_i = \sum_{j=1}^{14} w_j \times y_{ij}, \quad 0 \le F_i \le 1. \tag{7}$$

*3.3.3. Hierarchical Spatial Structure Computational Algorithm.* According to central place theory, regional spatial interactions have a hierarchical structure. Therefore, we believed that there is a city with the highest scale in the region, leading the regional development (i.e., the core city), and the surrounding cities have close social and economic ties with the core city. Each city has its own radiation area, thus forming a hierarchical regional spatial structure. This paper focuses on mining this structure, based on the above data. Therefore, a regional spatial structure analysis algorithm is constructed in this paper, and the regional spatial structure was determined by computationally implementing the algorithm. This algorithm identifies the radiation area of each city, based on its comprehensive power and the strength of intercity connections (i.e., population flows), and determines the hierarchical spatial structure of the region (Figure 2).

Network relationships are formed through interactions between cities located in the same region. We determined the hierarchical spatial structure by evaluating the urban comprehensive power and the size of population flows among cities. If the city most closely contacting with city B is city A and the comprehensive power of city A is stronger than that of city B, then it is said that city B is in the radiation area. Therefore, the radiation area of a stronger city will be composed of several weaker cities. Thus, a hierarchical spatial structure analysis algorithm was constructed. The specific steps are: select any city A, compare and screen the city B with the closest connection and higher comprehensive power one by one, and designate city B as the upper-level central city of city A (if city A has the highest comprehensive power, it is said that city A has no superior central city). Then, select the next city and repeat the above process, until all cities are compared.

## 4. Results

*4.1. Comprehensive Power of Cities in the UAYB.* According to formulas (2)–(7), the comprehensive power of the cities in UAYB was evaluated. The results are shown in Table 2 and Figure 3.

Ordos had the highest level of attractiveness in the UAYB, as well as the highest degree of association with the cities in the urban agglomeration. At the same time, Yinchuan and Yulin also had relatively high attractiveness, becoming the two subcenters with the second-highest attractiveness in the region. It is also worth mentioning that Alxa League also has a higher attractiveness. Unlike the abovementioned cities, the level of urban development in Alxa is not in line with the intensity of the city's attractiveness. This may be due to the better development of local tourism, and the study period happened to be the reason for the local tourist peak season.

Taiyuan is the capital of Shanxi Province and is the only type I city (The State Council's "Notice on Adjusting the Criteria for the Classification of Urban Scales." Type I city: the population size of built-up areas is above 3 million and below 5 million and type II city: the population size of built-up areas is above 1 million and below 3 million) in the UAYB. It also has the highest economic score. In addition, Yinchuan, Baotou, Ordos, and Hohhot also have the highest economic scores, where their economic development momentum is very strong. The two provincial capital cities, Hohhot and Taiyuan, achieved high scores in social science and technological culture. This was due to the relatively good foundations of the provincial capitals, in terms of society, culture, and public services.

In general, it can be seen that the comprehensive power level of cities in the UAYB presented a significant "center-periphery" spatial pattern, with Ordos, Yulin, Hohhot, Yinchuan, and so on, as the center, while the comprehensive power scores of the cities gradually decrease to the west and north. The comprehensive power scores of these center cities were all significantly higher than those of other cities, such that they were all absolute power cities in the region. The cities in the second echelon of comprehensive power included Baotou, Alxa League, Datong, and Yan'an, which were close to the first echelon of cities spatially and were obviously affected by their radiation and driving effect. These cities have gradually risen in recent years, especially Baotou and Datong, which have experienced rapid industrial development. Other cities, such as Baiyin in Ningxia province, Shuozhou in Shanxi province, and Ulanqab and other cities in Inner Mongolia, are located in the west and north of the UAYB, being geographically far away from the first echelon of cities. In addition to inconvenient transportation, they lack the necessary development conditions. In the short term, compared with the above cities, their comprehensive power is still weak, and the gap remains large.

*4.2. Regional Spatial Structure of the UAYB*

*4.2.1. The Network Structure Characteristics of Population Flows.* First, the overall network density describes the closeness of connections between nodes in a network. The greater the value, the more connection paths and interactions between nodes. When this value exceeds its threshold, the entire network will assume a completely continuous region, forming a huge spatial group. The thresholds of

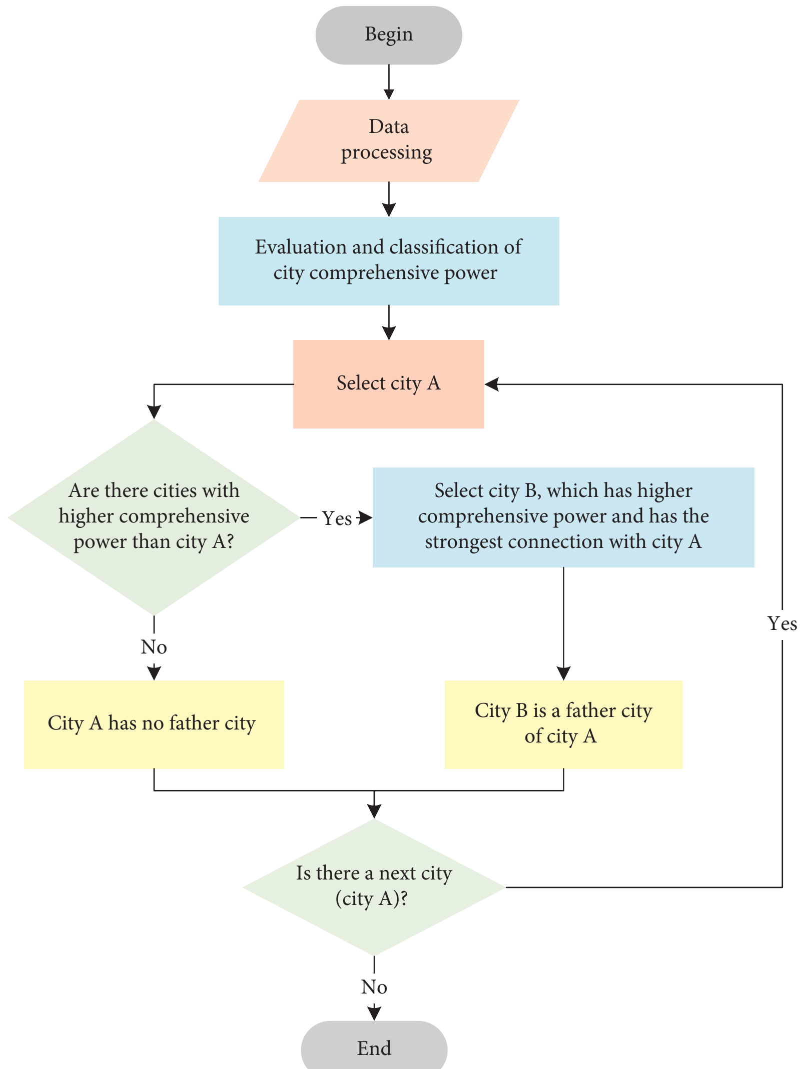



Figure 2: The computational algorithm for identifying the hierarchical spatial structure.

different network structures are different. The network structure of the UAYB can be abstracted into a triangular lattice; as such, the overall network density threshold is 0.5. According to formula (1), the overall network density of the UAVB was 0.409, which is close to, but has not yet reached, this threshold. This means that the intercity connection strength in the UAYB is moderate, and the channels for information circulation and population flow are relatively few. The reciprocity of information, capital, and technology needs to be further improved. Moreover, due to the relatively short development time of the transportation network that runs through the entire region in the UAYB, the integrated network of connections in the UAYB has not yet been formed.

Secondly, from the perspective of population flows between cities (Figures 4 and 5), we found that: (1) the population flows with the highest intensity level mostly occurred between cities in the same province. The total population flow (inflow and outflow) indices between Hohhot and Wulanchabu, Datong and Shuozhou, Taiyuan and Lvliang, Yinchuan and Wuzhong, and Yinchuan and Shizuishan all exceeded 3,500. These cities are close to each other spatially, have close economic ties, and have convenient transportation. (2) The interprovincial population flows are mainly between the cities with the strongest regional comprehensive power. Taiyuan, the only type I city in the region, has a high level of economic development and strong population attractiveness. Its total population flow (including inflow and outflow) index reached 23,754.48. Baotou, Hohhot, and Yinchuan are the three type II cities in the UAYB. Baotou is an important basic industrial base in China, with aluminum, copper, and rare Earth metal industries. As the capital of the Inner Mongolia Autonomous Region, Hohhot has political and economic advantages, in terms of urban development. Yinchuan takes new materials and high-end equipment manufacturing as its leading industries and has superior economic conditions. These cities occupy important economic and social positions in the UAYB and have become the main destinations of population flow. Significantly, the population flow in Taiyuan (the only

Table 2: The comprehensive power of the cities in the UAYB.

| City | Attractiveness score | Attractiveness rank | Economic score | Economic rank | Social science and technology culture score | Social science and technology culture rank | Total score | Total rank |
|---|---|---|---|---|---|---|---|---|
| Erdos | 0.333 | 1 | 0.182 | 4 | 0.153 | 4 | 0.669 | 1 |
| Taiyuan | 0.148 | 6 | 0.264 | 1 | 0.189 | 2 | 0.601 | 2 |
| Hohhot | 0.185 | 5 | 0.181 | 5 | 0.192 | 1 | 0.558 | 3 |
| Yinchuan | 0.222 | 2 | 0.185 | 2 | 0.148 | 5 | 0.555 | 4 |
| Yulin | 0.222 | 2 | 0.108 | 8 | 0.169 | 3 | 0.499 | 5 |
| Baotou | 0.148 | 6 | 0.183 | 3 | 0.116 | 6 | 0.448 | 6 |
| Alxa League | 0.222 | 2 | 0.089 | 10 | 0.062 | 12 | 0.374 | 7 |
| Datong | 0.111 | 10 | 0.115 | 7 | 0.077 | 9 | 0.303 | 8 |
| Yanan | 0.037 | 18 | 0.148 | 6 | 0.104 | 7 | 0.289 | 9 |
| Xinzhou | 0.148 | 6 | 0.061 | 13 | 0.055 | 14 | 0.258 | 10 |
| Bayannur | 0.148 | 6 | 0.052 | 16 | 0.053 | 16 | 0.254 | 11 |
| Linfen | 0.074 | 14 | 0.092 | 9 | 0.069 | 10 | 0.235 | 12 |
| Wuhai | 0.111 | 10 | 0.084 | 11 | 0.037 | 21 | 0.232 | 13 |
| Wuzhong | 0.111 | 10 | 0.050 | 17 | 0.067 | 11 | 0.228 | 14 |
| Lvliang | 0.074 | 14 | 0.047 | 19 | 0.089 | 8 | 0.210 | 15 |
| Zhongwei | 0.111 | 10 | 0.029 | 20 | 0.053 | 17 | 0.186 | 16 |
| Shizuishan | 0.074 | 14 | 0.056 | 14 | 0.054 | 15 | 0.184 | 17 |
| Ulanqab | 0.074 | 14 | 0.048 | 18 | 0.056 | 13 | 0.178 | 18 |
| Shuozhou | 0.037 | 18 | 0.076 | 12 | 0.041 | 20 | 0.155 | 19 |
| Baiyin | 0.037 | 18 | 0.054 | 15 | 0.044 | 19 | 0.135 | 20 |
| Qingyang | 0.000 | 21 | 0.027 | 21 | 0.051 | 18 | 0.078 | 21 |

type I city in the region and with the second strongest regional comprehensive power) is lower than that in Yinchuan, Ordos, Hohhot, and Yulin. This is mainly due to the fact that Taiyuan is located on the southeastern edge of the UAYB. Its social and economic ties with the cities in the UAYB are relatively weak, and most of them have relatively strong connections with other regions outside the UAYB. Therefore, Taiyuan's overall leading role in the UAYB is also relatively poor.

Finally, from the perspective of population inflows and outflows (Table 3), the top five cities, in terms of total inflow index, were Yinchuan, Ordos, Taiyuan, Hohhot, and Yulin, which also had relatively large net population inflow indices, as these cities have a higher level of economic and social development and a larger population attractiveness. Among them, Ordos also had a large total outflow index and positive net population inflow index, which means that Ordos had a high degree of social and economic activity. Except for the special case of Ordos, Wuhai, Shizuishan, Wuzhong, and Bayannaoer were the other major population outflow cities, with negative net flow population indices, and therefore, the associated population loss is serious. These cities are geographically far away from the regional center of the UAYB and are located in remote areas. They are less affected by the central city's economic radiation, and the economic development is weak in these cities.

*4.2.2. Hierarchical Spatial Structure of the UAYB.* Through the regional spatial structure computational algorithm, the multitree structure of the hierarchical structure of the UAYB was obtained (Figure 6). To further visualize the structure, the color of the patches represents the comprehensive power of cities (Figure 7). The color is divided into five levels, from light to dark. The heavier colors indicate stronger comprehensive power. The arrows point from child nodes to their father nodes. Each node represents a city, and all child nodes are attracted by their father node. The radiation area of a city depends on its own comprehensive power and intercity interaction strength (population flow).

As shown in Figures 6 and 7, Ordos became the root node by virtue of its strong comprehensive power, playing the leading role in the UAYB. The spatial interaction strength takes Ordos as the center in space and spreads radially outward in the UAYB. Hohhot, Yinchuan, Yulin, and Taiyuan were directly attracted by Ordos as the second-tier city. These four cities formed their own relatively small radiation areas. Among them, Hohhot, Yinchuan, and Taiyuan are the capitals of the Inner Mongolia Autonomous Region, Ningxia Hui Autonomous Region, and Shanxi Province, respectively, which were able to gather the resources of the entire province, maintaining a strong momentum of development. They can radiate and drive the neighboring cities in terms of the economic scale, industrial structure, infrastructure, culture, and other aspects. Yulin is rich in oil, coal, natural gas, and other resources, and it is also a well-known tourist city. Therefore, its economic development is better, and its status in the UAYB is also higher.

The main radiation areas of Hohhot and Taiyuan are in their own provinces and do not extend to other provinces. This may be due to the fact that the Yellow River, as a natural provincial boundary, has a hindering effect on its radiation capacity, in terms of the natural topography, as well as administrative, historical, transportation, and other factors. Yinchuan has a relatively large radiation range. In addition

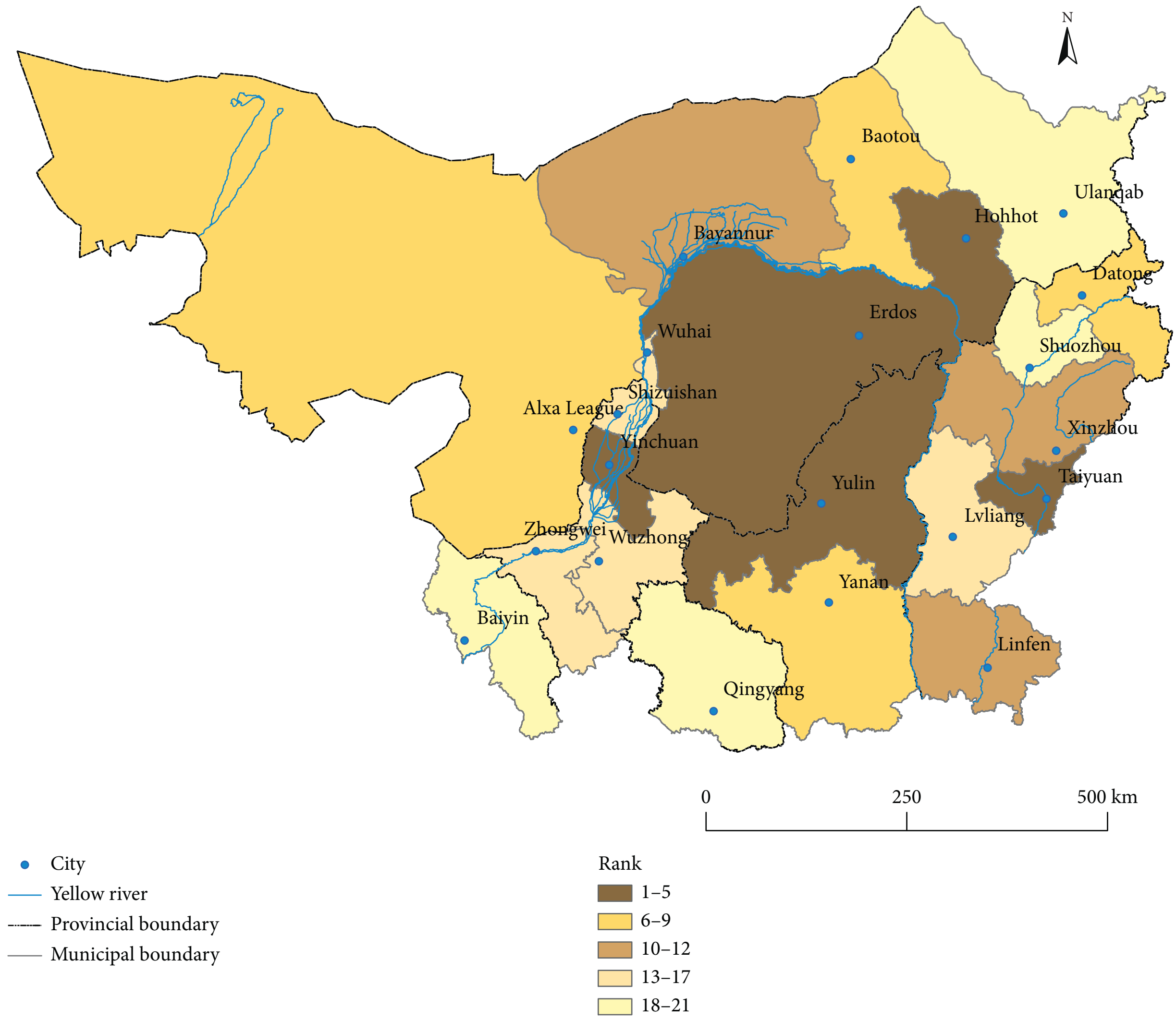



FIGURE 3: Spatial distribution of city comprehensive power in the UAYB.

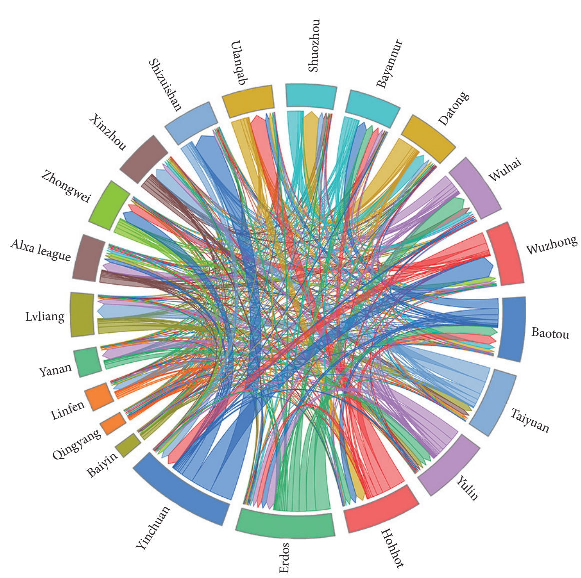



FIGURE 4: Population flows among cities in the UAYB.

to prefecture-level cities in Ningxia Autonomous Region, it also attracts Alxa League and Wuhai City in western Inner Mongolia. As a prefecture-level city in Shaanxi Province, Yulin belongs to the UAYB, and its economic development level is higher than that of its surrounding cities.

It is remarkable that the siphoning effect of higher-tier cities on the surrounding lower-tier cities was also very obvious. The development of higher-tier cities depends on the constant delivery of production factors from the surrounding lower-tier cities, which also limits the development of the lower-tier cities. It causes lower-tier cities to face the problems of talent loss and weak economic development. Such cities include Baiyin, Qingyang, Lvliang, Linfen, and Shuozhou, which are subject to far less radiation drive than the siphoning effect of Taiyuan city, leading such lower-tier cities to face development-related difficulties.

*4.3. The Connection Pattern between the UAYB and the External Regions in China.* As a national-level strategic urban agglomeration, the high-quality development not only

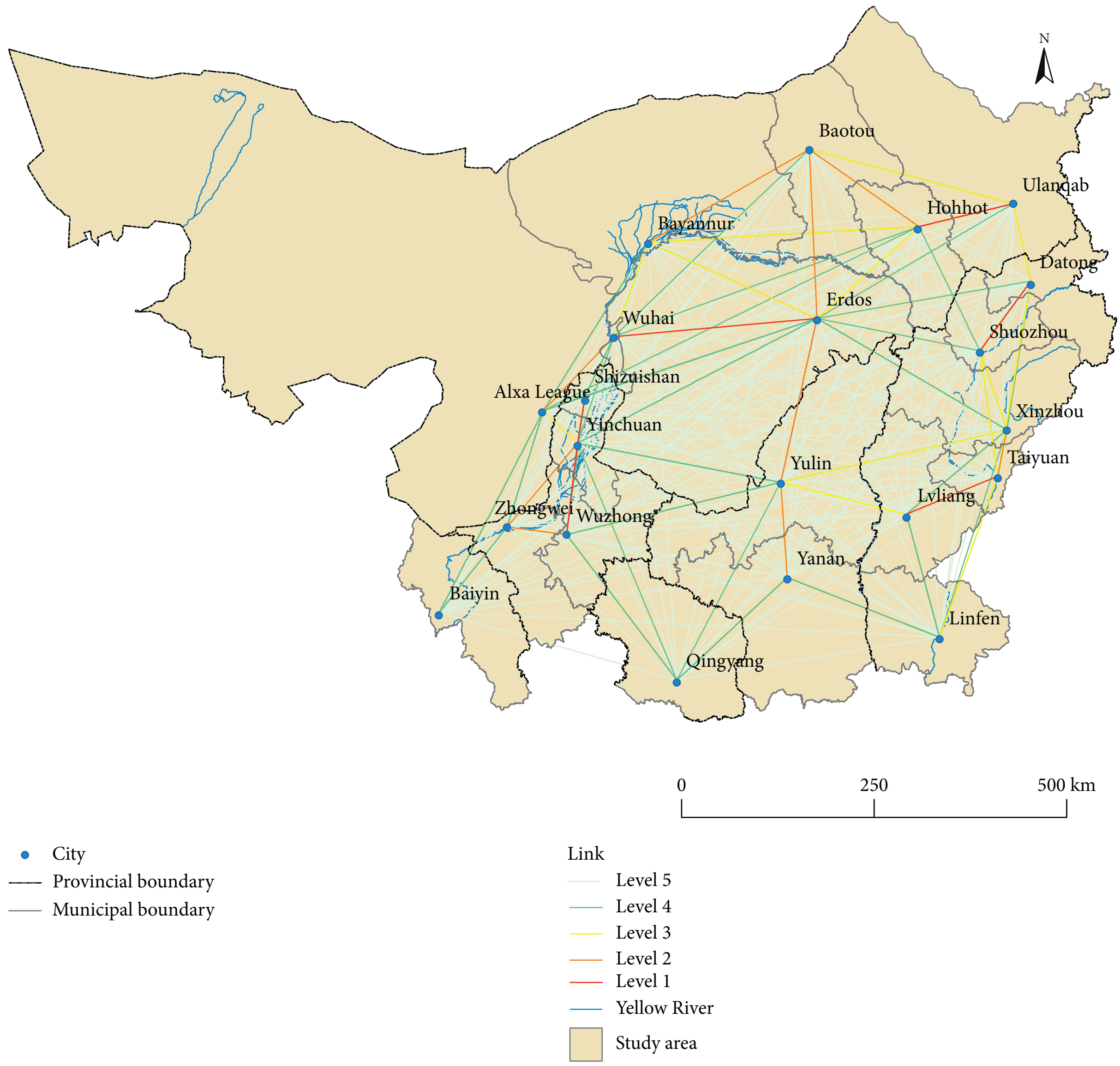



FIGURE 5: Spatial distribution of population flows among cities in the UAYB.

TABLE 3: Population inflows and outflows among cities in the UAYB.

| City | Top five cities of total inflow index | | City | Top five cities of the total outflow index | |
|---|---|---|---|---|---|
| | Total inflow index | Net population inflow index | | Total outflow index | Net population inflow index |
| Yinchuan | 17018.04 | 10281.6 | Wuhai | 8255.85 | −2530.76 |
| Ordos | 14477.44 | 6668.72 | Shizuishan | 7919.9 | −4593.92 |
| Taiyuan | 10922.97 | 6948.95 | Wuzhong | 7903.27 | −1772.46 |
| Hohhot | 10480.92 | 4040 | Ordos | 7808.72 | 6668.72 |
| Yulin | 9195.4 | 3105.18 | Bayannaoer | 7762.39 | −3841.39 |

requires the formation of a harmonious regional spatial structure internally but also close social and economic exchanges with external regions.

The top regions of the population outflow index from the UAYB are Ningxia Autonomous Region, Inner Mongolia Autonomous Region, Shaanxi, Shanxi, Gansu, Hebei, and Beijing (Figure 8(a)). The Beijing–Tianjin–Hebei metropolitan area is the largest urban agglomeration in northern China, with strong economic power and an obvious siphoning effect. The UAYB has a strong socioeconomic connection with this urban agglomeration, and a large number of people flow out to this area. During the study period, the population outflow indices to Beijing and Hebei were 4356.05 and 6309.41, respectively. At the same time, other areas in Ningxia, Inner Mongolia, Shaanxi, and Gansu that are not in the UAYB have close ties with the UAYB, as

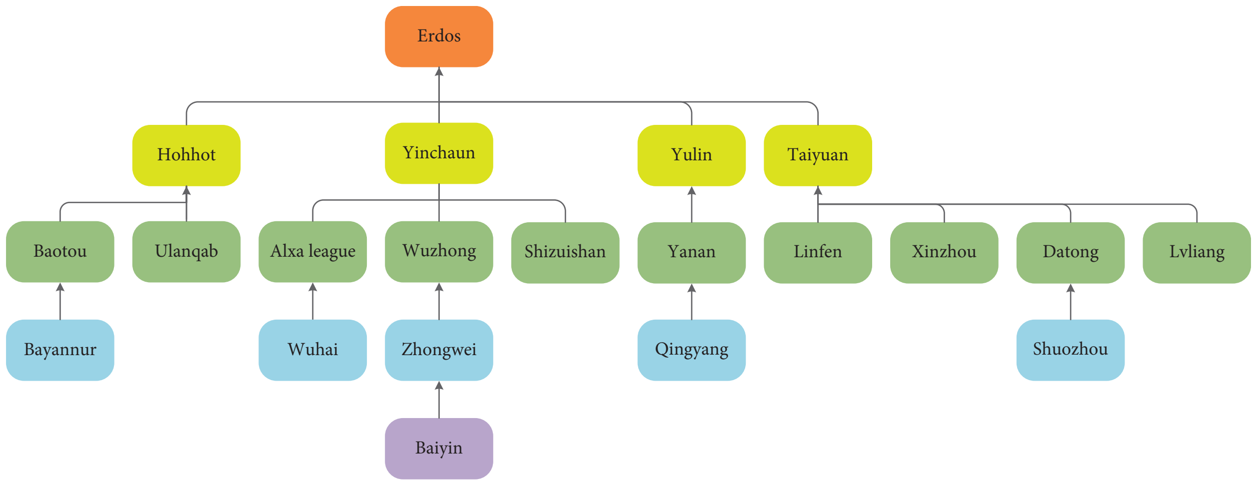



Figure 6: Regional ranking system of the UAYB.

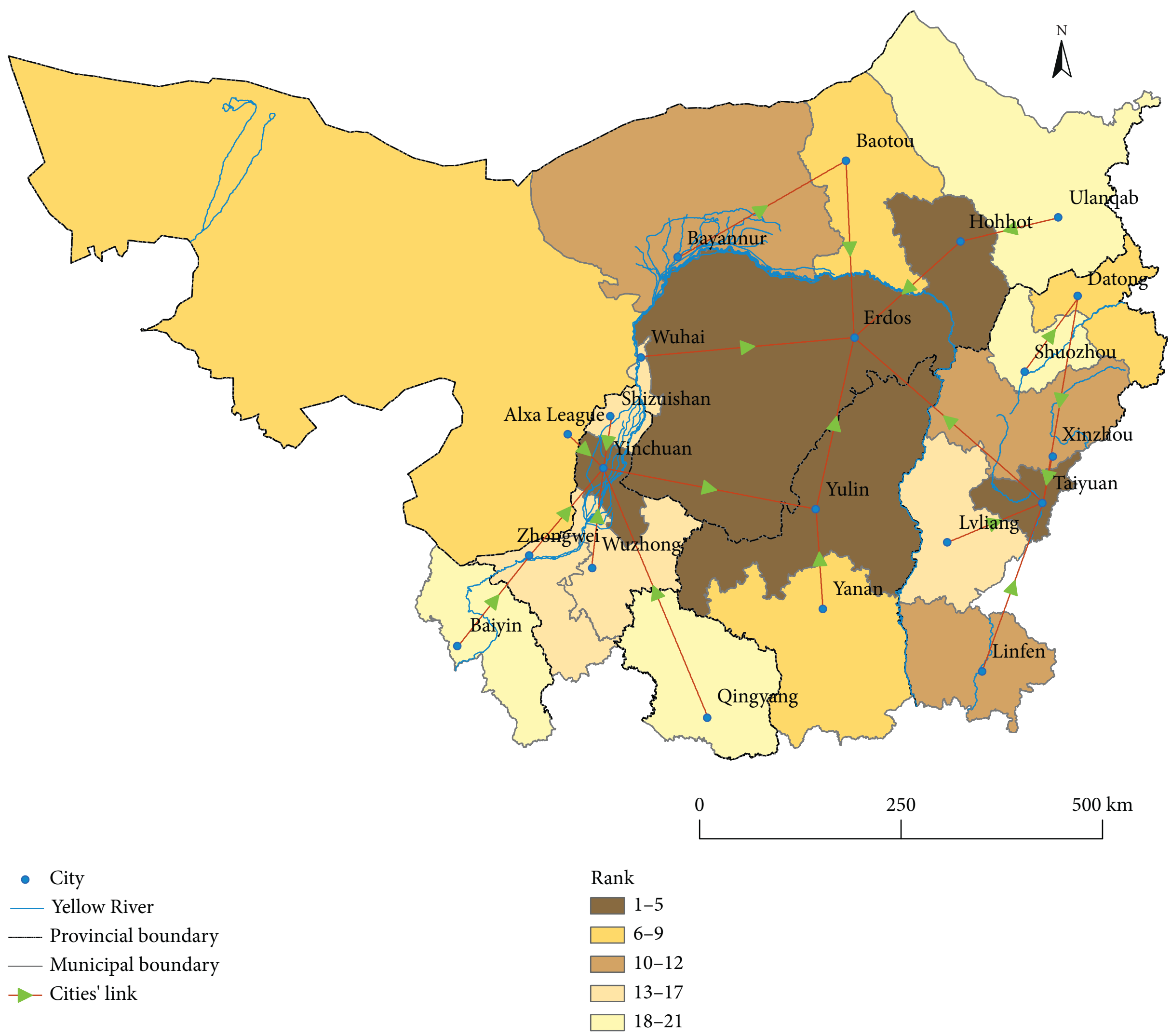
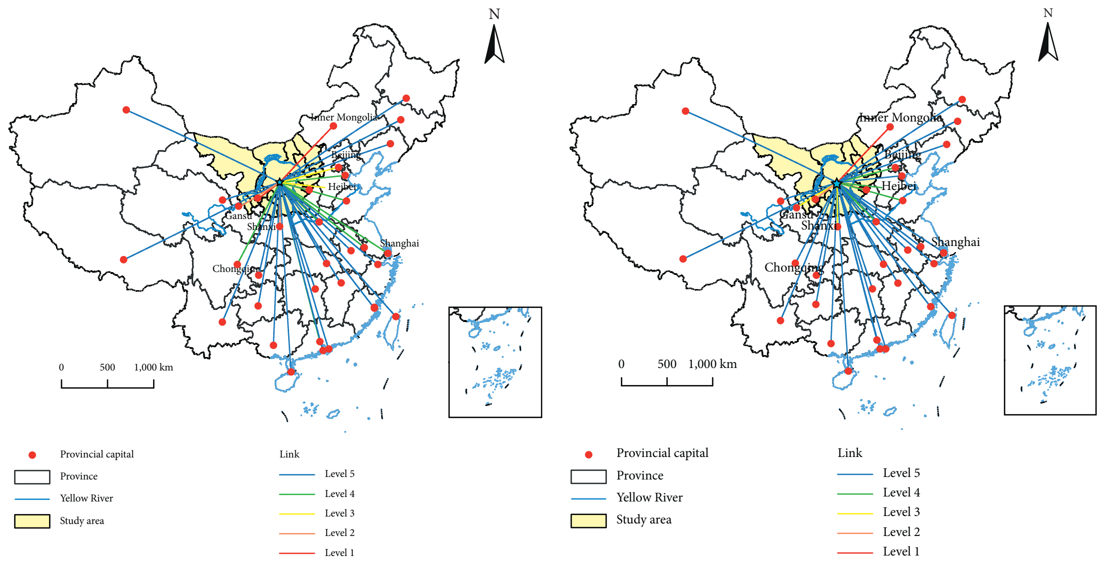



Figure 7: Hierarchical spatial structure of the UAYB.

these areas have similar natural conditions, cultural customs, and convenient transportation with the UAYB. In addition, the developed provinces along the coast and the Yangtze River also have close ties with the UAYB. Since the reform and opening up, these areas have experienced a high level of economic development. They have become China's main




(a) (b)

FIGURE 8: Spatial distribution of population flows between the UAYB and external regions: (a) population flows from the UAYB to the external region and (b) population flows from the external region to the UAYB.

population-carrying areas, attracting a large number of immigrants, including those from the UAYB. From the perspective of the population inflow index from external regions to the UAYB (Figure 8(b)), the areas with a higher population inflow index to the UAYB are basically the neighboring provinces, and most of these provinces have some regions (cities) included in the UAYB. Because these areas have similar natural conditions, cultural customs, and convenient transportation with the UAYB, the population inflow indices from these areas account for more than 85% of the total population inflows of the UAYB from all provinces across the country. It can be seen that the attractiveness of the UAYB to the external provinces and cities is far from sufficient, and therefore, the overall attractiveness of the entire UAYB needs to be improved.

Considering the comparison of population outflow and inflow indices, the outflow and inflow situation between most areas in China and the UAYB was basically the same. However, Shanxi, Inner Mongolia, and Hebei had much larger inflow indices into the UAYB than outflow indices from the UAYB. As the core cities of the UAYB are more attractive to other areas in the three neighboring provinces, they attract more people to move in. The indices of population outflow to Beijing, Guangdong, and Shanghai were much higher than the relative inflow indices, indicating that China's developed provinces and cities have a relatively high level of radiation and attractiveness to the whole country, where the UAYB is no exception.

## 5. Conclusions and Discussion

In this paper, we evaluated the comprehensive power of cities in the UAYB and analyzed the population flows in the UAYB. Based on these analyses, we explored the hierarchical spatial structure of the UAYB. Our contribution to the literature is twofold. First, we developed a new computational algorithm to assess the hierarchical spatial structure of urban agglomeration from the combination of traditional static data and population flow data, which makes the result closer to reality. This is not very common in urban agglomerations studies, most of which focus only on static data of each city or spatial flow data between cities, thereby making it one of the strong points of the paper in our view. Second, we provide scientific reference for the development of the UAYB, which is located in the fifth largest river basin in the world, and the national-level strategic urban agglomeration. The main conclusions of this paper are as follows. (1) The cities with the strongest regional comprehensive power were Ordos, Taiyuan, Yinchuan, Hohhot, and Yulin. They are spatially concentrated in the central area of the UAYB, and as such, the UAYB presented a significant "center-periphery" spatial pattern. (2) An integrated network of connections in the UAYB has not yet been formed. The highest population flows mostly occur between cities in the same province. The interprovincial population flows are mainly between the cities with the strongest regional comprehensive power. (3) The hierarchical spatial structure of the UAYB forms a multitree structure, with Ordos City as the core, which forms the largest urban radiation area. Hohhot, Yinchuan, Yulin, and Taiyuan also have secondary radiation areas in this structure. (4) The UAYB is the most attractive to the populations of the three provinces Shanxi, Inner Mongolia, and Hebei, which are adjacent to the UAYB. Economically developed areas in China, such as Beijing, Guangdong, and Shanghai, are the most attractive areas to the UAYB, and the population outflow indices to these areas were the highest.

Some policy implications based on our findings regarding population flows of the UAYB can be noted. First, we found that cities with strong comprehensive power have a strong control effect on the UAYB, while cities with backward development levels, such as Baiyin and Shuozhou, are in a subordinate position and have not yet entered a good coordinated development stage. However, weak cities in the periphery are key to the coordinated development of the region. Therefore, the government should promote the socioeconomic development of these relatively backward cities and improve their comprehensive power through the construction of infrastructure, technology, and culture. At the same time, the government also should strengthen social and economic ties, promote population mobility, and achieve the goal of regional coordinated development. Second, according to the results of this paper, the radiation areas of the second-tier cities are mainly inside their respective provinces, and the largest population flows mostly occur between cities in the same province. Under China's current administrative management system, the interprovincial administrative boundary has a certain obstructive effect on the economic integration of the administrative area of the UAYB, which is manifested as a significant shielding effect. The government should take the coordinated development of the UAYB as an opportunity to break through the barriers of the boundary shielding effect, strengthen the integrated construction of market systems, public services, industrial development, infrastructure, management systems, and so on, in order to achieve the actual needs of complementarity and speed up the urban network process from "point" to "axis" to "surface" of the UAYB. Finally, from a national perspective, the attractiveness of the UAYB is very weak (only slightly attractive to individual surrounding provinces), and a large number of people have migrated to the developed areas in China. Therefore, in the long run, the government needs to formulate sound regional development strategies to improve the overall power of the UAYB, in order to promote long-term development.

This paper also has some limitations and deficiencies: it did not involve a comparative study over multiple periods. Subsequent studies can conduct longitudinal temporal comparisons based on spatial analyses, thus potentially grasping the future development trend of the UAYB and making reasonable predictions. In addition, the data collection period of this paper had some special characteristics. The population flow data were collected during the peak season of Spring Festival tourism and the peak period of the COVID-19 outbreak in China, which led to some fluctuations and anomalies. These effects require more in-depth analysis in the follow-up research.

## Data Availability

This paper uses the "Baidu Migration" platform to obtain population migration data on Baidu Maps from January 1 to March 31, 2020 (http://qianxi.baidu.com/). The other data used in this study come from publicly published statistical yearbooks. A request for access to these data can be made to the corresponding authors.

## Conflicts of Interest

The authors declare that there are no conflicts of interest.

## Acknowledgments

This research was funded by the National Natural Science Foundation of China (grant nos. 41801149, 71864025, and 42071153) and Young Talents of Science and Technology in Universities of Inner Mongolia (grant no. NJTY-20-B09).

## References

[1] C. Fang, "Progress and the future direction of research into urban agglomeration in China," *Acta Geographica Sinica*, vol. 69, no. 8, pp. 1130–1144, 2014, in Chinese.

[2] C. Gu, "A prelimanary study on the division of urban ecomonic regions in China," *Acta Geographica Sinica*, vol. 46, no. 2, pp. 129–141, 1991, in Chinese.

[3] S. Wang, C. Lian, and Z. Zhao, "From central place to city network: a theoretical change in China's urban system study," *Geographical Research*, vol. 38, no. 1, pp. 64–74, 2019, in Chinese.

[4] P. J. Taylor and B. Derudder, *World-City Network: A Global Urban Analysis*, Routledge, London, UK, 2016.

[5] Y. Wang, X. Niu, and X. Song, "Research progress of regional spatial structure under the perspective of space of flow," *Urban Planning International*, vol. 32, no. 6, pp. 27–33, 2017, in Chinese.

[6] M. Castells, *The Rise of the Network Society*, Blackwell Publishers, London, UK, 1996.

[7] Yellow River Conservancy Commission, *Review of Yellow River Basin*, Henan People's Press, Zhengzhou, China, 1998.

[8] Ministry of Ecology and Environment of the People's Republic of China, "The government work report," 2020, http://www.mee.gov.cn.

[9] J. Friedmann, "The world city hypothesis," *Development and Change*, vol. 17, no. 1, pp. 69–83, 1986.

[10] J. U. Marshall, *The Structure of Urban Systems*, University of Toronto Press, Toronto, Canada, 1989.

[11] N. Gaitani, M. Santamouris, C. Cartalis et al., "Microclimatic analysis as a prerequisite for sustainable urbanisation: application for an urban regeneration project for a medium size city in the greater urban agglomeration of Athens, Greece," *Sustainable Cities and Society*, vol. 13, pp. 230–236, 2014.

[12] S. Y. Lee, R. J. K. Dunn, R. A. Young et al., "Impact of urbanization on coastal wetland structure and function," *Austral Ecology*, vol. 31, no. 2, pp. 149–163, 2006.

[13] P. Hall, K. Pain, and Z. Luo, "The polycentric metropolis, learning from megacity regions in Europe," *Shang-hai Urban Planning Review*, vol. 1, p. 80, 2011.

[14] R. L. Mitchelson and J. O. Wheeler, "The flow of information in a global economy: the role of the American urban system in 1990," *Annals of The Association of American Geographers*, vol. 84, no. 1, pp. 87–107, 1994.

[15] H. Matsumoto, "International urban systems and air passenger and cargo flows: some calculations," *Journal of Air Transport Management*, vol. 10, no. 4, pp. 239–247, 2004.

[16] L. Dai, B. Derudder, and X. Liu, "Transport network backbone extraction: a comparison of techniques," *Journal of Transport Geography*, vol. 69, pp. 271–281, 2018.

[17] Y. Long and J.-C. Thill, "Combining smart card data and household travel survey to analyze jobs-housing relationships in Beijing," *Computers, Environment and Urban Systems*, vol. 53, pp. 19–35, 2015.

[18] D. César and L. Beauguitte, "Spatial science and network science: review and outcomes of a complex relationship,"

*Networks and Spatial Economics*, vol. 14, no. 3-4, pp. 297–316, 2014.

[19] B. Derudder, F. Witlox, J. Faulconbridge, and J. Beaverstock, "Airline data for global city network research: reviewing and refining existing approaches," *GeoJournal*, vol. 71, no. 1, pp. 5–18, 2008.

[20] X. Ma and P. Dou, "Characteristics of polycentricity spatial structure of shangdong coastal city-region based on passenger traffic flow," *Modern Urban Research*, vol. 10, pp. 101–109, 2017.

[21] L. Cai, X. Ma, W. Chen et al., "Characteristics of functional polycentricity of PRD urban region based on passenger traffic flow," *Economic Geography*, vol. 33, no. 11, pp. 52–57, 2013, in Chinese.

[22] J. Qiu, Y. Liu, H. Chen et al., "Urban network structure of Guangdong-Hong Kong-Macao greater bay area with the view of space of flows: a comparison between information flow and transportation flow," *Economic Geography*, vol. 39, no. 6, pp. 7–15, 2019, in Chinese.

[23] H. Zhou and F. Wang, "Research on structure characteristics of the inter-provincial tourist flow spatial network in China based on the modified gravity model," *Geographical Research*, vol. 39, no. 3, pp. 669–681, 2020, in Chinese.

[24] Z. Wang, S. Yang, F. Gong et al., "Identification of urban agglomerations deformation structure based on urban-flow space: a case study of the Yangtze river delta urban agglomeration," *Scientia Geographica Sinica*, vol. 37, no. 9, pp. 1337–1344, 2017.

[25] F. Niu, X. Yang, and F. Wang, "Urban agglomeration formation and its spatiotemporal expansion process in China: from the perspective of industrial evolution," *Chinese Geographical Science*, vol. 30, no. 3, pp. 532–543, 2020.

[26] S. Wang, S. Gao, and Y. Wang, "Spatial structure of the urban agglomeration based on space of flows: the study of the Pearl river delta," *Geographical Research*, vol. 38, no. 8, pp. 1849–1861, 2019, in Chinese.

[27] F. Niu and W. Liu, "Identifying the hierarchical regional spatial structure using internet big data," *Journal of Geo-Information Science*, vol. 18, no. 6, pp. 719–726, 2016, in Chinese.

[28] W. Zhang and B. Derudder, "How sensitive are measures of polycentricity to the choice of "centres"? A methodological and empirical exploration," *Urban Studies*, vol. 56, no. 10, 2019.

[29] National Bureau of Statistics of the People's Republic of China, *China Statistical Yearbook*, China Statistics Press, Beijing, China, 2020.

[30] Editorial Board and Editorial Department of China City Statistical Yearbook, *China City Statistical Yearbook*, China Statistics Press, Beijing, China, 2020.

[31] E. Strano, S. Shai, S. Dobson, and M. Barthelemy, "Multiplex networks in metropolitan areas: generic features and local effects," *Journal of the Royal Society, Interface*, vol. 12, no. 111, Article ID 20150651, 2015.

[32] F. Wei, A. Pengli, and L. Siyao, "Evolution characteristics and regional roles' influencing factors of interprovincial population mobility network in China," *Complexity*, vol. 2021, Article ID 6679580, 11 pages, 2021.

[33] S. Xu and F. Zhen, "Measurement and comparison research of provincial urban network centrality: taking comparative analysis of jiangsu and Guangdong as example," *Human Geography*, vol. 1, pp. 135–144, 2021, in Chinese.

[34] L. C. Freeman, "Centrality in social networks: I. Conceptual clarification," *Social Networks*, vol. 1, no. 3, pp. 215–239, 1979.

Hindawi
Complexity
Volume 2021, Article ID 2193782, 15 pages
https://doi.org/10.1155/2021/2193782

*Research Article*

# Comparison of Intercity Travel Network Structure during Daily Time and Holiday in China

**Runze Qi, Jinghu Pan, and Rong Zhang**

*College of Geography and Environmental Science, Northwest Normal University, Lanzhou 730070, China*

Correspondence should be addressed to Jinghu Pan; panjh_nwnu@nwnu.edu.cn

Received 7 May 2021; Revised 12 July 2021; Accepted 29 July 2021; Published 6 August 2021

Academic Editor: Anirban Chakraborti



Intercity travel by residents promotes the regathering and dissemination of social and economic factors. Based on big data from Tencent's location-based service, 346 cities above the prefecture level in China were chosen as study objects, with 2018 as the study time node. To construct the intercity residents' travel network, complex network analysis and GIS spatial analysis methods were used. Furthermore, when analyzing the structural characteristics and spatial differences of Chinese residents' intercity travel from different time perspectives (the whole year, daily, Spring Festival travel rush, and special holidays), Gephi network analysis tools and ArcGIS spatial analysis software were used. The following are the major findings: daily and the whole year intercity travel by Chinese residents, as well as intercity travel during special holidays and the Spring Festival, all exhibit the "diamond" structure, with Beijing, Shanghai, Guangzhou-Shenzhen, and Chengdu-Chongqing at the core. The distribution of lines in and around the "diamond" is large and concentrated from the perspective of the hierarchical nature of the residents' intercity travel network. Significant increases in high-intensity population flow lines within the "diamond" can be seen during Spring Festival travel and holidays. The number of cities involved in the inflow line is significantly greater than that involved in the outflow line, as demonstrated by the number of residents in the first point of travel, indicating that there is a difference between the central cities flowing into and out of the network. The first flow of the central city is the most visible during the Spring Festival travel period. Most cities in the resident intercity travel network have relatively low degrees of centrality, closeness centrality, and betweenness centrality, and the number of cities with large values of the three is small, and they are concentrated in the apex and interior of the "diamond" structure.

## 1. Introduction

Travel migration and population flow are regarded as activities in which production factors are reallocated in space, thereby promoting the regathering and diffusion of social and economic factors [1]. Population travel embodies the interaction between their production and life styles and other production and life spaces; it reflects individuals' ability to participate in social and economic activities. The acceleration of globalization and regional integration has strengthened global ties even further. Regional communication and collaborative development are facilitated by an efficient, comprehensive transportation system, convenient communication services, and urban development modes. The current situation and prospects of urban development, as the output and receiving place of population migration, determine the scale of population inflow and outflow. Residents' travel characteristics include travel purpose, travel time, and travel mode, which are useful for studying residents' behavior and alleviating urban traffic problems. The study of residents' travel can serve as a reference for the rational formulation of policies and has far-reaching implications for guiding the healthy development of urban transportation in the future [2, 3].

Population flow has evolved into a spatial representation of social and economic development in China. The scale of population movement influences the level of social and economic development. In the process of globalization, information technology, and regional integration, the connection between different regions in China is becoming increasingly close, the trend of population flow is clear, the total amount of population travel is increasing year by year, and the travel purposes are becoming more diverse, such as studying, working, traveling, visiting relatives, and so on. Based on this, an in-depth study of the characteristics of China's residents' travel network, mastery of residents' travel rules, and comprehension of urban development differences reflected in residents' flow can provide a scientific theoretical foundation for the formulation of regional development planning.

"Flow space" was first proposed by Castells, an American sociologist, and detailed in his book *The Rise of Network Society* [4, 5]. He believes that "flow space" is the use of a specific medium across geographical distances in space to achieve the exchange of time and space elements. It depicts the transformation of traditional economies, societies, and cultures brought about by the advancement of information technology [4, 6]. Flows, in general, include information technology, people, things, transportation, and capital flow, and space is the material support for these flows. The rapid development of information technology allows cities to connect across physical distances, and the flow affects all aspects of residents' lives in various ways [7]. From the standpoint of geography, "flow space" is a new spatial logic established against the backdrop of the information age, supported by transportation infrastructure, and characterized by the continuous movement of people flow, logistics, traffic flow, information flow, technology flow, and capital flow [8]. The main components of "flow space" are various entity flows and virtual flows, and "flow's" high-speed characteristics make the circulation of various elements in space more convenient and efficient. "Flow" not only shortens the distance between cities but also serves as a carrier for residents' activities, cities, and regions while also promoting the development of urban spatial networks and reshaping geographical space. Because all kinds of elements flow at breakneck speed in today's information society, "flow space" is critical to the formation and development of urban and regional networks [8, 9]. "Flow space" data can now be accurately measured thanks to the rapid advancement of information technology and the efficient application of information data. Using various "flow" elements, residents' travel networks can be explored more dynamically, scientifically, and intuitively, and the essence of spatial relationships can be better approached.

A network is a carrier made up of abstract nodes and connecting node edges that can describe the complex and intertwined objective world [2]. Network or networked research perspectives have been introduced into geography research with the development of research theories and methods [8]. In urban geography, it is assumed that cities are not isolated in regional space but rather have complex interactions with one another, forming the urban network. This type of network has a distinct spatial structure as well as a functional organization [10]. The interaction and complementary advantages between and within cities are strengthened through the exchange of population, materials, information, finance, and other elements. As a result, complex networks of various scales and levels, such as transportation networks, population flow networks, and information networks, have evolved [11]. At the end of the twentieth century, complex network theory based on mathematical graph theory and statistical physics provided a theoretical foundation for the study of network system complexity [12, 13]. Following that, scholars from various fields attempt to use complex network theory to explore the real world by abstracting various complex relations in the objective world and then excavating and describing the complexity characteristics and structural relations of the complex world [14, 15]. Scholars have discovered that the real world network is a small world or scale-free complex network, rather than a regular network connected according to certain rules or a random network connected in a random way [14, 16]. Complex network theory and methods are becoming the primary research tools as network complexity increases [14, 16]. As a result, based on the theory of "flow space" and the network perspective, it is critical to study intercity residents' travel in order to master the law of residents' spatial movement and the dynamic characteristics of intercity spatial correlation on a regional scale [17].

Human activities, such as mobile trajectory, mobile signaling, location request, and so on, will generate a large amount of spatiotemporal information data. It contains rich semantic features and spatiotemporal dimension dynamic association information, which must be exploited and applied in a reasonable, efficient, and comprehensive manner [18]. With the rapid development and popularization of sensor networks, wireless communication, mobile positioning, and Internet technology in the Big Data era, it is now possible to research and obtain high-precision and massive individual movement trajectory data [16, 19]. Big data with geographic positioning information has numerous benefits, including broad coverage, high time precision, multiple attributes, and so on. Simultaneously, spatial big data mining exploration, computer science, geography, and other methods intersect to provide strong support for population flow and another possibility for quantitative research on the spatial characteristics of population flow. Furthermore, statistical analysis of massive data is very useful in discovering the implicit and robust laws of data, which greatly reduces the uncertainty caused by sample randomness [20]. Big data records a large number of people flow, logistics, and information flow data, which not only assists us in conducting fine research [21, 22] but also contributes significantly to improving the accuracy and comprehensiveness of the research. The use of geographic big data to perceive human social activities is currently a hot topic in academic research. The travel big data reflects the displacement and time-space trajectory of human activities, which provides data support for scholars to study deeply and carefully on a microscale [23]. Moreover, the observation and development of the continuity of multiperspectives and large scale bring new ideas and new technologies to geography [23–25].

People do not pay much attention to the intercity residents' travel network at the moment. Mastering the structure and rules of intercity residents' travel networks and reasonably planning the routes and flows of residents' travel has become one of the most effective ways to promote the long-term development of cities and regions. As a result, beginning with different travel time periods (the entire year, daily, special holidays, and the Spring Festival travel rush), this paper conducts a comprehensive evaluation of the structural characteristics and contact pattern of residents' travel networks in 346 Chinese cities, which has both theoretical and practical significance. On the one hand, it can help us understand and grasp the national residents' travel rules and status; on the other hand, it can provide scientific references for regional development planning.

The significance of this paper lies in the following aspects: First of all, the study of residents' intercity travel can provide scientific reference to guide the rational flow of population and the planning of regional town systems; second, we use the "flow space" research method instead of the traditional static local space; and third, we use big data of location instead of traditional static statistics to ensure the timeliness of the data. Fourth, we compare the similarities and differences in Chinese residents' intercity travel network structures over time.

## 2. Related Work

With the advancement of information processing technology and the popularity of smart phones, a number of businesses began to collect population positioning and travel data, which are widely used in scientific research. Scholars analyze the problem of residents' travel using a variety of data sources. Jiang investigated the primary reasons for residents to travel by taxi using Swedish urban taxi trajectory data [26]. Murakami obtained travel information from residents via telephone interview, recorded travel information from private cars combined with GPS, and then conducted a comparative analysis on resident behavior characteristics [21]. Limtanakool et al. discovered differences in behavior characteristics and network structure heterogeneity using data from a European long-distance interregional travel survey [27]. De Montis et al. built an intercity travel network based on intercity commuting data from Italy and studied the structural characteristics of the intercity travel system [28]. Western scholars frequently use Twitter data to examine the temporal and spatial distribution characteristics of population flow [23]. Chinese scholars have frequently used data from Sina Weibo [29, 30], Baidu migration data [31], and Tencent population migration data [32, 33] in recent years. For example, Li et al. [9] used Baidu migration data to examine the characteristics of population migration during the Spring Festival in China. Pan and Lai [34] collected Tencent's population migration data during the Mid Autumn Festival and National Day holidays, divided the holiday into travel, journey, and return periods, and analyzed the population migration in each period. These studies used big data to analyze individual and group behavior and reflect spatial behavior, spatial cognition, and connection mode. Such data can be used to reflect individual and group decision-making in terms of spatiotemporal behavior, and it is becoming a hot research frontier for studying travel and population flow.

## 3. Methods and Data Sources

Chinese residents travel at different times of the year [35], such as during the Spring Festival, Mid Autumn Festival, and National Day. To begin, we examine the proximity effect of a travel network using intercity correlation potential. Second, we investigate the urban hierarchy as reflected by the travel network using transition centrality and transition control.

*3.1. Complex Network Analysis.* Network analysis is a critical component of GIS spatial analysis. Network analysis is founded on graph theory, which encompasses graph theory analysis, optimization analysis, and dynamic analysis [36]. Erdos and Renyi, mathematicians, established a random graph [36] in 1960, which provided a new method for constructing networks. Then, after conducting extensive actual data research, scientists discovered that the majority of actual networks are not regular or random networks. It heralds the arrival of the stage of complex network research. The term "complex network" refers to a network that is somewhere between a completely regular network and a completely random network. It is a small world with a network that is not scaled [37]. This feature is common in population travel networks [38]. Degree and its distribution characteristics, agglomeration degree and its distribution characteristics, betweenness and its distribution characteristics, and so on are the fundamental measures of complex network research [39, 40]. In this paper, the following indicators are used.

*3.1.1. Degree Centrality.* Degree reflects the interaction of a node with other nodes and is an important expression of city interconnection. The higher a node's degree, the greater its degree centrality and the more important its position in the network [41]:

$$C_D(i) = \frac{K_i}{N-1}. \tag{1}$$

$C_D(i)$ is the degree centrality of node $i$, $K_i$ is the degree of node $i$, and $N$ is the number of nodes in the network [41].

*3.1.2. Closeness Centrality.* Closeness centrality measures the proximity of a node to other nodes in the network by using the shortest path distance. The closer a node is to its centrality, the more important it is in the network [42, 43].

$$C_C(i) = \frac{1}{\sum_j D_{ij}}. \tag{2}$$

$C_C(i)$ is the closeness centrality of node $i$; $D_{ij}$ is the shortest path distance between node $i$ and node $j$ [42].

*3.1.3. Betweenness Centrality.* The proportion of nodes passing through all of the network's shortest paths is described by betweenness centrality. If the shortest path between many nodes passes through a specific node, the node has high betweenness centrality; the larger the value, the greater the node's transfer and connection ability in the network [41]:

$$C_B(m) = \sum_{i\neq j} \frac{\sigma_{ij}(m)}{\sigma_{ij}}. \tag{3}$$

$C_B$ $(m)$ is the betweenness centrality of node $m$, $\sigma_{ij}$ is the number of all shortest paths from node $i$ to node $j$, and $\sigma_{ij}$ $(m)$ is the number of all shortest paths from node $i$ to node $j$ through node $m$ [41].

*3.1.4. Measurement of Intercity Correlation.* The proportion of the connection strength of a specific line in the network to the total connection strength [44] is referred to as the intercity correlation potential measure, and it represents the importance of a specific intercity connection edge in the network. The following is the precise formula:

$$\mathrm{RSI}_{ij} = \frac{t_{ij}}{\sum_{i=1}^{I}\sum_{j=1}^{J} t_{ij}}. \tag{4}$$

$t_{ij}$ is the travel population size between city $i$ and city $j$, $\mathrm{RSI}_{ij}$ is the intercity correlation superiority degree, and $0 \leq \mathrm{RSI}_{ij} \leq 1$. If the value is closer to 1, it means that the greater the proportion of lines connecting cities $j$ and $i$, the higher the degree of dominance [22, 45].

*3.1.5. Alter-Based Centrality and Alter-Based Power.* In the study of world city networks, Zachary proposed the concept of recursive centrality and recursive control, which he later changed to Alter-based centrality (AC) and Alter-based power (AP) [46, 47]. He believes that centrality refers to resource concentration and diffusion. The aggregation of resource elements (labor, capital, information, and so on) to world cities, as well as the outward diffusion of resource elements from world cities, is both manifestations of centrality [48]. The control power represents a city's influence and dominance in the resource circulation process. The location of the network and the role it plays determine the size of a city's control power [48].

*3.2. Data.* The data used in this paper is from the "Tencent location big data" platform's population migration data (heat.qq.com/qianxi.php), and the time node is from January 1, 2018, to December 31, 2018, a total of 365 days of more than 600,000 population migration data. The data attributes include the origin and destination cities' longitude and latitude coordinates, the total amount of migration, and the migration ratio of three different modes of transportation (plane, train, and car). The migration volume is determined by the migration ratio, and the population migration volume is used to determine the strength of the network connection between cities. We compiled data to determine the intensity of population flow between cities in 2018 and used it to build the residents' travel network.

Given the study's comprehensiveness and representativeness, we chose Tencent migration data from the entire year (2018) as the basic data and conducted in-depth analysis from various travel periods. In terms of the time period, four representative periods were chosen for the study: daily, Spring Festival travel rush, special holidays, and the entire year. (Daily time: 2018.11.01–2018.12.10, Spring Festival travel rush: 2018.02.01–2018.03.12, Special holidays: Other statutory holidays except for Spring Festival) In order to fully cover the trip, the data includes daily data for two days before and after the holiday. Figure 1 depicts the locations of major cities in China.

## 4. Results

*4.1. Hierarchical Characteristics of Residents' Intercity Travel.* Based on the natural fracture method, we use ArcGIS to classify Chinese residents' intercity travel routes and intensity (Figure 2). Across the four time periods, it is discovered that population travel routes are denser in the east than in the west. In all four time periods, population travel routes are denser in the east than in the west and denser in the south than in the north. The population of high-intensity travelers is concentrated in the east and south. The "Hu Huanyong Line" divides travel routes and the number of Chinese residents into two distinct parts, with more parts in the east and fewer parts in the west, forming a "diamond" structure with Beijing, Shanghai, Guangzhou-Shenzhen, and Chengdu-Chongqing as the core. There are obvious differences in residents' travel patterns across time periods, primarily as follows: According to the entire year residents travel routes and intensity map (Figure 2(a)), Beijing, Shanghai, Dongguan, Shenzhen, Chongqing, Chengdu, and Changsha are the major cities with more than 31.8 million residents; 3081–31.8 million lines are diffused and filled along the "diamond" structure. The line space between 12.68 million and 30.81 million people is mostly filled with "diamond" structures, and there is also a trend of northeast and southwest expansions. The number of visitors ranges between 5.16 and 12.68 million. There are numerous routes, the majority of which connect provincial capitals in the northwest, southwest, and northeast. Short routes connecting provincial capital cities and other prefecture-level cities characterize the travel routes of 1.34 to 5.16 million people. In order to more clearly show the difference in population travel routes and intensity during the three time periods of daily (Figure 2(b)), special holidays (Figure 2(c)), and Spring Festival (Figure 2(d)), we categorize them according to the standard of the daily time period, it shows that the number of lines with the high-intensity population during the three periods of daily, holidays, and Spring Festival transportation has increased significantly, and the carrying pressure on each line has generally increased. In the first level (population travel intensity >2.59 million), there are 25 lines in a daily time period, 52 in holidays, and 69 in Spring Festival travel. The number of daily, special holiday, and Spring Festival travel rush travel routes among the 1.11–2.59 million people is 97, 232, and 289, respectively. There are 440, 673, and 1022 population travel routes in the third level (450–1.11 million people). The number of lines at the fourth level (between 120,000 and 450,000 people) is 1517, 2018, and 2815, respectively. The number of routes at


Figure 1: Locations of major cities in China.

each level has increased from daily-special holiday-Spring Festival transportation, with the number of routes at the third and fourth levels increasing the most. This demonstrates that the increased routes are primarily concentrated in the "diamond" interior and the northeast and northwest regions. The number of people traveling increased significantly during the Spring Festival and special holidays. The purpose of travel during the Spring Festival transportation is complex (return home, study, tourism, etc.), which mostly involves medium and long-distance travel, including many cities in China. Travel on special holidays, on the other hand, is generally limited by time, so it tends to travel nearby.

According to the statistical table of residents' travel network scale in different periods (Table 1), the number of intercity trips of Chinese residents reached 14,947.96 million in 2018, with the Spring Festival travel rush accounting for 15.72 percent of the total trips. Special holiday trips accounted for 11.57 percent of total trips, while daily trips accounted for 7.35 percent. During the Spring Festival and special holidays, residents' trips are more concentrated, and the number of trips is greater. The average number of trips per day has surpassed 50 million. Spring Festival, China's most important traditional festival, holds a special place in the hearts of the Chinese people. Spring Festival is a day for family reunions, and most wanderers return home for the occasion. This unique sense of belonging to one's hometown increases the likelihood of traveling. During special holidays, many people choose to travel or engage in other activities to enrich their leisure time. Throughout the year, the contact line has the highest value, followed by the Spring Festival and daily periods, and the number of sides is the smallest during special holidays. The density shows that the whole year > - Spring Festival travel rush > special holidays > daily. In 2018, the links between each node city in the residents' travel network are closer than during daily and special holidays, and the links during the Spring Festival travel rush have increased, but there is still a gap when compared to the whole year.

We chose to characterize the dominant relationship in the network, i.e., the first flow, in order to more clearly and concisely demonstrate the travel of residents in each city in China. It is a method of revealing the important spatial structural characteristics of the entire network by reducing the amount of data, thereby reflecting the network status of cities on a macroscopic scale [49]. In the population mobility network, the first flow is the line of a city with the highest flow. The first flow consists of the first inflow and first outflow. The route with the most inbound traffic in a city is referred to as the first inflow, while the route with the most outbound traffic is referred to as the first outflow. Figure 3 depicts the first outflow and first inflow lines for each city (the lines with inflow and outflow degree values of 1 are hidden in the figure to facilitate a clear and intuitive display of the flow effect). And overall, in each of the four time periods, the number of cities involved in inflow lines is significantly greater than the number of cities involved in outflow lines, and there are differences in the cities with the highest number of inflow and outflow lines in the same time period:

(1) During the Spring Festival period, Chengdu has the highest first inflow, followed by Beijing, Wuhan, Urumqi, and Zhengzhou, among others. Each inflow center city is primarily the provincial capital city and receives inflow from surrounding cities, primarily prefecture-level cities in the province. Chengdu receives visitors not only from neighboring prefecture-level cities but also from provincial capital cities such as Nanjing, Jinan, and Lhasa. Beijing accepts inflows from cities such as Lanzhou, Changsha, Wuhan, Harbin, and Nanchang, encompassing a diverse range of cities with a high level of connectivity. Two cities flowed out of each other from the first outflow, such as Taiyuan-Jinzhong, Guangzhou-Foshan, Yuncheng-Linfen, and so on. During the Spring Festival, Chongqing has the most outflow routes, followed by Chengdu and Beijing. Lai conducted the same study [49]. Chongqing was found to be a net outflow city during the Spring Festival, while Shenzhen, Kunming, and Shanghai experienced greater outflows of residents, with outflow grades improving. During the Spring Festival, the first outflow routes are staggered and highly interconnected, whereas the first inflow has a gyratory distribution in the center.

(2) There is a clear trend of staggered links between the first inflow cities during special holiday periods, with some cities avoiding provincial capital cities and linking with other prefecture-level cities. Chengdu maintains first place, Wuhan has increased the number of contact lines and has surpassed Beijing, the number of cities with the first mutual inflow of the two cities has increased, and Chengdu has more contacts with Xi'an and Nanjing during this period.

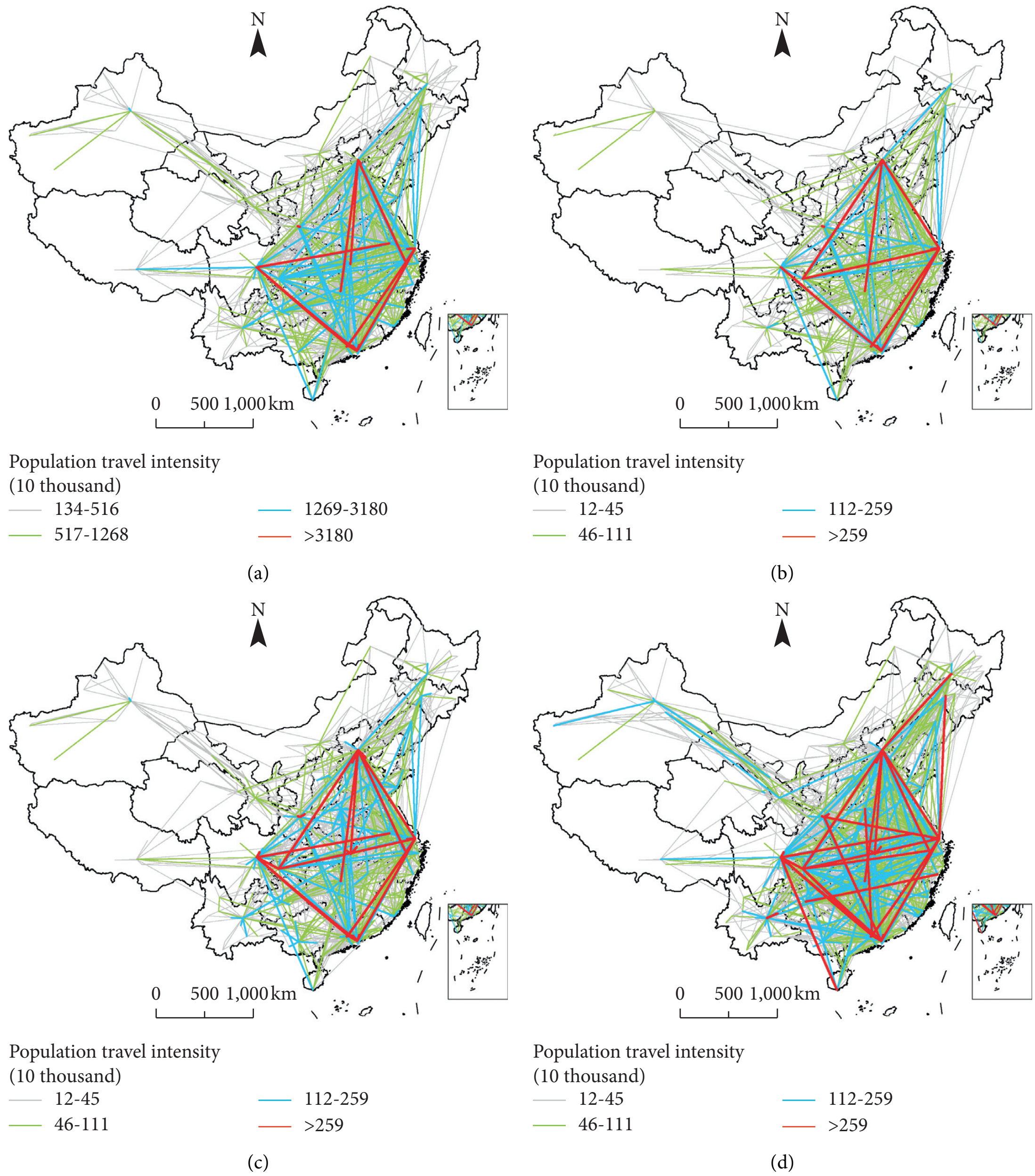


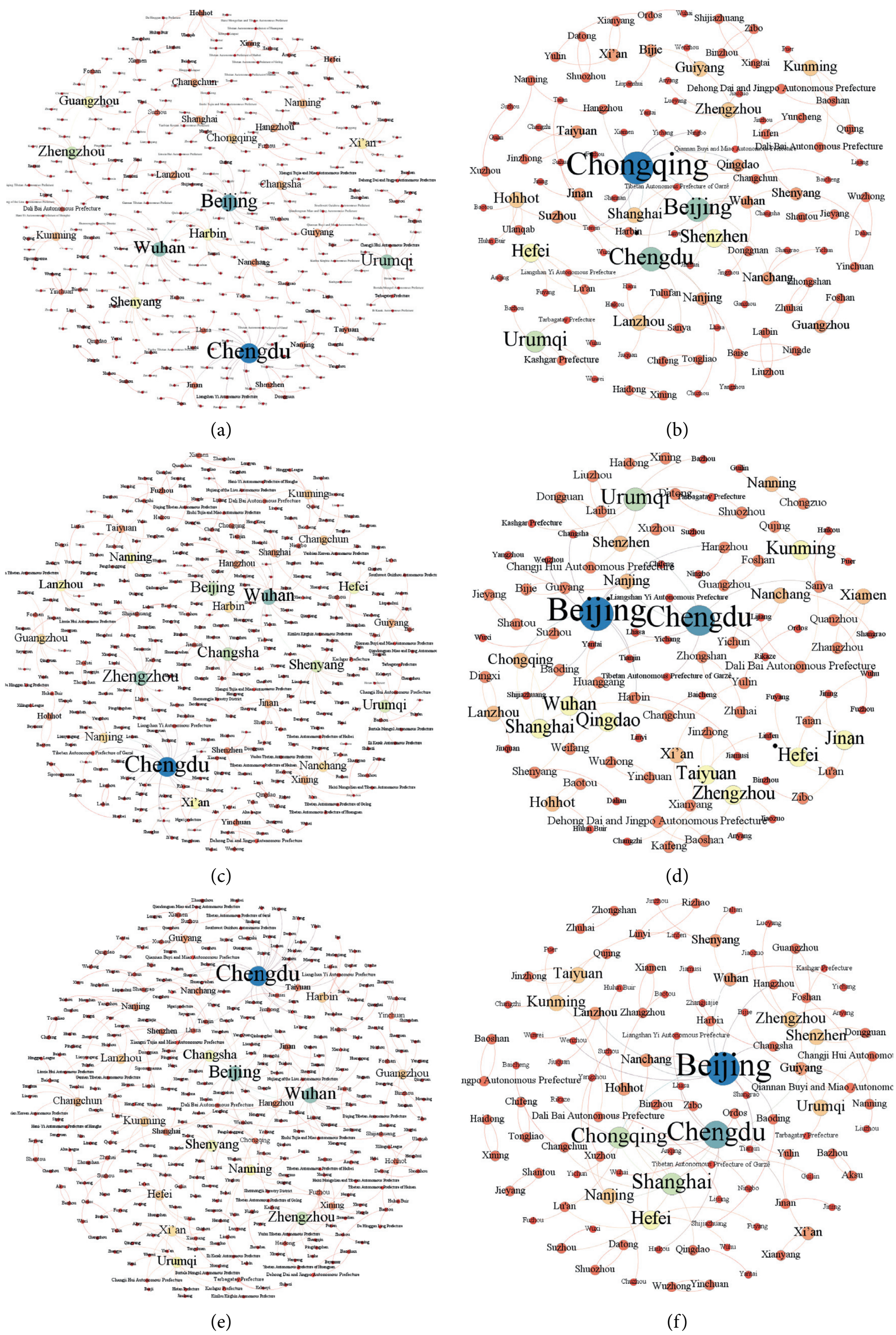

FIGURE 2: Intercity travel routes and intensity of Chinese residents in different periods. (a) The whole year period. (b) Daily time. (c) Special holidays. (d) Spring festival travel rush.

TABLE 1: Statistics of residents' travel network scale in different periods.

| Period | Total number of trips | Number of contact edges | Average daily trips | Proportion | Network density |
|---|---|---|---|---|---|
| The whole year | $1.49 * 10^{10}$ | 23759 | $4.09 * 10^{7}$ | 1.000 | 0.199 |
| Spring Festival travel rush | $2.34 * 10^{9}$ | 14063 | $5.87 * 10^{7}$ | 15.719 | 0.118 |
| Special holidays | $1.72 * 10^{9}$ | 11945 | $5.08 * 10^{7}$ | 11.566 | 0.100 |
| Daily time | $1.09 * 10^{9}$ | 12557 | $2.74 * 10^{7}$ | 7.347 | 0.105 |

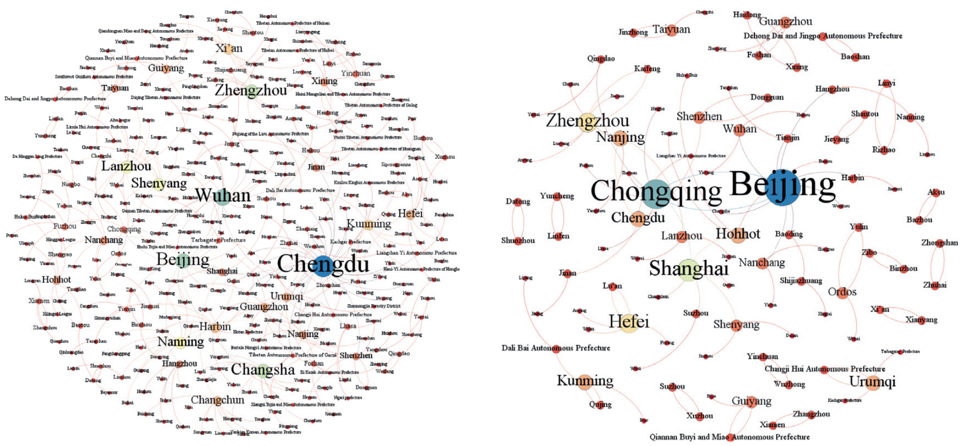
Beijing and Chengdu have the first outflow lines during the special holiday period, and tourist cities such as Urumqi, Jinan, and Kunming have risen in rank, with more travelers on special holidays and cities with more tourist resources more likely to attract tourists.

(3) Cities such as Chengdu, Beijing, Wuhan, and Zhengzhou continue to dominate the first inflows for the entire year, with significant interlocking links between cities and a more visible network intertwining phenomenon. Beijing is unquestionably the leader among the first outflows. Beijing, as China's




(a) (b) (c) (d) (e) (f)

Figure 3: Continued.




(g) (h)

Figure 3: The first flow (outflow and inflow) of urban residents at various times. (*Note.* The size of the circle indicates the number of connecting lines. The circle grows in size as the number of lines increases. The color of connecting lines matches that of nodes.) (a) First inflow during Spring Festival. (b) First outflow during Spring Festival. (c) First inflow during special holidays. (d) First outflow during special holidays. (e) First inflow during the whole year. (f) First outflow during the whole year. (g) First inflow during daily time. (h) First outflow during daily time.

capital, receives the first outflows from several provinces, which is the main concentration of the first outflows in northern China. Furthermore, the first outflow routes from Shanghai increase significantly throughout the year, particularly with cities such as Chengdu, Chongqing, Nanjing, and Hefei.

(4) The first inflow and first outflow in daily periods are comparable to the total year period. For example, the number of contact lines in Chengdu is consistent with the entire year period, but there are minor variations. Line intertwining is most visible in the first inflow, with the number of first inflow lines decreasing in Beijing, Wuhan, and other provincial capital cities while increasing in some lower-ranked cities. The number of cities with degree values greater than one decreases significantly in the first outflow, and Chongqing has a higher status than Chengdu and is directly linked to Beijing, Nanjing, Lanzhou, Shenzhen, and Hohhot, among others. The figure contains the most ellipses, indicating that the number of interconnected cities is the greatest.

*4.2. Centrality Characteristics.* There are three indicators in this paper: degree centrality, closeness centrality, and betweenness centrality. They are chosen to calculate the nodes of the residential travel network at various time intervals. The degree centrality of a node city reflects its ability to interact with other nodes, and the higher its value, the more important and powerful its position in the network [41]. The closeness of the value of centrality reflects a node's position in the network, and the higher its value is, the closer it is to the network [50]. Betweenness centrality refers to a city's ability to act as a bridge node in order to communicate with other node cities, and the higher its value, the greater the node's transit and bridging ability in the network [41].

In this paper, we use Gephi software to calculate the three types of indicators, which we then aggregate and express using origin. Figure 4 depicts the results of the centrality of nodes in the travel network of Chinese residents over the course of a year. Most cities, in particular, have low betweenness centrality, degree centrality, and closeness centrality, with 298 cities meeting betweenness centrality $<1000$, degree centrality $<200$, and closeness centrality $<0.8$ at the same time, indicating a relatively obvious gap in Chinese city development. The proportion of cities that are highly developed is small, while the proportion of cities that are still developing is large. In 3D space, all three indicators form an ascending curve, with Shanghai having the highest value, having not only a large number of cities connected to it, but also playing the most obvious bridging role. Shanghai has the greatest ability to monitor and control the "node pairs" of other cities, as well as the most important position and status in the travel network of Chinese residents. Following them are the cities of Beijing, Chongqing, Guangzhou, Shenzhen, and Chengdu. These cities are at the far end of the "diamond" in the Chinese residents' travel network. They have a pivotal position in building the travel network of China, which has a strong economic base to support these cities to take on a larger travel population. Provincial capital cities such as Wuhan, Xi'an, Tianjin, Changsha, Lanzhou, and Jinan exist between the low and high values, as do noncapital regional centers such as Xianyang, Qingdao, and Foshan. Overall, the top ranked cities are primarily developed cities in the east, whereas western cities are more frequently located in low-value areas.

Figure 5 depicts the spatial representation of the centrality of the entire year period. There are six cities in the first tier ($>0.723$) of the spatial distribution of closeness centrality, namely Beijing, Shanghai, Chongqing, Guangzhou, Shenzhen, and Chengdu, indicating that these cities are the

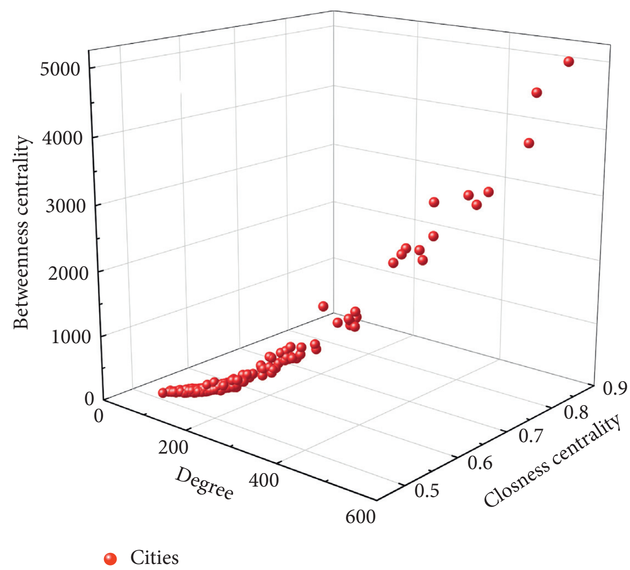



FIGURE 4: Analysis of residents' travel network in the whole year.

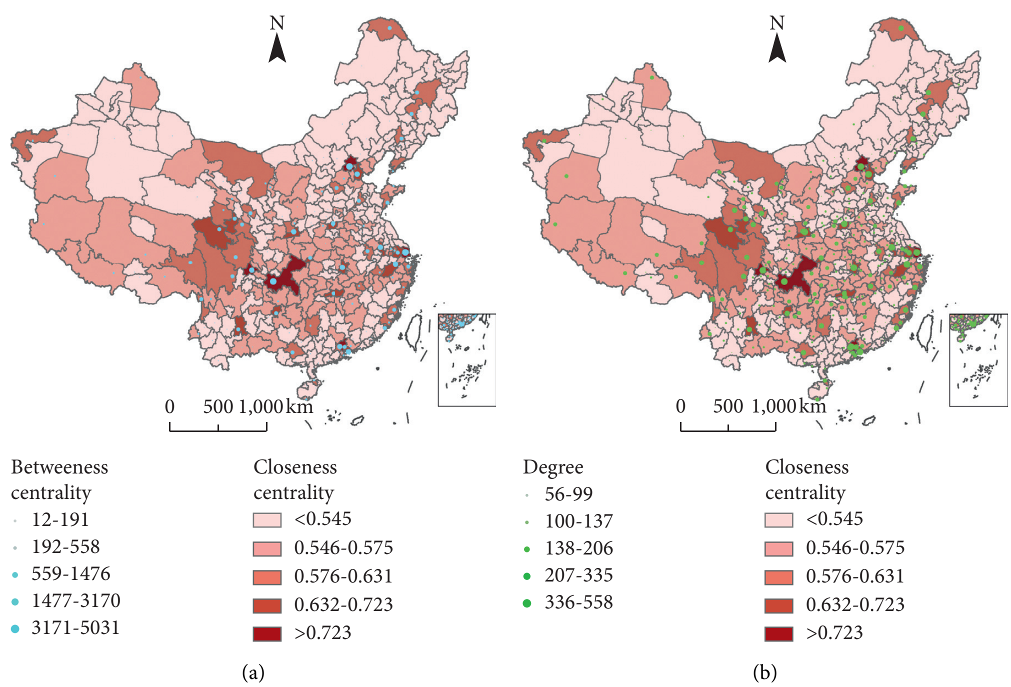



FIGURE 5: Spatial distribution of centrality in the whole year.

most centrally located in the entire residential travel network. The second tier includes 15 cities that are secondary important nodes in the travel network, such as Hangzhou, Wuhan, Xi'an, Nanjing, and Dongguan. The third-tier cities are primarily provincial capital cities in the west and northeast and noncapital cities in the developed eastern region. The fourth and fifth tiers jointly guard the most remote and underdeveloped areas, which are also the nodes in the travel network of Chinese residents that cannot be abandoned. The degree centrality and betweenness centrality are normalized and presented as a bar chart in the graph, demonstrating that cities with high degree centrality also have high betweenness centrality, indicating that they are consistent. The cities with a high degree of centrality are spatially distributed around the urban agglomerations of Beijing, Shanghai, Guangzhou, and Chongqing and the Beijing-Tianjin-Hebei, Pearl River Delta Chengdu-Chongqing, and Yangtze River's middle reaches. Furthermore, the capital cities of each province have significantly higher degree centrality values than neighboring cities, with an average degree centrality value of around 130 per city at the prefecture level and above, and there are more cases of multi-city connections. The betweenness centrality mainly demonstrates the transit and acceptance capacity of network nodes, reflecting the gap between cities. Shanghai, for example, has the highest value (up to 5031), whereas Taizhou only has 12. Unlike degree centrality, betweenness centrality appears in each provincial capital city with larger values, either the provincial capital city or a more developed city, and some provincial capital cities in the east even have multiple high value points in one provincial capital. For example, Zhengzhou accepts nationwide travel exchange, disseminating it to neighboring cities, such as Luoyang and Kaifeng, which not only accept Zhengzhou's population flow but also connect cities with lower levels, and cities with larger betweenness centrality radiating and driving cities with smaller values around them.

Table 2 shows the number of cities with varying degrees of centrality indicators for three time periods: daily, special holidays, and spring festival. Cities in the lower tiers of the three centrality indicators account for a large proportion of the total. Degree centrality and betweenness centrality have the most cities in the fifth tier, while closeness centrality has the most cities in the fourth tier. Because the study cities are prefecture-level and above, few nodes belong to the end nodes in the battle for the central position of the travel network, and most nodes are more or less connected to each other. Cities in Northwest China and some remote cities have low closeness centrality due to topographical constraints, whereas cities with high participation in national residents' travel have high centrality. For daily, special holidays, and spring festival periods, the number of cities in the first and second levels of degree centrality is 17, 15, and 18, respectively, accounting for about 5% of the total number of cities, and the proportion of cities in the higher levels is small. In the fourth level, the number of cities during the spring festival is twice that of the other two periods, and students and workers return to their hometowns, and travel direction flows from big cities to small and medium cities, indirectly increasing the degree of centrality of third and fourth tier cities. It is worth noting that during special holidays, the number of high-quality cities is the lowest, owing to the fact that people are limited by time during this period, and they will mostly choose short-distance self-driving trips or close trips around, and do not generally go to big cities like Beijing and Shanghai. Betweenness centrality is always stable across the entire resident travel network, and the gap is also the smallest across all three time periods. During the Spring Festival, the number of cities varies the most between each tier, with the second tier experiencing the greatest growth and the third tier experiencing the smallest. The difference between the daily and Spring Festival periods is small in the closeness centrality. The number of cities in the third and fourth tiers clearly increases during the Spring Festival period, whereas the number of cities in the first to fourth tiers remains at its lowest during special holidays. When compared to the daily and Spring Festival periods, the number of cities in the fifth tier is the highest.

*4.3. Spatial Proximity Effect.* We compare the superiority of intercity connections over four time periods, and we take the period of each line's maximum intercity connection superiority to obtain the superiority of different time periods. Figure 6 depicts each city's superiority relationship with other cities. A total of 7473 pairs of city connections are superior to other time periods during the Spring Festival, with the connections between cities in the south being spatially concentrated. The northern section consists primarily of a connection line centered on Beijing. During normal business hours, there are 4867 superior linkage lines, forming a diamond-shaped superior linkage structure with the cores of Guangzhou-Shenzhen, Beijing, Shanghai, and Chengdu-Chongqing. On special holidays, there are 4316 pairs of advantageous linkage lines that show a clear proximity effect in space, primarily connecting provincial capital cities with other secondary cities in the province. The proximity of superior linkage lines is the most important feature of special holidays, and residents will prefer to travel nearby during this period due to factors such as travel time, travel destination, and scale. Throughout the year, there are 7101 pairs of city connections that outperform other time periods, primarily in the form of core connection routes between Wuhan, Beijing, Chengdu, and Chongqing, such as Wuhan-Beijing, Chengdu-Lhasa, Wuhan-Kunming, and so on. Overall, the proximity effect is most visible in the four time periods for special holidays, which are primarily formed by connections between provincial capital cities and neighboring cities. The spring festival, daily life, and the entire year are based primarily on cross-regional connections with significant spatial differences.

*4.4. City Hierarchy.* AC and AP jointly determine the status of cities in urban network analysis [51, 52]. In this paper, we use population travel data from the annual time period to calculate the AC and AP in order to investigate the city hierarchy formed under the resident travel network. The results were graded by natural break using ArcGIS, and they are shown in Figure 7. The city hierarchy presented by both AC and AP values has a pyramidal structure, which means that the higher the AC and AP values, the fewer cities there are. The AP and AC values in most cities are in the same gradation tier. For example, in Beijing, Shanghai, Chongqing, Chengdu, Shenzhen, and Dongguan, the AP and AC values are in the highest tier, but there are some cities that differ. Guangzhou, for example, has an AC value in the second tier and an AP value in the third tier. Overall, cities with high AC values outnumber those with high AP values. Cities' AC values have a positive correlation with their AP values, and cities with high AC values also have high AP values. Their ability to gather and distribute resources is strong, as is their ability to dominate resources, but there are differences in some cities. According to Zachary et al.'s method of measuring world city information networks, cities in travel networks are classified into four types based on AC and control scores: high centrality-high control, high centrality-low control, low centrality-high control, and low centrality-low control. Furthermore, city types classified according to centrality and control power can better identify the city status and attribute characteristics.

Table 2: Urban numbers of centrality at different levels.

| Index | Period | First level | Second level | Third level | Fourth level | Fifth level |
|---|---|---|---|---|---|---|
| Degree centrality | | >288 | 176–288 | 103–175 | 60–102 | <59 |
| | Daily time | 7 | 10 | 27 | 80 | 222 |
| | Special holidays | 6 | 9 | 22 | 77 | 232 |
| | Spring Festival travel rush | 8 | 10 | 32 | 148 | 148 |
| Betweenness centrality | | >5559.4 | 1705.5–5559.4 | 670.7–1705.4 | 217.9–670.6 | <217.8 |
| | Daily time | 6 | 7 | 17 | 23 | 293 |
| | Special holidays | 5 | 7 | 12 | 22 | 300 |
| | Spring Festival travel rush | 4 | 10 | 11 | 22 | 299 |
| Closeness centrality | | >0.636 | 0.568–0.636 | 0.532–0.567 | 0.506–0.532 | <0.506 |
| | Daily time | 7 | 15 | 32 | 263 | 29 |
| | Special holidays | 6 | 9 | 31 | 225 | 75 |
| | Spring Festival travel rush | 7 | 13 | 49 | 239 | 38 |

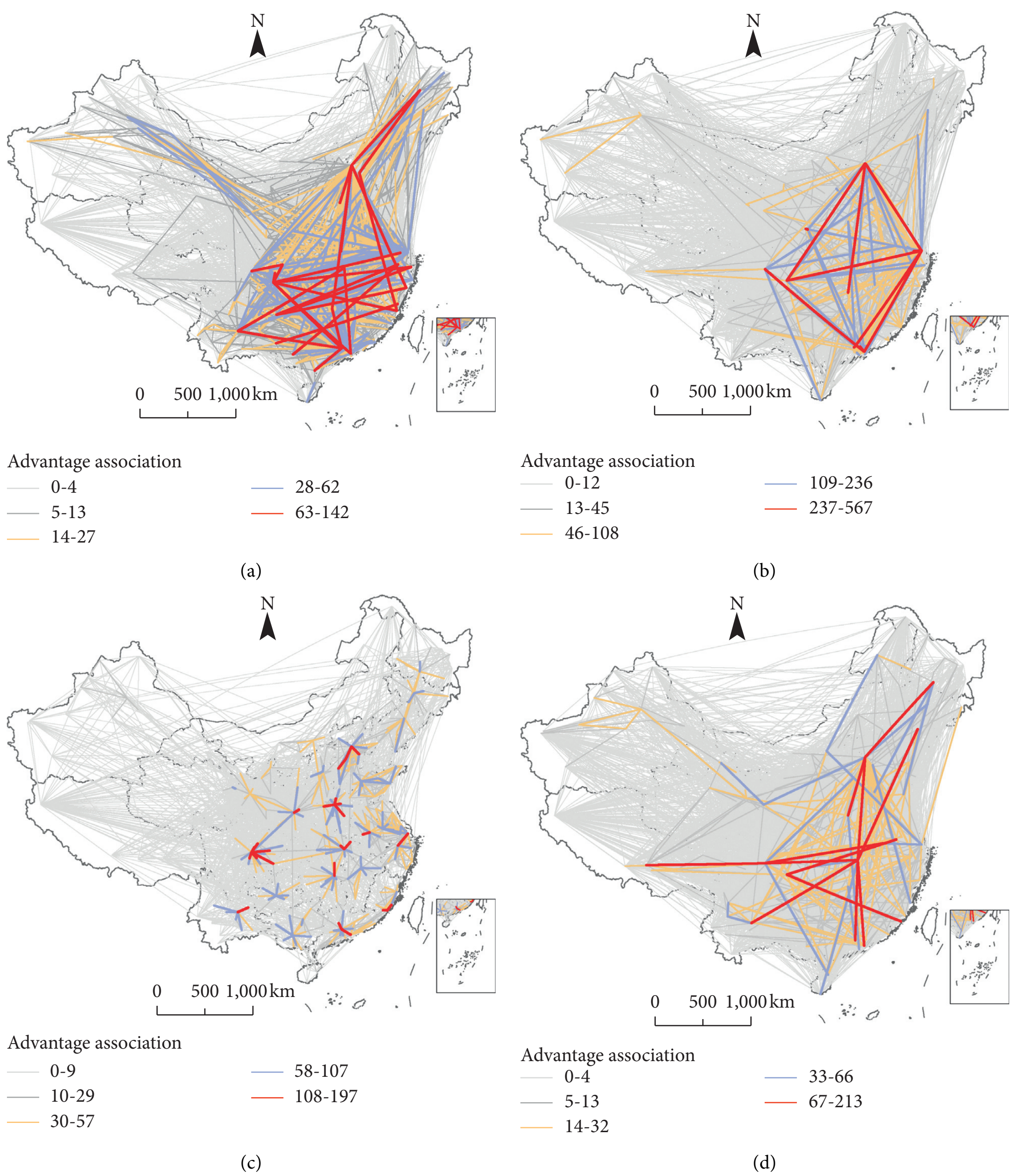
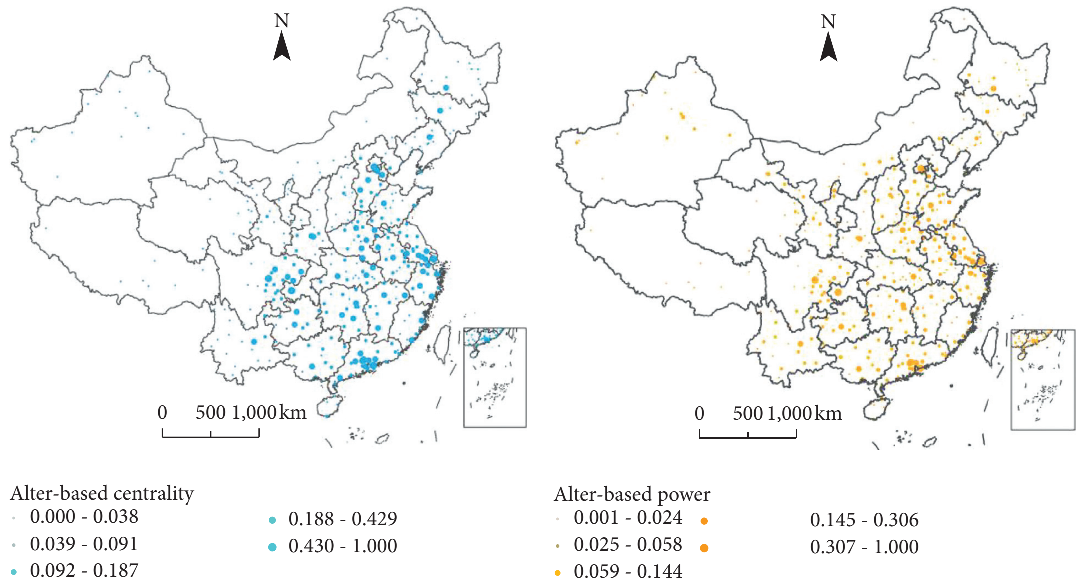



Figure 6: Comparison of intercity travel network advantages of Chinese residents in different periods.




(a) (b)

FIGURE 7: Classification of alter-based centrality and alter-based power in China.

In the residential travel network's city hierarchy (Figure 8), the ratio of the four city types: high centricity-high control, high centricity-low control, low centricity-high control, and low centricity-low control is 23 : 22 : 2 : 299. Low centricity-low control cities have the most people, low centricity-high control cities have the fewest, and high centricity-low control cities primarily serve as hubs. These cities have more opportunities for connection and convenient transportation and a suitable location provides more transit opportunities for small-scale cities to connect with large cities. Low centricity-high control cities serve as a gateway to neighboring cities, compressing opportunities and possibilities for regional residents to travel and demonstrating a monopoly on regional resource circulation [46]. As a result, a large number of small-scale cities (which can only exchange network resources through gateway cities) suffer from path dependence and lack of paths [46]. Cities with a high centricity-high control not only have a strong capacity to gather and spread resources, but they also have a strong dominance over resources and can be classified as typical cities.

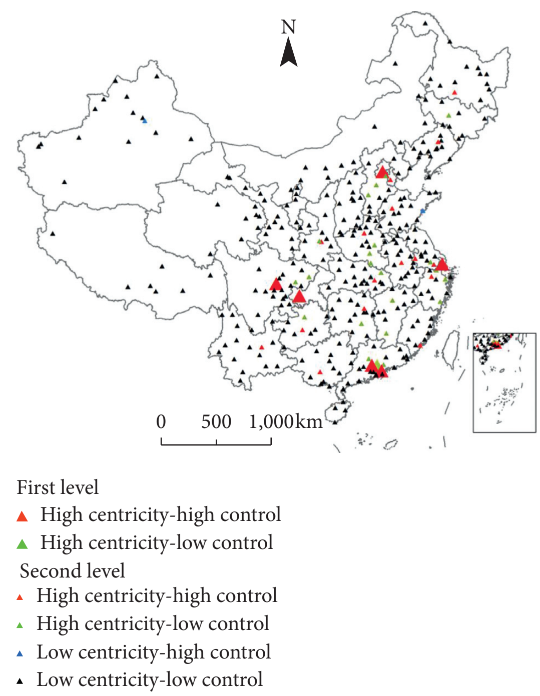



FIGURE 8: Hierarchical structure of Chinese cities.

## 5. Discussion

In terms of resident travel, we discover that there is a clear hierarchy. The first level is a "diamond" structure with nodes in Beijing, Shanghai, Guangzhou, Shenzhen, Chengdu, Chongqing, and Wuhan; this structure supports China's entire transportation network and plays an important role in transportation lines. The second level is the network of contacts between provincial capitals. There are numerous travel routes and a large number of people between cities, which supplement and support the overall "diamond" structure. The third level transportation network is most visible as a link between prefecture-level cities and provincial capital cities. The fourth level network is primarily a contact network between cities at the prefecture level. Whether we look at the entire year, daily, special holidays, or the more special spring festival, we can see the existence of this hierarchy, which is consistent with the findings of Lai and Pan [27]. The existence of hierarchy also demonstrates the existence of a hierarchical structure between cities in China.

In terms of urban spatial connection, we discover that there are differences in urban spatial connection across four time periods. During the Spring Festival travel rush, the space focuses on the connection between the southern cities, while the north primarily uses Beijing as the center of the connection line. In space, a rhombic dominant association structure with the cores of Guangzhou-Shenzhen, Beijing, Shanghai, and Chengdu-Chongqing forms in daily time. Special holidays exhibit an obvious proximity effect in space, with the majority of them focusing on the connection between provincial capital cities and other prefecture-level cities in the province. We discover that the proximity effect of special holidays is the most obvious when we compare the differences in urban spatial connection in four periods. The main benefit of special holidays is the short contact line distance. Residents in this period will choose to travel nearby based on travel time, destination, scale, and other factors. The Spring Festival travel rush, daily and year-round periods are primarily cross-regional connections with significant spatial differences.

We compute cities' AC and AP, as well as their hierarchical structure, which is characterized by a pyramid structure with a small top and a large bottom. Cities at the top include Beijing, Shanghai, Guangzhou, Shenzhen, Chengdu, and Chongqing, among others. They have a strong sense of centrality and control, and they are the city with the best resident travel connections. These cities are the most appealing to residents; regardless of the time period, the intensity of residents' travel between cities is at the highest level. It demonstrates that population travel between these cities has become routine, unaffected by time. More cities have low centrality and control power. It demonstrates that these cities are at a relatively low level in the transportation network, and their appeal to residents and network control is not strong. There is a significant difference in travel time between the Spring Festival travel rush and holidays. Residents primarily return to their hometowns during the Spring Festival travel rush. People who work outside of the city will attend the reunion. As can be seen from the result chart of travel intensity, both the travel routes and the travel intensity of residents have improved, with long-distance travel performing best. Residents mostly take short trips and play during the holidays, which are limited by the length of the holiday. The number of short-distance trips increases sharply, as does the number of travel routes, as can be seen intuitively in the advantage correlation results. We believe that most existing studies on residents' travel are limited to a single time period, such as working days or holidays, and that there is a lack of research on residents' travel status throughout the year. As a result, this paper divides the entire year based on the travel habits of Chinese residents and finally determines the research period as follows: Spring Festival travel rush, daily time, special holidays, and the entire year. We improve the comparison of residents' travel differences across time periods in the time period selection, and we enrich the research on residents' travel directions. Simultaneously, the analysis of multiple periods can help decision makers better understand the travel rules and travel characteristics of residents over the course of a year, as well as serve as a point of reference for the formulation of relevant policies.

In comparison to previous statistical data, Tencent migration data is updated at a rapid rate and with high timeliness, breaking through the lag effect of traditional data. Tencent has a large user base and can provide a large amount of data with greater accuracy. Tencent had more users and higher accuracy when compared to other data. However, due to the nature of data generation and acquisition and the protection of personal privacy, it was impossible to obtain social attributes such as travelers' occupation, gender, age, and travel purposes. It was impossible to delve deeper into the population's willingness and the group effect. Furthermore, some travel paths may be disassembled, and the points of origin and destination cannot be studied as network nodes, resulting in errors in the research results. In the future, it will be necessary to integrate various data sources in order to analyze the characteristics of residents' travel and to provide new research ideas to global researchers.

The majority of existing studies examine the characteristics of residents' travel networks over a specific time period or in a specific region. There are few studies on a national scale and even fewer on different time periods in the country. This study has enriched the research by analyzing the structure characteristics of the Chinese residents of the intercity travel network over different travel time periods. Our research yielded the following new findings: (1) There is a clear hierarchy in the spatial structure of Chinese residents' intercity travel network, which is not monotonous but dynamic for different travel periods. (2) The main origins and destinations of Chinese residents change from time to time, as evidenced by the first flow. (3) The centrality and control power of Chinese cities are not the same in terms of transportation. Cities at different levels in the urban hierarchy of residents' travel networks play different roles and functions in the transportation network. This research can assist policymakers in proposing appropriate management measures for various travel time periods in order to improve the efficiency of residents' travel and optimize the residential travel network. It will also assist readers all over the world in better understanding the travel habits of Chinese residents and deepening their understanding of China.

## 6. Conclusions

This paper comprehensively evaluates the travel network structure characteristics and connection patterns of 346 Chinese cities based on different travel periods (the entire year period, daily time, special holidays, and Spring Festival travel rush). We arrived at the following conclusions:

(1) In terms of routes and people, the travel network of Chinese residents along the "Hu Huanyong Line" showed an obvious trend of more in the east and less in the west during the four designated time periods.

(2) From a hierarchical standpoint, the "diamond" frame structure centralized travel areas with Beijing, Shanghai, Guangzhou-Shenzhen, and Chengdu-Chongqing as nodes formed during these four periods, and the internal routes of the "diamond"

structure increased significantly during the Spring Festival travel rush and holidays.

(3) Centrality depicts each city's position and status in the travel network, and the number of cities with low value is large. Centrality demonstrates that fewer cities have significant status, and they are primarily concentrated at the "diamond" structure's apex.

(4) The spatial proximity effect demonstrates that during special holidays, residents' travel advantages are primarily related to proximity, as represented by the connection between provincial capitals and surrounding cities. During the Spring Festival travel rush, throughout the year, and on a daily basis, residents' travel advantages are primarily related to cross-region travel, and the spatial distribution varies.

(5) The city level reflected by the travel network is primarily divided into four types: high centrality-high control, high centrality-low control, low centrality-high control, and low centrality-low control, with a city ratio of 23 : 22 : 2 : 299.

## Data Availability

The population migration data of Tencent in 2018 are the core data of this paper, which can be obtained from the location big data released by Tencent. The data come from https://heat.qq.com/qianxi.php. In order to protect users' privacy, Tencent has closed the data acquisition interface, and the latest data can no longer be obtained. Therefore, it is inconvenient for us to disclose the obtained historical data. Some sample data can be sent via e-mail (qirunze_gis@163.com) to Qi Runze.

## Conflicts of Interest

The authors declare that there are no conflicts of interest regarding the publication of this paper.

## Acknowledgments

This work was supported by the National Natural Science Foundation of China (Grant no. 42071216).

## References

[1] H. De Haas, "Migration and development: a theoretical perspective," *International Migration Review*, vol. 44, no. 1, pp. 227–264, 2010.

[2] B. J. L. Berry, "Cities as systems within systems of cities," *Papers in Regional Science*, vol. 13, no. 1, pp. 147–163, 1964.

[3] W. Zhang, B. Derudder, J. Wang, and F. Witlox, "An analysis of the determinants of the multiplex urban networks in the Yangtze River Delta," *Tijdschrift voor economische en sociale geografie*, vol. 111, no. 2, pp. 117–133, 2020.

[4] M. Castells, *The Informational City: Information Technology, Economic Restructuring, and the Urban-Regional Process*, Blackwell Oxford, Oxford, UK, 1989.

[5] M. Castells, *The Rise of the Network Society*, Blackwell Publishing, Oxford, UK, 1996.

[6] T. Pei, C. Song, S. Guo et al., "Big geodata mining: objective, connotations and research issues," *Journal of Geographical Sciences*, vol. 30, no. 2, pp. 251–266, 2020.

[7] J. Henderson, P. Dicken, M. Hess, N. Coe, and H. W.-C. Yeung, "Global production networks and the analysis of economic development," *Review of International Political Economy*, vol. 9, no. 3, pp. 436–464, 2002.

[8] C. Gan, M. Voda, K. Wang, L. Chen, and J. Ye, "Spatial network structure of the tourism economy in urban agglomeration: a social network analysis," *Journal of Hospitality and Tourism Management*, vol. 47, pp. 124–133, 2021.

[9] J. Li, Q. Ye, X. Deng, Y. Liu, and Y. Liu, "Spatial-temporal analysis on spring festival travel rush in China based on multisource big data," *Sustainability*, vol. 8, no. 11, p. 1184, 2016.

[10] M. Tao, Y. Huang, and H. Tao, "Urban network externalities, agglomeration economies and urban economic growth," *Cities*, vol. 107, Article ID 102882, 2020.

[11] X. Liu, B. Derudder, and K. Wu, "Measuring polycentric urban development in China: an intercity transportation network perspective," *Regional Studies*, vol. 50, no. 8, pp. 1302–1315, 2016.

[12] D. J. Watts and S. H. Strogatz, "Collective dynamics of 'small-world' networks," *Nature*, vol. 393, no. 6684, pp. 440–442, 1998.

[13] A.-L. Barabási and R. Albert, "Emergence of scaling in random networks," *Science*, vol. 286, no. 5439, pp. 509–512, 1999.

[14] Z. Jing and J. Wang, "Sustainable development evaluation of the society-economy-environment in a resource-based city of China: A complex network approach," *Journal of Cleaner Production*, vol. 263, Article ID 121510, 2020.

[15] D. A. Smith and M. F. Timberlake, "World city networks and hierarchies, 1977–1997," *American Behavioral Scientist*, vol. 44, no. 10, pp. 1656–1678, 2001.

[16] Z. Zhang, J. Wang, and B. Li, "Determining the influence factors of soil organic carbon stock in opencast coal-mine dumps based on complex network theory," *CATENA*, vol. 173, pp. 433–444, 2019.

[17] H. Yang, F. Dobruszkes, J. Wang, M. Dijst, and P. Witte, "Comparing China's urban systems in high-speed railway and airline networks," *Journal of Transport Geography*, vol. 68, pp. 233–244, 2018.

[18] T. Xia, X. Song, H. Zhang, X. Song, H. Kanasugi, and R. Shibasaki, "Measuring spatio-temporal accessibility to emergency medical services through big GPS data," *Health & Place*, vol. 56, pp. 53–62, 2019.

[19] Y. Zhao, H. Zhang, L. An, and Q. Liu, "Improving the approaches of traffic demand forecasting in the big data era," *Cities*, vol. 82, pp. 19–26, 2018.

[20] P. Zhao and H. Hu, "Geographical patterns of traffic congestion in growing megacities: big data analytics from Beijing," *Cities*, vol. 92, pp. 164–174, 2019.

[21] E. Murakami and D. P. Wagner, "Can using global positioning system (GPS) improve trip reporting?" *Transportation Research Part C: Emerging Technologies*, vol. 7, no. 2-3, pp. 149–165, 1999.

[22] Y. Liu, Z. Sui, C. Kang, and Y. Gao, "Uncovering patterns of inter-urban trip and spatial interaction from social media check-in data," *PLoS One*, vol. 9, no. 1, Article ID e86026, 2014.

[23] D. Wong and Q. Huang, ""Voting with their feet": delineating the sphere of influence using social media data," *ISPRS International Journal of Geo-Information*, vol. 6, no. 11, p. 325, 2017.

[24] H. Barbosa, M. Barthelemy, G. Ghoshal et al., "Human mobility: models and applications," *Physics Reports*, vol. 734, pp. 1–74, 2018.

[25] T. Schwanen, "Geographies of transport I," *Progress in Human Geography*, vol. 40, no. 1, pp. 126–137, 2016.

[26] B. Jiang, J. Yin, and S. Zhao, "Characterizing the human mobility pattern in a large street network," *Physical Review. E, Statistical, Nonlinear, and Soft Matter Physics*, vol. 80, no. 2, Article ID 021136, 2009.

[27] N. Limtanakool, M. Dijst, and T. Schwanen, "A theoretical framework and methodology for characterising national urban systems on the basis of flows of people: empirical evidence for France and Germany," *Urban Studies*, vol. 44, no. 11, pp. 2123–2145, 2007.

[28] A. De Montis, S. Caschili, and A. Chessa, "Time evolution of complex networks: commuting systems in insular Italy," *Journal of Geographical Systems*, vol. 13, no. 1, pp. 49–65, 2011.

[29] W. Zhang, B. Derudder, J. Wang, W. Shen, and F. Witlox, "Using location-based social media to chart the patterns of people moving between cities: the case of Weibo-users in the Yangtze River Delta," *Journal of Urban Technology*, vol. 23, no. 3, pp. 91–111, 2016.

[30] R. Zhu, Y. Wang, D. Lin et al., "Exploring the rich-club characteristic in internal migration: evidence from Chinese Chunyun migration," *Cities*, vol. 114, Article ID 103198, 2021.

[31] Y. Wei, W. Song, C. Xiu, and Z. Zhao, "The rich-club phenomenon of China's population flow network during the country's spring festival," *Applied Geography*, vol. 96, pp. 77–85, 2018.

[32] W. Zhang, Z. Chong, X. Li, and G. Nie, "Spatial patterns and determinant factors of population flow networks in China: analysis on Tencent Location Big Data," *Cities*, vol. 99, Article ID 102640, 2020.

[33] J. Xu, A. Li, D. Li et al., "Difference of urban development in China from the perspective of passenger transport around Spring Festival," *Applied Geography*, vol. 87, pp. 85–96, 2017.

[34] J. Pan and J. Lai, "Spatial pattern of population mobility among cities in China: case study of the National Day plus Mid-Autumn Festival based on Tencent migration data," *Cities*, vol. 94, pp. 55–69, 2019.

[35] W. Yang, H. Chen, and W. Wang, "The path and time efficiency of residents' trips of different purposes with different travel modes: an empirical study in Guangzhou, China," *Journal of Transport Geography*, vol. 88, Article ID 102829, 2020.

[36] P. Erdős and A. Rényi, "On the evolution of random graphs," *Publications of the Mathematical Institute of the Hungarian Academy of Sciences*, vol. 5, no. 1, pp. 17–60, 1960.

[37] D. He, Z. Liu, and B. Wang, *Complex System and Complex Network*, Higher Education Press, Beijing, China, 2009.

[38] C. Zhu, X. Wang, and L. Zhu, "A novel method of evaluating key nodes in complex networks," *Chaos, Solitons & Fractals*, vol. 96, pp. 43–50, 2017.

[39] P. Bonacich, "Power and centrality: a family of measures," *American Journal of Sociology*, vol. 92, no. 5, pp. 1170–1182, 1987.

[40] M. C. Mahutga, X. Ma, D. A. Smith, and M. Timberlake, "Economic globalisation and the structure of the world city system: the case of airline passenger data," *Urban Studies*, vol. 47, no. 9, pp. 1925–1947, 2010.

[41] N. Perra, B. Gonçalves, R. Pastor-Satorras, and A Vespignani, "Activity driven modeling of time varying networks," *Scientific Reports*, vol. 2, no. 1, 2012.

[42] U. Brandes, "A faster algorithm for betweenness centrality∗," *The Journal of Mathematical Sociology*, vol. 25, no. 2, pp. 163–177, 2001.

[43] G. Sabidussi, "The centrality index of a graph," *Psychometrika*, vol. 31, no. 4, pp. 581–603, 1966.

[44] H. Yang, M. Dijst, P. Witte, H. Van Ginkel, and W. Yang, "The spatial structure of high speed railways and urban networks in China: a flow approach," *Tijdschrift voor economische en sociale geografie*, vol. 109, no. 1, pp. 109–128, 2018.

[45] Y. Liu, C. Kang, S. Gao, Y. Xiao, and Y. Tian, "Understanding intra-urban trip patterns from taxi trajectory data," *Journal of Geographical Systems*, vol. 14, no. 4, pp. 463–483, 2012.

[46] Z. Neal, "Does world city network research need eigenvectors?" *Urban Studies*, vol. 50, no. 8, pp. 1648–1659, 2013.

[47] J. P. Boyd, M. C. Mahutga, and D. A. Smith, "Measuring centrality and power recursively in the world city network: a reply to Neal," *Urban Studies*, vol. 50, no. 8, pp. 1641–1647, 2013.

[48] Z. Zhao, Y. Wei, and R. Pang, "Research on the centrality and control of China's urban network transformation based on the inter-provincial mobility of population-also on the conditionality of recursive theory in urban network research," *Acta Geographica Sinica*, vol. 72, no. 6, pp. 1032–1048, 2017.

[49] J. Lai and J. Pan, "China's city network structural characteristics based on population flow during spring festival travel rush: empirical analysis of "tencent migration" big data," *Journal of Urban Planning and Development*, vol. 146, no. 2, Article ID 04020018, 2020.

[50] C. Fang, "Significant progress and future development directions in the study of urban agglomerations in China," *Acta Geographica Sinica*, vol. 69, no. 8, pp. 1130–1144, 2014.

[51] A. S. Alderson and J. Beckfield, "Power and position in the world city system," *American Journal of Sociology*, vol. 109, no. 4, pp. 811–851, 2004.

[52] P. J. Taylor, G. Catalano, and D. R. F. Walker, "Measurement of the world city network," *Urban Studies*, vol. 39, no. 13, pp. 2367–2376, 2002.

Hindawi
Complexity
Volume 2021, Article ID 9931985, 15 pages
https://doi.org/10.1155/2021/9931985

*Research Article*

# Spatiotemporal Evolution and Complexity of Urban Networks in China, 1978–2019: An Enterprise Linkages Perspective

**Jiayi Lu**[1,2] **and Dongqi Sun**[1]

[1]*Key Laboratory of Regional Sustainable Development Modeling, Institute of Geographic Sciences and Natural Resources Research, CAS, Beijing 100101, China*
[2]*College of Resources and Environment, University of Chinese Academy of Sciences, Beijing 100049, China*

Correspondence should be addressed to Dongqi Sun; sundq@igsnrr.ac.cn

Received 22 March 2021; Accepted 20 May 2021; Published 9 June 2021

Academic Editor: Jing-Hu Pan



With the development of globalization and informatization, the relationships among cities have become closer, and a "network" paradigm in urban studies is gaining attention. To examine China's urban network evolution in a long time series, we used flow-based data to measure enterprise linkages from 1978 to 2019. We investigated the spatiotemporal evolution and complexity characteristics of urban networks in China and arrived at the following conclusions. (1) Intercity enterprise linkages in China have been continuously strengthened. The scale and density of urban networks have increased rapidly. Although the distribution of node cities' importance and influence has been significantly unbalanced, the degree of which has lessened over time. (2) Network density has significantly improved since 1978, gradually forming a monocentric (Beijing) radial pattern. From the beginning of the twenty-first century, the status of core nodes (e.g., Shanghai) has gradually become prominent. Finally, four vertices stood out in 2019, forming a stable diamond structure. The spatial connection flows of enterprises constituted the core networks with Beijing as the center, skeleton networks with trunk lines formed by subnodes, and regional networks covering a wide range of peripheral areas. (3) China's urban networks were typically small-scale and scale-free. However, the scale-free characteristics were weakened after 2010. The overall scale gap of intercity enterprise linkages gradually narrowed, and the structure of urban networks became optimized. Meanwhile, the urban networks were heterogeneous. There were more cities with headquarter-branches and active investment behaviors, which had strong influence and control over networks, playing their functions of "broker" and "transfer."

## 1. Introduction

Since the 1990s, the rapid development of globalization and informatization has deeply influenced and been key in reconstructing the global urban system. The development of any region or city is no longer carried out in a closed system [1–3]. As Sassen, Castells, Derudder, and Taylor have pointed out, cities are increasingly involved in factor flows around the world [4–6], and their importance tends to be determined by their connections with other cities in the world rather than by their absolute size [7, 8]. Therefore, scholars have turned their attention from urban hierarchies to urban networks, forming a new paradigm that can better explain the structure of urban spatial organization [9].

Urban network research originated from western scholars' research on the network of world city systems. The research comprises three stages: world city, global city, and world city networks. In the 1960s, Hall first put forward a criterion for "world city" from the perspective of urban function [10]. Friedmann then proposed a "world city" hypothesis based on the theory of new international divisions of labor, but his focus remained on cities alone and did not involve measuring their interrelationships [11]. In the early 1990s, under the profound influence of globalization, Sassen proposed the concept of a "global city." She believed that the gathering of advanced producer services in global cities made them develop two core functions, financial centers and headquarters of multinational corporations, thus shaping the transnational city networks [4]. At the same time, the acceleration of informatization promoted the emergence of a "network" paradigm based on the complex relationships among cities [5]. Castells believed that the wide application of information network technologies has brought about space-time compression. "Space of flows" has gradually replaced "space of places" as the dominant form of spatial organization [12]. The proposal of "space of flows" provides a solid theoretical basis for the empirical study of world city networks [13]. Based on the work of Sassen and Castells, the Globalization and World Cities Study Group and Network (GaWC) headed by Peter Taylor began to study the influence of spatial organization networks of large enterprises on city regions and on the whole world. Taylor conceptualized the relationships among cities as an "interlocking network model" and considered the subnodal level represented by producer service firms to play a key role in world city networks. The "flows" formed by the global layout of producer service firms stimulated connections among cities and the formation of world city networks [14]. Research on urban systems has gradually turned its perspective from "attribute" to "relationship" and its paradigm from "hierarchies" to "networks" [15].

Today, urban network research has been gaining increased attention. Much research uses microdata, such as enterprise flows, traffic flows, information flows, and knowledge flows, to measure urban networks [16–20]. As economic globalization and regional economic integration continue to spread, the economic connections among cities, as well as intercity enterprise linkages, continue to strengthen [21]. Studies of urban networks from the perspective of enterprise linkages are becoming increasingly prevalent in western urban networks research [22]. There are three main reasons for this: (1) enterprise forms the main body of industry connections and the flows of multielements [23]; (2) enterprise linkage data are easier to obtain than other data and are more generalizable [24]; (3) urban network studies based on enterprise linkages can be applied at almost any spatial scale [22], such as the global scale [6, 25], national scale [26, 27], regional scale [28, 29], and city scale [30]. Representative studies mainly fall into two categories: one is based on advanced producer services (APS) and the other is based on headquarter-branch enterprises.

Urban network research based on APS enterprise linkages is mainly conducted at a global scale and a regional scale. At the global scale, GaWC used distribution data of APS enterprises to build world city networks and analyzed modern service relationships among cities. The findings had a far-reaching impact [31]. At the regional scale, POLYNET led by Peter Hall introduced the perspective of APS enterprise networks to study European urban systems and proposed the concept of a "megacity region" [32]. From then on, scholars have carried out a variety of empirical studies using APS enterprise data. For example, based on GaWC's research on the status of Chinese cities in world city networks, with the latest data from Chinese APS enterprises from 2010 to 2016, Derudder et al. found that the connectivity for all cities in China increased, except for that of Hong Kong, Macao, and Kaohsiung in Taiwan. However, this large-scale connectivity growth was geographically unbalanced [33]. Neal et al. predicted the possibility of 104 APS enterprises in 525 cities in the world expanding, shrinking, or maintaining the status quo and prospectively judged their location selection, thus revealing the transformation trend of world city networks [34]. Although extensive empirical research on urban networks based on APS enterprise linkages has been carried out [35], Krätke suggested that it has ignored the intercity linkages constructed by enterprise in other industries. This is especially with regard to developing countries and newly industrialized countries, where the real economy, e.g., manufacturing industry, is indispensable and should not be excluded. Hence, APS enterprise data alone cannot tell the whole story of urban networks [36].

Urban network research based on headquarter-branch enterprises is represented by the work of Alderson and Beckfield. Based on the location data of the headquarter-branches of 446 industry-wide multinational corporates in 3692 cities, they discussed the centrality of nodes, "core-periphery" structure, and other network characteristics [37]. They introduced social network analysis into urban network research field, which has been widely used in recent years. For example, Carroll applied social network analysis to study the world city networks formed by board members of the world's top 350 multinational corporations [38]; Pan et al. applied social network analysis to study the cooperation of China's advanced producer service enterprises in the Initial Public Offering (IPO) [39]. Wall and Knaap compared world city networks based on APS enterprises with those based on industry-wide enterprises, employing the distribution data of the headquarter-branches of the world's top 100 multinational corporates in 2259 cities in 2005. In general, urban network research based on headquarter-branch enterprises emphasized the network control of headquarters in global cities over the cities where the branches are located. Focusing on industry-wide enterprise networks can better reflect the panorama of city connections. Therefore, headquarter-branch enterprises have become important in measuring urban network characteristics [40].

Urban network research on China originated from the urban system theories of China put forward by domestic scholars in view of national conditions in the 1980s. "Three structures and one network" proposed by Song and Gu [41, 42], "pole-axis" theory, and "T-shaped" spatial structure strategy proposed by Lu [43–45] have always played an important role in the planning and development of territorial space and the construction of new urbanization in China [46, 47]. The "network" trend of China's urban spatial structure has been enhanced over time [48]. In this context, the empirical research of China's urban networks from the perspective of enterprise linkages roughly follows that of the western path and has the following main characteristics: the research mainly adopts a single indicator of headquarter-branch enterprises or APS enterprises [8, 49], focuses on the main developed urban agglomerations or regions in China [50, 51], and selects specific years or short time series [52, 53].

Therefore, compared with Leng et al.'s study using a gravitational method [54] and Wu et al.'s study based on a single year [27], both of which investigated the complexity of urban networks at the national level from an enterprise linkages perspective, we used flow-based data and mixed indicators. We applied GIS, Gephi, Matlab, and other technical means to explore the spatiotemporal evolution and complexity characteristics of Chinese urban networks since the Reform and Opening-Up. With the improved measurement indicators and long time series, we hope to better reflect the evolution and complexity of urban networks at the national scale. This will hopefully provide a scientific reference for the coordinated development of macro city regions and related policy formulation.

## 2. Materials and Methods

*2.1. Measurement Indicators.* In contrast with traditional measurement of intercity linkages using a gravity model improvement [55, 56], we gathered flow-based data to construct China's urban networks, to accurately reflect actual urban network connections [27]. In this study, headquarter-branches and enterprise investment were applied as measurement indicators of enterprise linkages.

*2.1.1. Headquarter-Branches.* As mentioned earlier, urban network research from an enterprise linkages perspective can be divided into two categories: one is based on APS [57, 58] and the other is based on headquarter-branches [37, 59]. The former has a certain explanatory power with regard to social and economic ties at the global scale, but not enough to truly reflect the actual structure of urban networks at the national scale [60]. Since connections lie in production, sales, supply, capital, and many other aspects between headquarters and branches, cross-regional headquarter-branch enterprises help to reveal the functional ties between cities [61]. Therefore, this article regarded industry-wide headquarter-branch enterprises as "agents" of cities, applying them as a measurement indicator in the study of urban networks in China.

*2.1.2. Enterprise Investment.* Enterprise investment, especially referring to off-site investment, which is an investment in regions away from the site of the enterprises' main business, is an effective way for enterprises to expand their market and improve the competitive advantages of products [62]. Enterprises choose places to invest after careful strategic analysis. They form association networks, which enable capital and other elements to flow fluently among cities, influencing the structure of urban systems, thus forming an important entry point of studying urban network [63]. In addition to "intraenterprise" organizational connections formed by headquarter-branches, "interenterprise" connections are also considered in enterprises' off-site investment. Therefore, this article not only used the data of headquarter-branches but also used enterprises' off-site investment indicators so as to reveal urban network connections more clearly.

*2.2. Research Period.* Globalization, marketization, and other factors have been the institutional basis of China's reform and development since 1978 and have profoundly affected the regional economic development pattern of China [64]. Based on this, we investigated the temporal and spatial evolution and complexity characteristics of China's urban networks from 1978 to 2019 and selected five typical years, 1978, 1992, 2001, 2010, and 2019. The reasons for this are as follows: (1) 1978 represents the early stage of China's Reform and Opening-Up, when a highly centralized planned economic system was implemented. In 1992, the goal of establishing a socialist market economic system was made clear, and marketization began to intensify and spread. The period 1978–1992 represents the planned economy period of China. (2) 2001 marked China's accession to the World Trade Organization (WTO) when China took the initiative in meeting the challenge of economic globalization and entered a new stage of opening up to the world. The period 1992–2001 represents the socialist market development period of China. (3) In 2010, to mitigate the impact of the international financial crisis, China accelerated the transformation of economic development, beginning to formulate the "Twelfth Five-Year Plan" for economic restructuring, optimization, and upgrading; thus, 2001–2010 represents the globalization period of China. (4) The period 2010–2019 represents the strategic period of China's economic restructuring and upgrading [65].

*2.3. Research Area.* We investigated 353 cities in mainland China (including four municipalities directly under central government, 292 prefecture-level cities, nine districts, 30 autonomous prefectures, three leagues, and 15 county-level cities directly under the government of provinces and autonomous regions). The number of cities involved in enterprise linkages has increased year by year (Figure 1 and Table 1). In 1978, only 221 cities established enterprise linkages, accounting for 62.6% of the total number of cities in the research area. In 1992, the number increased to 349, and the newly added cities were mainly located in Xinjiang, Tibet, Inner Mongolia, and central provinces. In 2001, another two cities were added, Nagqu and Ali in Tibet. In 2010, Kunyu City in Xinjiang was added. Cities involved in enterprise linkages remained unchanged in 2019. So far, only Huyanghe City in Xinjiang has not established enterprise contact with other cities in China (Figure 1).

*2.4. Data Processing and Research Methods.* Registration data for China's industrial and commercial enterprises in 1978, 1992, 2001, 2010, and 2019 were obtained, and an overlay analysis with 353 municipal administrative divisions in mainland China was carried out in ArcGIS 10.3. Thus, the enterprise database was converted to shapefile data in ArcGIS, and city pairs with enterprise flows in 1978, 1992, 2001, 2010, and 2019 were calculated (Table 1). From 1978 to 2019, the number of pieces of enterprise flows in Chinese cities increased greatly, from 598 in 1978 to 90003 in 2019, almost a 150-fold increase. The processing procedure of the 598, 6985, 27195, 48521, and 90003 enterprise flows in 1978,


Figure 1: Distribution map of (newly added) node cities from 1978 to 2019.

Table 1: Basic research data.

| Year | City pairs with enterprise flows (unit: pieces) | | | | Cities involved |
|---|---|---|---|---|---|
| | Headquarter-branches | Investor-investee | Total flows before processing | Total flows after processing | |
| 1978 | 229 | 369 | 598 | 527 | 221 |
| 1992 | 3438 | 3547 | 6985 | 5515 | 349 |
| 2001 | 14691 | 12504 | 27195 | 19264 | 351 |
| 2010 | 25161 | 23360 | 48521 | 33542 | 352 |
| 2019 | 42644 | 47359 | 90003 | 58700 | 352 |

1992, 2001, 2010, and 2019 is as follows: (1) the research focuses on intercity contacts, so the enterprise linkage samples within a city were eliminated. (2) The number of headquarter-branches and the amount of off-site investment were used as indicators to measure the strengths of intercity linkages. Since they are different in dimensions, we adopted an entropy method to ensure the accuracy and objectivity of the index weight determination; this method is also suitable for double indicators [66, 67]. Because the entropy method is an objective weighting method based on the degree of data dispersion, and there are natural differences in the degree of data dispersion over the five years, the weights of headquarter-branches and off-site investments may differ. In fact, the weights of headquarter-branches were 0.34, 0.32, 0.22, 0.24, and 0.28 in 1978, 1992, 2001, 2010, and 2019; the weights of off-site investments were 0.66, 0.68, 0.78, 0.76, and 0.72, respectively. (3) City pairs with linkage strengths were established, with 527, 5515, 19264, 33542, and 58700 groups of city pairs, respectively. Cities were abstracted as nodes, and the enterprise flows among cities were abstracted as edges to construct directed and weighted urban networks for China.

The study applied a complex network approach to measure and analyze the spatiotemporal evolution and complexity of urban networks in China. Since "networks" are becoming a new form of and new research paradigm for intercity spatial structure, a complex network constructed using the interaction among entities of a complex system provides a new means of studying network complexity. In recent years, the complex network approach has been widely used in urban economic network research and has achieved impressive results [68, 69]. Researchers used complexity theory and tools that directly describe the topology of complex networks, e.g., the degree centrality, average path length, and clustering coefficient, to study the nodes and edges of urban economic networks [27, 70] (Figure 2).

*2.4.1. Degree and Related Centrality Indicators*

(1) Degree, degree distribution, weighted degree, and neighborhood degree: "degree" refers to the number of cities connected to a given city, reflecting the centrality of the city in the network. The larger the value of the degree, the higher the centrality of the city [71]. The calculation formula of degree is as follows:

$$k_i = \sum_{j \in n} a_{ij}, \tag{1}$$

in which $a_{ij}$ is the number of cities $j$ connected to city $i$ and $n$ is the total number of cities in the network.

If the number of cities with degree $k$ in the network is $n_k$, then the probability distribution $p(k)$ is used to describe the degree distribution in the network:

$$p(k) = \frac{n_k}{n}. \tag{2}$$

Considering the small scale of some actual networks and the discontinuity of the degree distribution, the cumulative probability distribution $P(k)$ is often used to reduce errors [27]:

$$P(k) = \sum_{k'=k} p(k'). \tag{3}$$

In directed networks, the degree can also be divided into "indegree" and "outdegree." The indegree of a city refers to the pieces of flow ending in a city, while the outdegree refers to the pieces of flow starting from a city. In weighted networks, the weighted degree refers to the average weight of intercity flows, i.e., the average strength value of linkages between a city and other cities. The calculation formula is as follows:

$$S_i = \sum_{j \in n} W_{ij} a_{ij}, \tag{4}$$

in which $W_{ij}$ is the weight of the connection edge between cities $i$ and $j$.

Degree correlation analysis is used to describe the matching characteristics of networks. The formula for calculating the neighborhood degree of each node city is as follows [27, 71]:

$$k_{nn,i} = \frac{1}{k_i} \sum_{j \in N_i} k_j = \frac{1}{k_i} \sum_{j=1}^{n} a_{ij} k_j, \tag{5}$$

in which $N_i$ represents the set of cities connected with city $i$.

(2) Degree centrality, neighborhood centrality, and betweenness centrality are used to measure the centrality of nodes in a network. Neighborhood centrality refers to the sum of the shortest path distances from a given city to other cities, reflecting the relative accessibility of a city in the network. Betweenness centrality refers to the number of shortest paths through a city, which reflects the transit and convergence function of the city in the network [72]. The calculation formulas of degree centrality, neighborhood centrality, and betweenness centrality are described by the following:

$$\mathrm{DC} = \frac{1}{n-1} k_i, \tag{6}$$

$$\mathrm{CC}_i = \left[ \frac{1}{n-1} \sum_{j=1, j \neq i}^{n} d_{ij} \right]^{-1}, \tag{7}$$

$$\mathrm{BC}_k = \frac{2}{n^2 - 3n + 2} \sum_{i=1, j \neq k}^{n} \sum_{j \neq k}^{n} \frac{\delta^k{}_{ij}}{\delta_{ij}}, \tag{8}$$

where $n$ is the total number of cities in the network, $k_i$ is the degree of city $i$, $\delta_{ij}^k$ is the number of shortest paths passing through city $k$ from city $i$ to city $j$, $\delta_{ij}$ is the number of shortest paths from city $i$ to city $j$, and $d_{ij}$ is the number of the shortest paths between any two cities in the network.

*2.4.2. Average Path Length.* Average path length refers to the average number of shortest paths between any two cities in the network, which reflects the overall nature of the network [72]. The calculation formula is as follows:

$$\mathrm{L} = \frac{1}{1/2n(n-1)} \sum_{i>j} d_{ij}, \tag{9}$$

in which $L$ is the average path length, $n$ is the total number of cities in the network, and $d_{ij}$ is the number of shortest paths from city $i$ to city $j$. The smaller $L$ is, the better the connectivity of the network is, and the better the performance and efficiency of the spatial network organization are [72].

*2.4.3. Clustering Coefficient.* The clustering coefficient is used to measure the degree of network agglomeration, i.e., the closeness between cities and neighboring cities in the network, and reflects the local attribute of the network [72]. The calculation formula is as follows:

$$C_i = \frac{2E_i}{k_i(k_i-1)}, \tag{10}$$

in which $C_i$ is the clustering coefficient of city $i$, $E_i$ is the actual number of edges between city $i$ and its neighboring cities, and $k_i$ is the degree of city $i$. $C_i$ is between 0 and 1, and

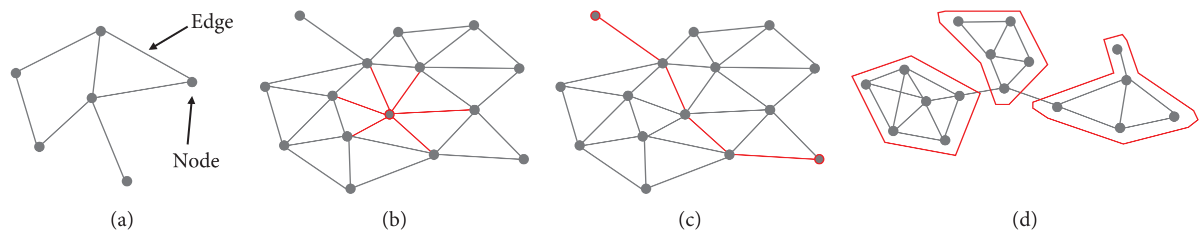



FIGURE 2: Models of some statistic tools for complex network topology. (a) Topology. (b) Degree Centrality. (c) Path length. (d) Clustering.

the larger the value of $C_i$, the closer the connection between city $i$ and its neighboring cities.

The average clustering coefficient reflects the closeness of cities in the whole network. The calculation formula is as follows:

$$C = \frac{1}{n}\sum_{i=1}^{n} C_i, \tag{11}$$

in which $C$ is the average clustering coefficient of the network, $C_i$ is the clustering coefficient of city $i$, and $n$ is the total number of cities in the network. The larger the value of $C$, the closer the local connections of the whole network.

## 3. Spatiotemporal Evolution of Urban Networks in China

*3.1. Topology Evolution of Urban Networks in China.* From 1978 to 2019, the scale of China's urban networks expanded rapidly. The number of nodes and edges in the network increased from 211 to 527 in 1978 to 352 and 58700 in 2019, respectively. The network density increased from 0.011 in 1978 to 0.475 in 2019, and the network diameter decreased from 8 in 1978 to 3 in 2019, indicating that the urban network density of China increased greatly. However, during this process, there were significant differences in the evolution of nodes in China's urban networks (Table 2).

Degree centrality and betweenness centrality reflect the importance and influence of nodes in the network. From 1978 to 2019, the coefficient of variation of degree centrality remained above 3.3, and the Gini coefficient remained above 0.6. The coefficient of variation of betweenness centrality decreased year by year, but the Gini coefficient also remained above 0.6. This illustrates the discrete and unbalanced distribution of the importance and influence of China's node cities. However, the degree of unbalances has lessened over time.

*3.2. Spatial Structure Evolution of Urban Networks in China.* Using the ArcGIS spatial analysis platform, the topology of China's urban networks was transformed into spatial connections, whose strengths were divided into four grades. The spatial structure of China's urban networks from 1978 to 2019 is shown in Figure 3. From 1978 to 2019, China's urban networks roughly formed a core diamond structure. The four vertices of the diamond were Beijing (north), Shanghai (east), Guangzhou-Shenzhen (south), and Chengdu (west), which confirmed the findings of Wu et al. [27], Ma et al. [73], and Zhang et al. [74].

The diamond structure was gradually formed in the continuous evolution of the urban networks. In 1978, China's urban networks were relatively sparse. At this time, linkage flows were constrained by distance. Node cities with the highest levels of connection strength were Kunming-Yuxi (0.665), Beihai-Nanning (0.335), and Changsha-Zhuzhou (0.309), all of which were within provincial boundaries. Nanning and Beijing were the most important urban network nodes, and their weighted degrees reached 3.387 and 1.793, respectively, while the values of other cities were below 1 at this time. Nanning gained a high priority in China's node cities. This was mainly due to the "third-line" construction that begun in the mid-1960s, which has led to the formation of a number of industrial cities and bases with military, machinery, and electronics in southwest China. Large- and medium-sized enterprises have formed in these industrial cities [75]. These enterprises have a good industrial base and strong technical force and have established strong enterprise ties with other cities in the province and with important node cities, such as Guangzhou and Beijing.

In 1992, China's urban networks became significantly denser (Table 1) and began to present a monocentric radial pattern. Beijing stood out as the core node of the urban networks and began to focus on developing linkages with cities in the Pearl River Delta, Yangtze River Delta, and Chengdu-Chongqing region, with Beijing-Guangzhou (0.980), Beijing-Shanghai (0.331), and Beijing-Chongqing (0.329) becoming the highest-level linkages.

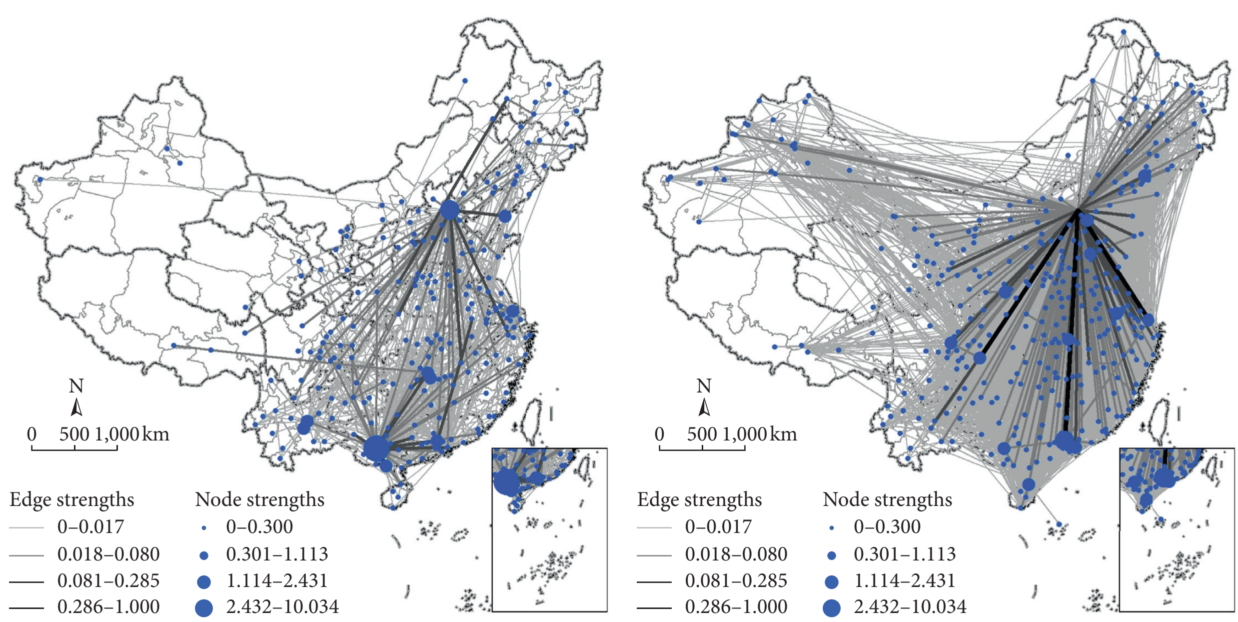
Since the beginning of the twenty-first century, the linkages in China's urban networks have become gradually closer. In 2001, although the strongest flow in the urban networks was Shenzhen-Haikou (0.788), the core status of Beijing's monocentric radiation remained unchanged, and the southeast and southwest edges of the diamond structure began to strengthen. Shanghai-Guangzhou (0.043) ranked 47th, Shenzhen-Shanghai (0.032) ranked 85th, Shenzhen-Chengdu (0.028) ranked 105th, and Guangzhou-Chengdu (0.016) ranked 250th in all city linkages in terms of the strength. Notably, the four city pairs did not enter the top 300 city linkages in 1992 in terms of connection strength.

After 2010, Beijing's position as the core node became consolidated. Beijing-Nantong (0.832) has become the strongest flow in China's urban networks, more than three times that of Beijing-Shanghai (0.239), in 2nd place. In 2019, Beijing-Nantong (0.792) and Beijing-Shanghai (0.297) still maintained the highest places in the urban networks in

Table 2: Characteristics statistics of urban networks in China from 1978 to 2019.

| Statistics | Indicators | 1978 | 1992 | 2001 | 2010 | 2019 |
|---|---|---|---|---|---|---|
| Network size | Number of nodes | 221 | 349 | 351 | 352 | 352 |
| | Number of edges | 527 | 5515 | 19264 | 33542 | 58700 |
| | Density | 0.011 | 0.045 | 0.155 | 0.271 | 0.475 |
| | Diameter | 8 (9) | 5 (9) | 4 (9) | 3 (9) | 3 (9) |
| Degree centrality | Average degree | 2.385 | 15.802 | 54.572 | 95.29 | 166.761 |
| | Average weighted degree | 0.075 | 0.102 | 0.108 | 0.143 | 0.244 |
| | Variable coefficient | 3.640 | 6.858 | 4.583 | 4.536 | 3.303 |
| | Gini coefficient | 0.802 | 0.727 | 0.717 | 0.680 | 0.660 |
| Neighborhood centrality | Average neighborhood centrality | 0.257 | 0.420 | 0.535 | 0.587 | 0.678 |
| | Variable coefficient | 0.734 | 0.250 | 0.152 | 0.162 | 0.193 |
| | Gini coefficient | 0.333 | 0.122 | 0.066 | 0.079 | 0.106 |
| Betweenness centrality | Average betweenness centrality | 248.439 | 465.900 | 310.164 | 260.264 | 184.247 |
| | Variable coefficient | 5.800 | 5.492 | 3.800 | 2.447 | 1.449 |
| | Gini coefficient | 0.958 | 0.901 | 0.855 | 0.760 | 0.616 |
| Small-world | Average clustering coefficient | 0.268 (0.027) | 0.476 (0.024) | 0.551 (0.025) | 0.563 (0.024) | 0.658 (0.026) |
| | Average path length | 3.645 (2.808) | 2.382 (2.723) | 1.891 (2.714) | 1.741 (2.709) | 1.525 (2.678) |
| Scale-free | Weighted degree distribution Power law fitting | $P(k) = 0.1636k^{-2.46}$ ($R^2 = 0.9996$) | $P(k) = 0.9422k^{-1.84}$ ($R^2 = 0.9742$) | $P(k) = 0.9365k^{-1.84}$ ($R^2 = 0.9789$) | $P(k) = 2.195k^{-1.34}$ ($R^2 = 0.8074$) | $P(k) = 3.051k^{-1.02}$ ($R^2 = 0.4983$) |

*Note.* The values in brackets are the statistics of random networks of the same scale.




(a) (b)

Figure 3: Continued.

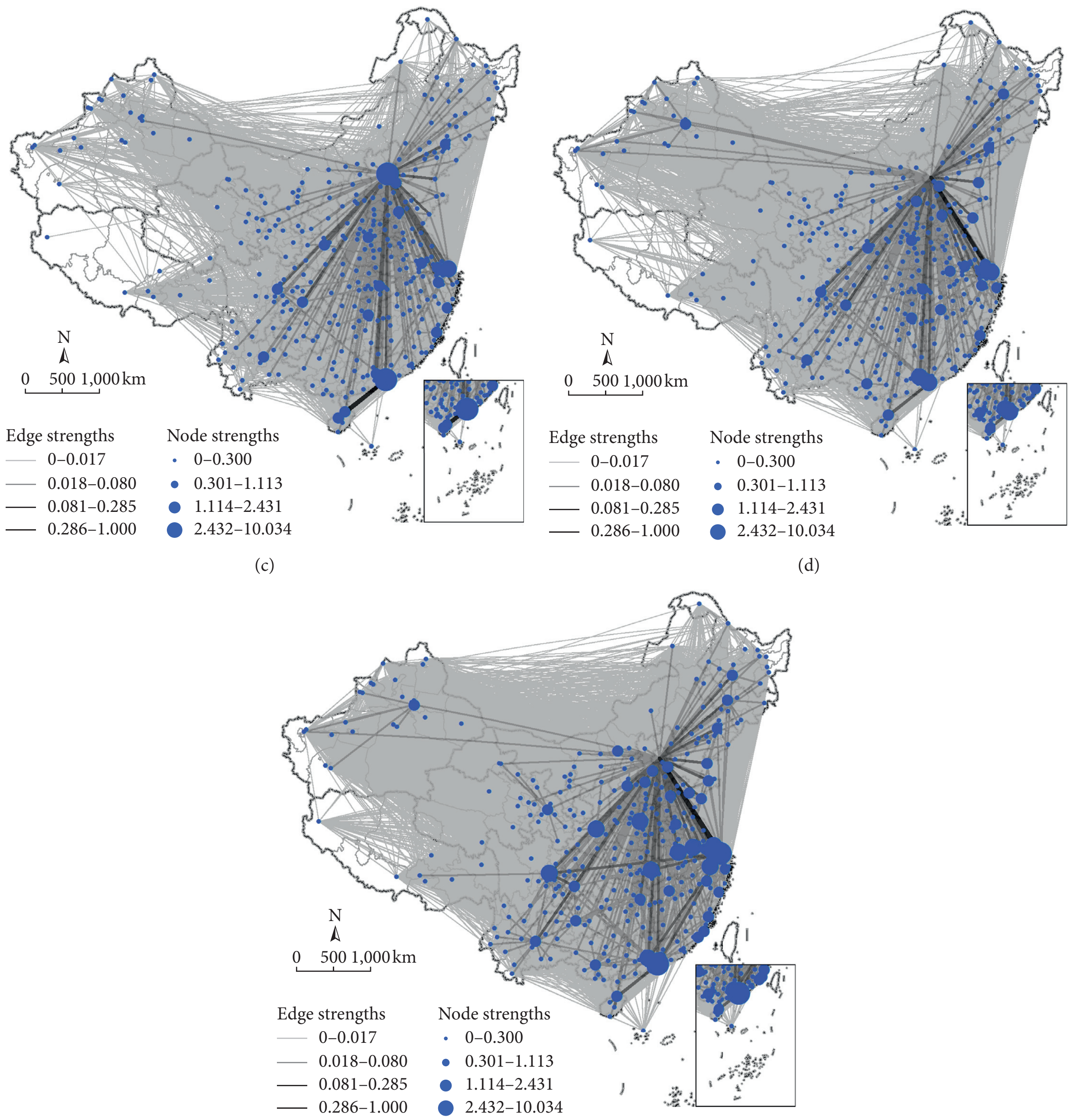



Figure 3: Spatial structure of urban networks in China from 1978 to 2019. (a) 1978; (b) 1992; (c) 2001; (d) 2010; (e) 2019.

terms of linkage strength. Hefei-Nanjing (0.391) also ranked among the highest-level flows. At the same time, Shanghai, as one of the vertices of the diamond, strengthened its ties with other important node cities in the network, with Shanghai-Chengdu (0.100) in 36th place, Shenzhen-Shanghai (0.092) in 42nd place, Shanghai-Wuhan (0.081) in 51st place, and Shanghai-Chongqing (0.0715) in 60th place. Shanghai's core status has gradually become prominent, and the diamond structure of urban networks in China has taken shape.

*3.3. Classification of Urban Network Structures in China.* To discuss the status quo of China's urban networks in detail, 58700 flows among 352 node cities in 2019 were sorted according to linkage strengths from high to low, and the top 0.1‰, 0.1‰–0.25‰, 0.25‰–0.5‰, 0.5‰–1‰, and 1 ‰ flows were extracted using the relevant threshold division method [27, 59] (Figure 4).

The core networks contained six pieces of the top 0.1‰ flows, involving core node cities, such as Beijing, Nantong, Hefei, Nanjing, Shanghai, Tianjin, and Chongqing. Beijing-

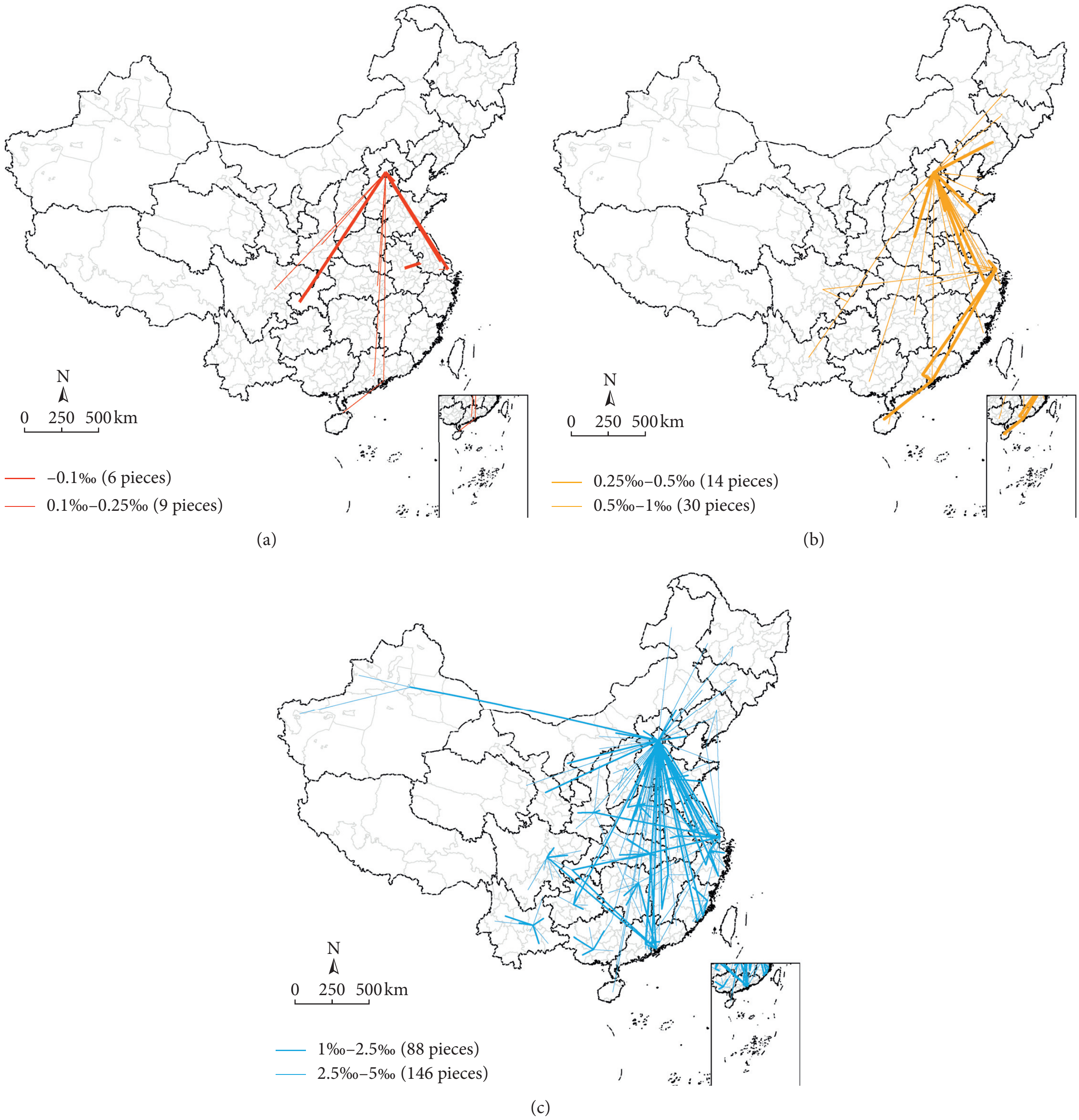



Figure 4: Three types of urban network structure in 2019. (a) Core networks in 2019. (b) Skeleton networks in 2019. (c) Regional networks in 2019.

Nantong (0.792) ranked first in the whole city network, more than twice the linkage strength of Hefei-Nanjing (0.391), followed by Beijing-Shanghai (0.297), Beijing-Tianjin (0.257), Beijing-Chongqing (0.238), and Shanghai-Beijing (0.228). These six pieces of flow were mainly centered in Beijing, radiating to the Yangtze River Delta and Chengdu-Chongqing region. The weighted outdegree of Beijing (12.710) was nearly four times that of Shanghai (3.296). Beijing has become the most important node city in the networks with an absolute high primacy. There were nine pieces of the top 0.1‰–0.25‰ flows, which started from Beijing, Guangzhou, and Shenzhen and pointed to nine newly added important node cities: Chengdu, Wuhan, Guangzhou, Foshan, Shenzhen, Haikou, Suzhou, Xi'an, and Shijiazhuang. The nine cities were provincial capitals and important cities, located in the central and western regions of China and in the southeast coastal areas, which constituted the diamond structure of the urban networks.

The skeleton networks contained 14 pieces of the top 0.25%–0.5% flows. Besides 16 important node cities in the core networks, Danzhou, Hangzhou, Jinan, Qingdao, Zhengzhou, Baoding, and Shenyang were added to the important node cities. Cities in the Bohai Rim-Yangtze River Delta and in the Yangtze River Delta-Pearl River Delta were closely connected, forming the right half of the diamond structure. There were 30 pieces of the top 0.5‰–1‰ flows. Dongguan, Changsha, Dalian, Harbin, and other cities entered the skeleton networks, and the subnodal linkages in the diamond structure were enhanced.

The regional networks included the top 1%–2.5% and 2.5%–5% flows, which increased by 88 and 146 pieces of flows, respectively, compared with the core and skeleton networks. Except for Tibet, all provinces in mainland China participated in the regional networks, which covered a wide area in China. Compared with core networks and skeleton networks, the diamond structure of regional networks expanded to the peripheral regions in the northeast, northwest, and southwest of China. Connections between node cities and the capital cities in the peripheral regions have strengthened. After sorting the weighted indegree of nodes from high to low, Kunming (0.331), Shenyang (0.322), and Urumqi (0.201) ranked 25th, 26th, and 46th among 352 node cities in China, respectively, becoming subnodes in the urban networks. Subnodes focused on developing connections with neighboring cities. For example, the top 15 flows starting from Kunming were all connected with cities within Yunnan Province, among which seven pieces entered the top 1%–5% flows in the regional networks.

## 4. Complexity Analysis of Urban Networks in China

*4.1. Small-World and Scale-Free Urban Networks in China.* The average path lengths of China's urban networks ranged from 1.525 to 3.645 from 1978 to 2019 and decreased year by year. Except for 1978, the values of average path lengths were slightly lower than the theoretical values of the average path lengths for random networks of the same scale. The clustering coefficients ranged from 0.268 to 0.658 and increased year by year, which were much higher than the theoretical values (Table 1). The average path lengths of China's urban networks were relatively short, while the clustering coefficients were relatively high, indicating the typical small-world characteristics of the network.

The distribution of the weighted degree of the urban networks followed a typical "long-tail distribution" (Figure 5), which was fitted by a power law—the goodness of fit remained above 0.97, illustrating that a large number of nodes had a small degree; only a few nodes were large nodes. The network showed a polarization trend and had significant scale-free network characteristics. In 2010, the goodness of fit decreased to 0.807, while in 2019, it was only 0.498, indicating that the scale-free characteristics of China's urban networks gradually weakened. The gap of intercity enterprise linkages narrowed. The excessive concentration of headquarter-branches and investment power in some cities was alleviated, and the structure of urban networks became optimized.

*4.2. The Correlation of Urban Networks in China.* The degree distribution reflects the probability of different values of node degree in a network. However, even networks with the same degree distribution may show different properties or behaviors because internode correlation is important and requires consideration. Thus, we further explored the structural characteristics of China's urban networks using correlation analysis. The results showed that 1978–2010 had similar correlation characteristics to those of 2019. Due to space limitations, only a correlation analysis of 2019 is presented here. The results for other years will not be described in this article.

Firstly, we investigated the degree-degree correlation of the network, which describes the joint probability of node degree at both ends of an edge randomly selected in a network. If the node degree at both ends of an edge is completely random, whether or not there is a connected edge has nothing to do with the degree values of the two associated nodes, meaning the network does not have degree correlation. Otherwise, the network has degree correlation. If the overall network is positively correlated (a "homology network"), this means that large nodes in the network are more inclined to connect with similar nodes. In contrast, the network is negatively correlated (a "heterogeneous network") [76].

Using equation (5), the average neighborhood centrality of the network was calculated. Correlation analysis between the average neighborhood centrality and node degree showed a significant negative correlation, with a correlation coefficient of -0.921 (Figure 6(a)). This indicates that the larger the degree of the node, the smaller the average degree of its connected nodes, meaning the whole network connection is heterogeneous. In China's urban networks, Beijing, Shanghai, Shenzhen, and Guangzhou, which were of high degree, had enterprise linkages with most node cities. The degree distribution showed that the degree values of most nodes in the network were quite small, which made the cities connected with those important ones such as Beijing and Shanghai, not only large in quantity but small in average degree value. The average neighborhood degrees of Beijing (212.766), Shanghai (212.766), Shenzhen (213.271), and Guangzhou (215.713) were all lower than the average value of the network (243.850). On the other hand, although Sansha City in Hainan, Hainan Tibetan Autonomous Prefecture in Qinghai, and Tumushuke City in Xinjiang are located in remote areas, cities which had their corresponding headquarters and investors were generally the above-mentioned cities with higher degrees, such as Beijing, Shanghai, Guangzhou, and Shenzhen, making these remote cities higher in average neighborhood degree, despite these cities' own degrees not being high. In addition, the correlation coefficient between degree and weighted degree was 0.423, showing a significant positive but nonlinear correlation (Figure 6(b)). The nodes with small degrees grew slowly in weighted degree. With increasing node degree, the weighted degree increased more significantly. When both the degree and weighted degree of the nodes were converted into the double logarithm, a positive linear correlation was obvious, with a correlation coefficient of 0.915. This indicates that the relationship between degree and weighted degree generally obeyed a power law.

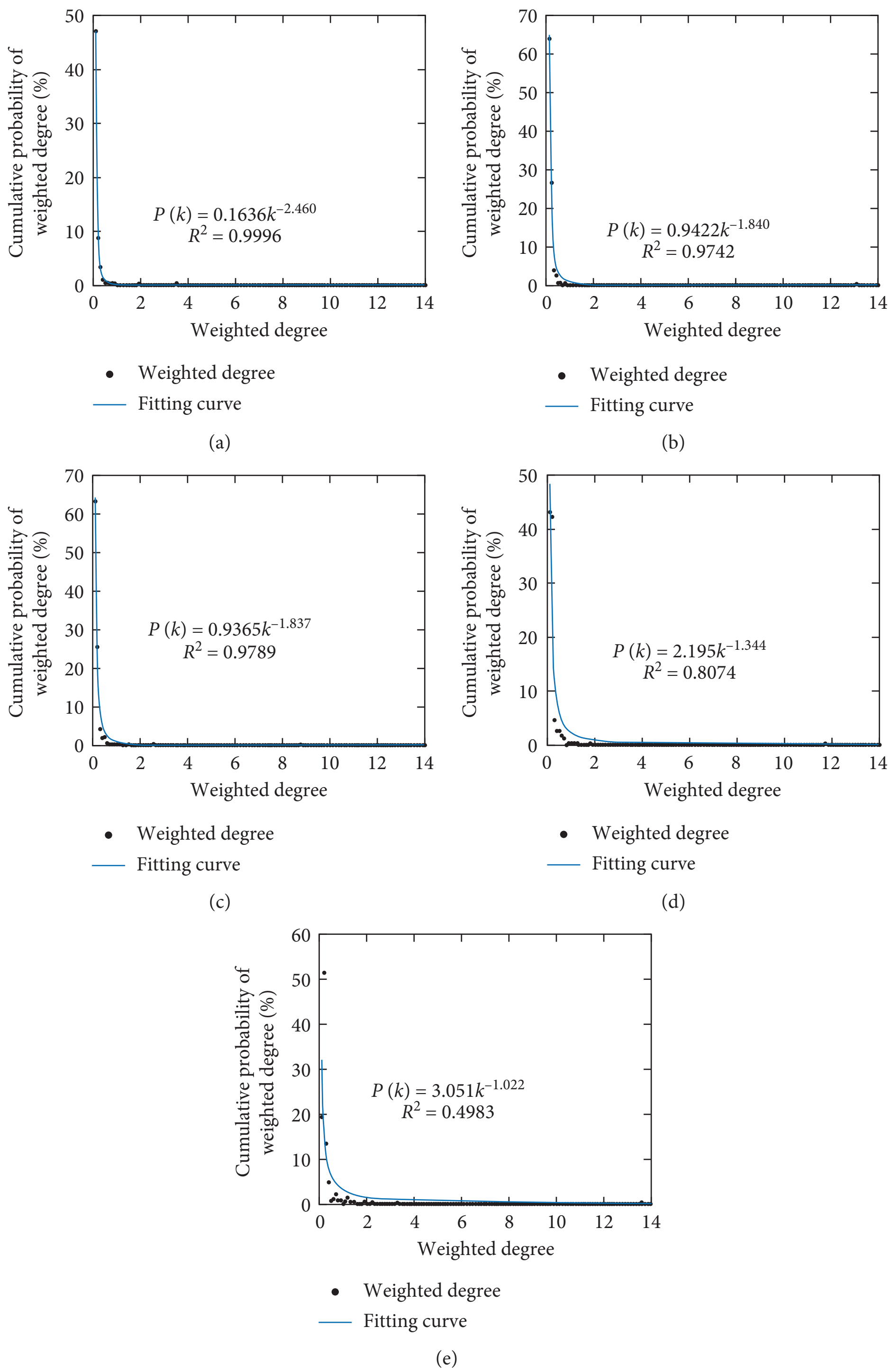



Figure 5: Cumulative probability of weighted degree of urban networks in China from 1978 to 2019. (a) 1978; (b) 1992; (c) 2001; (d) 2010; (e) 2019.

The relationship between degree and clustering coefficient is called the "cluster-degree correlation," which is also an important aspect of network structure. If there is an approximate relationship $C(k) \sim k^{-1}$between the clustering coefficient and node degree, then the network is hierarchical [27]. In China's urban networks, the correlation coefficient between node degree and clustering coefficient was -0.946, showing a significant negative correlation (Figure 6(c)). We found that when the degree value was less than 300, the negative cluster-degree correlation was not significant; otherwise, the negative correlation was significant. Because low-value nodes tended to be connected with high-degree nodes in heterogeneous networks, the connections among the latter were high; thus, low-value nodes presented high-

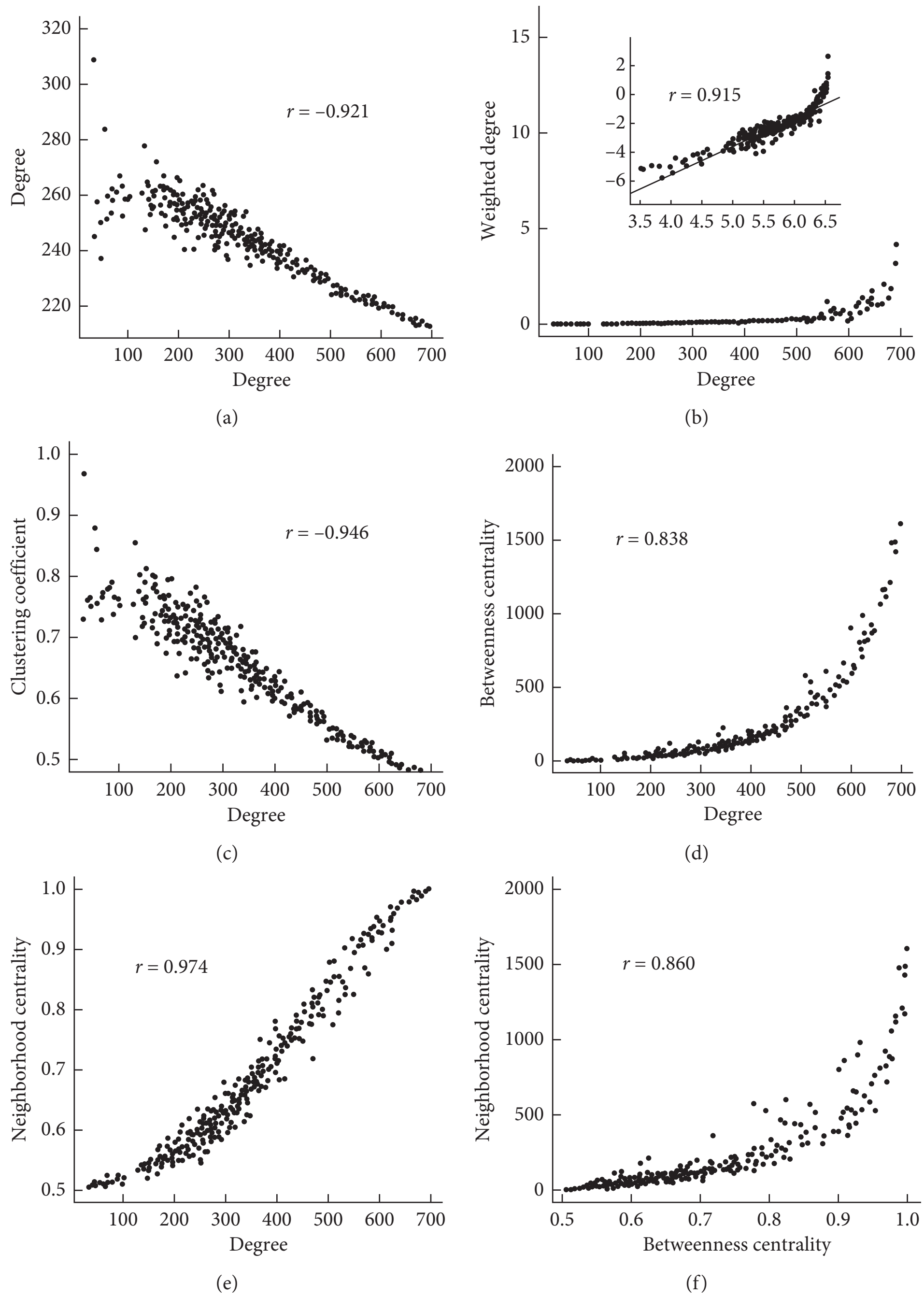



Figure 6: Correlation analysis of urban networks in China in 2019.

level aggregation. In contrast, as it was connected to a large number of low-degree nodes, the agglomeration of high-degree nodes was weakened [27]. The fitting results show that the cluster-degree correlation of China's urban networks was $C(k) \sim k^{-0.1835}$, the power value was less than 1, and the hierarchy of the whole network was not obvious, confirmed by the "long-tail" feature of the degree distribution already noted. On the whole, China's urban networks had a small number of core nodes but large degree values. Most nodes were of a small degree.

Degree had a significant positive correlation with betweenness centrality and neighborhood centrality, with correlation coefficients of 0.838 and 0.974, respectively (Figures 6(d) and 6(e)). Betweenness centrality and neighborhood centrality also had a significant positive correlation, with a correlation coefficient of 0.860 (Figure 6(f)). This shows that with increasing node degree, betweenness centrality and neighborhood centrality increased. Moreover, a higher neighborhood centrality gave rise to higher betweenness centrality. Cities with more headquarter-branches and more investment activity had a larger centrality, equating to a higher status and stronger influence over the networks. These core cities are closer to other nodes, ensuring their functions of "broker" and "transfer"can be better fulfilled.

## 5. Conclusion and Discussion

From the perspective of "space of flows," flow-based data were used to measure the enterprise linkages of 353 cities in China from 1978 to 2019. We constructed directed and weighted urban networks for China and investigated their spatiotemporal evolution and complexity characteristics. The main conclusions are as follows:

(1) Intercity enterprise linkages in China have been continuously strengthened from 1978 to 2019. The scale and density of urban networks have increased rapidly. At the same time, the distribution of importance and influence of node cities has been significantly unbalanced but has become more balanced over time.

(2) The diamond structure was the core of China's urban networks from 1978 to 2019. Urban networks were sparse in 1978, and enterprise flows were restricted by distance attenuation. Network density then significantly increased, forming a monocentric (Beijing) radial pattern. Since the beginning of the twenty-first century, the status of core nodes, e.g., Shanghai, has gradually become prominent. In 2019, four vertices stood out, composing a stable diamond structure. The spatial flows of enterprises constituted core networks with Beijing as the center, skeleton networks with trunk lines of subnodes, and the regional networks covering a wide range of peripheral areas.

(3) Complexity analysis shows that China's urban networks were typically small-scale and scale-free. However, the scale-free characteristics of China's urban networks became weakened after 2010. The overall scale gap of intercity enterprise linkages gradually narrowed, and the structure of urban networks became optimized. Meanwhile, China's urban networks featured as heterogeneous, without an obvious hierarchy. There were more cities with headquarter-branches and active investment behaviors. This imposed a strong influence on and control over networks, ensuring their functions of "broker" and "transfer" were fulfilled.

Using real flow-based data of enterprises to build a directed and weighted network can better reflect objective and realistic urban networks. By investigating a long time series, this article described and analyzed the topology, spatial structure evolution, and the complexity of China's urban networks. However, this work possesses several limitations: (1) The linkage strengths were calculated using an entropy method. Since this is an objective weighting method based on the dispersion of the data itself, an influence of data structure differences on the comparison results of each year cannot be ruled out. (2) Although the data obtained cover a wide range of years and cities, they only include the amount of investment and the number of branches of the whole industry and do not involve subindustry and specific category information of investment and branches. Thus, the analysis of different activity types and industries is restricted to a certain extent. (3) In this article, we focused on the topology, spatial structure, and complexity characteristics of China's urban networks in a long time series. The formation and evolution mechanism behind them is also worth discussion in the future.

## Data Availability

Data of registered enterprises in China that we used to support the findings were supplied by State Administration for Industry and Commerce of People's Republic of China under license, which cannot be made freely available.

## Conflicts of Interest

The authors declare that they have no conflicts of interest.

## Acknowledgments

This study was supported by the National Natural Science Foundation of China (Grant no. 41971162).

## References

[1] P. M. Hohenberg and L. H. Lees, *The Making of Urban Europe, 1000–1994*, Harvard University Press, Cambridge, MA, USA, 1985.
[2] P. Dicken, *Global Shift: Mapping the Changing Contours of the World Economy*, Sage, London, UK, 2007.
[3] J. Lu, D. Sun, J. Yu, J. Li, and F. Niu, ""Local versus nonlocal" enterprise linkages of global cities: a comparison between beijing and Shanghai, China," *Complexity*, vol. 2020, Article ID 8918534, 13 pages, 2020.
[4] S. Sassen, *The Global City*, Princeton University Press, New York, NY, USA, 1991.
[5] M. Castells, *The Rise of the Network Society*, Blackwell, Oxford, UK, 1996.
[6] B. Derudder and P. J. Taylor, "Central flow theory: comparative connectivities in the world-city network," *Regional Studies*, vol. 52, no. 8, pp. 1029–1040, 2018.
[7] J. V. Beaverstock, R. G. Smith, P. J. Taylor, and J. Taylor, "World-city network: a new metageography?" *Annals Of the Association Of American Geographers*, vol. 90, no. 1, pp. 123–134, 2000.
[8] F. Pan, "Mapping urban networks through inter-firm service relationships: the case of China," *Urban Studies*, vol. 54, no. 16, pp. 3639–3654, 2017.

[9] R. Capello, "The city network paradigm: theory and empirical evidence," *Contributions to Economic Analysis*, vol. 266, pp. 495–529, 2004.
[10] P. Hall, *The World Cities*, Heinemann, London, UK, 1966.
[11] J. Friedmann, "The world city hypothesis," *Development and Change*, vol. 17, no. 1, pp. 69–83, 1986.
[12] M. Castells, *The Information City*, Basil Blackwell, Oxford, UK, 1989.
[13] M. Castells, *The Space of Flows*, vol. 1, pp. 376–482, Blackwell, Oxford, UK, 1996.
[14] P. J. Taylor, "Specification of the world city network," *Geographical Analysis*, vol. 33, no. 2, pp. 181–194, 2001.
[15] Z. L. Tang, T. Li, and C. Li, "Research on the interlocking network of major cities in China," *City Planning Review*, vol. 41, no. 1, pp. 28–39+82, 2017.
[16] R. Capello, "The city network paradigm: measuring urban network externalities," *Urban Studies*, vol. 37, no. 11, pp. 1925–1945, 2000.
[17] B. Derudder, "On conceptual confusion in empirical analyses of a transnational urban network," *Urban Studies*, vol. 43, no. 11, pp. 2027–2046, 2006.
[18] P. J. Taylor, M. Hoyler, and R. Verbruggen, "External urban relational process: introducing central flow theory to complement central place theory," *Urban Studies*, vol. 47, no. 13, pp. 2803–2818, 2010.
[19] G. L. Xi, F. Zhen, J. L. He, and Y. H. Gong, "City networks of online commodity services in China: empirical analysis of tmall clothing and electronic retailers," *Chinese Geographical Science*, vol. 28, no. 2, pp. 231–246, 2018.
[20] J. Huang, D. Levinson, and J. Wang, "Tracking job and housing dynamics with smartcard data," *Proceedings of the National Academy of Sciences*, vol. 115, no. 50, 2018.
[21] M. X. Zhao, B. Derudder, and J. Huang, "Examining the transition processes in the Pearl River Delta polycentric mega-city region through the lens of corporate networks," *Cities*, vol. 60, pp. 147–155, 2017.
[22] R. Wall, "Corporate profiles and networks of gulf cities," *Al Manakh*, vol. 2, 2010.
[23] P. J. Taylor and D. R. F. Walker, "World cities: a first multivariate analysis of their service complexes," *Urban Studies*, vol. 38, no. 1, pp. 23–47, 2001.
[24] P. J. Taylor, *World-city Network: A Global Urban Analysis*, Routledge, London, UK, 2004.
[25] D. Bassens, B. Derudder, and F. Witlox, "Setting Shari'a standards: on the role, power and spatialities of interlocking Shari'a boards in Islamic financial services," *Geoforum*, vol. 42, no. 1, pp. 94–103, 2011.
[26] M. Hoyler, "External relations of German cities through intra-firm networks-A global perspective," *Raumforschung Und Raumordnung-Spatial Research And Planning*, vol. 69, no. 3, pp. 147–159, 2011.
[27] K. Wu, C. L. Fang, and M. X. Zhao, "The spatial organization and structure complexity of Chinese intercity networks," *Geographical Research*, vol. 34, no. 4, pp. 711–728, 2015.
[28] Z. Cao and B. Z. Peng, "Comparing the physical, functional and knowledge integration of the Yangtze River Delta city-region through the lens of inter-city networks," *Cities*, vol. 82, pp. 119–126, 2018.
[29] M. X. Zhao, B. Derudder, and J. H. Huang, "Polycentric development in China's mega-city regions, 2001-08: a comparison of the Yangtze and Pearl River Deltas," *DIE ERDE*, vol. 148, no. 1, pp. 1–13, 2017.
[30] X. Wang, Z. D. Luo, and C. S. Zhu, "Dialectics of the center and the hinterland: a comparative study on regional spatial relations between Suzhou and ningbo based on corporation relation," *Urban Planning Forum*, vol. 5, pp. 79–85, 2014.
[31] B. Derudder, M. Hoyler, P. J. Taylor et al., *International Handbook of Globalization and World Cities*, Edward Elgar Publishing, Cheltenham, UK, 2012.
[32] P. Hall and K. Hall, *The Polycentric Metropolis - Learning from Mega-City Regions in Europe*, Earthscan, London, UK, 2006.
[33] B. Derudder, Z. Cao, X. J. Liu et al., "Changing connectivities of Chinese cities in the world city network," *Chinese Geographical Science*, vol. 28, no. 2, pp. 183–201, 2018.
[34] Z. P. Neal, B. Derudder, and P. J. Taylor, "Should I stay or should I go: predicting advanced producer services firm expansion and contraction," *International Regional Science Review*, vol. 42, no. 2, pp. 207–229, 2019.
[35] R. S. Wall and G. A. van der Knaap, "Sectoral differentiation and network structure within contemporary worldwide corporate networks," *Economic Geography*, vol. 87, no. 3, pp. 267–308, 2011.
[36] S. Kratke, "Network analysis of production clusters: the Potsdam/Babelsberg film industry as an example," *European Planning Studies*, vol. 10, no. 1, pp. 27–54, 2002.
[37] A. S. Alderson and J. Beckfield, "Power and position in the world city system," *American Journal of Sociology*, vol. 109, no. 4, pp. 811–851, 2004.
[38] W. K. Carroll, "Global cities in the global corporate network," *Environment And Planning A-Economy And Space*, vol. 39, no. 10, pp. 2297–2323, 2007.
[39] F. H. Pan, W. K. Bi, X. J. Liu, and T. Sigler, "Exploring financial centre networks through inter-urban collaboration in high-end financial transactions in China," *Regional Studies*, vol. 54, no. 2, pp. 162–172, 2020.
[40] P. J. Taylor, "Cities in globalization: practices, policies and theories," *Applied Cognitive Psychology*, vol. 2, no. 2, pp. 1–334, 2007.
[41] J. T. Song and C. L. Gu, "The theory and method of urban system planning," *Journal of Geography*, vol. 2, 1988.
[42] Q. Zhang and J. Liu, "The relationship between urban system planning reform and innovation and "three regulations in one"-from "three structures and one network"," *Urban Planning*, vol. 38, no. 10, pp. 13–27, 2014.
[43] D. D. Lu, "Macro strategy of regional development in China," *Journal of Geography*, vol. 2, 1987.
[44] D. D. Lu, *Location Theory and Regional Research Methods*, Science Press, "Beijing, China, 1988.
[45] D. D. Lu, ""On the best structure and development of the region - review and reanalysis since the "point-axis system" and "*T*" structure were put forward," *Journal of Geography*, vol. 2, 2001.
[46] L. Zhou, X. W. Dang, H. W. Mu, B. Wang, and S. H. Wang, "Cities are going uphill: slope gradient analysis of urban expansion and its driving factors in China," *Science of the Total Environment*, vol. 775, pp. 1–10, 2020.
[47] L. Zhou, X. W. Dang, Q. K. Sun, and S. H. Wang, "Multi-scenario simulation of urban land change in Shanghai by random forest and CA-Markov model," *Sustainable Cities and Society*, vol. 55, pp. 1–10, 2020.
[48] C. Bao and X. J. Chen, "Review and prospect of research on spatial pattern of urban system in China," *Advances in Geographical Sciences*, vol. 33, no. 10, pp. 1300–1311, 2014.
[49] Y. Sun, L. C. Zhang, and S. M. Yao, "Spatial structure of urban network in China based on the headquarters and branches of express enterprises," *Journal of University of Chinese Academy of Sciences*, vol. 34, no. 5, pp. 91–597, 2017.

[50] Y. R. Wang and R. X. Gu, "The spatial structure and evolution of Yangtze River Delta urban network: analysis based on enterprise connection," *Urban Studies*, vol. 26, no. 6, pp. 21–29+78, 2019.
[51] X. Jin, G. J. Hu, H. L. Ding, S. L. Ye, Y. Q. Lu, and J. H. Lin, "Evolution of spatial structure patterns of city networks in the Yangtze River Economic Belt from the perspective of corporate pledge linkage," *Growth And Change*, vol. 51, no. 2, pp. 833–851, 2020.
[52] Y. Li and N. A. Phelps, "Megalopolitan glocalization: the evolving relational economic geography of intercity knowledge linkages within and beyond China's Yangtze River Delta region, 2004-2014," *Urban Geography*, vol. 40, no. 9, pp. 1310–1334, 2019.
[53] H. M. Zong and R. H. Lyu, "The spatial characteristics and evolution of Chinese urban network based on logistics enterprise data in 2007-2017," *Scientia GeographicaSinica*, vol. 40, no. 5, pp. 760–767, 2020.
[54] B. R. Leng, Y. C. Yang, H. Xing et al., "The complexity analysis of urban network structure in China since 1990s based on gravity method," *Regional Research and Development*, vol. 5, pp. 68–73, 2011.
[55] Y. Dong, J. Luo, W. S. Zheng et al., "The urban network structure and complexity in hubei province," *Economic Geography*, vol. 7, pp. 76–84, 2019.
[56] G. Xian, G. Zeng, and X. Z. Cao, "Structural feature and proximity mechanism of Chinese intercity innovation network," *World Regional Studies*, vol. 27, no. 5, pp. 136–146, 2018.
[57] B. Derudder and P. Taylor, "The cliquishness of world cities," *Global Networks-A Journal Of Transnational Affairs*, vol. 1, pp. 71–91, 2005.
[58] P. J. Taylor, B. Derudder, M. Hoyler, and P. F. Ni, "New regional geographies of the world as practised by leading advanced producer service firms in 2010," *Transactions Of the Institute Of British Geographers*, vol. 38, no. 3, pp. 497–511, 2013.
[59] C. Rozenblat, "Opening the black box of agglomeration economies for measuring cities' competitiveness through international firm networks," *Urban Studies*, vol. 47, no. 13, pp. 2841–2865, 2010.
[60] L. Resmini, "Cities in globalization: practices, policies and theories," *Journal of Regional Science*, vol. 49, no. 1, pp. 228–230, 2009.
[61] C. Q. Nie, *Study on Urban Network Spatial Structure Evolution of the Fan-Yangtze River Delta Based on the Corporate Organization Network*, East China Normal University, Shanghai, China, 2018.
[62] L. Yi, *"E-commerce and Enterprises' Trans-provincial Investment–Imperial Research of Chinese Listened Companies,"* Nanjing University, Nanjing, China, 2017.
[63] Z. R. Li, F. Zhen, and H. H. Fu, "Mapping urban network through inter-firm investment relationship: a case study of Yangtze River Delta," *Scientia GeographicaSinica*, vol. 39, no. 11, pp. 1763–1770, 2019.
[64] Y. D. Wei, "Decentralization, marketization, and globalization: the triple processes underlying regional development in China," *Asian Geographer*, vol. 20, pp. 7–23, 2001.
[65] K. Huang, G. Tian, and Y. Wang, "Tackle China's economic complexities by deepening reform and opening up: macroeconomic outlook, policy simulations, and reform implementation: a summary of the annual SUFE macroeconomic report (2018-2019)," *Frontiers of Economics in China*, vol. 14, 2019.
[66] K. H. Zhang, Q. L. Qian, and Q. S. Yang, "Analysis of multilevel influencing factors of urbanization of urban land in China," *Journal of Geography*, vol. 75, no. 1, pp. 179–193, 2020.
[67] J. Li and J. H. Long, "Research on the precise subsidy education model based on big data analysis," *Information and Computer (Theoretical Edition)*, vol. 5, pp. 166–168, 2019.
[68] Y. Wang, W. Zhouhao, and C. Li, "The complexity of large-scale urban networks: a comparative study in china," *Transportation Research Board Annual Meeting*, 2015.
[69] W. Liu, Q. Hou, Z. Xie et al., "Urban network and regions in China: an analysis of daily migration with complex networks model," *Sustainability*, vol. 12, 2020.
[70] S. Boccaletti, V. Latorabc, Y. Morenode, M. Chavezf, and D.-U. Hwanga, "Complex networks: structure and dynamics," *Physics Reports-Review Section of Physics Letters*, vol. 424, pp. 175–308, 2006.
[71] M. E. J. Newman, *Networks: An Introduction*, Oxford University Press, Oxford, UK, 2010.
[72] D. J. Watts and S. H. Strogatz, "Collective dynamics of 'small-world' networks," *Nature*, vol. 393, no. 6684, pp. 440–442, 1998.
[73] X. G. Ma and Y. Lu, "Spatial structure and connection of cities in China based on air passenger transport flow," *Economic Geography*, vol. 38, no. 8, pp. 47–57, 2018.
[74] X. Zhang, F. Z. Yu, and L. J. Xu, "Spatial structure of urban networks in China based on the perspective of cultural industry enterprise networks," *Progress in Geography*, vol. 39, no. 1, pp. 78–90, 2020.
[75] B. X. Li, "Some thoughts on the development strategy and policy of southwest china," *Review of Economic Research*, vol. z2, pp. 1264–1274, 1993.
[76] M. Newman and M. Girvan, "Mixing patterns and community structure in networks," *Statistical Mechanics Of Complex Networks*, vol. 625, 1970.

Hindawi
Complexity
Volume 2021, Article ID 5538872, 14 pages
https://doi.org/10.1155/2021/5538872

*Research Article*

# Hierarchical Characteristics and Proximity Mechanism of Intercity Innovation Networks: A Case of 290 Cities in China

**Xianzhong Cao**,[1] **Gang Zeng**,[1] **Lan Lin**,[2] **and Lin Zou**[1,3]

[1]*Center for Modern Chinese City Studies and Institute of Urban Development, East China Normal University, Shanghai 200062, China*
[2]*Institute of Urban and Population Development, Shanghai Academy of Social Sciences, Shanghai 200020, China*
[3]*Shanghai University of Engineering Science, Shanghai 201620, China*

Correspondence should be addressed to Lan Lin; linlan@sass.org.cn and Lin Zou; zoulin200618@163.com

Received 1 February 2021; Revised 8 April 2021; Accepted 18 May 2021; Published 25 May 2021

Academic Editor: Jing-Hu Pan



The formation mechanism of innovation networks is one of the core issues in the current research of innovation networks, and proximity plays an important role in the formation and development of innovation networks; however, which proximity is more important and how different proximities interact remain to be further researched. This study conducts a social network analysis and adopts a spatial interaction model to examine innovation networks among 290 Chinese cities. The results reveal that, first, the hierarchical characteristics of Chinese cities' innovation networks reflect a core periphery structure and the spatial patterns of large dispersion and small agglomeration. Further, bound by the Hu line, the hierarchy is high in the east and low in the west. Second, geographical, institutional, and cognitive proximities positively affect Chinese cities' innovation networking. Cognitive proximity, particularly, has the highest impact. Geographical proximity reinforces the effect of institutional proximity, and thus, their interactions are complementary.

## 1. Introduction

Knowledge and technological innovations have become the key driving forces of regional economic development. Innovation includes the production of new products, the adoption of new production methods, the development of new markets, and the expansion of new sources of supply or new forms of organization. Acquiring knowledge is a prerequisite for the implementation of innovation activities [1]. In fact, they are considered more important than traditional material capital and are essential strategic resources for sustained economic growth [2]. China's central government believes that innovations play a critical role in leading development and are strategic support in building modern economic systems. Cities accumulate talent, capital, information, enterprises, and other innovative elements and are an important platform for innovation-driven development [3]. The traditional model for closed innovation, along with regional innovation systems and learning areas, can no longer match the needs of rapid economic development or effectively cope with dynamic competitive environments. Thus, local governments are increasingly pursuing innovative cooperation and knowledge promotion through innovative partnerships with other cities and the optimised allocation of innovative resources to achieve complementary advantages and win-win situations. The expected outcomes are the improved efficiency of city innovations, reduced innovative costs and risks, and enhanced city competitiveness. In this new era, intercity networking and collaborative innovations have become a new and frequent trend associated with city development [4].

The mode of regional innovation has changed from single-actor independent innovation to multiactor collaborative innovation. Moreover, the paradigm of innovation research in economic geography and regional economics has shifted from traditional locations to modern flow space.

Thus, the perspective of networks based on relationships has become an important starting point to analyse regional innovation and city development [5, 6]. It has been a common innovation mode for firms to innovate through cooperative networks; the innovation mode has successively gone through technology promotion, demand pull, interaction, and comprehensive; now, it has entered the fifth generation innovation mode with innovation networks, which is also the leading direction of future innovation research [7]. The most fundamental reason for innovation networking lies in the limited innovation ability of a single actor and the scarcity of resources; individuals can acquire more external knowledge and resources in the innovation networks [8]. The literature of economic geography focuses on the structure and formation mechanisms of innovation networks, particularly those formed by microinnovation actors such as enterprises, universities, and research institutions. Some scholars examine the spatial and topological structure and dynamic evolution processes of innovation networks [9], while others discuss the impact of proximity on innovation cooperation, knowledge flow, dynamic evolution, internal impact mechanisms, and innovation performance [10]. The formation of innovation networks is affected by many factors; first, the endogenous effect of network structure, including the embeddedness, externality, absorptive capacity, small world, and technical goalkeeper. Second, the characteristics of network organizational elements, including the nature, scale, and status of organizational elements; third, the perspective of multidimensional proximity, which mainly includes geographical proximity, social proximity, cognitive proximity, institutional proximity, and cultural proximity [11]. The role of proximity in innovation and interorganizational networks has received increasing attention over the past decade [12]. Bergé [13] investigates that how network proximity influences the structure of interregional collaborations and how it interacts with geography. However, few studies examine city structures from an innovation perspective, whereas a majority of the studies are limited to innovation network structures that fail to detail why such network structures are formed, which is a research-worthy topic. Which proximity has greater influence on the formation of innovation networks and how different proximities interact are still unsolved questions, which is the purpose of this paper.

Adopting an innovation networks' perspective, this study fills the gap in the literature by analysing the hierarchical structural characteristics of 290 Chinese cities, including the locational and spatial attributes of their innovation networks, and incorporating multidimensional proximity in a mechanism analysis of proximity. In doing so, it aims to understand the current collaborative innovation situation in Chinese cities and clarify the relationship between the formation of intercity innovation networks and proximity. The findings will serve as a reference to formulate more accurate regional innovation policies and a basis to enrich relevant innovation studies in economic geography.

The remainder of this paper is organised as follows. Section 2 presents the origin of innovation networks and the proximity framework. Section 3 describes the data and empirical variables used to analyse the structural and proximity mechanisms of innovation networks. Section 4 conducts a social network analysis (SNA) on the network structure of 290 Chinese cities. Section 5 discusses the influence mechanisms of geographical, institutional, and cognitive proximities on the cities' innovation networks. Section 6 concludes the paper.

## 2. Theoretical Analytical Framework

Open collaborative innovations have become a new mode of city innovation and have further given rise to intercity innovation networks. The structure of intercity innovation networks not only reflects the contact way of each city in a network and but also determines the mutual position and relationship of each city. In other words, it directly affects knowledge exchanges and interactions among cities in a network [14], which are associated with the depth of integration and utilisation of innovation resources as key factors in determining the performance of city innovation [15]. The concept of innovation networks was originally proposed in the sociology literature, and later, economic geographers applied it to innovation research. Freeman [16] defines innovation networks as a new institutional arrangement that breaks away from the previous innovation model to realise systematic innovation that can also be recognised in academic circles. In essence, an innovation network is a closely related system formed by various formal and informal linkages among innovative actors including firms, universities and research institutions, governments, capital markets, and intermediaries. Boix and Trullén [17] measured the factors that affect the evolution of different intensities of knowledge in a region's cities. Researchers have defined innovation networks at the city level [18]. They believe it is a strategic collaborative process for a city to realise the joint complementation and optimal allocation of knowledge spillovers and innovative elements (i.e., talents, funds, and information), both of which can be responses to rapidly changing innovation demands in a knowledge economy era. Intercity innovation networks can also be viewed as an interactive form of city space.

The concept of proximity originated in Marshall's pioneering research on industrial cluster economies. It was originally defined as the spatial colocation of economic activity actors within the same cluster [19], thus emphasising geographical proximity in a given period. Economic geographers are currently extending the proximity perspective to analyse the impact factors, evolution dynamics, and mechanisms of innovation networks. Many studies focus on the relationship between geographical proximity and knowledge innovation and show that the former not only promotes firms' agglomeration but also affects the structure of innovation networks [20]. Notably, single-dimensional geographical proximity does not sufficiently explain interactive learning and cooperative innovation among innovation actors. To this effect, the French School of Proximity Dynamics considers a multidimensional proximity framework comprising geographical, organizational, institutional, cognitive, and social proximities [21, 22], thus shifting the proximity discussion from single-dimensional proximity to multidimensional proximity. The Netherlands School of Utrecht has conducted an in-depth analysis of interactions among geographical, organizational, institutional, cognitive, and social proximities and their impact on firms' innovation cooperation [23], which is popular in multidimensional proximity analysis [24]. Capone and Lazzeretti [12] investigated the role of various forms of proximity in multiple informal interorganizational relationships. The literature has both qualitative and empirical studies that emphasise the importance of geographical, cognitive, and institutional proximities in knowledge flow and firms' innovation networking [25].

It is clear from the discussion above that most of the literature examines the independent effect of proximity on the various dimensions of innovation networks. However, the proximity of different dimensions is not divided and independent; rather, it is interactive [26]. Recently, empirical studies have examined the impact of multidimensional proximity interactions [13] although they remain in their infancy, thus warranting the strengthening of related empirical tests and theoretical discussions. The proximity framework generally includes three dimensions, geographical dimension (characterized by the difference of physical distance between members of innovation networks), cognitive dimension (characterized by the degree of knowledge similarity), and institutional dimension (degree of common ownership, strength of social ties, degree of sharing standards, habits, regulations, and laws), and is widely used in related research such as regional coordinated development and innovation networks. This study also employs the three proximities to explain structural differences in intercity innovation networks.

## 3. Methodology and Data

*3.1. Data Sources.* Given the difficulties associated with collecting large amounts of network data, most scholars apply a modified spatial interaction model (i.e., gravity or gravity model) with a city output or comprehensive index system to measure innovation linkages between cities. Joint patent applications represent knowledge flow among various innovation actors in different cities and more effectively reflect the innovative relationship between cities. Thus, an increasing number of scholars are using data on joint patent applications to measure innovation networks. Among various patents, such as appearance design and utility model patents, invention patents are more representative of original technology and technological innovation performance [27]. Therefore, this study uses data on joint invention patent applications to examine intercity innovation networks in China.

Patent data are obtained from the patent retrieval and analysis system of the National Intellectual Property Administration (http://www.pss-system.gov.cn/sipopublicsearch/patentsearch/table Search-showTableSearchIndex.shtml). In China, it takes 18 months for patents to be publicly released. This study uses data on invention patent applications by two or more actors for 290 cities in 2014. The data of patent application in 2014 was selected because it was published in 2016, but the new patent law was implemented in 2017, and there were new restrictions and regulations on patent application. The data before 2017 was stable, and the data after that was unstable, so the data of patent application in 2014 was selected.

First, the researchers logged into the patent retrieval and analysis system and searched for Chinese invention patents for the scope, "201401 20141231" for application date, and "?; ?; " for applicant names of the 290 cities for applicant address (e.g., Beijing, Shanghai, Guangzhou, and Shenzhen). Second, data for the following criteria were deleted to ensure reliable network data: (i) less than two joint applications for institution patents by applicants including individuals or joint applications between individuals and one institution, (ii) individual applications in which it is difficult to determine applicants' city, and (iii) foreign applicants since this study focuses on innovation linkage among 290 cities in Mainland China. Finally, the analysis conducts a two-to-two estimation for invention patent applications by three or more institutions. In addition, information on applicants' city is extracted to establish an omnidirectional intercity innovation network. Data on 63,108 joint invention patents have been collected since April 2017. Following a screening and processing, data for 42,921 patents, including 286 city nodes, were retained.

## 4. Methodology

*4.1. Social Network Analysis.* The SNA is commonly used to describe innovation network structures. Ter Wal and Boschma [28] conducted an SNA on innovation networks and portrayed and visualized network structures and evolution. SNA has been deemed "the most promising empirical analysis tool."

This study uses centrality and network density to measure the local and overall structure of intercity innovation networks. In addition, it uses the UCINET software to analyse the local and overall structures.

(i) *Network Density.* Network density reflects the degree of node connections in networks: the higher the network density, the closer the node connections. The formula is as follows:

$$D = \sum_{i=1}^{k} \sum_{j=1}^{k} \frac{\mathrm{d}(ni, nj)}{k(k-1)}, \quad (1)$$

where $D$ is network density, $k$ is a node, and $d(n_i, n_j)$ is the connection between nodes $i$ and $j$.

*Network Centrality.* Network centrality measures mainly the degree of node centrality in the network and can be formulated as follows:

$$CD(n_i) = \sum_{j=1}^{n} Xji, \quad (2)$$

where $CD(n_i)$ is network centrality and $X_{ji}$ is the contact strength between nodes $i$ and $j$.

*Network Centrality Potential.* Network centrality potential measures the network's degree of centralisation and reflects the degree of deviation in a network by examining the overall structural characteristics of the network. The formula is as follows:

$$C = \frac{\sum_{i=1}^{n} (C\max - Ci)}{\max\left[\sum_{i=1}^{n} (C\max - Ci)\right]}, \tag{3}$$

where $C_{\max}$ is the maximum degree of centrality for the nodes in the network and $C_i$ is the degree of centrality for node $i$.

*4.2. Spatial Interaction Model.* Early proximity studies have ignored the effect of spatial interaction in a multiregional context. Wanzenböck [29] proposed a new measure for assessing the network proximity between aggregated units, based on disaggregated information on the network distance of actors. Scherngell and Barber [30] categorised factors influencing regional interactions into scale factors and distance variables that symbolise tension and resistance using a gravity model. Scherngell and Hu [31] proposed a spatial interaction model that addresses limitations in previous research. Since then, their model has been widely applied in innovation research by economic geographers. Montobbio and Sterzi [32], for example, analysed factors influencing international technical cooperation. Kunze [33] explored the relationship between innovation and trade in Europe. Gui et al. [34] discussed the proximity mechanism of the global cooperation network for scientific research papers. This study adopts this model, which is based on the conceptual framework of multidimensional proximity, to develop a proximity mechanism model for intercity innovation networks as follows: R&D expenditure, human capital, and GDP per capita:

$$\begin{aligned} \mathrm{COL}_{ij} = {} & \alpha + \beta_1 \mathrm{PAT}_i + \beta_2 \mathrm{PAT}_j + \beta_3 \mathrm{RDE}_i + \beta_4 \mathrm{RDE}_j + \beta_5 \mathrm{HUC}_i + \beta_6 \mathrm{HUC}_j + \beta_7 \mathrm{GPC}_i + \beta_8 \mathrm{GPC}_j + \beta_9 \mathrm{GEO}_{ij} \\ & + \beta_4 \mathrm{INS}_{ij} + \beta_5 \mathrm{COG}_{ij} + \varepsilon, \end{aligned} \tag{4}$$

where explained variable $\mathrm{COL}_{ij}$ is innovation cooperation between cities measured by the number of joint applications for invention patents by cities $i$ and $j$ in 2014. Control variables conclude patent applications ($\mathrm{PAT}_i$ and $\mathrm{PAT}_j$), R&D expenditure ($\mathrm{RDE}_i$ and $\mathrm{RDE}_j$), human capital ($\mathrm{HUC}_i$ and $\mathrm{HUC}_j$), and GDP per capita ($\mathrm{GPC}_i$ and $\mathrm{GPC}_j$). $\mathrm{PAT}_i$ and $\mathrm{PAT}_j$ are the total number of invention patent applications by cities $i$ and $j$ in 2014 and are used to measure a city's technological innovation capability. This study incorporates these variables in its model to examine the spatial interaction effect between the two cities.$\mathrm{DIS}_{ij}$ is the geographical proximity between city $i$ and city $j$. According to Balland et al. [10]; $\mathrm{GEO}_{ij} = 10 - \ln(\mathrm{DIS}_{ij} + 1)$, where $\mathrm{DIS}_{ij}$ is the spatial spherical distance between city i and city j and is calculated using ArcGIS. $\mathrm{INS}_{ij}$ is the degree of similarity in institutional environments—this includes informal systems such as culture, language, social values, norms, and formal systems; for example, law and regulations and regional development policies—between city $i$ and $j$. As in the study by Ejermo and Karlsson [35], this study sets $\mathrm{INS}_{ij}$ as a virtual variable that takes the value of 1 if the two cities belong to the same province; otherwise, 0. $\mathrm{COG}_{ij}$ denotes similarity in the technical knowledge structure between two cities. In the study by Jaffe [36], $\mathrm{COG}_{ij} = (\sum_{m=1}^{8} \mathrm{PAT}_{im} \mathrm{PAT}_{jm} / \sqrt{\sum_{m=1}^{8} \mathrm{PAT}_{im}^2 \mathrm{PAT}_{jm}^2})$, $\mathrm{PAT}_{im}$, and $\mathrm{PAT}_{jm}$ are the number of invention patent applications under the m-th number of the International Patent Classification for cities i and j. Data on classified patents can be collected from the patent retrieval and analysis system of the National Intellectual Property Administration (see Table 1).

## 5. Empirical Results

The empirical results highlight that 286 Chinese cities (or 98.6%) established innovation networks and, in particular, 42,921 innovation linkages, in 2014 (see Figure 1). However, innovation cooperation is relatively weak in China's intercity innovation networks, reporting a density of only 0.05. In other words, it is difficult to initiate innovation cooperation among cities because the average path length is 2.07, which is greater than that of random networks of the same scale. Moreover, it necessary to strengthen the intensity of innovation cooperation among cities where the average degree of centrality is 15.

This study uses ArcGIS 10.2 and employs Jenks natural breaks' optimisation and further divides the link intensity of China's intercity innovation networks into four grades: high (1,149–3,109), medium (362–1,149), low (89–362), and lower (1–89) intensities. Figure 2 shows that an increase in connection strength causes a rapid decline in the number of intercity linkages. Moreover, a majority of the city innovation linkages demonstrate low strength; that is, 1,626 groups (76.99 percent of total linkages) report less than 10 linkages. Only 3 and 16 groups show high and medium strength, which is less than 1% of total linkages. These results highlight that most cities in China have a relatively low degree of innovation cooperation. While intercity innovation networks are large in scale and wide in coverage, there are obvious problems such as loose links, poor accessibility, and knowledge spillovers, and thus, the functions of intercity innovation networks need to be further enhanced. Boix and

Table 1: Definitions of variables in the proximity mechanism model for intercity innovation networks.

| Variable | Index | Definition | Measure method |
|---|---|---|---|
| Explained variable | Innovation cooperation $(COL_{ij})$ | Ties for innovation cooperation between two cities | Number of joint invention patents for cities $i$ and $j$ |
| | Geographical proximity $(GEO_{ij})$ | Geographical proximity between two cities | $GEO_{ij} = 10 - \ln(DIS_{ij} + 1)$ |
| Explanatory variable | Institutional proximity $(INS_{ij})$ | Degree of similarity in institutional environments between two cities | 1 or 0 |
| | Cognitive proximity $(COG_{ij})$ | Similarity in technical knowledge structure between two cities | $COG_{ij} = \left(\sum_{m=1}^{8} PAT_{im}PAT_{jm} / \sqrt{\sum_{m=1}^{8} PAT_{im}^2 PAT_{jm}^2}\right)$ |
| | Innovation performance (PAT) | City's technological innovation capability | Number of invention patents |
| Control variable | Innovation input (RDE) | | R&D expenditure |
| | Innovation talent input (HUC) | Human capital | Number of personnel with bachelor degree or above |
| | Economic development (GPC) | | GDP per capital |

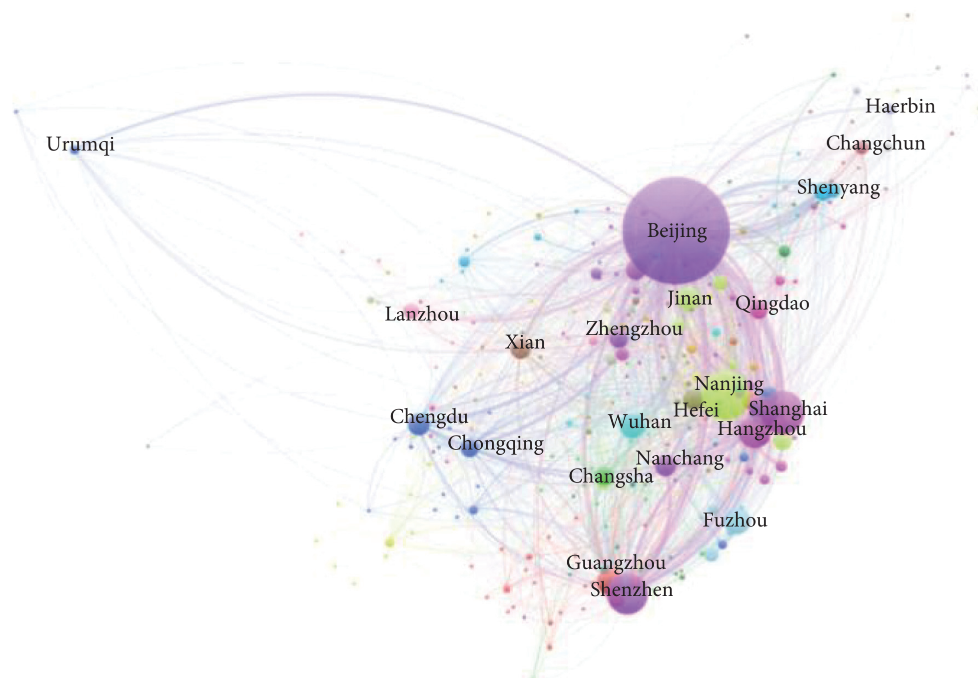



Figure 1: Topological structure of China's intercity innovation networks in 2014.

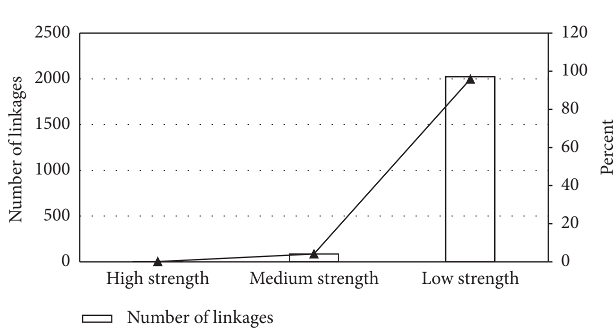



Figure 2: Linkage strength in China's intercity innovation networks for 2014.

Trullén [17] also found that higher growth rates are associated with higher levels of technology and knowledge, and the growth of the different kinds of knowledge is related to local and spatial factors (agglomeration and network externalities), and each knowledge intensity shows a particular response to these factors. Bettencourt et al. [37] found that the structure of the patent coauthorship network weakly correlated to increasing rates of patenting.

The Hu line has largely contributed to the east-west inequality in innovation linkages between cities. The linkages in eastern cities are greater than those in western cities. High-intensity innovation linkages are observed mainly among few big cities, such as Beijing-Nanjing, Beijing-Tianjin, and Beijing-Shanghai. The geographical spatial pattern of medium-intensity innovation linkages demonstrates a “single-centre radiation” feature, in which Beijing is at the centre and radiating outward. Low-intensity innovation linkages, on the other hand, report “single-centre radiation and local networking,” where Beijing is at the centre. Complex and interactive network linkages can also be observed in Yangtze River Delta and Pearl River Delta. The researchers found that network linkages to the east of the Hu line tend to have varying intensities, while those in the western cities have low intensity.

The network centrality potential of China’s intercity innovation networks is 0.95, indicating a “core edge” structure. Using the Pajek 4.08 software and a hierarchical clustering algorithm for the block model, this study divides China’s intercity innovation networks into five levels: strong core, core, strong semi-edge, weak semi-edge, and edge (see Table 2). Next, it uses the network map drawing tool in VOSviewer 1.6.5 to draw the hierarchical structure diagram of China’s intercity innovation networks (see Figure 3). The results show that only Beijing is at the strong core of the network; 17 cities including Shanghai, Nanjing, and Wuhan are at the core; 30 cities consisting of Hefei, Wuxi, and Dalian are at the strong semi-edge; 75 cities comprising Yueyang, Beihai, and Jinmen are at the weak semi-edge; and, 163 cities including Guyuan, Sanmenxia, and Bazhong are at the edge. Figure 3 shows that cities in the strong core of the network are deeply connected with cities that are at the core, are lower in the hierarchy, and have fewer innovation linkages with other cities.

According to Table 2 and Figure 4, the case cities can be divided into five types: strong core city, core city, subcore city, subperiphery city, and periphery city. See Figure 4 for a visualisation using ArcGIS 10.2. Similar to the distribution of network linkages, the city hierarchy is bound by the Hu line. That is, a majority of the core and subcore cities are located to the east of the Hu line, and the overall distribution is relatively scattered. Few cities concentrated in the Yangtze River Delta and Pearl River Delta report the geographical distribution characteristics of “large dispersion and small agglomeration.” By contrast, most of the cities to the west of the Hu line are subperiphery or periphery cities. Cities at the top of China’s intercity innovation networks are mainly administrative such as Beijing. Shenzhen and Suzhou, which are provincial capitals or municipalities, are core cities possibly because China’s national or regional innovation system is significantly influenced by the government. The research of van der Wouden and Rigby [18] shows that metropolitan regions with more local and nonlocal network ties outperform cities where economic agents are isolated, and coinventor networks differ between cities that produce specialized and diversified knowledge. Many important scientific research institutions and universities are located in administratively central cities, such as Beijing, Shanghai, Nanjing, Wuhan, Guangzhou, and Hangzhou, and this increases the possibility of innovation networking [38,39]. Lobo and Strumsky [40] also found that agglomerative features of metropolitan areas are more important determinants of metropolitan patenting productivity than the structural feature of the inventive networks.

## 6. Mechanism for City Innovation Networking in China

Using formula (4), a model for the relationship between intercity innovation networks and geographical, institutional, and cognitive proximities is established. The correlation coefficients for the explanatory variables indicate that, except institutional and geographical proximities, which are 0.707, the other variables are less than 0.2. The variance inflation factor (VIF) for all explanatory variables is less than 3. Therefore, there is no multiple collinearity problem among the explanatory variables in the relationship model, and their reliability is high. For the explained variables, which report a nonnegative integer, it is necessary to use a discrete counting model such as the Poisson regression or negative binomial regression models. The data of the explained variables are excessively dispersed since the variance of the explained variables is 10739.98, while the expected value is only 20.32, which is significantly less than the variance. This study employs a negative binomial regression model to test the proximity mechanism of China’s intercity innovation networking (see Table 3).

A stepwise regression test was conducted on the models, and the results show that Models 1–3 reject the original hypothesis ($\alpha = 0$) at the 5% level, thus reiterating the rationality of the negative binomial regression model. The $p$ values for the three models are zero, and the results do not reject the null hypothesis. A majority of the estimated parameters for the explanatory variables are significant at the 1% level, demonstrating the reliability of the test results above (see Table 4).

Models 1–3 show that geographical proximity is significantly positive at the 1% level, suggesting that greater geographical proximity between two cities renders innovation networking more likely. The contrary also holds true. Innovation is a process of strong interactive learning and communication along with the obvious characteristics of knowledge exchange and transfer, particularly tacit knowledge, and requires frequent face-to-face interactions among innovation actors [41, 42]. Huggins et al. [43] also pointed out that the springboard effect and the geography of external knowledge networks are associated with the regional economic context, Capone and Lazzeretti [12]; Knoben [44] underlined the heterogeneous impact of various forms of proximity on the different relationships and the strong impact of social ties on innovation. This also explains the criticality of geographical proximity. Increasing geographical distances between cities will increase the time and economic cost of face-to-face interactions and create obstacles in cooperative efforts. Moreover, it significantly decreases the potential for innovation networking among cities. This finding strongly refutes the views of “the death of geographical proximity” and “geographical death” proposed by Cairncross [45] and Friedman [46]. Further, it proves that geographical proximity continues to play a vital role in China’s intercity innovation networks, even though transportation and information technology are highly developed,

Table 2: Hierarchy of China's intercity innovation networks.

| Network hierarchy | Number of cities | Network centrality | Average of network centrality potential | Cities |
|---|---|---|---|---|
| Strong core | 1 | 255 | 255 | Beijing |
| Core | 17 | 54–140 | 77 | Shanghai, Nanjing, Suzhou, Wuhan, Zhengzhou, Tianjin, Xian, Jinan, Chongqing, Chengdu, Changsha, Shenzhen, Hangzhou, Guangzhou, Shenyang, Qingdao, and Changchun |
| Strong semiperiphery | 30 | 23–52 | 34 | Hefei, Wuxi, Dalian, Dongguan, Ningbo, Kunming, Xiamen, Lanzhou, Foshan, Taiyuan, Changzhou, Nanchang, Xuzhou, Yantai, Guiyang, Nanning, Zhenjiang, Urumqi, Nantong, Yinchuan, Fuzhou, Zhuhai, Yangzhou, Luoyang, Taizhou (Jiangsu), Yichang, Yancheng, Shijiazhuang, Haerbin, and Lianyungang |
| Weak semiperiphery | 75 | 9–22 | 13 | Yueyang, Beihai, Jinmen, Jian, Kaifeng, Yichun, Zhongwei, Huizhou, Jiangmen, Chuanzhou, Anyang, Qinhuangdao, Chifeng, Jiaozuo, Huainan, Fuxin, Dezhou, Jilin, Luan, Huaibei, Suqian, Huangshan, Tongling, Bengbu, Changde, Zhaoqing, Langfang, Qingyuan, Shantou, Sanya, Guilin, Liuzhou, Shizuishan, Hengyang, Shaoguan, Huzhou, Heyuan, Zhangzhou, Zhongshan, Zhuzhou, Anqing, Tongliao, Handan, Weifang, Weihai, Taian, Chuzhou, Jiujiang, Mianyang, Jinhua, Zibo, Lishui, Zhoushan, Quzhou, Wenzhou, Taizhou (Zhejiang), Xiangtan, Wuhu, Jining, Shaoxing, Jiaxing, Zhanjiang, Haikou, Xuchang, Xining, Pingdingshan, Maanshan, Changzhi, Dongying, Huaian, Tangshan, Zhangjiakou, Baoding, and Anshan |
| Periphery | 163 | 1–8 | 4 | Remaining 163 cities |

and geographical distance is a crucial factor in innovation actors choosing innovative partners.

Models 2 and 3 show that institutional proximity also significantly promotes innovation networking among Chinese cities, and the probability of innovation cooperation among cities with institutional proximity is higher. A city is not a simple geospatial unit but a comprehensive carrier of various natural and human elements. In the course of their historical development, cities develop different humanistic qualities that contribute to varying institutional environments, including informal institutional environments such as culture, language, and norms. However, these differences create institutional barriers in intercity innovation networking. Thus, cities with similar institutional environment factors create more conducive environments to enhance mutual trust between their innovation actors. This reduces uncertainty in the innovation networking process, decreases the costs of exchange, and enhances the possibility of innovation networking between cities. This finding is consistent with the viewpoint of Maillat and Kebir [47], who highlighted that special informal norms play a crucial role in forming innovation networks among regional innovative actors and similar institutions facilitate interregional innovation networking.

Model 3 shows that cognitive proximity is significantly positive at the 1% level, indicating that cognitive proximity promotes innovative cooperation among Chinese cities. In other words, cities with similar cognitive levels are highly likely to become innovative partners. Cognitive convergence renders two cities similar in technology specialisation, thus ensuring effective communication in innovation networking and the identification and absorption of new knowledge. This finding is consistent with those of Guellec [48], Ponds et al. [49], and Guellec [48], who found that countries with similar technical expertise are more likely to engage in innovation cooperation. Ponds et al. [49] examined the industry-university-research cooperation in the Netherlands and showed that technological proximity is a necessary prerequisite for innovative cooperation among cross-regional enterprises, scientific research institutions, and universities.

Some scholars believe that the proximity of various dimensions affects and interacts with each other, and there may be some substitution or complementary effects [50, 51]. This study conducts a variable interactions' test to further examine the interaction among geographical, cognitive, and institutional proximities in China's intercity innovation networks. Because geographical and cognitive proximities are continuous variables, institutional proximity is considered a category variable. To ensure that the method is correct and the results are reliable, the study uses one proximity type among geographical, cognitive, or institutional proximities as the main variable, and the other two are considered regulatory variables to examine interactions between two proximities (see Table 5).

Using geographical proximity as the main variable and cognitive proximity as the regulatory variable, Model 4 is constructed to test the regulatory effect of cognitive proximity on geographical proximity. Similarly, Model 5 is constructed to test the regulatory effect of geographical proximity on cognitive proximity. Since both geographical and cognitive proximities are continuous variables, an

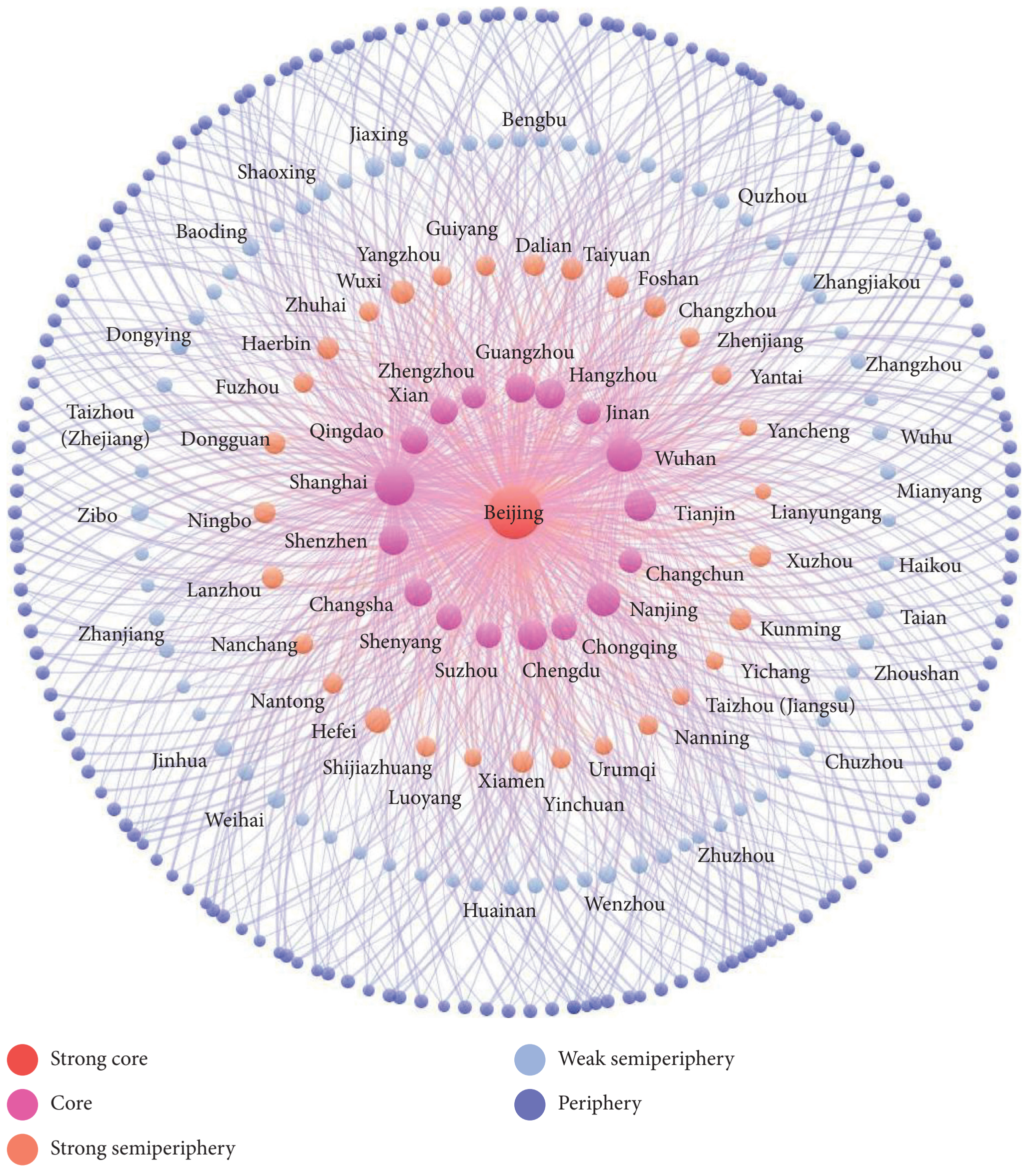



FIGURE 3: Hierarchical structure diagram of China's intercity innovation networks.

interaction term for the centralisation of geographical and cognitive proximities is added to the basic test model. Models 4 and 5 show that the interactive coefficient between geographical and cognitive proximities is significant at the 1% level, indicating significant interactions between both proximities. However, the significantly negative coefficient suggests that cognitive proximity will weaken the impact of geographical proximity on intercity innovation networking and vice versa. On the one hand, this substitution effect can be attributed to innovative cooperation with partners that are both cognitively and geographically close and that would otherwise be unable to access new information and knowledge, resulting in "excessive embedding" and "cognitive locking." On the other hand, the scope for learning from each other considerably reduces for regions with similar geographical and cognitive proximity, and the risk of unconscious knowledge spillovers increases [52].

Next, institutional proximity is treated as the main variable, while geographical and cognitive proximities are considered regulatory variables. The interactive terms of $INS_{ij} * GEO_{ij}$ and $INS_{ij} * COG_{ij}$ are added to the basic model. Model 6 shows that the interactive coefficient for institutional and geographical proximities is significantly positive at the 5% level. In other words, the impact of institutional proximity on intercity innovation networking is positively regulated by geographical proximity and that of institutional proximity on intercity innovation networking will strengthen when the two cities are geographically close; this is a complementary effect. Model 7 shows that the interactive coefficient for institutional and cognitive proximities is also significant at the 5% level, but it is negative. This finding indicates that cognitive proximity will negatively regulate the relationship between institutional proximity and intercity innovation networking, which is a substitution effect.

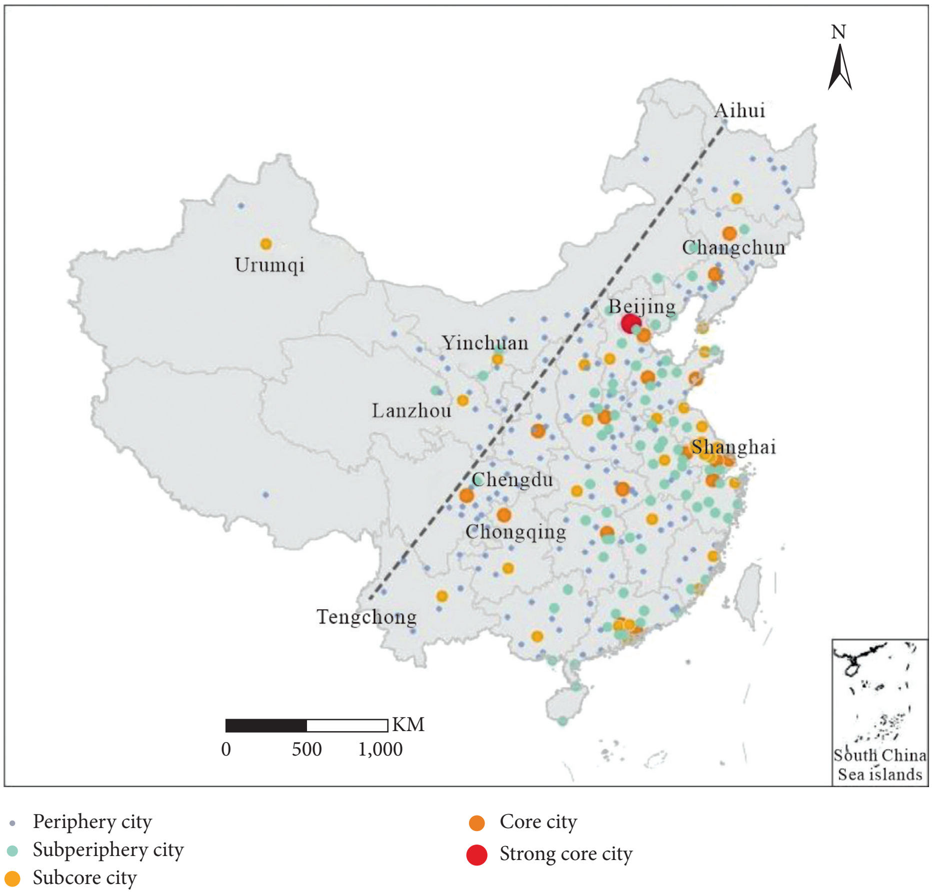



Figure 4: Hierarchical spatial distribution of China's cities.

Table 3: Descriptive statistical and correlation coefficient matrix of variables ($N = 2{,}112$).

| Variables | Expected value | Variance | Max | Min | VIF | $COL_{ij}$ | $DIS_{ij}$ | $INS_{ij}$ | $COG_{ij}$ |
|---|---|---|---|---|---|---|---|---|---|
| $COL_{ij}$ | 20.32 | 10739.98 | 3109.0 | 1.00 | | 1.00 | | | |
| $GEO_{ij}$ | 3.57 | 0.88 | 7.00 | 1.85 | 2.020 | 0.016 | 1.00 | | |
| $INS_{ij}$ | 0.18 | 0.15 | 1.00 | 0.00 | 2.000 | −0.024 | $0.707^{***}$ | 1.00 | |
| $COG_{ij}$ | 0.81 | 0.02 | 1.00 | 0.25 | 1.035 | −0.020 | $0.181^{***}$ | $0.153^{***}$ | 1.00 |

*Notes.* $^{*}p < 0.10$, $^{**}p < 0.05$, and $^{***}p < 0.01$.

Next, this study uses both geographical and cognitive proximities as the main variables and institutional proximity as a regulatory variable to test the regulatory effect of institutional proximity on geographical and cognitive proximities. First, a group regression is conducted to divide the sample data into two groups depending on institutional proximity, and then, a negative binomial regression analysis is performed on the two sample groups (see Table 6). Models 8 and 9 show that geographical proximity actively promotes intercity innovation networking at the 1% level. However, the influence coefficient in the sample group for institutional proximity is larger, and the promotion effect is stronger, indicating that institutional proximity can enhance the positive influence of geographical proximity on intercity innovation networking, which is a complementary effect. The coefficient for cognitive proximity fails the significance test on the sample group for institutional proximity, although it is significantly positive in the sample group for institutional nonproximity. This means institutional proximity will weaken the positive impact of cognitive proximity on intercity innovation networking, which is a substitution effect.

Figure 5 is a graph of the variables' regulatory effect and is created to better understand interactions between the different proximities. It shows the regulatory effect of geographical, cognitive, and institutional proximities. For cities with greater geographical proximity, the positive impact of cognitive proximity on intercity innovation networking weakens or even has a negative impact. In [13], the author points out that interregional network proximity is important in determining future collaborations but its effect is mediated by geography. However, the impact of institutional proximity on intercity innovation networking significantly increases. For cities with higher cognitive

Table 4: Regression test results for impact of three proximities on intercity innovation networking.

| Variables | Model 1 | Model 2 | Model 3 |
|---|---|---|---|
| $PAT_i$ | 0.0000412*** (0.000) | 0.000042*** (0.000) | 0.000044*** (0.000) |
| $PAT_j$ | 0.0000432*** (0.000) | 0.0000455*** (0.000) | 0.0000464*** (0.000) |
| $RDE_i$ | 0.0000222** (0.000) | 0.0000021*** (0.000) | 0.0000004*** (0.000) |
| $RDE_j$ | 0.0000231** (0.000) | 0.0000025*** (0.000) | 0.0000014*** (0.000) |
| $HUC_i$ | 0.0000082 * (0.000) | 0.0000079** (0.000) | 0.0000052*** (0.000) |
| $HUC_j$ | 0.0000063 * (0.000) | 0.0000065** (0.000) | 0.0000050*** (0.000) |
| $GPC_i$ | 0.0000118** (0.000) | 0.0000226*** (0.000) | 0.0000076** (0.000) |
| $GPC_j$ | 0.0000137** (0.000) | 0.0000253*** (0.000) | 0.0000069** (0.000) |
| $GEO_{ij}$ | 0.4528901*** (0.000) | 0.2658215*** (0.000) | 0.252175*** (0.000) |
| $INS_{ij}$ | | 0.5930518*** (0.000) | 0.6121559*** (0.000) |
| $COG_{ij}$ | | | 1.338232*** (0.001) |
| CONS | 1.01691*** (0.000) | 0.9689111*** (0.000) | 0.9120978*** (0.000) |
| N | 2112 | 2112 | 2112 |
| Prob. > chi2 | 0.0000 | 0.0000 | 0.0000 |
| Log pseudolikelihood | −6791.1238 | −6775.763 | −6754.5513 |

*Note.* $^{*}p < 0.10$, $^{**}p < 0.05$, and $^{***}p < 0.01$. $N$ represents the sample size in the model.

Table 5: Regression results of the interactive model for different proximities.

| Variables | Model 4 | Model 5 | Model 6 | Model 7 |
|---|---|---|---|---|
| $PAT_i$ | 0.0000437*** (0.000) | 0.0000437*** (0.000) | 0.0000441*** (0.000) | 0.0000445*** (0.000) |
| $PAT_j$ | 0.0000469*** (0.000) | 0.0000469*** (0.000) | 0.0000463*** (0.000) | 0.0000469*** (0.000) |
| $RDE_i$ | 0.0000232** (0.000) | 0.0000221*** (0.000) | 0.0000204*** (0.000) | 0.0000204*** (0.000) |
| $RDE_j$ | 0.0000252** (0.000) | 0.0000225*** (0.000) | 0.0000214*** (0.000) | 0.0000234*** (0.000) |
| $HUC_i$ | 0.0000462 * (0.000) | 0.0000469** (0.000) | 0.0000452*** (0.000) | 0.0000433*** (0.000) |
| $HUC_j$ | 0.0000336 * (0.000) | 0.0000356** (0.000) | 0.0000305*** (0.000) | 0.0000325*** (0.000) |
| $GPC_i$ | 0.0000118** (0.000) | 0.0000122*** (0.000) | 0.0000112** (0.000) | 0.0000176** (0.000) |
| $GPC_j$ | 0.0000168** (0.000) | 0.0000132*** (0.000) | 0.0000155** (0.000) | 0.0000143** (0.000) |
| $GEO_{ij}$ | 0.2566725*** (0.000) | 0.2566725*** (0.000) | 0.2548991*** (0.000) | 0.2383039*** (0.000) |
| $INS_{ij}$ | 0.6575141*** (0.000) | 0.6575141*** (0.000) | 0.1419111 (0.000) | 0.7544208*** (0.000) |
| $COG_{ij}$ | 1.029992** (0.012) | 1.029992** (0.012) | 1.378283*** (0.000) | 1.097303*** (0.007) |
| $GEO_{ij}*COG_{ij}$ | −1.195481*** (0.005) | | | |
| $COG_{ij} * GEO_{ij}$ | | −1.195481*** (0.005) | | |
| $INS_{ij}*GEO_{ij}$ | | | 0.3993853** (0.022) | |

Table 5: Continued.

| Variables | Model 4 | Model 5 | Model 6 | Model 7 |
|---|---|---|---|---|
| $INS_{ij} * COG_{ij}$ | | | | −3.389262** (0.014) |
| CONS | 0.933344*** (0.000) | 0.933344*** (0.000) | 0.806204*** (0.000) | 0.9181173*** (0.000) |
| N | 2112 | 2112 | 2112 | 2112 |
| Prob. > chi2 | 0.0000 | 0.0000 | 0.0000 | 0.0000 |
| Log pseudolikelihood | −6741.4869 | −6741.4869 | −6748.3122 | −6741.4968 |

*Note.* * $p < 0.10$, ** $p < 0.05$, and *** $p < 0.01$. $N$ represents the sample size in the model.

Table 6: Regulatory effect of institutional proximity on geographical and cognitive proximities.

| Variables | Sample group of institutional proximity (Model 8) | Sample group of institutional nonproximity (Model 9) |
|---|---|---|
| $PAT_i$ | 0.0000783*** (0.000) | 0.0000439*** (0.000) |
| $PAT_j$ | 0.0000494*** (0.000) | 0.0000463*** (0.000) |
| $RDE_i$ | 0.0000652*** (0.000) | 0.0000632*** (0.000) |
| $RDE_j$ | 0.0000572*** (0.000) | 0.0000545*** (0.000) |
| $HUC_i$ | 0.0000486** (0.000) | 0.0000415*** (0.000) |
| $HUC_j$ | 0.0000669 * (0.000) | 0.0000681** (0.000) |
| $GPC_i$ | 0.0000529** (0.000) | 0.0000531*** (0.000) |
| $GPC_j$ | 0.0000547** (0.000) | 0.0000533*** (0.000) |
| $GEO_{ij}$ | 0.5462855*** (0.001) | 0.1794366*** (0.000) |
| $COG_{ij}$ | −1.189488 (0.341) | 1.682881*** (0.000) |
| CONS | 0.7722817*** (0.001) | 0.7867398*** (0.000) |
| $N$ | 376 | 1,736 |
| Prob. > chi2 | 0.0000 | 0.0000 |
| Log pseudolikelihood | −1250.723 | −5475.5285 |

Note: * $p < 0.10$, ** $p < 0.05$, and *** $p < 0.01$. $N$ represents the sample size in the model.

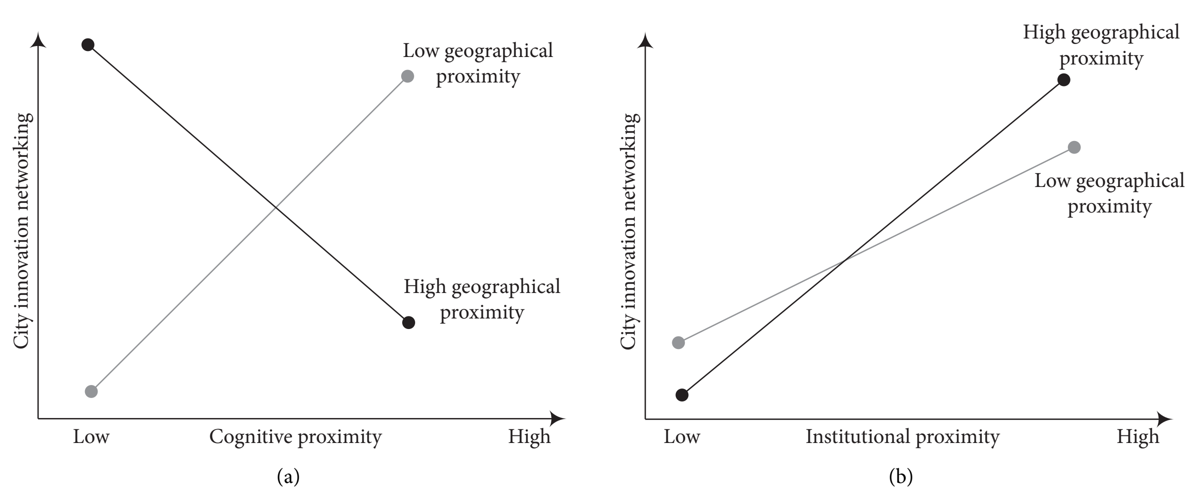



Figure 5: Continued.

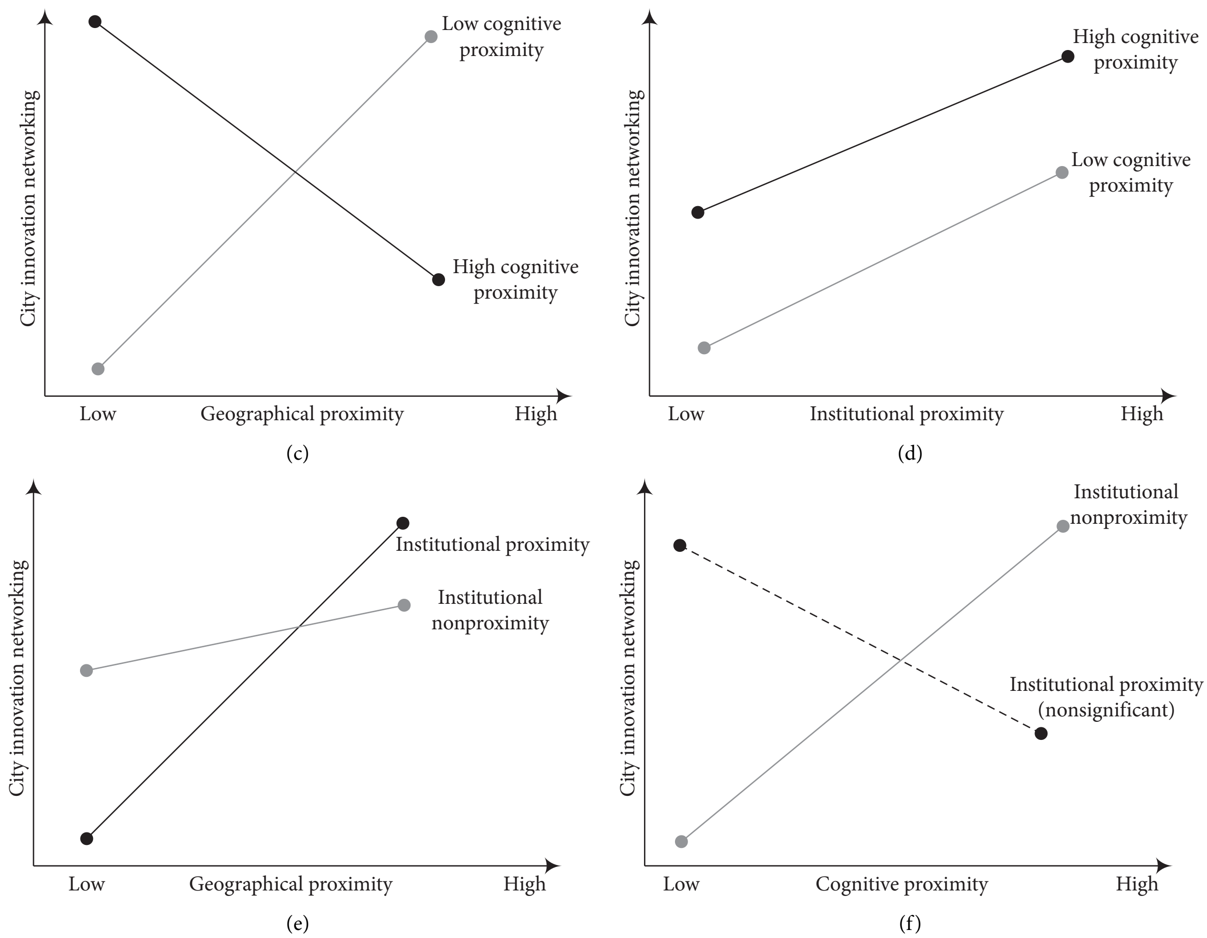



Figure 5: Regulatory effects of geographical, cognitive, and institutional proximities.

proximity, the positive impact of both geographical and institutional proximities considerably weakens. Finally, for cities with greater institutional proximity, the impact of geographical proximity strengthens and becomes more positive, while that of cognitive proximity is no longer significant.

## 7. Conclusions

There is growing concern regarding regional development and innovation in the economic geography literature [53–55]. In numerous recent studies on regional economic development, spatial proximity, density, and localised processes, which cannot explain new economic phenomena, global and local proximities have become increasingly important [50, 56, 57]. Adopting the innovation networks' perspective and conducting an SNA, this study analyses the structural characteristics, that is, the topology and spatial pattern of city networks in China. Drawing on Scherngell and Barber's [30] spatial interaction model, a model for the relationship between proximity and intercity innovation networks is established. The model is then employed to explain the interaction and impact of geographic, institutional, and cognitive proximities on China's intercity innovation networks.

The findings reveal that while the scale and scope of Chinese intercity innovation networks have been increasing, their density is lower and most cities are loosely connected. At the network level, the results suggest that the intercity innovation networks are more centred. That is, the agglomeration and hierarchy of innovation activities are prominent and develop into a core-periphery structure, which can be further divided into strong core, core, strong semiperiphery, weak semiperiphery, and periphery. Only Beijing is at the strong core of the network. Shanghai, Nanjing, and Wuhan are among the 14 cities at the core of network, and the remaining are at the strong semiperiphery, weak semiperiphery, or periphery of the innovation networks. According to the structure of innovation networks in the Chinese cities, we can divide the cities into five levels: strong core, core, subcore, subperiphery, and periphery. We find that the hierarchy of Chinese cities is high in the east and low in the west, with the core and subcore cities demonstrating a spatial pattern of "large dispersion and small agglomeration" and bound by the Hu line. To the east of the Hu line, the density and strength of innovation linkages between cities are higher, and thus, the networks are complex and efficient with Beijing as the strong core city. However, to the west of the Hu line, the innovation linkages are weaker and most cities are subperiphery or periphery cities.

## 8. Discussion

This study analyses mechanisms for intercity innovation networking in China and offers interesting conclusions consistent with the viewpoints of Maillat and Kebir [47], Guellec [48], Ponds et al. [49], and van der Wouden and Rigby [18], who argue that inventors in specialized cities value spatial proximity less and cognitive proximity more than inventors in diversified cities as they partner with nonlocal inventors. However, it refutes Cairncross [45] and Friedman's [46] views of "the death of geographical proximity" and "geographical death." The analysis reveals that geographical, institutional, and cognitive proximities positively impact innovation networks in Chinese cities. In particular, cognitive proximity has the strongest impact, followed by institutional and geographical proximities. The test results for variable interactions show significant interactions among geographical, cognitive, and institutional proximities. Geographical and institutional proximities positively regulate each other's relationship with intercity innovation networking, and the relationship between geographical and institutional proximities reports a complementary effect. However, there is a substitutive relationship between cognitive proximity and geographical and institutional proximities, and as a result, their role in promoting intercity innovation networking is weakened. This further confirms that cognitive proximity may result in increasingly weaker networks and is not conducive to the transformation of product structures in developing countries and regions [26].

Future research should consider further exploring data and methods to measure intercity innovation networks, introducing more proximity types, exploring the dynamics of the same proximity over time, and summarizing the regularity of each proximity type within the dynamic evolution of innovation networks.

## Data Availability

The data used to support the findings of the study are available from the corresponding author upon reasonable request.

## Conflicts of Interest

The authors declare that they have no conflicts of interest.

## Acknowledgments

This work was supported by the National Natural Science Foundation of China (41801109 and 41771143) and Project of Science and Technology Commission of Shanghai Municipality (19DZ1203801).

## References

[1] P. M. Cooke, *Regional Innovation System Barriers and the Rise of Boundary-Crossing Institutions*, Academia-Business Links, Palgrave, Macmillan, UK, 2004.

[2] F. Quatraro and S. Usai, "Knowledge flows, externalities and innovation networks," *Regional Studies*, vol. 51, no. 8, pp. 1133–1137, 2017.

[3] L. Lyu, Z. Liang, and R. Huang, "The innovation linkage among Chinese major cities," *Scientia Geographica Sinica*, vol. 35, no. 1, pp. 30–37, 2015.

[4] F. Zhang and F. &Wu, "Rethinking the city and innovation: a political economic view from China's biotech," *Cities*, vol. 85, pp. 150–155, 2018.

[5] R. Huggins and P. Thompson, "A network-based view of regional growth," *Journal of Economic Geography*, vol. 14, no. 3, pp. 511–545, 2014.

[6] R. Huggins and P. Thompson, "Networks and regional economic growth: a spatial analysis of knowledge ties," *Environment and Planning A*, vol. 49, no. 2, pp. 1–19, 2017.

[7] R. Rothwell, "Successful industrial innovation: critical factors for the 1990s," *R&D Management*, vol. 22, no. 3, pp. 221–240, 1992.

[8] X. Cao, G. Zeng, T. Teng, and Y. Si, "The best spatial scale of firm innovation networks: evidence from Shanghai high-tech firms," *Growth and Change*, vol. 49, no. 4, pp. 696–711, 2018.

[9] I. Liefner and S. Hennemann, "Structural holes and new dimensions of distance: the spatial configuration of the scientific knowledge network of China's optical technology sector," *Environment and Planning A: Economy and Space*, vol. 43, no. 4, pp. 810–829, 2011.

[10] P.-A. Balland, M. De Vaan, and R. Boschma, "The dynamics of interfirm networks along the industry life cycle: the case of the global video game industry, 1987–2007," *Journal of Economic Geography*, vol. 13, no. 5, pp. 741–765, 2013.

[11] P. A. Balland, "Proximity and the evolution of collaboration networks: evidence from research and development projects within the global navigation satellite system industry," *Regional Studies*, vol. 46, no. 6, pp. 741–756, 2011.

[12] F. Capone and L. Lazzeretti, "The different roles of proximity in multiple informal network relationships: evidence from the cluster of high technology applied to cultural goods in Tuscany," *Industry and Innovation*, vol. 25, no. 9, pp. 897–917, 2018.

[13] L. R. Bergé, "Network proximity in the geography of research collaboration," *Papers in Regional Science*, vol. 96, no. 4, pp. 785–816, 2017.

[14] D. Sudharshan, R. Madhavan, and S. C. Liu, "DiffuNET: the impact of network structure on diffusion of innovation," *European Journal of Innovation Management*, vol. 8, no. 2, pp. 240–262, 2005.

[15] M. A. Schilling and C. C. Phelps, "Interfirm collaboration networks: the impact of large-scale network structure on firm innovation," *Management Science*, vol. 53, no. 7, pp. 1113–1126, 2007.

[16] C. Freeman, "Networks of innovators: a synthesis of research issues," *Research Policy*, vol. 20, no. 5, pp. 499–514, 1991.

[17] R. Boix and J. Trullén, "Knowledge, networks of cities and growth in regional urban systems," *Papers in Regional Science*, vol. 86, no. 4, pp. 551–574, 2007.

[18] F. van der Wouden and D. L. Rigby, "Co-inventor networks and knowledge production in specialized and diversified cities," *Papers in Regional Science*, vol. 10, no. 98, pp. 1833–1853, 2019.

[19] A. Marshall, *Principles of Economics*, Macmillan, London, UK, 1890.

[20] J. Knoben, "Localized inter-organizational linkages, agglomeration effects, and the innovative performance of firms,"

*The Annals of Regional Science*, vol. 43, no. 3, pp. 757–779, 2009.
[21] M. Bradshaw, “Multiple proximities: culture and geography in the transport logistics of newsprint manufactured in Australia,” *Environment and Planning A: Economy and Space*, vol. 33, no. 10, pp. 1717–1739, 2001.
[22] J. Knoben and L. A. G. Oerlemans, “Proximity and inter-organizational collaboration: a literature review,” *International Journal of Management Reviews*, vol. 8, no. 2, pp. 71–89, 2006.
[23] R. Boschma, “Proximity and innovation: a critical assessment,” *Regional Studies*, vol. 39, no. 1, pp. 61–74, 2005.
[24] P.-A. Balland, J. A. Belso-Martínez, and A. Morrison, “The dynamics of technical and business knowledge networks in industrial clusters: embeddedness, status, or proximity?” *Economic Geography*, vol. 92, no. 1, pp. 35–60, 2016.
[25] T. Broekel and R. Boschma, “Knowledge networks in the Dutch aviation industry: the proximity paradox,” *Journal of Economic Geography*, vol. 12, no. 2, pp. 409–433, 2012.
[26] C. He, L. Jin, and Y. Liu, “How does multi-proximity affect the evolution of export product space in China?” *Geographical Research*, vol. 36, no. 9, pp. 1613–1626, 2017.
[27] C. Zhou, G. Zeng, and X. Cao, “Chinese inter-city innovation networks structure and city innovation capability,” *Geographical Research*, vol. 36, no. 7, pp. 1297–1308, 2017.
[28] A. L. J. Ter Wal and R. A. Boschma, “Applying social network analysis in economic geography: framing some key analytic issues,” *The Annals of Regional Science*, vol. 43, no. 3, pp. 739–756, 2009.
[29] I. Wanzenböck, “A concept for measuring network proximity of regions in R&D networks,” *Social Networks*, vol. 54, no. 54, pp. 314–325, 2018.
[30] T. Scherngell and M. J. Barber, “Spatial interaction modelling of cross-region R&D collaborations: empirical evidence from the 5th EU framework programme,” *Papers in Regional Science*, vol. 88, no. 3, pp. 531–546, 2009.
[31] T. Scherngell and Y. Hu, “Collaborative knowledge production in China: regional evidence from a gravity model approach,” *Regional Studies*, vol. 45, no. 6, pp. 755–772, 2011.
[32] F. Montobbio and V. Sterzi, “The globalization of technology in emerging markets: a gravity model on the determinants of international patent collaborations,” *World Development*, vol. 44, no. 4, pp. 281–299, 2013.
[33] K. Kunze, “Empirical analysis of innovation and trade in Europe: a gravity model approach,” *International Journal of Trade and Global Markets*, vol. 9, no. 3, pp. 197–211, 2016.
[34] Q. Gui, C. Liu, and D. Du, “Does network position foster knowledge production? Evidence from international scientific collaboration network,” *Growth and Change*, vol. 49, no. 4, pp. 594–611, 2018.
[35] O. Ejermo and C. Karlsson, “Interregional inventor networks as studied by patent coinventorships,” *Research Policy*, vol. 35, no. 3, pp. 412–430, 2006.
[36] A. B. Jaffe, “Technological opportunity and spillovers of R&D: evidence from firms’ patents, profits, and market value,” *American Economic Review*, vol. 76, no. 5, pp. 984–1001, 1986.
[37] L. M. A. Bettencourt, J. Lobo, and D. Strumsky, “Invention in the city: increasing returns to patenting as a scaling function of metropolitan size,” *Research Policy*, vol. 36, no. 1, pp. 107–120, 2007.
[38] D. E. Andersson, S. Gunessee, C. W. Matthiessen, and S. Find, “The geography of Chinese science,” *Environment and Planning A: Economy and Space*, vol. 46, no. 12, pp. 2950–2971, 2014.
[39] Q. Wang, G. Zeng, and G. Lyu, “Structural evolution of innovation networks of China’s equipment manufacturing industry,” *Acta Geographica Sinica*, vol. 71, no. 2, pp. 251–264, 2016.
[40] J. Lobo and D. Strumsky, “Metropolitan patenting, inventor agglomeration and social networks: a tale of two effects,” *Journal of Urban Economics*, vol. 63, no. 3, pp. 871–884, 2008.
[41] B. Veelen, “The new oxford handbook of economic geography,” *Local Environment*, vol. 23, no. 8, pp. 900-901, 2018.
[42] G. L. Clark, M. S. Gertler, and M. P. Feldman, “The oxford handbook of economic geography,” *Economic Geography*, vol. 78, no. 3, pp. 387–392, 2002.
[43] R. Huggins, H. Izushi, and D. Prokop, “Regional advantage and the geography of networks: explaining global-local knowledge sourcing patterns,” *Papers in Regional Science*, vol. 60, no. 98, pp. 1567–1584, 2019.
[44] J. Knoben, *Firm Mobility and Organizational Networks: Innovation, Embeddedness and Economic Geography*, Edward Elgar Publishing, Cheltenham, UK, 2008.
[45] F. Cairncross, *The Death of Distance 2.0: How the Communications Revolution Will Change Our Lives*, Texere, Knutsford, England, UK, 2001.
[46] T. L. Friedman, “The world is flat: a brief history of the twenty-first century,” *International Journal*, vol. 9, no. 1, pp. 67–69, 2005.
[47] D. Maillat and L. Kebir, *The Learning Region and Territorial Production Systems*, Springer, Berlin, Heidelberg, 2001.
[48] D. Guellec and B. van Pottelsberghe de la Potterie, “The internationalisation of technology analysed with patent data,” *Research Policy*, vol. 30, no. 8, pp. 1253–1266, 2001.
[49] R. Ponds, F. van Oort, and K. Frenken, “The geographical and institutional proximity of research collaboration,” *Papers in Regional Science*, vol. 86, no. 3, pp. 423–443, 2007.
[50] R. Crescenzi, M. Nathan, and A. Rodríguez-Pose, “Do inventors talk to strangers? On proximity and collaborative knowledge creation,” *Research Policy*, vol. 45, no. 1, pp. 177–194, 2016.
[51] P. D’Este, F. Guy, and S. Iammarino, “Shaping the formation of university–industry research collaborations: what type of proximity does really matter?” *Journal of Economic Geography*, vol. 13, no. 4, pp. 537–558, 2013.
[52] R. Crescenzi and S. Iammarino, “Global investments and regional development trajectories: the missing links,” *Regional Studies*, vol. 51, no. 1, pp. 97–115, 2017.
[53] H. Bathelt, P. Li, and Y. Zhu, “Geographies of temporary markets: an anatomy of the canton fair,” *European Planning Studies*, vol. 5, pp. 1–19, 2017.
[54] R. Martin, “Roepke lecture in economic geography-rethinking regional path dependence: beyond lock-in to evolution,” *Economic Geography*, vol. 86, no. 1, pp. 1–27, 2010.
[55] S. Zhu, C. He, and Y. Zhou, “How to jump further and catch up? Path-breaking in an uneven industry space,” *Journal of Economic Geography*, vol. 17, no. 3, pp. 521–545, 2017.
[56] C. He and H. Chen, “Participation in global production networks and export product upgrading,” *Acta Geographica Sinica*, vol. 72, no. 8, pp. 1331–1346, 2017.
[57] E. Uyarra, “Regional innovation systems revisited: networks, institutions, policy and complexity,” *Openloc Working Papers*, vol. 27, no. 6, pp. 697–711, 2011.

Hindawi
Complexity
Volume 2021, Article ID 9927397, 10 pages
https://doi.org/10.1155/2021/9927397

*Research Article*

# Simulating the Sustainability of Xiong'an New Area Undertaking the Industrial Transfer from Beijing

**Fangqu Niu** [1,2]

[1] *Key Laboratory of Regional Sustainable Development Modeling, Institute of Geographic Sciences and Natural Resources Research, Chinese Academy of Sciences, Beijing 100101, China*
[2] *Collaborative Innovation Center for Geopolitical Setting of Southwest China and Borderland Development, Yunnan Normal University, Kunming 650500, China*

Correspondence should be addressed to Fangqu Niu; niufq@lreis.ac.cn

Received 10 March 2021; Revised 15 April 2021; Accepted 3 May 2021; Published 17 May 2021

Academic Editor: Jing-Hu Pan



The national new area Xiong'an has been established to take over the noncapital functions of Beijing in China. In light of local resources and environmental constraints, it is important to clarify the mode of industrial transfer for Xiong'an new area (XNA) to achieve the developmental goals. This study simulates and analyzes the speed of industrial transfer and the capacities of XNA in light of resource and environmental constraints. The results show that, to just realize modernization in 2035, the transfer rate of the secondary industry and tertiary industry is annually 3.7 billion yuan and 6.4 billion yuan, respectively. But at this speed, the atmospheric environment will be overloaded by 2028. The minimum and maximum transfer rate to realize modernization in 2035 without overloading resources and environment are also specified as well as the total size of economy and population. The results indicate that land for the construction of XNA is rich, but local water is not sufficient to support industrial growth. The atmospheric environment of XNA is also an important limiting factor. It is thus necessary to reduce air pollution and increase the population and industrial scale of the area by increasing the transfer rate of the tertiary industry.

## 1. Introduction

The population and scale of large cities are constantly increasing with the economic development and urbanization. Urban space is rapidly expanding outward, leading to a series of problems, such as improper land use, traffic congestion, and environmental pollution [1]. In this case, new areas emerge and break the sprawl in the city [2–4]. These peripheral areas take on part of the functions of the central city and have frequent and intensive contact with the entire city at the same time [5]. The construction of new areas is usually carried out with a clear direction for development and special supporting policies and has become an important means for the government to promote regional economic development.

In April 2017, the Central Committee of the Communist Party of China proposed establishing the Xiong'an new area (XNA), Hebei Province. The establishment of the XNA is an important measure to implement the strategy of the coordinated development of the Beijing-Tianjin-Hebei region and to move over the noncore functions of Beijing as the capital. The aim is to convert XNA into a green, low-carbon, competitive, and influential modern city by 2035. On the contrary, China's rapid economic growth in the past 40 years, since the reform and opening-up in 1978, has led to a series of serious imbalances and disharmony [6]. They are prominently manifested in the excessive consumption of resources and increased environmental pollution. The rapid expansion of the capital, Beijing, as one of China's megacities, has led to congestion and pollution. The establishment of the XNA is highly targeted, and it is necessary to draw lessons from the past in order to maintain high-quality development.

Therefore, in light of limited resources and environmental constraints, the government and researchers are faced with the following questions to achieve the development goals: what kind of mode of industrial transfer should be chosen, what should the expected speed of industrial transfer be, and what population and economic scale should be adopted for XNA?

## 2. Literature Review

Research on new areas can be traced back to the garden city proposed by the British urbanologist Ebenezer Howard at the end of the 19th century [7] to resettle the population outside big cities, set up hospitals, industries, and cultural, recreation, and commercial centers, and form a new and relatively independent society. He advocated for the establishment of new cities to solve urban problems, and this has had an important impact on the subsequent construction of satellite cities and research on urban function dispersion theory. Since the 1990s, with the emergence of suburbanization and counterurbanization, low-carbon space [8], fringe cities [9], and smart growth [10] have become major issues of research on new areas, and more attention has been paid to the emergence of urban problems and coordinated regional development in the context of urban sprawl and expansion. The new area is considered an important carrier of urban spatial growth that can share functions of the central city, thus providing an important means to redistribute urban population and realize industrial transfer [11]. The focus of the research has gradually shifted from the social and economic needs of the new area to its ecological needs.

In general, different from Western countries, where new areas are mainly decentralized, China's new areas are planned and designed at the national level under the leadership of the government, with the aim of developing them into poles of regional economic growth [12]. Therefore, the construction and development of China's new areas have distinctive characteristics. Research on new areas in China began in the 1990s. In this period, the scale of cities expanded rapidly, and "big city disease" became more and more serious. This made it increasingly important to study the distribution of urban functions. Studies on the relevant functions have been carried out from the two perspectives of population and industry. Xie [13] analyzed the plane and vertical modes of population evacuation in megacities, which caused researchers to attend to population evacuation. Taking Shanghai and Changsha as examples, Shi and Liu [14] and Xu [15] analyzed the characteristics of the urban industrial layout and proposed measures to optimize the industrial distribution in downtown areas. Meng [16] studied the layout of the commercial medical industry in Beijing with the help of GIS technology and proposed that medical resources in Dongcheng and Xicheng districts are distributed externally to reduce congestion. On this basis, other scholars have claimed that land price, labor cost, transportation factors, and government guidance are the main driving forces for the migration of the population and industry to the outskirts of cities [17–20]. They have also proposed that industrial relocation can drive population relocation [21, 22]. With the successive establishment of new areas at the national and regional levels, the academic circle has begun conducting more and more empirical studies on their planning and construction.

The research has focused on three perspectives: (1) Strategic positioning and functions of new areas. New areas at different scales often have diversified functional positioning, including regional open doors, economic growth poles, and inland open economic highlands [23, 24]. (2) Spatial pattern and management system. The research has focused on the location selection and spatial pattern of the new area [25–28], and scholars have explored management systems for it [29]. (3) Relations between new and old urban areas. This is mainly the problem of industrial transfer and production-city integration in old and new urban areas [30, 31]. The literature has not only analyzed the developmental path and problems of individual new areas [32] but also compared the functions and advantages of several new areas [33] to understand the relationships of heterogeneity and coopetition relationship among them. The aim is to determine the driving mechanisms for the development of different new areas and their effects on the regional economy.

It could be concluded that few dynamic simulations and predictions of the impact of industrial transfer on economic development are available concerning the resource and environment carrying capacity. This study simulates and analyzes the industrial transfer from Beijing to XNA and determines the appropriate speed of this transfer and the maximum scale of it under different modes. Through a comparison of scenarios, the appropriate plan is determined to provide decision-making support for realizing the sustainable development of XNA. This study contributes to work on new areas in terms of models of industrial undertaking. It is also a response to the framework document of "Future Earth" initiative.

## 3. Study Area and Data

Based on the specific needs for economic and social development at various stages, China (as of March 2019) has established 19 state-level new areas, 7 special economic zones, and 12 free trade zones since the establishment of the Shenzhen special economic zone in 1980. Although various zones have different names, they are all subject to special policies and thus are collectively called "new areas." Of them, state-level new areas are established with the approval of the State Council of China. Based on the relevant administrative regions and special functional zones, they undertake the major tasks of national development and the reform and opening-up strategy [34]. They have the general characteristics of new space, new function, and new mode, different from those of the old city. This highlights the strategic direction of the country and the unique characteristics of "strategy and leading" [35]. The spatiotemporal distribution of state-level new areas is shown in Figure 1.

XNA is located in Hebei Province and includes the three counties of Xiongxian, Rongcheng, and Anxin as well as part of the surrounding areas (Figure 2). It covers an area of 1770 $km^2$. According to the Hebei Economic Yearbook 2018, the population of the area in 2017 was 1.109 million and its GDP was 2.77 billion USD. The multiyear average volume of water

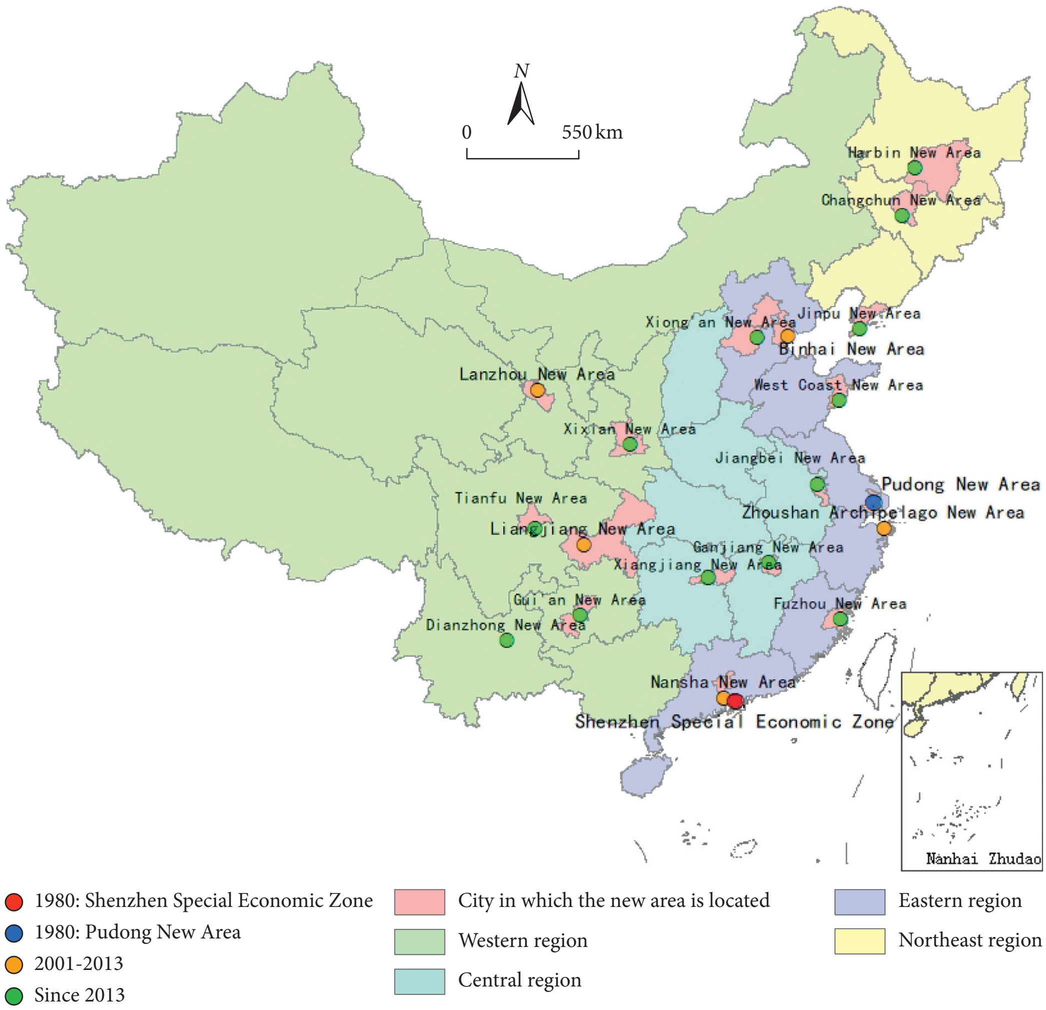



FIGURE 1: Spatiotemporal distribution of China's new area.

resources of XNA was 173 million cubic meters [36]. According to its developmental location, it will undertake the capital's emerging industries as well as its science and education industries and establish an ecologically livable and innovative leading zone. The region covers approximately 2000 km$^2$. The long-term plan is for the area to have a population of 2 to 2.5 million.

The data were drawn from various statistical yearbooks, including Hebei Economic Yearbook (2018), Beijing Statistical Yearbook (2018), and Statistical Yearbook of Urban Construction in China (2017). The data source for each index is explained below in the corresponding section.

## 4. Methodology

In the process of regional socioeconomic development, the interaction of such factors as the industry, population, resources, and the environment forms a complex dynamic system. Changes in each factor affect other factors. The system dynamics model is suitable for simulating the process of change in nonlinear systems [37, 38] and has been widely used in simulations of the relationship between the social economy and the resource environment [7, 39]. This paper constructs a system dynamics model called the Xiong'an Model (XAM) to simulate the industrial structure and scale, population scale, changes in urbanization as well as its impact on the water and atmospheric environment, and the consumption of land and water resources in XNA during the industrial transfer from Beijing. In this way, XAM helps determine suitable speeds for the industrial transfer and the upper limits of scale under different industrial transfer modes to achieve the goals set out for it by the government.

XAM includes an industrial module, a population module, an urbanization module, an environmental capacity module, a land demand module, and a water demand module. Using the available data and planning targets for XNA, with 2017 as the base year, a year-by-year simulation of XNA is conducted until 2035. Vensim DSS was used as a development platform.

*4.1. Industry and Population Modules.* People go where the industry is, because of which the population and industry modules considered here are intimately linked. The transfer

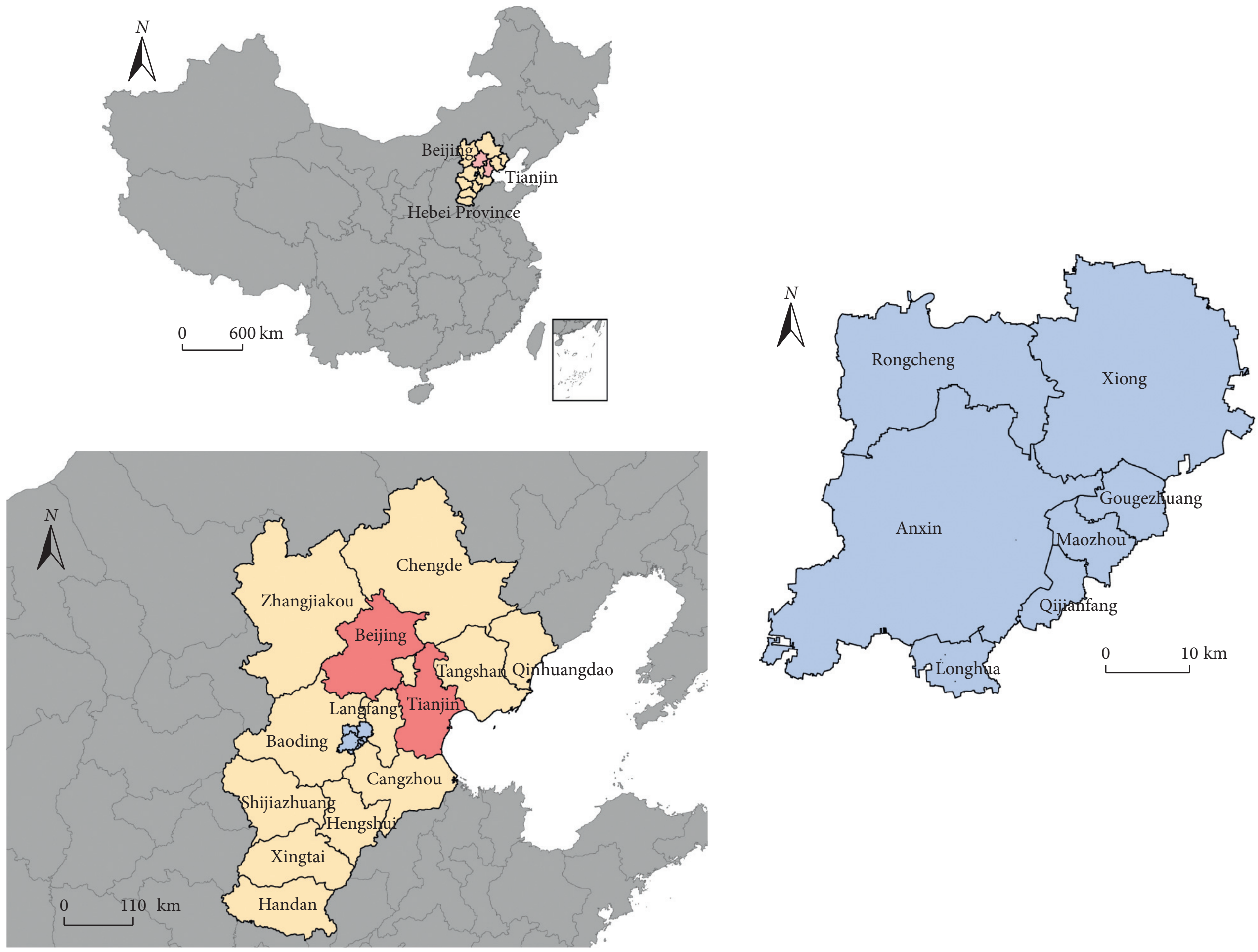



Figure 2: Study area: Xiong'an, China.

of Beijing's industries to XNA will inevitably be accompanied by population migration. We calculate the population corresponding to the transferred industries based on the number of employees in each.

By the end of 2017, Beijing had a permanent population of 21.71 million, with 12.47 million employees, including 488,000 in the primary industry, 1.93 million in the secondary industry, and 10.05 million in the tertiary industry. The ratio of the permanent population to the number of employees was 1.74 : 1, so the populations corresponding to the above three industries were thus 0.85 million, 3.35 million, and 17.49 million people, respectively. The GDP values of Beijing's primary, secondary, and tertiary industries in 2017 were 12.04 billion, 532.68 billion, and 2256.78 billion yuan, respectively, and their per capita GDP values were 14,200 yuan/person, 158,800 yuan/person, and 129,900 yuan/person, respectively. It is reasonable to assume that the same industry corresponds to the same number of employees under the same technical level no matter wherever they are located.

Based on the scale of the transferred industries and per capita GDP, the population corresponding to the transferred industries can be calculated. The migrant population plus the local population is the total population:

$$\mathrm{pop}_n = \frac{\mathrm{GDP}_s}{15.88} + \frac{\mathrm{GDP}_t}{12.90}, \qquad \mathrm{pop} = \mathrm{pop}_l + \mathrm{pop}_n. \tag{1}$$

In the above, $pop_n$ is the total population corresponding to the transferred industries, $GDP_s$ and $GDP_t$ are the transfer scales of the second and tertiary industries, respectively, and $pop$ and $pop_l$ are the total population and local population, respectively. The population of XNA will continue to grow with industrial transfer. The variables that we need to determine are the speed of industrial transfer, its scale, and the corresponding total population.

*4.2. Urbanization Module.* According to the Hebei Economic Yearbook 2018, the rate of urbanization of XNA in 2017 was 45.23% and its urban population was 501,600. The population transferred for the secondary and tertiary industries was regarded as the urban population, and that corresponding to the local original industry was added to obtain the total urban population, as shown in equation (2).

$$\mathrm{UrbanRatio} = \frac{\mathrm{pop}_{\mathrm{localUrban}} + \mathrm{pop}_n}{\mathrm{ttlPop}}. \tag{2}$$

In the above, *UrbanRatio* is the urbanization rate, $pop_{localUrban}$ is the orginal local urban population, $pop_n$ is the urban population corresponding to the transferred industry, and *ttlPop* is the total population. The population parameter is obtained from the output of the population module.

*4.3. Water Demand Module.* According to the Beijing Statistical Yearbook 2018, the city's per capita GDP is 129,000 yuan. The total water consumption was 3.95 billion cubic meters, of which the first industry used 510 million cubic meters at an intensity of water use of 423.24 $m^3$/10,000 yuan. The secondary industry used 350 million cubic meters at an intensity of 6.59 $m^3$/10,000 yuan; domestic water consumption was 1.83 billion $m^3$ (including water used by the tertiary industries) at an intensity of 84.31 $m^3$ per person; in addition, the ecological water consumption was 1.26 billion $m^3$, accounting for 31.9% of the total. Given the same technical level, we assume that the intensity of industrial water use remains the same.

A region with water consumption exceeding 70% of the total local water resources is usually regarded as suffering from severe water shortage [40, 41]. Therefore, we set 70% of the water resources as the upper limit of water use. The total available water resource in XNA was 156 million cubic meters, and the usable volume was 109 million $m^3$. If the total amount of water used for production and domestic use as well as ecological water exceeded this value, a serious water shortage was considered to have occurred. The total water consumption was calculated as

$$\begin{aligned} w = \mathrm{GDP} * r_{\mathrm{pri}} * 423.24 + \mathrm{GDP} * r_{\mathrm{sec}} * 0.026 \\ + \mathrm{pop} * 0.006 + w_{\mathrm{eco}}, \end{aligned} \tag{3}$$

where $w$ is the total water consumption, $r_{pri}$ and $r_{sec}$ are the ratios of the primary and secondary industries, respectively, *pop* is the population, and $w_{\mathrm{eco}}$ is the ecological water use. The values 423.24, 0.026, and 0.006 represent the intensities of use of the primary and secondary industries and those of domestic water use, respectively.

*4.4. Water Environment Capacity Module.* The quality of the water environment is affected by many factors, and it is usually not necessary to consider all of them. The commonly used COD measure was used to characterize the quality of the water environment in this study. According to China's surface water environmental quality standards (COD capacities of Class I and Class II water ≤ 15 mg/L) [42] and the total available water resources in XNA, the total capacity of its water environment was 2,340 tons.

According to public information [43], in 2017, COD emissions by Beijing's primary industry amounted to 8799 tons. Combined with the GDP of the primary industry, its intensity of COD emissions was calculated at 73.08 tons/100 million yuan. The COD emission of the secondary industry was 2232 tons, and the intensity of its emissions was 0.42 tons/100 million yuan. In urban areas, 70,312 tons of COD was emitted at an intensity of 32.4 tons per 10,000 people. Similarly, we assume that the intensities of COD emissions would not change given the same technical conditions.

Because the treated sewage usually enters natural water, the total amount of COD entering natural water could not exceed the capacity of the water environment. COD emissions from production and living are shown as

$$\mathrm{ttlCOD} = \mathrm{GD}P_{\mathrm{pri}} * 73.08 + \mathrm{GD}P_{\mathrm{sec}} * 0.42 + \mathrm{pop} * 32.40. \tag{4}$$

*ttlCOD* is the total COD emissions from production and living, $GDP_{pri}$ and $GDP_{sec}$ are GDP of the primary and secondary industries, respectively, and *pop* is the total population. To maintain the quality of the water environment, the total COD emission *ttlCOD* cannot exceed the total capacity of the water environment.

*4.5. Land Demand Module.* Because China has not described the detailed industrial land use of each city, this study refers to the national industrial land intensity as a measure (discussed later). According to the Statistical Yearbook of Urban Construction in China 2017, its urban population in that year was 813.47 million, and the area of land used for urban construction was 55,155.47 $km^2$, including 11,083.70 $km^2$ of industrial land, 5508 $km^2$ occupied by commercial services, logistics and warehousing, and other tertiary industries, 16,979.27 $km^2$ of residential land, and 21,584.5 $km^2$ of other lands (land for public management and services, road and traffic facilities, general facilities, and green squares, accounting for 39.1%). In 2017, China's industrial and tertiary industry GDP values were 27,999.69 billion yuan and 42,703.15 billion yuan, respectively. The intensities of use of various types of construction land were obtained from this: the intensity of use of industrial land was 0.0396 $km^2$/100 million yuan, that of land for tertiary industry was 0.0129 $km^2$/100 million yuan, and that of residential land was 0.2087 $km^2$/10,000 people. Based on this, the increase in the land used for construction corresponding to the industry and population growth in XNA was calculated:

$$\begin{aligned} \mathrm{Land}_c = \mathrm{GD}P_s * 0.0396 + \mathrm{GD}P_t * 0.0129 \\ + \mathrm{Urba}n_{\mathrm{pop}} * 0.2087 + \mathrm{lan}d_{\mathrm{other}}, \end{aligned} \tag{5}$$

where $Land_c$ is the total land used for construction, $GDP_S$ and $GDP_t$ are the GDP values of the secondary and tertiary industries, respectively, $Urban_{pop}$ is the urban population, and $land_{other}$ represents other types of land. Residential and industrial land continues to grow, but the total area cannot exceed the total planned area for construction. According to the planning outline of XNA, the intensity of its long-term development is controlled at 30% of land for construction, that is, 530 $km^2$.

*4.6. Atmospheric Environmental Capacity Module.* The increase in industry and population usually leads to an increase in the volume of exhaust emissions. The atmospheric environmental capacity module calculates the social and economic carrying capacities of XNA based on the intensity of pollution emissions of its industries and population and its atmospheric environmental capacity. With the development of the new area, the use of motor vehicles will continue to increase, and emissions from them will become an important source of air pollution. The main pollutants from motor vehicles are nitrogen oxides [43]. Considering emissions from production, living, and motor vehicle emissions, we use nitrogen oxide content to evaluate the environmental atmospheric quality.

*4.6.1. Atmospheric Environmental Capacity.* The *A*-value method was used to measure the atmospheric environmental capacity. Based on the basic assumptions of the box model, it divides a given area into several functional areas and calculates the total allowable pollutant discharge according to the total area, the area of each functional partition, and the local total quantity control coefficient $A$ [44]. The formula is as follows:

$$Q_{ak} = \sum_{i=1}^{n} Q_{aki} = \sum_{i=1}^{n} A\left(C_{ki} - C_0\right)\frac{S_i}{\sqrt{S}}. \quad (6)$$

$Q_{ak}$ is the total annual allowable discharge of $k$-th pollutant in the total control area, that is, the atmospheric environmental capacity, $Q_{aki}$ is the annual allowable discharge of the $k$-th pollutant in the $i$-th functional zone, which is also the ideal air capacity, $C_{ki}$ is the annual average concentration limit of the $k$-th pollutant in the $i$-th functional area, $C_0$ is the background concentration value of the area, $S$ is the total area, $S_i$ is the city's $i$-th functional area, and $A$ is the control coefficient, which is mainly determined by the volume of local ventilation.

According to the above formula, the calculation of atmospheric capacity requires determining the local total volume control coefficient $A$. According to the Chinese national "Technical Methods for Establishing Local Air Pollutant Emission Standards (GB/T3840-91)," the cumulative probability of $A$ is 0.9, and the $A$-value of each region is unified according to $A = A_{\min} + 0.1 \times (A_{\max} - A_{\min})$. XNA is in Hebei Province, the $A$-value of which ranges from $4.2 \times 10^4$ to $5.6 \times 10^4$ ($km^2/a$). Based on this, the $A$-value of the XNA was determined to be $4.34 \times 10^4\ km^2/a$.

Using the total control coefficient, background concentration, and the area of each functional area, the total atmospheric environment capacity of the XNA was calculated to be 36,500 tons.

*4.6.2. Air Pollution Emission.* According to the 2017 Beijing Environmental Statistics Annual Report [43], Beijing emitted 144,514 tons of nitrogen oxides, mainly from industrial production, everyday living, and motor vehicles. In 2017, the volume of emission of nitrogen oxides from the secondary industry in Beijing was 15,405 tons. Combined with the GDP, the intensity of emission of pollutants by the secondary industry was calculated to be 12.784 tons/billion yuan. Urban household nitrogen oxide emissions were 7510 tons, and their intensity of emission was 3.46 tons/10,000 people. Nitrogen oxide emissions from motor vehicles were 121,564 tons at an intensity of 205.73 tons/10,000 vehicles. The total nitrogen oxide emissions from production, living, and motor vehicles should not have exceeded the atmospheric capacity. The total nitrogen oxide emissions (*ttlNO*) were calculated as

$$\text{ttlNO} = \text{GDP}_{\text{sec}} * 127.84 + Pop * 3.46 + \text{Pop} * 0.27 * 205.73. \quad (7)$$

In the above, $GDP_{sec}$ is the GDP of the secondary industry, *Pop* is its population, and 0.27 is the car ownership per capita. The values 127.84, 3.46, and 205.73 are, respectively, the intensities of industrial emission, emissions by the population, and those by motor vehicles.

*4.7. Construction of XAM Model.* The above modules were integrated to form the XAM, as shown in Figure 3. With the gradual transfer of industries, the population, GDP, and urbanization rate of XNA will change accordingly. The change in the industrial population will lead to changes in the total land for construction and water consumption as well as increases in COD emissions and nitrogen oxide emissions. This will affect the water environment and atmospheric environment. The process is a systematic one of element linkage.

## 5. Scenario Setting and Results of Prediction

By combining the developmental positioning of XNA and China's goal of modernization by 2035, this study set three development scenarios to help determine the appropriate speeds and scales of transfer under different industrial transfer modes. With reference to the situation of some developed countries, that is, US, UK, Korea, Japan, Germany, and so on [45, 46], modernization requires that the tertiary industry account for more than 60% of all industry, the rate of urbanization be more than 75%, and a high income (per capita GDP is higher than US$12,235, equivalent to RMB 81,974.5, at the exchange rate in 2017: 1 US dollar is approximately RMB 6.7).

(1) *Scenario 1.* XNA will be just modernized in 2035; that is, the tertiary industry will just account for 60% of the total, the rate of urbanization will just reach 75%, and the per capita GDP will be higher than 12,235 yuan.

The results of the XAM simulation (Table 1) show that, to achieve the above development goals, the rate of transfer of the secondary industry needed is 3.7 billion yuan per year and that of the tertiary industry is 6.4 billion yuan per year. At this speed, the tertiary industry will account for 60% of the total in 2035, the urbanization rate will reach 75%, and the per capita GDP will reach 83,900 yuan. By then, the GDP and population of the area will be 202.8 billion yuan and 2.416 million people, respectively, so the population will not exceed thelimitof the long-term development plan.

Under this development scenario, the construction land will not exceed the limit but the atmospheric environment will be

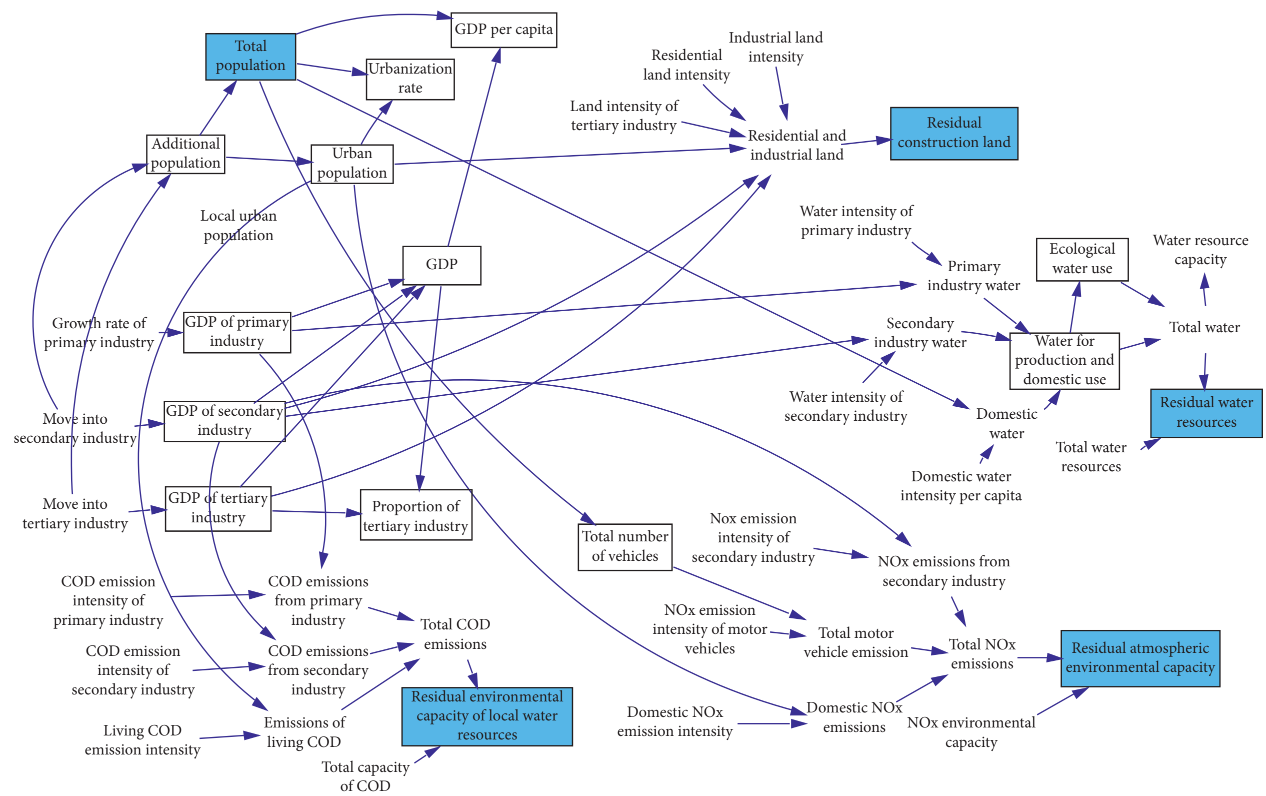



Figure 3: Xiong'an model (XAM).

Table 1: Simulation results by scenario.

| Industrial transfer scenario | Scenario 1 | Scenario 2 | Scenario 3 |
|---|---|---|---|
| Transfer speed of second industry (billion yuan/year) | 3.7 | 1.5 | 1.9 |
| Transfer speed of tertiary industry (billion yuan/year) | 6.4 | 8.6 | 10.8 |
| Water resources | Overloaded | Overloaded | Overloaded |
| Water environment | Overloaded | Overloaded | Overloaded |
| Atmospheric environment | Overload in 2028 | Not overload | Not overloaded |
| Land resources | Not overloaded | Not overload | Not overloaded |
| GDP (billion yuan) | 202.8 | 202.8 | 249.6 |
| Population (10,000 persons) | 241.6 | 247.2 | 282.2 (beyond planned population) |
| Per capita GDP (million yuan/person) | 839 | 820 | 884 |
| Ratio of tertiary industry | 60% | 80% | 81% |
| Urbanization rate | 75% | 75% | 78% |

*Note.* Water resources and the water environment carrying status are calculated based on local water resources (excluding inbound water resources).

overloaded by 2028. In addition, water resources were already overloaded in 2017, indicating that the original local water resources cannot support the local social economy. Therefore, the sustainable development of XNA requires the supply of inbound water. According to the 2017 Hebei Water Resources Bulletin, the inbound volume of water of Hebei Province was 3.545 billion $m^3$. Owing to the inbound water supply, the following simulation no longer considers restrictions of water resources.

(2) *Scenario 2*. We achieve the goal of modernization by 2035 while ensuring that resources and the environment are not exhausted. We calculate the minimum speed of transfer as well as the corresponding scale economy and population.

The atmospheric capacity in scenario 1 was overloaded by 2028. Since the tertiary industry has less pollution emissions, this scenario requires increasing the speed of transfer of the tertiary industry until modernization is achieved in 2035 without exhausting resources and the environment. The results shown in Table 1 show that the minimum required transfer rate of the industry was 1.5 billion yuan per year for the secondary industry and 8.6 billion yuan for the tertiary industry. According to this rate, the ratio of the tertiary industry to the total will reach 80% in 2035, the rate of urbanization will reach 75%, and the per capita GDP will be 82,000 yuan. The total GDP and population will be 202.8 billion yuan and 2.472 million, respectively. The population will not exceed the limit on the population in the long-term plan. Compared with scenario 1, the per capita GDP has declined, because the tertiary industry is more labor intensive compared with the manufacturing.

(3) *Scenario 3*. On the basis of scenario 2, we expand the growth rate to find the maximum transfer rate that can be set as well as the corresponding economic and population-related scale without exhausting resources and the environment.

The results in Table 1 show that while ensuring that the resources and environment were not exhausted, the maximum transfer rates were 1.9 billion yuan per year for secondary production and 10.8 billion yuan per year for tertiary production. At this rate, the GDP and population in 2035 will be 249.6 billion yuan and 2.822 million, respectively, and the per capita GDP will reach 88400 yuan. The total population will exceed the planned population (2.5 million). Therefore, although this transfer model can achieve the largest population and economic scale without exhausting resources and the environment, it is limited by the planned population. These results provide an important reference to determine whether the planned population needs to be adjusted according to the industrial transfer mode.

In summary, the planned construction land in XNA is relatively rich in resources and can meet the needs of future development. The local water resources are relatively scarce and are insufficient to support the demand for water for industrial growth. However, inbound water resources are available. Development will be limited by the supply of inbound water resources, and the atmospheric environment will become an important limiting factor. To increase the carrying capacity of the atmosphere of XNA, it is necessary to increase the transfer ratio of tertiary industry. However, the tertiary industry has a greater demand for human resources. Increasing the volume of transfer of the tertiary industry also increases the total population. A larger scale of the economy and population transfer can be pursued while maintaining a healthy ecosystem but will be limited by the planned population. The results here can provide an important basis for industrial transfer and population planning.

## 6. Discussion

*6.1. Validity and Usefulness of the Model.* There are many other factors affecting the accuracy of the model. Among them, the industrial scale is an important factor that affects the effects of the industry on resources and the environment. In particular, the relation between industry and land use is complex, and the growth in the industrial scale is not linear with the demand for land use. Technological progress is another important factor that affects the intensity of resource consumption and pollution emission. However, the impact of technology is difficult to quantify, and in a short time, it is reasonable to assume that the technical level remains the same.

Even with all these possible uncertainties, the XAM can still help decision-making in this context. The results simulated by the XAM do not exactly represent the situation in the future but provide a tool to compare different scenarios and provide a basis for decisions on the developmental planning of new areas. The relative values of the model's predictions are more meaningful than their absolute values.

*6.2. Model Optimization and Improvement.* The industrial transfer is not determined unilaterally. The transfer of industries in Beijing will be restricted by its industrial planning. According to the "Outline of the Beijing-Tianjin-Hebei Coordinated Development Plan," Beijing's development is positioned as a "political center, a cultural center, a center for international exchanges, and a center for scientific and technological innovation." What needs to be addressed is the "noncore function of the capital," that is, the industries needing to be transferred. Therefore, the accurate differentiation of noncore functions in combination with Beijing's developmental positioning on the basis of further industrial segmentation should be examined in future work. In addition, this model focused only on the three common elements of water, soil, and the atmosphere as resource-related and environmental factors. In the future, more elements can be included, such as biodiversity and the soil environment.

*6.3. Policy Implications.* Economic growth normally follows an exponential law. This research focused on the Beijing-XNA industrial transfer by setting a fixed speed of industrial transfer. In practice, the speed of industrial transfer is usually not constant. Compared with the speed of industrial transfer, the upper limit of the scale of transfer under various transfer modes is insurmountable. It is the ceiling set by local resources and the environment for industrial development and provides an important reference for decision-making.

## 7. Conclusions

The XNA was set up to relieve Beijing of its noncapital core functions and promote the coordinated development of the Beijing-Tianjin-Hebei region. The new area's undertaking in terms of industry and population is limited by its resource and environmental carrying capacity. This study simulated and analyzed the speed and maximum scale of industrial transfer to the new area under different industrial transfer modes. The results indicate that the carrying capacity of original local water resources in XNA has been overloaded, and it currently relies on incoming water resources to support its social economy. In addition to water resources, the primary limiting factor is the atmospheric environment, whereas the land is relatively rich in resources. To improve the economic scale of XNA, it is necessary to strengthen the transfer of the tertiary industry, which is labor intensive and will lead to greater population transfer and increased motor vehicle ownership. Therefore, to reduce air pollution, it is necessary to optimize motor vehicle control policies. The results specifically showed the following:

(i) To just achieve the goal of modernization by 2035 (the tertiary industry will just account for 60% of the total and the rate of urbanization will just reach 75%), the transfer rates of the secondary and the tertiary industries should be 3.7 billion yuan/year and 6.4 billion yuan/year, respectively; then, the GDP and population in 2035 will be 202.8 billion yuan and 2.416 million people, respectively. In this mode, the atmospheric environment will be overloaded by 2028.

(ii) To achieve modernization by 2035 and avoid the overloading of the atmospheric environment, the minimum speeds of transfer of secondary and tertiary industries needed to be 1.5 billion yuan per year and 8.6 billion yuan per year, respectively, and the total size of the economy and the population in 2035 would then be 202.8 billion yuan and 2.472 million people, respectively.

(iii) To determine the highest speed of industrial transfer under the condition that resources and the environment are not overloaded, the rates of transfer of the secondary and tertiary industries should be 1.9 billion yuan per year and 10.8 billion yuan per year, respectively. Then, the economic scale and population will reach 249.6 billion yuan and 2.822 million people in 2035, respectively. The population here exceeds the planned population. This shows that although this transfer mode can maximize the economic growth of XNA, it will be limited by the planned population.

Finally, we highlight several aspects of this research which need to be improved. First, this study analyzed industrial transfer based on three major industry classifications, but differences occur in the resource consumption and intensity of pollution emission of various sectors within the same industry. In the future, it is necessary to further divide these industries. Second, this paper focused on resource-related and environmental constraints of XNA without analyzing the impact of industrial transfer on Beijing. Analyzing the impact of industrial transfer on Beijing's economy and population in combination with its industrial development and controlled planning is another area that needs to be explored.

## Data Availability

The data used to support the results of this study come from various publicly published statistical yearbooks, which are described in detail in the article. A request for access to these data can be made to the corresponding author.

## Conflicts of Interest

The authors declare that there are no conflicts of interest.

## Acknowledgments

We are grateful to the two anonymous referees for their constructive comments and suggestions on the earlier versions of the paper. This study was supported by the Strategic Priority Research Program of the Chinese Academy of Sciences (Grant no. XDA19040401) and the National Natural Science Foundation of China (Grant no. 42071153).

## References

[1] L. Zhou, X. Dang, Q. Sun, and S. Wang, "Multi-scenario simulation of urban land change in shanghai by random forest and CA-Markov model," *Sustainable Cities and Society*, vol. 55, pp. 1–10, 2020.

[2] S. Su, Z. Jiang, Q. Zhang, and Y. Zhang, "Transformation of agricultural landscapes under rapid urbanization: a threat to sustainability in hang-jia-hu region, China," *Applied Geography*, vol. 31, no. 2, pp. 439–449, 2011.

[3] T. N. Chase, R. A. Pielke, T. G. F. Kittel, R. R. Nemani, and S. W. Running, "Simulated impacts of historical land cover changes on global climate in northern winter," *Climate Dynamics*, vol. 16, no. 2-3, pp. 93–105, 2000.

[4] C. Weber and A. Puissant, "Urbanization pressure and modeling of urban growth: example of the tunis metropolitan area," *Remote Sensing of Environment*, vol. 86, no. 3, pp. 341–352, 2003.

[5] S. Yao, P. Zhang, C. Yu, G. Li, and C. Wang, "The theory and practice of new urbanization in china," *Scientia Geographica Sinica*, vol. 34, no. 6, pp. 641–647, 2014, in Chinese.

[6] F. Niu, X. Yang, and X. Zhang, "Application of an evaluation method of resource and environment carrying capacity in the adjustment of industrial structure in tibet," *Journal of Geographical Sciences*, vol. 30, no. 2, pp. 319–332, 2020.

[7] E. Pasqualini and M. Witkowski, "System dynamics approach for critical infrastructure and decision support: a model for a potable water system," *Human Reproduction*, vol. 25, no. 3, pp. 799–804, 2010.

[8] G.-E. Shim, S.-M. Rhee, K.-H. Ahn, and S.-B. Chung, "The relationship between the characteristics of transportation energy consumption and urban form," *The Annals of Regional Science*, vol. 40, no. 2, pp. 351–367, 2006.

[9] V. Henderson and A. Mitra, "The new urban landscape: developers and edge cities," *Regional Science and Urban Economics*, vol. 26, no. 6, pp. 613–643, 1996.

[10] R. Biles, "Thinking the unthinkable about our cities," *Journal of Urban History*, vol. 25, no. 1, pp. 57–74, 1998.

[11] C. Fang, S. Wang, and Y. Wang, "Current situation and countermeasures of Chinese low-carbon eco-metro area development," *Geographical Research*, vol. 35, no. 9, pp. 1601–1614, 2016, in Chinese.

[12] Y. Cao, "The hidden worries of the construction of state-level new areas," *Decision-Making*, vol. 12, pp. 56–58, 2014, in Chinese.

[13] J. Xie, "Population decentralization of the megalopolis in China," *Urban Problems*, vol. 3, no. 12–15, 1992, in Chinese.

[14] Y. Shi and D. Liu, "Excessive urbanization and rejuvenation in the metropolitan city center in shanghai, China," *Journal of Urban Planning and Development*, vol. 144, no. 4, pp. 1–11, 2018.

[15] W. Xu, Z. Zeng, Z. Xu et al., "Public health benefits of optimizing urban industrial land layout - the case of Changsha, China," *Environmental Pollution*, vol. 263, pp. 1–10, 2020.

[16] H. Meng, "Study on the present situation, problems and solutions of medical decentralization in central city of Beijing," *Capital University of Economics and Business*, vol. 3, pp. 19–27, 2014, in Chinese.

[17] P. Zhao and S. Li, "Suburbanization, land use of TOD and lifestyle mobility in the suburbs: an examination of passengers' choice to live, shop and entertain in the metro station areas of Beijing," *Journal of Transport & Land Use*, vol. 11, no. 1, pp. 195–215, 2018.

[18] C. Zhao, J. Jensen, and B. Zhan, "A comparison of urban growth and their influencing factors of two border cities: laredo in the US and nuevo laredo in mexico," *Applied Geography*, vol. 79, pp. 223–234, 2017.

[19] X. Lin, Y. Wang, S. Wang, and D. Wang, "Spatial differences and driving forces of land urbanization in China," *Journal of Geographical Sciences*, vol. 25, no. 5, pp. 545–558, 2015.

[20] J. Liu, J. Zhan, and X. Deng, "Spatio-temporal patterns and driving forces of urban land expansion in China during the economic reform era," *AMBIO: A Journal of the Human Environment*, vol. 34, no. 6, pp. 450–455, 2005.

[21] K.-y. Wu and H. Zhang, "Land use dynamics, built-up land expansion patterns, and driving forces analysis of the fast-growing hangzhou metropolitan area, eastern china (1978-2008)," *Applied Geography*, vol. 34, pp. 137–145, 2012.

[22] X. Shen and L. J. C. Ma, "Privatization of rural industry and de facto urbanization from below in southern Jiangsu, China," *Geoforum*, vol. 36, no. 6, pp. 761–777, 2005.

[23] Y. Li and D. Jia, "The spatial-temporal distribution, strategic positioning and policy characteristics of state-level new areas," *Journal of Beijing Administrative College*, vol. 3, no. 22–31, 2016, in Chinese.

[24] L. Li, "Development potential and future direction of national new areas–A GRNN approach," *Economic Geography*, vol. 2, pp. 92–99, 2015, in Chinese.

[25] W. Kuang, W. Chi, D. Lu, and Y. Dou, "A comparative analysis of megacity expansions in China and the U.S.: patterns, rates and driving forces," *Landscape and Urban Planning*, vol. 132, pp. 121–135, 2014.

[26] Q. Weng, "Land use change analysis in the zhujiang delta of china using satellite remote sensing, GIS and stochastic modelling," *Journal of Environmental Management*, vol. 64, no. 3, pp. 273–284, 2002.

[27] J. Luo and Y. H. D. Wei, "Modeling spatial variations of urban growth patterns in Chinese cities: the case of Nanjing," *Landscape and Urban Planning*, vol. 91, no. 2, pp. 51–64, 2009.

[28] J. Cheng and I. Masser, "Modelling urban growth patterns: a multiscale perspective," *Environment and Planning A: Economy and Space*, vol. 35, no. 4, pp. 679–704, 2003.

[29] J. N. Wang and C. P. Luo, "Management system of national new areas and domain status and strategic tropism," *Reform*, vol. 2012, no. 03, pp. 21–36, 2012, in Chinese.

[30] C. Liu, X. Li, and X. Hang, "Path toward city-industry integration in new urban development zones," *Urban Planning Forum*, vol. S1, pp. 104–109, 2012, in Chinese.

[31] Y. Ma and R. Xu, "Remote sensing monitoring and driving force analysis of urban expansion in Guangzhou City, China," *Habitat International*, vol. 34, no. 2, pp. 228–235, 2010.

[32] G. Meng and M. Liu, "Evaluation on the establishment of free trade zones in tianjin binhai new area," *Acta Geographica Sinica*, vol. 66, no. 2, pp. 223–234, 2011, in Chinese.

[33] R. Xing, J. Chen, and L. Tian, "Development heterogeneity and dynamic mechanism of state-level new area—a case study of Pudong new area of Shanghai and Jiangbei new area of Nanjing," *Resources and Environment in the Yangtze Basin*, vol. 25, no. 6, pp. 859–867, 2016, in Chinese.

[34] National Development and Reform Commission (Chief Editor), *Report on the Development of State-Level New Areas*, China Planning Press, Beijing, China, 2015, in Chinese.

[35] D. Chen and W. F. Kong, "Discussion on characteristics and problems of state-level new area," *Bulletin of the Chinese Academy of Sciences*, vol. 31, no. 1, pp. 118–125, 2016, in Chinese.

[36] R. Yang, B. Hou, F. Lu et al., "Preliminary analysis and promotion of water resources carrying capacity for Xiongan New Area," *Water Resources and Hydropower Engineering*, vol. 50, no. 10, pp. 1–9, 2019, in Chinese.

[37] J. A. Macedo, "Reference approach for policy optimization in system dynamics models," *System Dynamics Review*, vol. 5, no. 2, pp. 148–175, 1989.

[38] F. Castellacci, "Co-evolutionary growth: a system dynamics model," *Economic Modelling*, vol. 70, pp. 272–287, 2018.

[39] A. Macmillan, J. Connor, K. Witten, R. Kearns, D. Rees, and A. Woodward, "The societal costs and benefits of commuter bicycling: simulating the effects of specific policies using system dynamics modeling," *Environmental Health Perspectives*, vol. 122, no. 4, pp. 335–344, 2014.

[40] A. Y. Hoekstra, M. M. Mekonnen, A. K. Chapagain, R. E Mathews, and B. D Richter, "Global monthly water scarcity: blue water footprints versus blue water availability," *PLoS ONE*, vol. 7, no. 2, Article ID e32688, 2012.

[41] Y. Wada, L. P. H. van Beek, D. Viviroli, H. H. Dürr, R. Weingartner, and M. F. P. Bierkens, "Global monthly water stress: 2. Water demand and severity of water stress," *Water Resources Research*, vol. 47, no. 7, Article ID W07518, 2011.

[42] Ministry of Ecology and Environment of the People's Republic of China, *Environmental Quality Standards for Surface Water (GB 3838–2002)*, Ministry of Ecology and Environment of the People's Republic of China, Beijing, China, 2002, http://kjs.mee.gov.cn/hjbhbz/bzwb/shjbh/shjzlbz/200206/t20020601_66497.shtml in Chinese.

[43] Beijing Municipal Ecological Environment Bureau, *Beijing 2017 Environmental Statistics Annual Report*, Beijing Municipal Ecological Environment Bureau, Beijing, China, 2018, http://sthjj.beijing.gov.cn/bjhrb/xxgk/ywdt/zlkz/hjtj37/842547/index.html in Chinese.

[44] D. Xu, Y. Wang, and R. Zhu, "Atmospheric environmental capacity and urban atmospheric load in mainland China," *Science China Earth Sciences*, vol. 61, no. 1, pp. 33–46, 2018.

[45] World Bank. (2017). World Development Indicators: Structure of Output. http://wdi.worldbank.org/table/4.2. (Accessed in Jan., 2021).

[46] World Bank. (2019). World Development Indicators: Structure of Output, https://databank.worldbank.org/source/population-estimates-and-projections. (Accessed in March., 2021).