|The following publication Gu, X., Tang, X., Chen, T., & Liu, X. (2024). Predicting the network shift of large urban agglomerations in China using the deep-learning gravity model: A perspective of population migration. Cities, 145, 104680 is available at https://doi.org/10.1016/j.cities.2023.104680.

Predicting the network shift of large urban agglomerations in China using the deep-learning gravity model: A perspective of population migration

Abstract:

The demands of socioeconomic development frequently lead to large-scale population migration among cities. While complex network and population migration algorithms have been employed to evaluate this phenomenon, predicting the future shift of urban networks has remained challenging. In this study, we expend the conventional two-dimensional perception of urban structure, projecting geographic information of cities into a high-dimensional future dimension to forecast changes in the network structure with deep learning algorithms. Using the population migration data from 362 Chinese cities, we employed multivariate and non-linear layers to construct a deep learning model that exhibits good geographic and temporal generalization across major metropolitan regions in China, enabling us to forecast the urban network for the year 2025. The result shows that the urban network becomes more equitable and less concentrated in a few dominant cities. This shift suggests a more balanced distribution of resources, opportunities, and development across the urban agglomerations. Understanding the urban structure from the lens of future mobile networks offers deeper insight and perception of its future dimensional nature. By embracing this paradigm shift, we can retain knowledge about urban dynamics and pave the way for more effective urban management.

Keywords: Urban network, urban population migration, mobility prediction, deep learning model, urban agglomerations

1. Introduction

The spatial organization of regional cities becomes increasingly networked as the relationships between regional cities transform a vertical hierarchy of administration into a complex network (Sun & Hou, 2020). Existing studies over the past two decades have proved that a networked urban structure, characterized by the expansion of urban incomes, the dispersion of economic variables, and the frequency of inter-city exchanges, is beneficial for regional development (Bai et al., 2019; Cao et al., 2023). From a practical point of view, the study of urban networks can help planners conceptualize the spatial structure of urban agglomerations, find effective ways to improve the regional spatial layout, and optimize regional resource allocation and infrastructure construction (Burger & Meijers, 2016; Song et al., 2022). From a theoretical point of view, further refinement of urban network analysis and exploration of diverse perspectives and methods can provide valuable support for subsequent studies on urban structure.

Geographical and planning researchers have increasingly focused on urban networks, shifting their attention from attributed networks to multi-relational networks, particularly emphasizing large metropolitan areas as the primary research context (Cheng et al., 2013). An urban network is a collection of interconnected cities that exchange resources, compete for talent, and shift labor among them (Glaeser et al., 2016). Urban networks can be characterized morphologically as significant hubs or concentrations of activity within a given region and the physical and functional links that connect these nodes. In addition to this, networks within a large urban area can often be characterized by the number and diversity of activities carried out by transport or human flow networks (Cheng et al., 2013; De Goei et al., 2010).

Traditionally, studies on urban networks have primarily focused on one- and two-dimensional spaces. The analysis of the static spatial organization of cities frequently employed the earliest version of Christaller central place theory, which focused on vertical hierarchical relationships among cities (Getis & Getis, 1966). While civilizations comprise diverse mobile resources, such as information, capital, technology, and human resources, they rely on mobile spaces to connect with other cities in the region, leading to complex and varied interactions among cities. Static urban scales cannot adequately capture the intricate relationships among cities (Godfrey & Zhou, 1999; Palla et al., 2005), and regional spatial organization is also difficult to grasp. A dynamic "mobile space" perspective, which can better depict the external relations and network structure of cities, has started to take center stage in the study of regional spatial organization (De Goei et al., 2010; Kamarianakis & Prastacos, 2003). Although considerable research has been carried out on the urban network, most still need more than a two-dimensional exploration. They only describe and analyze the historical or current urban network but do not pay enough attention to predict the future changes in urban networks in a forward-looking manner.

Thanks to the diffusion of new algorithms, such as machine learning, this study could break the limitations of past-present dimensions. We choose inter-city population migration, often used to characterize urban networks as the main objective, introducing a future dimension to map the possible urban networks. Combining open-source geographic information and population migration datasets, this study uses deep neural networks to extract non-linear relationships between urban information and mobility elements to study the urban network shifts in 2025 in three of the essential large urban agglomerations, i.e., Beijing-Tianjin-Hebei (BTH), Yangtze River Delta (YRD), and Pearl River Delta (PRD), which are the most densely populated and mobile areas in China. All three urban agglomerations are important industrial and economic development regions, and their inter-city linkages and population migration have been the focus of scholarly and governmental attention and are strategically important for Chinese future urban planning and development. However, related research still needs to be completed.

Overall, the main contributions of this study are an essential advancement of urban structure research to a future dimension, a population migration prediction model that can be extended to all major metropolitan areas in China, and an urban planning solution for future rapid urban mobility. Moreover, the structure of the full text is organized as follows: The second part first gives a brief overview of the recent literature on the dimensions, measures, and perspectives of urban network structure. The third part focuses on this study's population migration prediction model. The fourth section presents the experiments and findings of this study, focusing on the urban network and inter-city linkages shift. Finally, we discuss how the results indicate future human mobility for government planning and management in 2025.

2. Literature Review

2.1 Dimensions and Measures of Urban Structure and Network

Urban network research has expanded its scope to encompass various disciplines, including geography, sociology, engineering, physics, and different scales of analysis (Derudder & Taylor, 2018; Kong, Chen et al., 2022; Taylor et al., 2010). Spatial network research has also gradually changed from focusing on geographic regions with similar spatial attributes to emphasizing the interactive relationships between these regions (Derudder & Taylor, 2018; W. Zhang et al., 2020). This shift encompasses investigations into various levels of network formation, ranging from the local social networks within neighborhoods to regional transportation networks and even global, transnational networks. In addition, urban network research has rapidly expanded in size, scope, and scale, energizing the field and giving administrators access to various empirical studies (Cheng et al., 2013). The different analytical dimensions in urban network studies, each with a unique theoretical perspective and explanatory focus, are critically summarized in this section.

Early research on urban networks interpreted regional spatial organization regarding socioeconomic, political, and cultural characteristics, grouping regions with comparable attributes (Anas, 1990; Getis & Getis, 1966). Subsequently, researchers utilized ideal surfaces and the presumption of an economic man to explain human mobility patterns, employing concepts such as central function and position (Dacey, 1965; Mulligan, 1984). For a long time afterward, urban planning and geography used the main theory to investigate urban systems, which led to the extension of the world city hypothesis (Pacione, 2002), the urban economic control function (Sassen, 2013), and other fundamental theories. While these studies provided valuable insights into regional disparities and the inherent structure of socio-spatial organization, they could have been more extensive in their focus. They primarily considered a vertical, one-level linear relationship between regions and metropolitan centers, presenting somewhat monolithic perspectives (Sun & Hou, 2020).

Later, with the rapid development of the global economy and technology and the growing economic and information exchange between cities, the urban structure has concentrated more on patterns of aggregation and dispersion of human activities or spatial attributes (W. Zhang et al., 2020). Urban networks have changed from one-dimensional hierarchical structures to a tetrahedral structure. Urban areas are complex systems comprising various flows, including information, capital, and people. These flows serve as the means through which cities communicate with one another and establish connections with other cities within the region (Palla et al., 2005; Taylor et al., 2010). Consequently, the spatial structure and functional areas of urban regions are not solely determined by geographic factors in a one-dimensional manner. Instead, they are shaped by local connections such as transportation commuting, industrial structure, and economic networks (Batty, 2008; Zhong et al., 2014). To better describe the internal network structure of cities, the external morphology of cities, the networks of urban groups, and the networks of urban functions, two-dimensional flow space studies need to choose the appropriate relational data, such as population migration, business investment (De Goei et al., 2010; Sun & Hou, 2020).

Although current one- and two-dimensional studies offer a multi-scale and varied perspective as well as empirical research for comprehending the spatial and network structure of urban regions, which can qualitatively inform city managers (Cheng et al., 2013), they are unable to predict changes in urban networks in the upcoming years prospectively. This may be because of the limited availability of sophisticated network information generation models and assessment tools for forecasting systems. It is worth noting that more and more prediction studies have been conducted in the fields of urban and geography as a result of the development of scientific techniques like deep learning, such as land expansion (M. Zhang et al., 2023), land use types (J. Wang et al., 2022), and business linkages (L. Li et al., 2021). However, there still needs to be more experience in the field when it comes to predicting more complex urban spatial networks. As a result, this study aims to address this gap by estimating how urban networks will evolve in the coming years, leveraging future forecasts of urban migration networks. This is a fresh way of looking at the large-scale urban structure and a helpful resource that can aid urban managers in making decisions.

2.2 Understanding Urban Network from an Urban Mobility Perspective

Urban network analysis uses inter-city interaction modeling as a critical tool to quantify population (L. Li et al., 2021; Y. Zhang et al., 2022), logistical, financial, informational, and technological flows, particularly in the context of population migration (De Goei et al., 2010; Sun & Hou, 2020). Understanding and interpreting geographical structures and networks from the perspectives of urban mobility and population migration has become common and effective with spatial data flows represented by open-source data (Z. Li et al., 2021). A popular and successful method for studying spatial structure and networks is based on urban mobility and population migration (L. Chen et al., 2022; Liu, 2021).

Urban mobility means the movement of humans in both space and time, influenced by capital imbalances, differences in well-being and living conditions, and socio-economic factors (Barbosa et al., 2018). Human mobility plays a significant role in shaping spatial networks, and population migration quantifies the strength of connections between geographic entities, giving rise to these interrelated spatially linked networks (Y. Zhang et al., 2022). These networks can analyze cities' spatial structure and changing trends. For instance, the spatial structure of urban agglomerations is measured using commuter flows (W. Zhang et al., 2020), the polycentric structure of urban agglomerations is evaluated using data from mobile phone signaling (Liu et al., 2021), and the fine-grained spatial structure of cities is examined using population trajectory data (Y. Zhang et al., 2021). This pedestrian flow data captures the origin (O) and destination (D) of users (citizens), activity time, and number of activities to carry significant geospatial interactions. They can be used to calculate spatial attractiveness based on geographic selection and intensity of flows of people, based on inferring the geographic characteristics of cities and accurately reflecting urban networks (Cao et al., 2023).

Human mobility modeling has received increasing attention in recent decades (Kong, Li et al., 2022; Luca et al., 2021). Relevant studies use gravity model (Barbosa et al., 2018; Lenormand et al., 2016), network analysis (P. Zhang et al., 2020) to measure traffic data, often assigning travel probabilities between any two locations, even at a large scale level, to estimate origin-destination matrices (Erlander & Stewart, 1990; Y. Lin et al., 2022; Simini et al., 2012), reflecting the urban structure and network. These approaches assume that the population flow between two locations increases with the population and decreases with the distance between the locations (Barbosa et al., 2018; Lenormand et al., 2016). Based on this assumption, it is possible to measure population migration between areas as long as the population and distance are known, which applies to multiple application scenarios such as transportation planning, spatial economics, and urban networks (Balcan et al., 2010; Erlander & Stewart, 1990; Karemera et al., 2000).

These models have undergone significant improvements and optimizations to facilitate the generation and prediction of mobility flows. For example, Beiró et al. (2016) combine extensive publicly available data from a popular photo-sharing system with the gravity model to propose a multi-scale enhanced mobility prediction model. Wang & Chen (2022) used a dual gravity model to predict human flows in cities within urban agglomerations and reveal the deep structure of urban and transport networks. Notably, Simini et al. (2021) proposed the Deep Gravity (DG) model, which pioneers the integration of multivariate data (such as road networks, traffic, facilities, and other geographic information) and nonlinear layers through deep neural networks (DNNs) into the gravity model. This innovation allows for the discovering of nonlinear relationships between city features and mobility patterns. These advanced models and methods have facilitated population migration simulation and allowed researchers to delve deeper into the intricate connections between cities, including historical urban interactions and geographical elements reflecting urban characteristics (e.g., poi, land use). Through the lens of population migration, these approaches facilitate the construction and modeling of urban networks, providing valuable insights for urban planning and analysis.

3. Materials and Methods

3.1 Research Area

In this study, we take 362 cities in China as the training area (except for strategic national regions such as Hong Kong, Macau, Hainan, and Taiwan) and focus on the three most important urban agglomerations, namely Beijing-Tianjin-Hebei (BTH), the Yangtze River Delta (YRD) and the Pearl River Delta (PRD) as testbeds, as described in Figure 1. As the top three urban economic zones in China, their urban networks, centrality, labor migration, industrial activities, and transport logistics have been the focus of domestic and international researchers and city managers (X. Li et al., 2016; F. Wang et al., 2019; W. Zhang et al., 2020).

Specifically, the BTH agglomeration is located in the heart of the Bohai Sea Rim in Northeast Asia, serving as the capital region of China. Comprising 14 cities in Beijing, Tianjin, and Hebei Province (see Figure 1-b), this region is characterized by geographical integration and synergistic development. It has emerged as the largest and most vibrant economic region in northern China, increasingly attracting the attention of both China and the whole world (F. Li et al., 2021). Encompassing a total area of approximately 218,000 km2, the BTH agglomeration is home to a resident population exceeding 112 million, accounting for 7.24% of the Chinese population (China Statistics Press, 2019). Next, the YRD agglomeration is situated in the alluvial plain at the mouth of the Yangtze River in central China (see Figure 1-c). It is recognized as one of the best urbanization-based regions in China, as well as a vital intersection of the Belt and Road Initiative and the Yangtze River Economic Belt. It mainly includes 26 cities in Shanghai, Jiangsu Province, and Zhejiang Province (*Yangtze River Delta City Cluster Development Plan*, 2016). With a total area of about 211,700 km2, this region has a resident population of more than 225 million, accounting for 14.54% of the national population (China Statistics Press, 2019). Finally, the PRD agglomeration is located in southern China (see Figure 1-d) and is an economic powerhouse driving the national socioeconomic development (Yeh & Chen, 2020). Covering a total area of about 42,200 km2, this agglomeration has a resident population of over 57.2 million. Although the PRD agglomeration accounts for only 0.6% of the national land area, it accommodates 4.3% of the population. It contributes 10.03% of the total GDP in China (*Statistical Bulletin of National Economic and Social Development*, 2021).

The city interaction of these three urban agglomerations has grown steadily since the opening-up of China, showing characteristics of high mobility, compound activities, and frequent regional interactions (G. C. S. Lin, 1999; Yang et al., 2011). However, along with the rapid urbanization, these three agglomerations also present problems such as unbalanced regional development, occupational and residential segregation, and population loss from legging cities (Z. Chen & Yeh, 2022). Therefore, it is of national strategic importance to forecast the future urban network of these three agglomerations, which can help to prospectively understand the population migration of Chinese mega-city regions and potentially provide new ideas for further research on planning and management in the large-scale areas of the Global South.

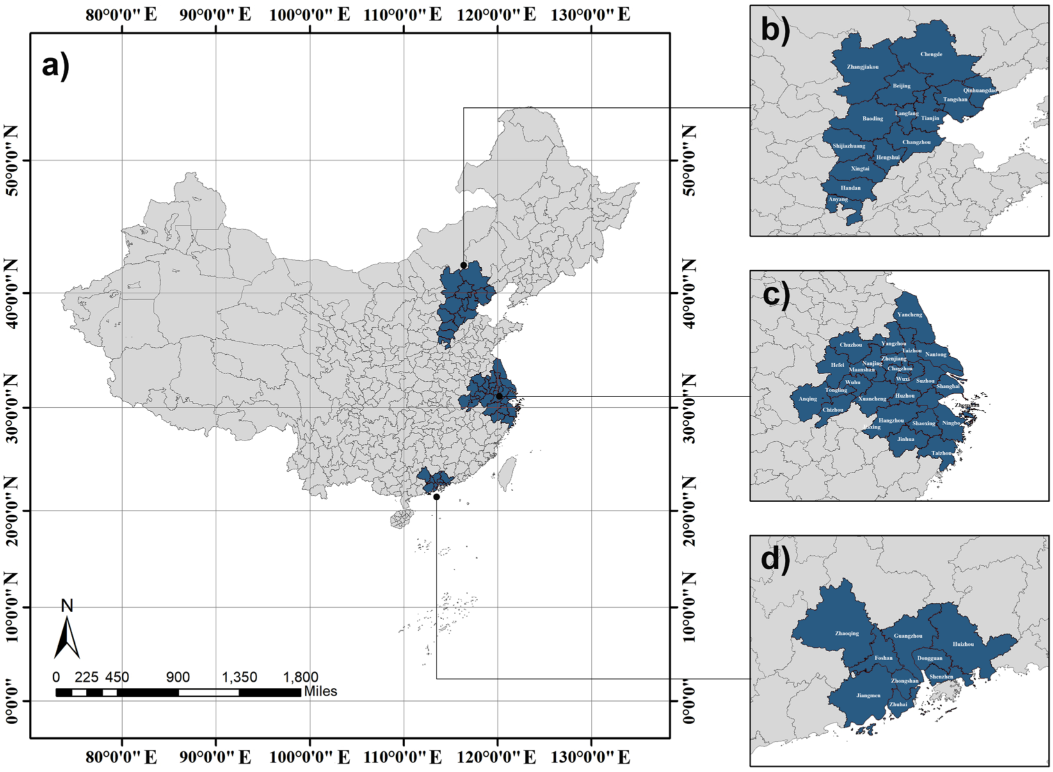

Figure 1. The geographical locations of the study areas. a) The whole region of China. b) The BTH (Beijing-Tianjin-Hebei) region and the cities included in the study. c) The YRD (Yangtze River Delta) region and the cities. d) The PRD (Pearl River Delta) region and the cities.

3.2 Data Source

The data used in this study can be categorized into three main categories: geo-information data, regional population data, and migration data. These data sources provide valuable information for analyzing and modeling the urban networks of the study areas.

The first data category used in this study is historical, geographical information obtained from OpenStreetMap (https://download.geofabrik.de/asia/china). This dataset includes eight features: land use, railways, roads, buildings, points of interest (pois), transports, traffic, and places. It is collected on the 1st of January from 2014 to 2022. These features provide comprehensive information about the development and infrastructure of cities, as well as the attractions and amenities that may influence the movement of migrant populations.

The second data category used in this study consists of official population projections for 2025 and historical regional population data for 2020. The former is from the China Population Projections 2025 report. The latter resident population data of each city from the 2020 China Census Yearbook and statistical yearbook data of the relevant cities. These data sources provide information on the population size and distribution in the study areas, which also act as one of the indispensable input features of our DG model; a detailed description can be found in section 3.3.

The third category is human mobility data, which uses the Gaude migration data from January to March 2020 and 2021 (https://trp.autonavi.com/migrate/page.do) as historical mobility data between cities. The data is recorded daily and includes information on cities that experienced population inflows, outflows, and the corresponding migration indices. By utilizing this dataset, the study gains insights into the patterns and dynamics of human mobility between cities during the specified periods.

These three types of data utilized in our study are all online open-source geographic data, offering several advantages: a) they provide a broad temporal coverage spanning multiple years, which enables a comprehensive analysis of changes over time. b) They encompass a wide study area, allowing for a holistic examination of the regions under investigation. c) These data sources are relatively easier to access compared to high-precision mobile signaling data and GPS data, which are often required by specific prediction models (Do & Gatica-Perez, 2012; X. Li et al., 2012; Simini et al., 2021). This accessibility makes the data more applicable in studies focused on urban regions, where acquiring fine-grained data can be challenging and costly. Moreover, using open-source geographic data enhances the generalizability of the findings, as it overcomes the limitations imposed by data availability and acquisition thresholds in urban research.

3.3 Research Framework for Predicting Urban Network Shifting in 2025

This study aims to predict the urban network 2025 by leveraging population migration perspectives in three major urban agglomerations in China. To achieve this, the study generalized the innovative deep learning gravity (DG) model proposed by Simini et al. (2021), along with open-source geo-information data covering the period from 2014 to 2022. The research framework of the DG model used in this study is depicted in Figure 2. This framework outlines the approach and methodology employed to analyze and predict the urban network based on population migration data. By integrating the DG model with the available data, the study seeks to gain insights into the future spatial connections and dynamics among cities in the selected regions.

The model can be considered as three processes from input to output. First, mainland China (except Hainan Island) is partitioned by 300km×300km squares, which without city centroids is removed, and finally, 126 valid regions ($R_q, q = 1,2,\ldots,126$) are obtained, including a total of 364 cities. Since the geographic boundaries of cities are irregular, the model symbolizes cities with city centroids and can randomly divide the centroids into different squares. Each square encompasses more than two neighboring cities and can be considered a cluster of cities. Besides, the geographic resolution of 300km was set concerning the parameters in the original DG model (Simini et al., 2021) to have more than two cities in each R_q to form city pairs/networks. Also, each city centroid will be given geographic features $\boldsymbol{x_i}$ that include nine features from OSM and populations, and all values have been logarithmically operated. Then, the R_q is randomly divided into two groups: one half is used as the training set of the model, and the other half is used as the validation set of the model (see Figure 2-a). The actual migrant data in 2020 provided by Gaude is used as the ground truth for model training. Moreover, the DG Model is chosen as 15 hidden layers of dimensions 256 (the bottom six layers) and 128 (the other layers) with the LeakyReLu activation function. Thus, the DG Model is trained to obtain a prediction model that can be extended to all regions in China, with model convergence in 60 epochs (see Figure 2-c).

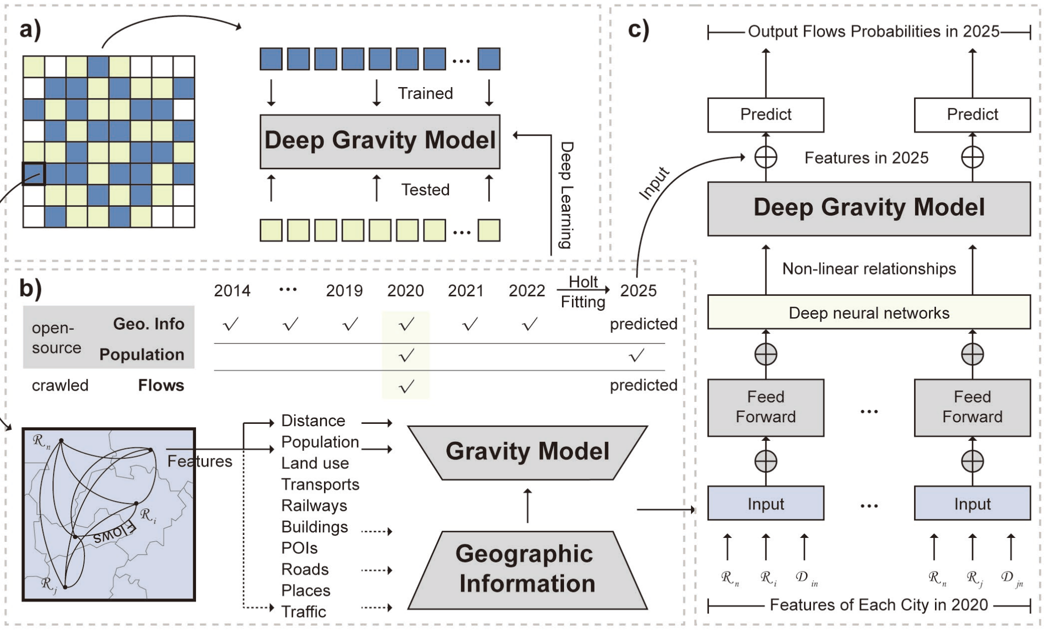

Figure 2. The research framework of this study. a) The study area is divided into a grid of squares using a random division approach. However, only squares that contain more than two city centroids are included in the model. b) For each city pair, eighteen features and the distance between two cities are used as inputs for the model. These features are derived from the geo-information data and population data. c) The architecture of the Deep Gravity Model includes input, hidden, and output layers. These layers enable the model to learn complex relationships between the input features and population migration patterns.

Second, we focus on the three major urban agglomerations, construct the Holt-Winter model on the eight urban features of each city, and predict the geographical parameters in 2025 using the geographic data from 2014-2022, described in section 4.5. The geographic features $\boldsymbol{x_i}$ of each city to be studied consist of the calculated urban features and the collected urban population in 2025. Finally, the urban network within the three urban agglomerations in 2025 is generated based on the predicted dataset and trained model.

3.4 Improved and Applied Deep Gravity Model in China

Human flow generation has been a vital issue in urban studies, usually defined R as a region where flows need to be generated. R is divided into n locations, which are called l_i, $i = 1,2,\dots,n$. Then we define O_i, $i = 1,2,\dots,n$ as the total outflow of location l_i, and $f(l_i, l_j)$, $i = 1,2,\dots,n$, $i \neq j$ as flows from origin location l_i to destination location l_j. For this study, location can be considered as the Chinese cities. The flow generation can be reformulated as knowing O_i, the total outflow of l_i, generating the flows $f(l_i, l_j)$ whose origin is l_i and destination is l_j in R. If assumed $O_i = 1$, the $f(l_i, l_j)$ represents the probability that the flow of people from l_i goes to l_j. However, in the actual study, the O_i is costly data to obtain and does not meet the requirement of "open-source data" in this paper, so the O_i is assumed in the following.

To evaluate the performance of flow generation models, the most common metric used is the Common Part of Commuters (CPC):

$$CPC = \frac{2\sum_{i,j} min\left(y^g(l_i, l_j), y^r(l_i, l_j)\right)}{\sum_{i,j} y^g(l_i, l_j) + \sum_{i,j} y^r(l_i, l_j)} \tag{1}$$

where $y^g(l_i, l_j)$ means the generated flows, and $y^r(l_i, l_j)$ means the real flows. CPC is always positive and contained in the closed interval [0, 1], with 1 indicating a perfect match between the generated flows and the ground truth and 0 highlighting the lousy performance with no overlap.

The gravity model is a well-accepted model for predicting mobility flows, inspired by the gravity model, which states that the magnitude of the gravitational force between two objects is proportional to the mass of those two objects and inversely proportional to the distance. Similarly, the gravity model considers the mobility flows between two cities to be proportional to the population of the two cities and inversely proportional to the distance between the two cities. The equation of the single-constrained gravity model is:

$$y^g(l_i, l_j) = O_i p_{ij} = O_i \frac{m_j^{\beta_1} f(r_{ij})}{\sum_k {m_k}^{\beta_1} f(r_{ik})} \tag{2}$$

$$f(r_{ij}) = e^{\beta_2 r} \; or \; r^{\beta_2} \tag{3}$$

In which, $y^g(l_i, l_j)$ represents the flows generated by the model from l_i to l_j. The p_{ij} represents the probability that the outflow of people from l_i goes to l_j. The m_j represents the population of the destination l_j. The r_{ij} represents the distance from the geometric center of mass of l_i to the geometric center of mass of l_j. The $f(r_{ij})$ is the deterrence function. Therefore, the idea of solving the classical Gravity model is to find the parameter $\beta_1$and β_2 that makes $y^g(l_i, l_j)$ closer to the actual value $y^r(l_i, l_j)$.

To further extend the applicability of the gravity model, the DG model introduced hidden and nonlinear layers, which can take on multivariate inputs and have better generalization capabilities. The model takes the geographical features set, describing the origin and destination city as input vectors, and outputs a scalar related to the value of the flows between the two places. In a larger range C (e.g., countries), divide the m types of R_q, $q = 1,2,3, \ldots, m$ as a given partitioning rule (e.g., grid, or boundary of cities). The data set of each city is represented by its geographic features $\boldsymbol{x_i}$, where $\boldsymbol{x_i}$ are vector containing geographic features such as land use, commerce, latitude, and longitude, etc. In generating $f(l_i, l_j)$, the input to the DG model is $\boldsymbol{S}(l_i, l_j) = (\boldsymbol{x_i}, \boldsymbol{x_j}, r_{ij})$, in which the $\boldsymbol{x_i}$ is the geographic feature vector of starting point l_i while the $\boldsymbol{x_j}$ is the geographic feature vector of ending point l_j, and the r_{ij} is the distance between the geometric centers of mass of l_i and l_j. The output of the DG model is a scalar. The larger the value of which, the more probability that outflows of l_i will move to l_j. If there are n cities in an R_q, there will be $n-1$ possible destinations l_j from l_i. Thus, it is necessary to input n-1 $\boldsymbol{S}(l_i, l_j)$ into DG Model in parallel, and the model outputs n-1 scalar values in parallel, which are then input into the SoftMax layer to obtain the distribution of flows from l_i to other cities in R_q.

Indeed, in addition to mining the relationship between geographic features and mobility flows, the DG model also aims to explore the temporal generalization capability. This involves analyzing the future dimension and understanding how the relationships between cities and mobility flows evolve. By considering temporal factors, the model can capture the changing patterns of population migration and predict future mobility flows based on historical data. This temporal generalization capability allows a deeper understanding of urban networks and enables more accurate predictions of future urban spatial patterns.

There are two improvements in this study. First, we choose different partitioning rules for C in the training and generation stages. The traditional DG model uses a tessellation to partition C in both training and generation stages, from which a random portion is used for the training and validation sets of the model, and another portion is used as the test set. However, we tried to improve this partitioning to be more flexible. In the training, C can be partitioned using tessellation to obtain more random partitioning results to help the model learn universal patterns. In the generation, C can be partitioned according to research interests, such as urban agglomerations and economic zones. The second improvement is that the model training and generation are not necessarily on the same time slice. Specifically, it can be trained based on the geographical features set and the actual flows data at the t_1 moment, but generates the flows based on geographic features set at the t_2 moment, $t_2 \neq t_1$, and t_2 is a future time slice.

3.5 Predicting Urban Features using Holt-Winters Exponential Smoothing

The Holt exponential smoothing method and its extension, Holt-Winters, are widely used techniques for forecasting time series data. Holt-Winters is particularly useful when the data exhibits a trend and seasonality (Holt, 1957; Winters, 1960). This study applied the Holt-Winters method to the time series datasets obtained from OSM from 2014 to 2022. The eight geographical urban features, including land use, railways, roads, buildings, pois, transports, traffic, and places, were aggregated within each geographic region R_q by spatially linking and statistically aggregating the values on GIS. By fitting the Holt-Winters model to the aggregated data, the study was able to capture the underlying trends in the urban features. This allowed for the prediction of values for the year 2025 using the model's forecasting capabilities. The implementation of Holt-Winters was carried out using the statsmodels library in Python.

4. Experiments and Findings

4.1 Model Experiment and Accuracy

In the first quarter of 2020, which marked the onset of the COVID-19 pandemic, the urban network among Chinese cities might have experienced inevitable disruptions, resulting in a decrease in overall population migration compared to previous years. To assess the potential negative impact of the outbreak data on future predictions and validate the DG model's predictive capabilities, this study experiments to predict the urban network from the population migration perspectives of three city agglomerations in the first quarter of 2021.

As discussed in section 3.3, the DG model underwent training for 60 epochs using inputs from 2020. The validation datasets produced a CPC (Classification Precision Coefficient) value of 0.76, which significantly exceeds the performance of the original DG experiment conducted by Simini et al. (2021). A series of projected evaluation indexes are also performed to examine the disparity between the actual and predicted data from January to March 2021, considering both features and migration possibilities.

As presented in Table 1, the predicted outcomes of the eight features utilizing the Holt-Winters method exhibited minimal deviations from the actual feature data in 2021. Similarly, as depicted in Table 2, the predicted results of the model demonstrated minimal variations from the actual population outflow data. These findings indicate that our predicted results accurately represent the intricate migration network between cities, with little impact from external factors such as epidemics or home segregation policies. Moreover, these results affirm the model's robust geographic and temporal transferability when forecasting urban networks in Chinese cities.

Table 1. The projected evaluation results between the actual and predicted input features (2021) by the Holt-Winter.

Features	CPC	R2	MAPE	RMSE
Land use	0.8784	0.8509	27.1806	1185.8476
Transport	0.9024	0.9419	30.5553	458.8102
Railway	0.9206	0.9639	20.3472	252.5057
Building	0.8802	0.8983	26.1089	5220.8783
Pois	0.9578	0.9900	18.3362	223.5306
Roads	0.9319	0.9303	14.5972	5332.6432
Places	0.7567	0.5366	33.4884	1394.7871
Traffic	0.8677	0.9024	32.0445	92.8436

Note: The closer CPC/R2 is to 1, the stronger the model's ability to explain the observed data and the better the fit. While MAPE/ RMSE usually ranges from 0 to positive infinity. A smaller MAPE indicates a lower relative error in the model and a higher prediction accuracy. This rule applies equally to Table 2.

Table 2. The projected evaluation results between the actual and predicted mobility flows (2021) by the DG model.

Evaluation	With the non-linear layer	Without the non-linear layer
CPC	0.7039	0.4213
R2	0.5819	0.0174
MAPE	162.7467	946.6406
RMSE	0.0731	0.1121

In addition, we conducted an additional experiment by removing the non-linear layer from the DG model and re-performing the 2021 mobility forecast based on the same dataset and conditions.

However, the results yielded a lower CPC value of 0.48 for validation datasets. Not only that, but we also compared the variability between the actual and predicted values in 2021. This difference was further analyzed using a series of projected evaluation indexes, as shown in Table 2, indicating more minor significance results than the original DG model, which included a non-linear layer. Hence, adding a non-linear layer to the forecasting model can significantly improve the prediction performance. Besides, the CPCs of the 2020 and 2021 years of data are very similar, suggesting that the model is robust enough to adapt to significant changes over time and still produce reliable predictions for the future. This also implies that the interrelationships between the cities are stable despite the impact of the epidemic suffered in 2020.

4.2 The Shifting of the Urban Network in 2025

Based on the tested DG model, this study generated the population migration of the three urban agglomerations in the first quarter of 2025 and visualized the urban network. As shown in Figure 3 and Figure 4, we compare the predicted network in 2025 with the actual patterns in 2020 to interpret the gravitational effect of the urban network between different cities and how it will shift in the future. Besides, it is essential to highlight that due to the extensive size and complexity of the city network in three agglomerations, we opted to extract and display only the top 35% of the migrating population to achieve a more precise visualization.

In 2020, there was a high level of human mobility between cities within each urban agglomeration, characterized by a prevalence of migrants traveling to nearby cities. This resulted in a distinct polycentric structure within the agglomerations. The central cities of the urban agglomerations (Beijing, Shanghai, Guangzhou, Shenzhen) formed strong network links with neighboring cities, accompanied by noticeable local co-location effects. However, the overall co-location effect across large urban agglomerations was less pronounced.

As for the BTH agglomeration in 2020 (see Figure 3-a-1 and 4-a-1), it presented a more robust monocentric urban network with Beijing as the central city, exerting a significant attraction force on the surrounding cities. This resulted in a closely connected center of the agglomeration. The external agglomeration areas mainly displayed activated mobility among themselves but relatively weak interaction with the central city. In the YRD agglomeration 2020 (see Figure 3-b-1 and 4-b-1), the urban network showcased strong mobility among the peripheral cities but weaker mobility within the internal cities. The most developed city in the Yangtze River Delta is Shanghai, but it is located in the eastern coastal region, which makes Shanghai quite tricky to spread out and drive the surrounding cities. In contrast, the PRD agglomeration demonstrated a clear polycentric structure in 2020 (see Figure 3-c-1 and 4-c-1). Specifically, Guangzhou, Foshan, and Zhaoqing in the northwest were closely linked and exhibited significant interaction. Similarly, Shenzhen, Dongguan, and Huizhou in the northeast had strong population mobility. Additionally, Zhuhai and Zhongshan on the southern coast also had strong interaction, forming a mobility pattern that aligned well with the actual urban development of the PRD region.

By examining the population migration network in the three urban agglomerations in 2020 and 2025, as illustrated in Figure 4, it becomes evident that long-distance cross-city migration is expected to increase in each agglomeration. This trend indicates a weakening polarizing effect within the urban network, meaning that larger cities will become less attractive to the surrounding small and medium-sized cities. As a result, the interaction between cities will become more balanced in the future.

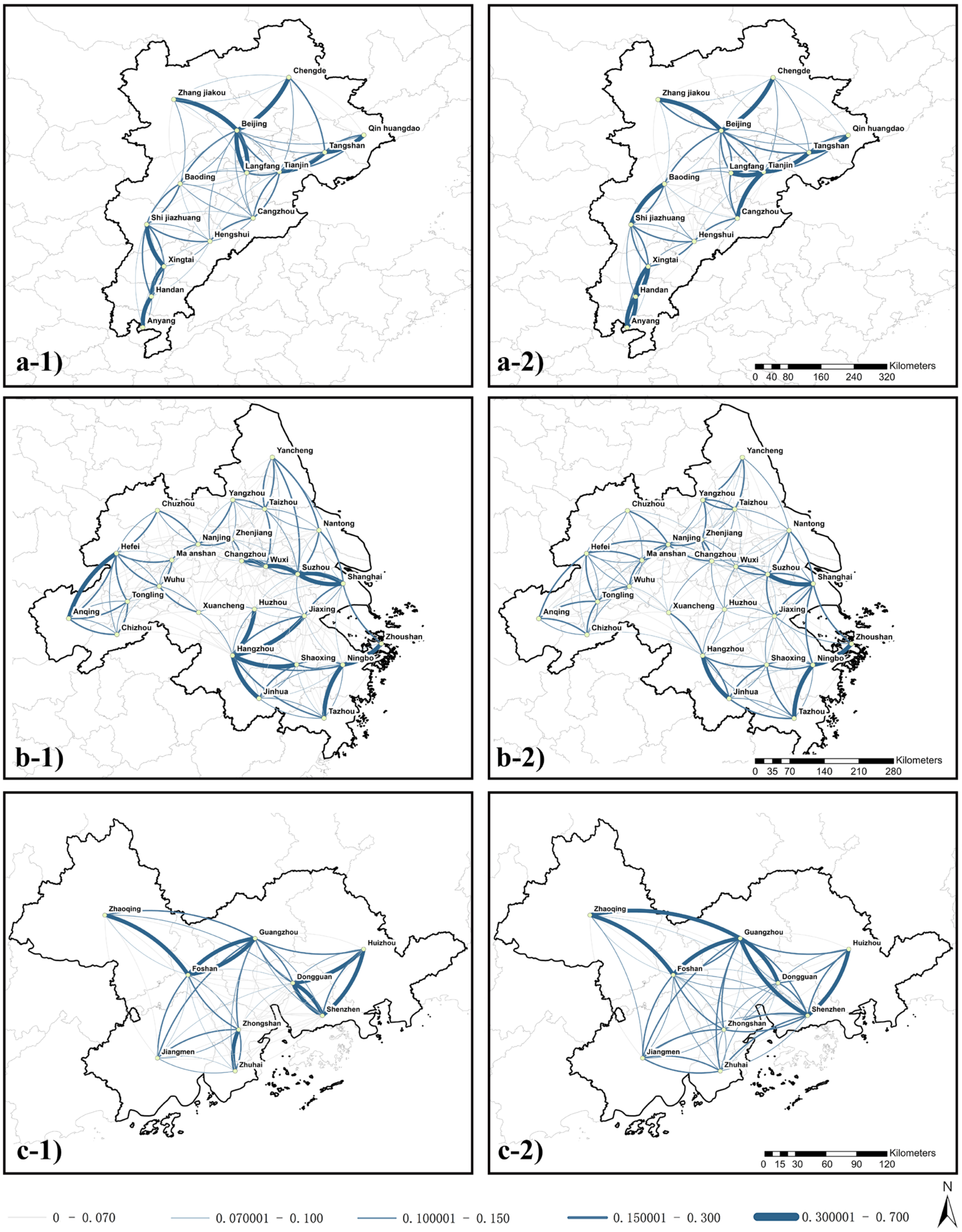

Figure 3. The large-scale mobility pattern in China’s three urban agglomerations in 2020 and 2025. The greater the curvature of the arc and the darker the color, the greater the proportion of the outflow population to the area's total population. a-1) The real mobility pattern in the BTH agglomeration in 2020. a-2) The predicted mobility pattern in the BTH agglomeration in 2025. b-1) The real mobility pattern in the YRD agglomeration in 2020. b-2) The predicted mobility pattern in the YRD agglomeration in 2025. c-1) The real mobility pattern in the PRD agglomeration in 2020. c-2) The predicted mobility pattern in the PRD agglomeration in 2025.

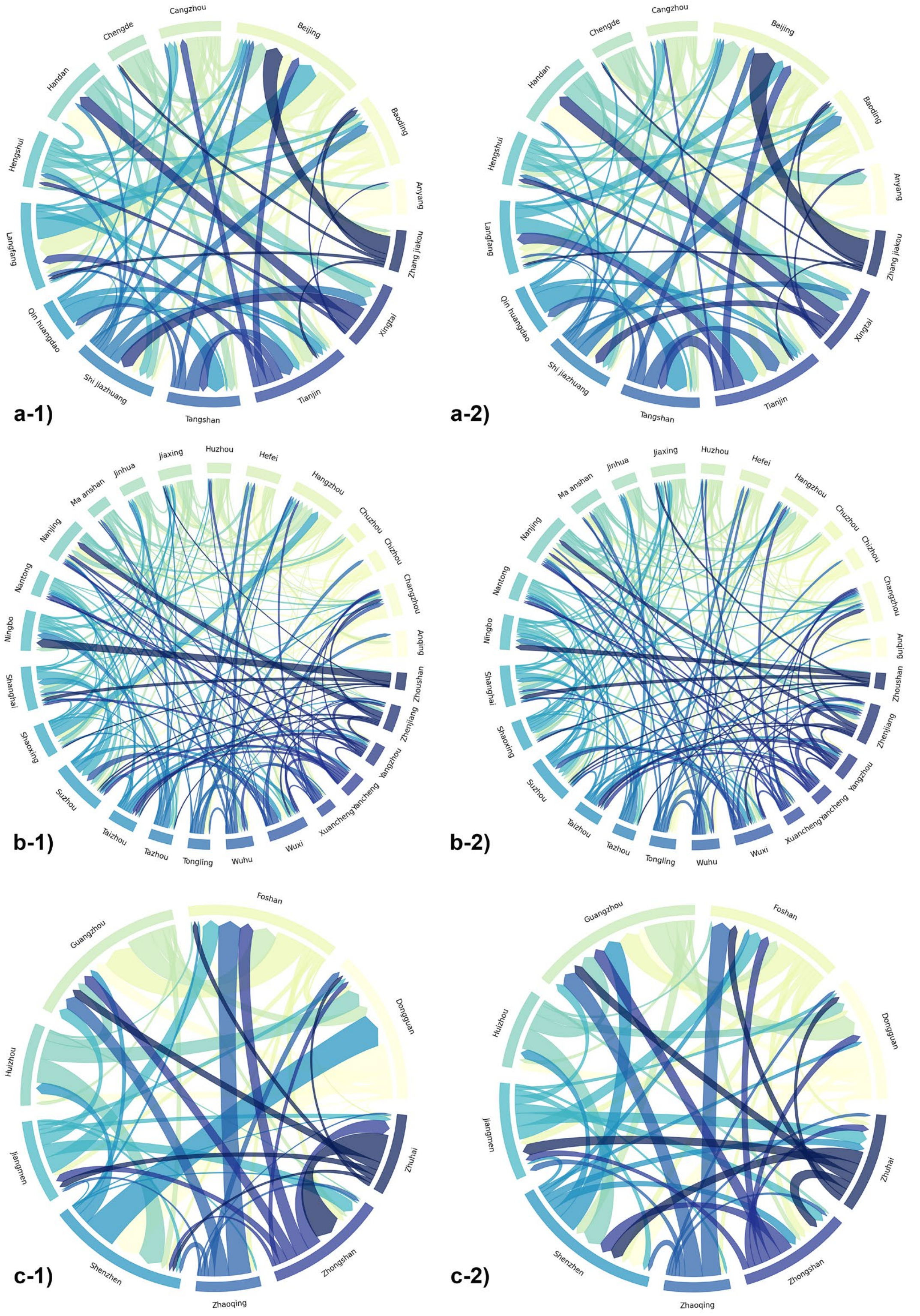

Figure 4. The human mobility network in China's three urban agglomerations in 2020 and 2025 hides the weak links. a-1) The real mobility network in BTH agglomeration in 2020. a-2) The predicted mobility network in BTH agglomeration in 2025. b-1) The real mobility network in YRD agglomeration in 2020. b-2) The predicted mobility network in YRD agglomeration in 2025. c-1) The real mobility network in PRD agglomeration in 2020. c-2) The predicted mobility network in PRD agglomeration in 2025.

The analysis of the BTH agglomeration reveals that both in 2020 and 2025, the urban network is influenced by the central city of Beijing, driving the development of peripheral cities. However, in 2025, this influence is projected to extend even further to the peripheral cities (see Figure 3-a-2 and 4-a-2). Notably, the urban network links between cities such as Anyang and Handan are expected to experience significant growth. Additionally, the interactions between cities such as Shijiazhuang and Baoding, Tianjin and Langfang, and Changzhou are also anticipated to increase. These findings suggest that the BTH agglomeration is still undergoing development and construction. By 2025, Tianjin will emerge as another important central city, and the surrounding cities will gradually reduce their dependence on Beijing for growth. Overall, the structure of the BTH agglomeration is evolving from monocentric to polycentric.

In the urban network of the YRD agglomeration, the cities on the region's periphery are projected to have weaker connections in 2025 compared to 2020 (see Figure 3-b-2 and 4-b-2). For instance, the links between Hefei and Anqing in the west are expected to weaken significantly, and Hangzhou in the south will also exhibit reduced connectivity with its peripheral cities. This indicates a diminishing polarization of polycentric roles and a flattering urban network within the urban agglomeration in 2025. Furthermore, a mobility pattern emerges where stronger mobility is observed in the eastern coastal cities and weaker mobility in the central and western inland cities.

Regarding the PRD agglomeration, the two central cities of Guangzhou and Shenzhen are projected to interact strongly with each other in 2025 (see Figure 3-c-2 and 4-c-2). Specifically, Zhaoqing and Guangzhou will establish a stronger urban connection, forming the Guangzhou-Foshan-Zhao city cluster. On the other hand, as Dongguan experiences development and growth, its links with Shenzhen and Huizhou are expected to weaken, while its connections with Guangzhou are projected to strengthen gradually. The southern "Pearl River Estuary West Coast Metropolitan Area" is also anticipated to disintegrate and form an urban network centered around the central city. Overall, the PRD agglomeration in 2025 is projected to become more centralized, transitioning from a separated multi-center cluster in 2020. The strong regional core formed by Guangzhou and Shenzhen will continue to drive the surrounding cities closer together within the region.

5. Discussion and Reflection

5.1 Future Policy Implications and Planning Response

According to the results, it is evident that the urban network in 2025 will exhibit a more equitable distribution, with reduced concentration in a few dominant cities. This shift signifies a transition towards a more balanced allocation of resources, opportunities, and development across China's three major urban agglomerations. As a result, it is expected to foster a more sustainable and inclusive urban landscape. These findings hold significant implications for the three agglomerations regarding regional integration, cooperative programs, and spatial planning initiatives. The shift towards a more balanced urban network calls for enhanced regional integration efforts, encouraging collaboration and cooperation among cities. It also highlights the need for strategic cooperative programs to leverage the collective strengths of the agglomerations.

For instance, the BTH agglomeration exhibits a polycentric structure characterized by multiple city pairings in 2025. This spatial organization closely aligns with the BTH Synergistic

Development Planning Outline, which Beijing has implemented to establish growth poles outside the city, including Tianjin, Baoding, and Langfang. This approach aims to alleviate the congestion resulting from the substantial population in Beijing. Despite this policy's long-standing implementation, Beijing is prominent in the BTH urban agglomeration. However, the situation is anticipated to significantly improve in 2025 as population shifts occur and other cities undergo development and expansion. This highlights the importance for urban planning managers to actively promote the development of neighboring cities and maintain a consistent pace of progress to align with planning expectations. By doing so, the agglomeration can diversify its economic and demographic structure effectively, reducing the dependency on Beijing as the sole dominant city.

Similarly, the YRD urban agglomeration is expected to foster new regional centers in the coming years. This urban network within this agglomeration adhered to the "multi-level and multi-class networked spatial pattern" outlined in the Yangtze River Delta City Cluster Development Plan. As a result, the overall structure of the agglomeration exhibits a more balanced and equitable arrangement. By 2025, the spatial structure observed will be consistent with policy expectations as the central cities within the agglomeration engage in cooperative development and establish dynamic urban connections with their neighboring cities. This augurs well for relevant urban planning and policy management. Efforts should be made to encourage and attract high-tech talent migration, facilitating the flow of developed resources from the peripheral cities to the less developed ones. This will contribute to a more inclusive and balanced development across the YRD agglomeration.

Regarding the PRD, there are notable mobile flows in the organization. As the major core cities, Shenzhen and Guangzhou will continue to draw people from the nearby small and medium-sized cities in 2025. However, the city cluster will adopt a more cooperative and mutually beneficial approach, particularly in conjunction with the anticipated development of the Guangdong-Hong Kong-Macao Greater Bay Area. The close collaboration among the leading cities will contribute to the establishment of a world-class bay area. City managers should consider these changes and plan for a more dynamic urban network while implementing effective policies to attract talent to their cities. Additionally, the city will have a greater flow of elements, leading to further resource allocation optimization.

Our findings underscore the importance of revisiting spatial planning initiatives to ensure they align with the evolving urban network, promoting harmonious and inclusive development across the regions. Moreover, these provide valuable insights for guiding future urban development policies and strategies to create resilient and thriving urban environments within China's three major urban agglomerations.

5.2 Limitations and Perspectives

This study utilizes the DG model, an effective model for generating flow probabilities, to predict human mobility. Compared to traditional gravity models, this model enhances non-linear relationship fitting performance through deep neural networks and demonstrates good geographical generalization. The model's excellent performance in England, Italy, and New York State has been discussed by Simini et al. (2021). Furthermore, this study optimizes the data partitioning pattern of the model, broadens its temporal generalization capability, and proves its good performance in densely populated urban agglomerations in developing countries such as China. To the best of our knowledge, few articles combine a state-of-the-art mobility prediction model with open-source data to achieve such high accuracy in mega-scale regions. Our paper introduces new research ideas and analytical perspectives for related studies. However, most existing case studies using the DG model still focus on large regions such as cities or districts, and its predictive performance at the million level is still unknown, heavily relying on a large amount of historical mobility data.

Due to the model's characteristics, this study chooses to construct and predict the future urban network from the perspective of population migration. Future research can use more data sources, introduce more diverse perspectives, and more advanced models to explore the high-dimensional urban structure. We welcome similar practices in other Chinese cities and other countries.

In terms of approach, we demonstrated the ability of the DG model to generate human flows in the future dimension, providing a worthwhile reference for future research. Future work can be extended in the following areas. First, the current work focuses on a city-level spatial scale and a year-level temporal scale. This choice of scale is appropriate for studying future patterns of human flows in urban agglomerations. However, we expect more refined spatio-temporal scale studies, even down to the street or grid level, supported by more adequate data. In addition, this study proposes to predict unknown urban features by Holt-Winter fitting, a simple and easy-use method that has also proven effective. Otherwise, there is a wide range of options for fitting and predicting urban features, and the right fitting way will be more effective and accurate when working on more specific problems. Last but not least, although the DG model has been greatly improved compared to traditional models, it still has some deficiencies in handling spatio-temporal sequential data. The model still receives static data as input instead of sequential (time-series) data. This makes it necessary to predict the features before future mobility can be predicted. It will be a greater breakthrough for spatio-temporal prediction models if the model could be enhanced from the input side in the future.

6. Conclusion

This study expands the conventional understanding of urban structure, moving beyond a two-dimensional to incorporate a future dimension in 2025. By examining the relationship between geographic information data of Chinese cities and the actual urban migration network, we provide evidence of the gravitational effect within the urban network and highlight the regularity of urban development. Furthermore, we demonstrate that the DG model effectively captures the population migration patterns between cities, aligning well with empirical data. Based on our analysis, the following main conclusions are drawn:

Firstly, the DG model demonstrates strong geographical and temporal generalization capabilities within China. Our simulations using the DG model closely align with actual urban mobility patterns, indicating its ability to capture changes in the mobility network over time. Secondly, the training model requires low training parameters and can be flexibly adapted and extended according to the city planning and research questions. This flexibility makes it well-suitable for areas where data availability may be limited. It can be seen that the open-source geographic data combined with a powerful deep learning model can achieve more accurate mobility pattern prediction. Finally, the predicted results highlight the three urban agglomerations' distinct development trajectories and spatial structures. The urban networks are exhibiting a trend towards flattening, indicating a shift towards a more equitable and sustainable distribution of resources.

Based on the above models and conclusions, the government can plan urban networks based on developing population migration to better cope with the pressure of large mobile population in major urban agglomerations, which is important for planning urban systems in the context of rapid population mobility.

Data Availability Statement: Publicly available datasets have been analyzed in this study. The model code and sample data can be found here: https://github.com/carolgxy/DGonChina.

Reference

Anas, A. (1990). Taste heterogeneity and urban spatial structure: The logit model and monocentric theory reconciled. *Journal of Urban Economics*, *28*(3), 318–335. https://doi.org/10.1016/0094-1190(90)90031-H

Bai, X., Colbert, M., McPhearson, T., Roberts, D., Siri, J., Walsh, B., & Webb, B. (2019). Networking urban science, policy and practice for sustainability. *Current Opinion in Environmental Sustainability*, *39*, 114–122. https://doi.org/10.1016/j.cosust.2019.08.002

Balcan, D., Gonçalves, B., Hu, H., Ramasco, J. J., Colizza, V., & Vespignani, A. (2010). Modeling the spatial spread of infectious diseases: The GLobal Epidemic and Mobility computational model. *Journal of Computational Science*, *1*(3), 132–145. https://doi.org/10.1016/j.jocs.2010.07.002

Barbosa, H., Barthelemy, M., Ghoshal, G., James, C. R., Lenormand, M., Louail, T., Menezes, R., Ramasco, J. J., Simini, F., & Tomasini, M. (2018). Human mobility: Models and applications. *Physics Reports*, *734*, 1–74. https://doi.org/10.1016/j.physrep.2018.01.001

Batty, M. (2008). The Size, Scale, and Shape of Cities. *Science*, *319*(5864), 769–771. https://doi.org/10.1126/science.1151419

Beiró, M. G., Panisson, A., Tizzoni, M., & Cattuto, C. (2016). Predicting human mobility through the assimilation of social media traces into mobility models. *EPJ Data Science*, *5*(1), 30. https://doi.org/10.1140/epjds/s13688-016-0092-2

Burger M. J., & Meijers E. J. (2016). Agglomerations and the rise of urban network externalities. *Papers in Regional Science*, *95*(1), 5–15. https://doi.org/10.1111/pirs.12223

Cao, Y., Hua, Z., Chen, T., Li, X., Li, H., & Tao, D. (2023). Understanding population movement and the evolution of urban spatial patterns: An empirical study on social network fusion data. *Land Use Policy*, *125*, 106454. https://doi.org/10.1016/j.landusepol.2022.106454

Castells, M. (1989). *THE INFORMATIONAL CITY, INFORMATION TECHNOLOGY, ECONOMIC RESTRUCTURING, AND URBAN-REGIONAL PROCESS,*. 5.

Chen, L., Zhao, L., Xiao, Y., & Lu, Y. (2022). Investigating the spatiotemporal pattern between the built environment and urban vibrancy using big data in Shenzhen, China. *Computers, Environment* *and Urban Systems*, *95*, 101827. https://doi.org/10.1016/j.compenvurbsys.2022.101827

Chen, Z., & Yeh, A. G.-O. (2022). Delineating functional urban areas in Chinese mega city regions using fine-grained population data and cellphone location data: A case of Pearl River Delta. *Computers, Environment and Urban Systems*, *93*, 101771. https://doi.org/10.1016/j.compenvurbsys.2022.101771

Cheng, J., Bertolini, L., Clercq, F. le, & Kapoen, L. (2013). Understanding urban networks: Comparing a node-, a density- and an accessibility-based view. *Cities*, *31*, 165–176. https://doi.org/10.1016/j.cities.2012.04.005

China Statistics Press. (2019). *National Bureau of Statistics of China, 2019*. http://www.stats.gov.cn/tjsj/ndsj/2019/indexch.htm

Dacey, M. F. (1965). The Geometry of Central Place Theory. *Geografiska Annaler: Series B, Human Geography*, *47*(2), 111–124. https://doi.org/10.1080/04353684.1965.11879288

De Goei, B., Burger, M. J., Van Oort, F. G., & Kitson, M. (2010). Functional Polycentrism and Urban Network Development in the Greater South East, United Kingdom: Evidence from Commuting Patterns, 1981–2001. *Regional Studies*, *44*(9). https://doi.org/10.1080/00343400903365102

Derudder, B., & Taylor, P. J. (2018). Central flow theory: Comparative connectivities in the world-city network. *Regional Studies*, *52*(8), 1029–1040. https://doi.org/10.1080/00343404.2017.1330538

Do, T. M. T., & Gatica-Perez, D. (2012). Contextual conditional models for smartphone-based human mobility prediction. *Proceedings of the 2012 ACM Conference on Ubiquitous Computing*, 163–172. https://doi.org/10.1145/2370216.2370242

Erlander, S., & Stewart, N. F. (1990). *The Gravity Model in Transportation Analysis: Theory and Extensions*. VSP.

Getis, A., & Getis, J. (1966). Christaller's Central Place Theory. *Journal of Geography*, *65*(5), 220–226. https://doi.org/10.1080/00221346608982415

Glaeser E. L., Ponzetto G. A. M., & Zou Y. (2016). Urban networks: Connecting markets, people, and ideas. *Papers in Regional Science*, *95*(1), 17–59. https://doi.org/10.1111/pirs.12216

Godfrey, B. J., & Zhou, Y. (1999). Ranking World Cities: Multinational Corporations and the Global Urban Hierarchy. *Urban Geography*, *20*(3), 268–281. https://doi.org/10.2747/0272-3638.20.3.268

Holt, C. C. (1957). Forecasting seasonals and trends by exponentially weighted moving averages. *International Journal of Forecasting*, *20*(1), 5–10. https://doi.org/10.1016/j.ijforecast.2003.09.015

Kamarianakis, Y., & Prastacos, P. (2003). Forecasting Traffic Flow Conditions in an Urban Network: Comparison of Multivariate and Univariate Approaches. *Transportation Research Record*, *1857*(1), 74–84. https://doi.org/10.3141/1857-09

Karemera, D., Oguledo, V. I., & Davis, B. (2000). A gravity model analysis of international migration to North America. *Applied Economics*, *32*(13), 1745–1755. https://doi.org/10.1080/000368400421093

Kong, X., Chen, Q., Hou, M., Rahim, A., Ma, K., & Xia, F. (2022). RMGen: A Tri-Layer Vehicular Trajectory Data Generation Model Exploring Urban Region Division and Mobility Pattern. *IEEE Transactions on Vehicular Technology*, *71*(9), 9225–9238. https://doi.org/10.1109/TVT.2022.3176243

Kong, X., Li, N., Zhang, C., Shen, G., Ning, Z., & Qiu, T. (2022). Multi-Feature Representation Based COVID-19 Risk Stage Evaluation With Transfer Learning. *IEEE Transactions on Network Science and Engineering*, *9*(3), 1359–1375. https://doi.org/10.1109/TNSE.2022.3142316

Lenormand, M., Bassolas, A., & Ramasco, J. J. (2016). Systematic comparison of trip distribution laws and models. *Journal of Transport Geography*, *51*, 158–169. https://doi.org/10.1016/j.jtrangeo.2015.12.008

Li, F., Yao, N., Liu, D., Liu, W., Sun, Y., Cheng, W., Li, X., Wang, X., & Zhao, Y. (2021). Explore the recreational service of large urban parks and its influential factors in city clusters – Experiments from 11 cities in the Beijing-Tianjin-Hebei region. *Journal of Cleaner Production*, *314*, 128261. https://doi.org/10.1016/j.jclepro.2021.128261

Li, L., Derudder, B., & Kong, X. (2021). A machine learning approach to the simulation of intercity corporate networks in mainland China. *Computers, Environment and Urban Systems*, *87*, 101598. https://doi.org/10.1016/j.compenvurbsys.2021.101598

Li, X., Huang, B., Li, R., & Zhang, Y. (2016). Exploring the impact of high speed railways on the spatial redistribution of economic activities—Yangtze River Delta urban agglomeration as a case study. *Journal of Transport Geography*, *57*, 194–206. https://doi.org/10.1016/j.jtrangeo.2016.10.011

Li, X., Pan, G., Wu, Z., Qi, G., Li, S., Zhang, D., Zhang, W., & Wang, Z. (2012). Prediction of urban human mobility using large-scale taxi traces and its applications. *Frontiers of Computer Science*, *6*(1), 111–121. https://doi.org/10.1007/s11704-011-1192-6

Li, Z., Huang, X., Ye, X., Jiang, Y., Martin, Y., Ning, H., Hodgson, M. E., & Li, X. (2021). Measuring global multi-scale place connectivity using geotagged social media data. *Scientific Reports*, *11*(1), Article 1. https://doi.org/10.1038/s41598-021-94300-7

Lin, G. C. S. (1999). Transportation and Metropolitan Development in China's Pearl River Delta: The Experience of Panyu. *Habitat International*, *23*(2), 249–270. https://doi.org/10.1016/S0197-3975(98)00049-6

Lin, Y., Huang, H., Gu, X., Xie, Y., Li, J., & Wang, C. (2022). *Identification and Evaluation of Urban Public Centers from Temporal and Spatial Perspectives: A Case Study of Six Districts in Guangzhou*. 594–607. https://doi.org/10.1061/9780784484562.065

Liu, K. (2021). *Exploring the relationship between functional urban polycentricity and the regional characteristics of human mobility: A multi-view analysis in the Tokyo metropolitan area*. 15.

Liu, K., Murayama, Y., & Ichinose, T. (2021). Exploring the relationship between functional urban polycentricity and the regional characteristics of human mobility: A multi-view analysis in the Tokyo metropolitan area. *Cities*, *111*, 103109. https://doi.org/10.1016/j.cities.2021.103109

Luca, M., Barlacchi, G., Lepri, B., & Pappalardo, L. (2021). *A Survey on Deep Learning for Human Mobility* (arXiv:2012.02825). arXiv. http://arxiv.org/abs/2012.02825

Mulligan, G. F. (1984). Agglomeration and Central Place Theory: A Review of the Literature. *International Regional Science Review*, *9*(1), 1–42. https://doi.org/10.1177/016001768400900101

Pacione, M. (2002). *The City: The city in global context*. Psychology Press.

Palla, G., Derényi, I., Farkas, I., & Vicsek, T. (2005). Uncovering the overlapping community structure of complex networks in nature and society. *Nature*, *435*(7043), 814–818. https://doi.org/10.1038/nature03607

Sassen, S. (2013). *The Global City: New York, London, Tokyo*. Princeton University Press.

Simini, F., Barlacchi, G., Luca, M., & Pappalardo, L. (2021). A Deep Gravity model for mobility flows generation. *Nature Communications*, *12*(1), Article 1. https://doi.org/10.1038/s41467-021-26752-4

Simini, F., González, M. C., Maritan, A., & Barabási, A.-L. (2012). A universal model for mobility and migration patterns. *Nature*, *484*(7392), 96–100. https://doi.org/10.1038/nature10856

Song, H., Xie, Y., Gu, X., & Wang, C. (2022). *Land Reserve Potential Evaluation Based on GIS and Big Data Analysis: A Case Study of Nansha District, Guangzhou, China*. 496–508. https://doi.org/10.1061/9780784484562.054

Statistical Bulletin of National Economic and Social Development. (2021).

Sun, J., & Hou, L. (2020). A Review for Urban Network Research. *Open Journal of Social Sciences*, *8*(6), Article 6. https://doi.org/10.4236/jss.2020.86031

Taylor, P. J., Hoyler, M., & Verbruggen, R. (2010). External Urban Relational Process: Introducing Central Flow Theory to Complement Central Place Theory. *Urban Studies*, *47*(13), 2803–2818. https://doi.org/10.1177/0042098010377367

Wang, F., Wei, X., Liu, J., He, L., & Gao, M. (2019). Impact of high-speed rail on population mobility and urbanisation: A case study on Yangtze River Delta urban agglomeration, China. *Transportation Research Part A: Policy and Practice*, *127*, 99–114. https://doi.org/10.1016/j.tra.2019.06.018

Wang, J., Hadjikakou, M., Hewitt, R. J., & Bryan, B. A. (2022). Simulating large-scale urban land-use patterns and dynamics using the U-Net deep learning architecture. *Computers, Environment and Urban Systems*, *97*, 101855. https://doi.org/10.1016/j.compenvurbsys.2022.101855

Wang, Z., & Chen, Y. (2022). *Exploring Spatial Patterns of Interurban Passenger Flows Using Dual Gravity Models* (arXiv:2201.04013). arXiv. https://doi.org/10.48550/arXiv.2201.04013

Winters, P. R. (1960). Forecasting Sales by Exponentially Weighted Moving Averages.

Management Science, *6*(3), 324–342. https://doi.org/10.1287/mnsc.6.3.324

Yang, J., Fang, C., Ross, C., & Song, G. (2011). Assessing China's Megaregional Mobility in a Comparative Context. *Transportation Research Record: Journal of the Transportation Research Board*, *2244*(1), 61–68. https://doi.org/10.3141/2244-08

Yangtze River Delta City Cluster Development Plan. (2016). https://www.ndrc.gov.cn/xxgk/zcfb/ghwb/201606/W020190905497826154295.pdf

Yeh, A. G.-O., & Chen, Z. (2020). From cities to super mega city regions in China in a new wave of urbanisation and economic transition: Issues and challenges. *Urban Studies*, *57*(3), 636–654. https://doi.org/10.1177/0042098019879566

Zhang, M., Kafy, A.-A., Xiao, P., Han, S., Zou, S., Saha, M., Zhang, C., & Tan, S. (2023). Impact of urban expansion on land surface temperature and carbon emissions using machine learning algorithms in Wuhan, China. *Urban Climate*, *47*, 101347. https://doi.org/10.1016/j.uclim.2022.101347

Zhang, P., Zhao, Y., Zhu, X., Cai, Z., Xu, J., & Shi, S. (2020). Spatial structure of urban agglomeration under the impact of high-speed railway construction: Based on the social network analysis. *Sustainable Cities and Society*, *62*, 102404. https://doi.org/10.1016/j.scs.2020.102404

Zhang, W., Fang, C., Zhou, L., & Zhu, J. (2020). Measuring megaregional structure in the Pearl River Delta by mobile phone signaling data: A complex network approach. *Cities*, *104*, 102809. https://doi.org/10.1016/j.cities.2020.102809

Zhang, Y., Zheng, X., Chen, M., Li, Y., Yan, Y., & Wang, P. (2021). Urban Fine-Grained Spatial Structure Detection Based on a New Traffic Flow Interaction Analysis Framework. *ISPRS International Journal of Geo-Information*, *10*(4), Article 4. https://doi.org/10.3390/ijgi10040227

Zhang, Y., Zheng, X., Helbich, M., Chen, N., & Chen, Z. (2022). City2vec: Urban knowledge discovery based on population mobile network. *Sustainable Cities and Society*, *85*, 104000. https://doi.org/10.1016/j.scs.2022.104000

Zhong, C., Arisona, S. M., Huang, X., Batty, M., & Schmitt, G. (2014). Detecting the dynamics of urban structure through spatial network analysis. *International Journal of Geographical Information Science*, *28*(11), 2178–2199. https://doi.org/10.1080/13658816.2014.914521