Functional urban area delineations of cities on the Chinese mainland using massive Didi ride-hailing records

Shuang Ma[a], Ying Long[b],*

[a] School of Architecture, Tsinghua University, Beijing 100084, China
[b] School of Architecture and Hang Lung Center for Real Estate, Tsinghua University, Beijing 100084, China

ABSTRACT

The problem associated with a city's administrative boundary being "under-" or "over-bounded" has become a global phenomenon. A city's administrative boundary city does not effectively represent the actual size and impact of its labor force and economic activity. While many existing case studies have investigated the functional urban areas of single cities, the problem of how to delineate urban areas in geographic space relating to large bodies of cities or at the scale of an entire country has not been investigated. This study proposed a method for FUA identification that relies on ride-hailing big data. In this study, over 43 million anonymized 2016 car-hailing records were collected from Didi Chuxing, the largest car-hailing online platform in the world (to the best of our knowledge). A core-periphery approach is then proposed that uses nationwide and fine-grained trips to understand functional urban areas in Mainland China. This study examined 4456 out of all 39,007 townships in an attempt to provide a new method for the definition of urban functional areas in Chinese Mainland. In addition, four types of cities are identified using a comparison of functional urban areas with their administrative limits, and a further evaluation is conducted using 23 Chinese urban agglomerations. With the rapidly increasing use of internet-based ride-hailing services, such as Didi, Grab, Lyft, and Uber, globally, this study provides a practical benchmark for the delineation of functional urban areas at larger scales.

1. Introduction

The problem associated with a city's administrative boundary being "under-" or "over-bounded" has become a global phenomenon. A city's administrative boundary city does not effectively represent the actual size and impact of its labor force and economic activity. For instance, a large number of residents live in Yanjiao and Sanhe in Hebei Province, China, but they commute to central Beijing for work. Such a commuting pattern inevitably generates substantial influences on the economy, housing, and the environment, and these influences frequently take place beyond the traditional administrative boundary of a city. The concept of functional urban areas (FUAs or metropolitan areas or functional regions), where cities, towns, and rural areas are socio-economically tied to the urban core, is an important tool for understanding economic and political regions of urban areas to address the aforementioned challenges and problems. FUAs offer more precise boundaries for urban planning and management that will assist with infrastructure construction, regional coordinated development, and environmental governance of projects, such as highway construction and other essential public services. FUAs are also imperative to digest

statistical information. For example, the U.S. Office of Management and Budget uses FUAs to delineate housing subsidies, wage levels, and medical subsidies (Zhang, 2010). In addition, understanding the geographic distribution of FUAs will clarify human interactions with the environment and present a new relationship between urbanization and the biosphere (Reba, Reitsma, & Seto, 2016).

The delineation of FUAs is a hot topic, and economics, human geography, and remote-sensing researchers have attempted to develop alternative approaches to define FUAs (Bode, 2008; Chen et al., 2016; Chi, Jiao, Dong, Gu, & Ma, 2016; Duranton, 2015; Newman, 2004; Xie & Ning, 2005). International organizations, such as the European Spatial Planning Observation Network (ESPON), the Organization for Economic Cooperation and Development (OECD), and the Word Bank, also delineate FUAs using global data sets of cities. Economists believe that cities contribute economic benefits because of their labor markets (Marshall, 1890), thus commuting patterns would provide a better picture of the dominant urban economic geography for the delineation of FUAs (Duranton, 2015; Newman, 2004; Ratti et al., 2010; Shen & Batty, 2019). For example, the U.S. Census Bureau (USCB) explicitly defines FUAs as counties where at least 25% of the workers living in the

---

Table 1
Key publications that describe FUA delineations.

Publication	Study area	Method	Data	Spatial unit
Shen & Batty, 2019	London	Multilevel modularity optimization algorithm that detects community structures within weighted flow graphs.	Disaggregate commuting data	Middle-layer Super Output Area
Bosker et al., 2018	Indonesia	Systematic comparisons of these approaches using different data.	Commuting flows, remotely sensed nighttime lights, and population	Subnational administrative units
Houssou, Guillaume, & Prigent, 2019	Porto	A graph-based approach with community detections that combine the use of POIs.	Taxi flows and POI data sets	A cartographic division using road networks.
Zhang, Wang, Liu, Li, & Pang, 2016	Shenyang and Beijing	A Convolutional Neural Network (CNN) based method used to classify aerial images.	Remotely sensed images	Disjoint regions using major roads
Arcaute et al., 2015	21 cities in England and Wales	Considers density-based cities and then adds areas to cities according to a commuting threshold.	Commuting and population density	Wards
Yuan et al., 2015	Beijing	A topic model-based approach to fuse POIs and mobility patterns.	Data sets for POIs, taxi trajectories, buses, and subways	Disjoint regions using major roads
Zhu, Yang, Zhong, Seiter, & Tröster, 2015	Seattle	A supervised learning via multi-output regression for travel survey data.	Crowd-Augmented Travel Survey Data	Census cells
Demsar, Reades, Manley, & Batty, 2014	Greater London	Edge-based community detection algorithm that uses the calculation of the similarity between each pair of vertices.	Taxi flows	Traffic Analysis Zones (TAZ)
Dijkstra & Poelman, 2014	European Union	Approach associated with dense clusters of populations.	Population grid	LAU2s
Gajović, 2013	Kraljevo in Serbia	Two methods of unsupervised learning: k-means clustering and self-organizing maps in machine learning.	Distances, census and some calculated indexes	Administrative or statistical territorial units
Farmer & Fotheringham, 2011	Ireland	Network-based method for network of travel-to-work flows with geographic weighting.	Travel-to-work data	Electoral district (EDs)
Ratti et al., 2010	Great Britain	A fine-grained approach based on analyzing networks of individual human transactions.	Telephone calls	9.5 km by 9.5 km pixels
Karlsson & Olsson, 2006	Fyrstad in Sweden	The local labor market, the commuting zone, and the accessibility approach	Commuting flows	Municipality

county work in the central county or counties of the Core Base Statistical Areas (USCB, 2014). ESPON used 10% as the commuting threshold to delineate FUAs in Germany (ESPON, 2007). However, for many countries, especially developing countries, commuting data are difficult to gather. Although a few indicators that reflect functional influences, such as travel distance, travel time, and road networks, can imply commuting conditions (Korcelli, 2008, OMB, 2010; ESPON, 2007), commuting flow data are the preferred approach used to delineate FUAs as functional local labor markets (Bosker, Park, & Roberts, 2018).

The development of the internet of things (IOT), the continued development of information and communications technology (ICT), and the expansion of big data create new opportunities to observe the socioeconomic composition of the territories and track human movements (Shen, Sun, & Che, 2017; Yuan et al., 2015). In particular, the fast growth of mobile internet, cloud computing, and location-based services make the ride-hailing companies, Didi Chuxing, Grab, Lyft, and others, scale users around the world, and these companies serve millions of users per day (Pham et al., 2017). Data sets from platforms such as Uber, API, and Lyft have become increasingly popular in the recent literature (Chaudhari, Byers, & Terzi, 2018; Dong, Wang, & Zhang; Schaller, 2017; Agatz, Erera, Savelsbergh, & Wang, 2011). They collect and share vast amounts of information, and these platforms can provide a new lens to delineate FUAs using a community detection algorithm. In 2018, Ford, Uber, and Lyft joined the SharedStreets project to provide a cloud platform to share data from public and private entities to help cities make more informed decisions.

This study proposed a method for FUA identification that relies on ride-hailing big data. Access was provided to ride-hailing data (including both privately owned and corporate taxis registered on the Didi Chuxing platform) from three working days by the Didi Chuxing Company (the largest online ride-hailing platform in China) in 2016. We aggregated orientation (O) township and destination (D) township of every anonymous traveler into a Geographic Information System (GIS) platform. For identification of FUAs, the following was concluded: (1) The core areas should be defined on the basis of urban built-up areas. Thus, a township with a sufficient urban built-up area should be a candidate for a core FUA. (2) If a township had high flow density, this township should be a considered part of the core FUA. (3) If an adjacent township had high interaction with the core areas, there should be a large percentage of residents commuting to core areas. Specifically, this work (1) analyzed FUAs for all of the Chinese mainland based on massive Didi ride-hailing records; (2) proposed suggestions to adjust the administrative boundaries of cities on the Chinese mainland by comparing with defined FUAs; (3) evaluated urban agglomeration development by evaluating the FUA rate in every urban agglomeration; and (4) compared FUA delineations using Didi ride-hailing records with FUA delineations using previous approaches.

This approach provided a straightforward way to identify FUAs. With the increasing availability of open travel data sources and the increasing possibility of shared ride-hailing data, this approach can be applied for the widespread delineation of FUAs.

This paper is organized as follows. Section 2 reviews the primary literature regarding approaches for FUA delineations. In Section 3, the study area, data, and methodology used to identify FUAs are discussed. Results are analyzed in Section 4, including a delineation of FUA patterns in China. The potential applications are discussed in Section 5, including suggestions for administrative boundary adjustments and an evaluation of urban agglomeration development. The final section compares the delineation results using previous approaches and discusses the Didi records, research biases, and potential policy applications of this study.

2. Literature review

The literature has several definitions of FUAs. Generally, FUAs are regions where socioeconomic interactions are stronger within their

boundary than other boundaries. According to the OECD, an FUA is a territorial unit that results from the organization of social and economic relations in that its boundaries do not reflect geographic particularities or historic events (OECD, 2002). The FUA is also described as a geographic region in which within-region interactions in terms of commuter travel and work flows are maximized, and between-region interactions are minimized (Farmer & Fotheringham, 2011). As such, commuting patterns over space are fundamental when delineating FUAs (Fotheringham & O'Kelly, 1989). The study areas, data used, methods, and spatial units used in the current literature are summarized in Table 1.

The literature proposes several approaches to capture FUAs by inferred commuting (Cörvers, Hensen, & Bongaerts, 2009; Ratti et al., 2010) or observed commuting (Arcaute et al., 2015; Houssou et al., 2019; Shen & Batty, 2019). Among the inferred community approaches, local economic outcomes, telephone calls, Twitter networks, the accessibility of public transport, and other indicators have been used to indicate inferred commuting patterns (Korcelli, 2008; Mulíček, 2013; Takhteyev, Gruzd, & Wellman, 2012). Observed commuting data are better and more straightforward to accurately capture interactions. Farmer and Fotheringham (2011) delineated the FUAs in Ireland by observing a travel-to-work flow network. They maximized the modularity of a network of travel-to-work flow to produce a regional boundary that exhibited less interaction than expected between regions. Inspired by Goddard (1973), Demšar investigated the possibility of using an edge-based community detection algorithm (Demsar et al., 2014) to identify overlapping FUAs using taxi flows in the greater London area. Other popular approaches have been based on satellite imagery of urban built-up areas or light nights (Zhang et al., 2016; Elvidge, Kihn, & Davis, 1996; Danko, 1992;) or a cluster algorithm based on population data sets, such as LandScan or GHS-Pop (Bosker et al., 2018; Dijkstra & Poelman's, 2014). The use of these more accessible data sets as an alternative method, however, is not sufficient as the direct use of commuting patterns, because the final aim of an FUA delineation is to capture the social and economic effects using commuting patterns.

Most current publications have focused on one or two cities or a typical region as the spatial unit of small administrative unit or road network. To delineate FUAs at a broader level, Dijkstra and Poelman (2014) developed the cluster algorithm to identify a metro area as a dense population cluster. The benefits of this approach include greater comparability among different countries in the European Union and data availability. The collection of commuting data is needed, however, before studying the interactions among boundaries of cities.

In this study, we proposed a new approach for FUA delineations for the Chinese mainland, which currently lacks well-defined FUAs that use massive ride-hailing data. This study addressed current limitations to conduct delineations for the entire country with the use of available ride-hailing data at the township scale, which is the smallest scale that has been used to delineate FUAs in China. With the rapidly increasing use of internet-based ride-hailing services, such as Didi, Grab, Lyft, and Uber, globally, this study provides a practical benchmark for city planning and statistical analyses.

3. Methods

3.1. Study area and data

The Chinese city system has long been defined from an administrative view, and most of the statistical data that have been gathered have corresponded to administrative cities (Long, 2016). According to the Ministry of Housing and Urban-rural Development of the People's Republic of China (MOHURD, 2014), China has a total of 654 cities. The study area chosen for this investigation covered the entire Chinese mainland territory. The study area was not limited to administrative areas of cities (city proper). Instead, administrative boundaries were

used to compare the delineated FUAs for classification and boundary adjustments. The resolution of an administrative division according to the definition of FUAs for the Chinese mainland was a township.

We collected anonymized internet-based ride-hailing records from Didi Chuxing for three workdays (Wednesday to Friday) from August 24 to 26, 2016. Didi is the largest ride-hailing service company in China and is one of the largest on-demand ride-sourcing service platforms in the world (Dong, Wang, Li, & Zhang, 2018; Shih, 2015). With continued increases in innovative transportation services, internet-based ride-hailing services are competitive with traditional taxi services. Previous research has found that the travel proportion of ride-hailing services occupies 5.5% of all forms of travel modes in China, which is 1.7% higher than traditional taxi services. The Didi ride-hailing service occupies 85.7% of the total ride-hailing trips, with other services, such as Shezhou and Yidao, occupying the remaining 14.3%. In addition, metro and bus information do not directly reflect the socioeconomic interactions between cities, because there is no consistent long-distance and intercity metro and only a few intercity transits during certain periods of time. In contrast, the average number of ride-hailing requests per day during the commuting period is approximately 130,096, which accounts for almost one-third of the total requests per day. Thus, the internet-based ride-hailing records from Didi Chuxing are representative of actual travel trips. Hence, they can be used to delineate FUAs according to their massive commuting data to accurately reflect the socioeconomic interactions across cities, which other travel data cannot reflect. Didi Chuxing now has approximately 4.5 billion registered users and 25 million daily rides, with a huge capacity for growth (Didi Chuxing, 2018; Dong et al., 2018). Currently, the Didi company provides four main types of travel services: taxi, express, private car, and hitch. The ride-hailing data used for this study included all four of these modes. The data reliably demonstrate the commuting patterns from a new perspective, even though it is not a thorough perspective.

The data set used in this study consisted of 43,846,160 records that covered 39,007 townships in a total of 654 cities on the Chinese mainland. These data further distinguished between commuting time records and total records using a defined commute time of 7:30 to 9:00, which has been the practice in many previous studies (Wang, Huo, He, Yao, & Zhang, 2008; Wöhrnschimmel et al., 2008; Zong, Juan, & Liu, 2007). Each record contained an origination township and a destination township. The following figure illustrates the ride-hailing records (traffic flows) at commuting time throughout the entire study area (Fig. 1).

China has three forms of township-level administrative units: Jiedao (subdistrict), Zhen (town), and Xiang (township). Jiedao is the basic administrative unit of a city proper and is used primarily for cities. Its counterpart in rural areas is Zhen or Xiang. In this study, we used the term "township" to represent all types of township-level administrative units in China. The township is the basic unit used to delineate FUAs to maintain a continuity of statistics used in population, economic, and social development analyses. We recorded the orientation and destination of every travel flow in all 39,007 townships in China. Fig. 2 illustrates the townships.

We gathered a data set of 23 officially recognized urban agglomerations in China from the 2010 Urban Agglomeration Development Report in China issued by the Institute of Geographic Sciences and Natural Resources Research (IGSNRR). We used the data to evaluate the development quality of the urban agglomerations. Fig. 3 illustrates the location and pattern of all the urban agglomerations.

We retrieved the 2015 urban built-up areas from the VIIRS/Day Night Band (DNB) (the updated version of DMSP/OLS) with a spatial resolution of 500 m. The urban area for the entire Chinese mainland was 72,208 km^2 in 2015, and this included 28,416 patches (Fig. 4). The recommendation of this study is that urban built-up areas should be used as the basis for identifying the core areas of FUAs in all territories of the Chinese mainland.
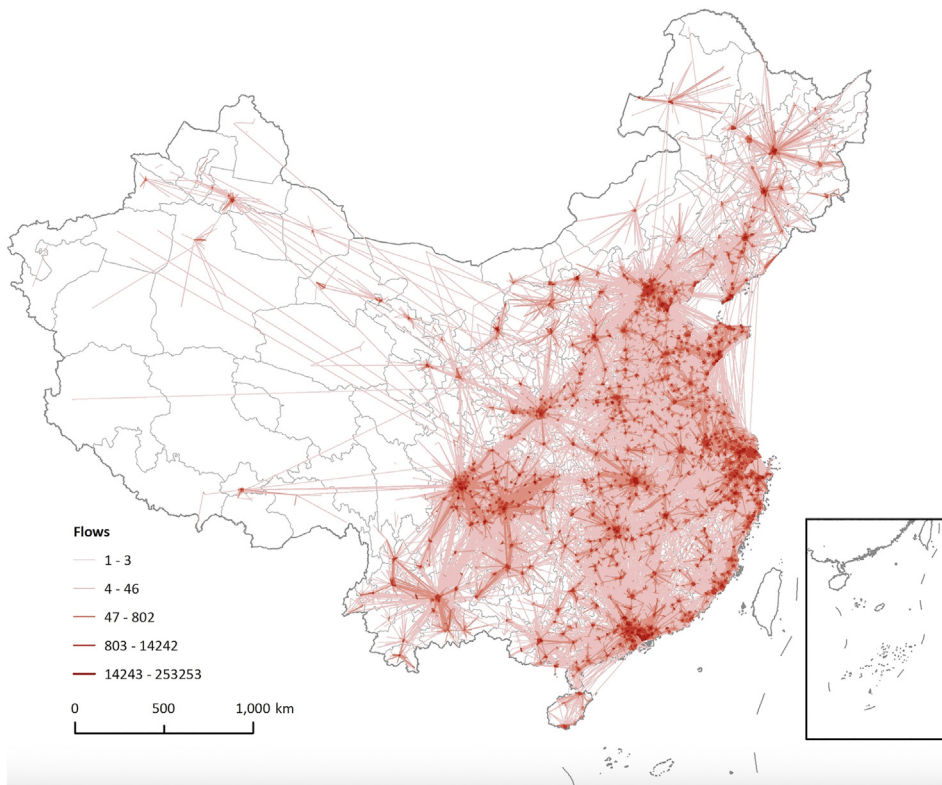
Fig. 1. The traffic flows throughout China according to Didi ride-hailing records.

3.2. Delineation method

In general, FUAs are composed of an economic unit characterized by densely populated "urban cores" and less-populated neighboring "commuting zones" that are socioeconomically tied to the urban cores (Arcaute et al., 2015; Gajović, 2013). In this study, we proposed a straightforward method for delineating FUAs that consists of pre-determined core areas and aggregated peripheral areas. Fig. 5 shows
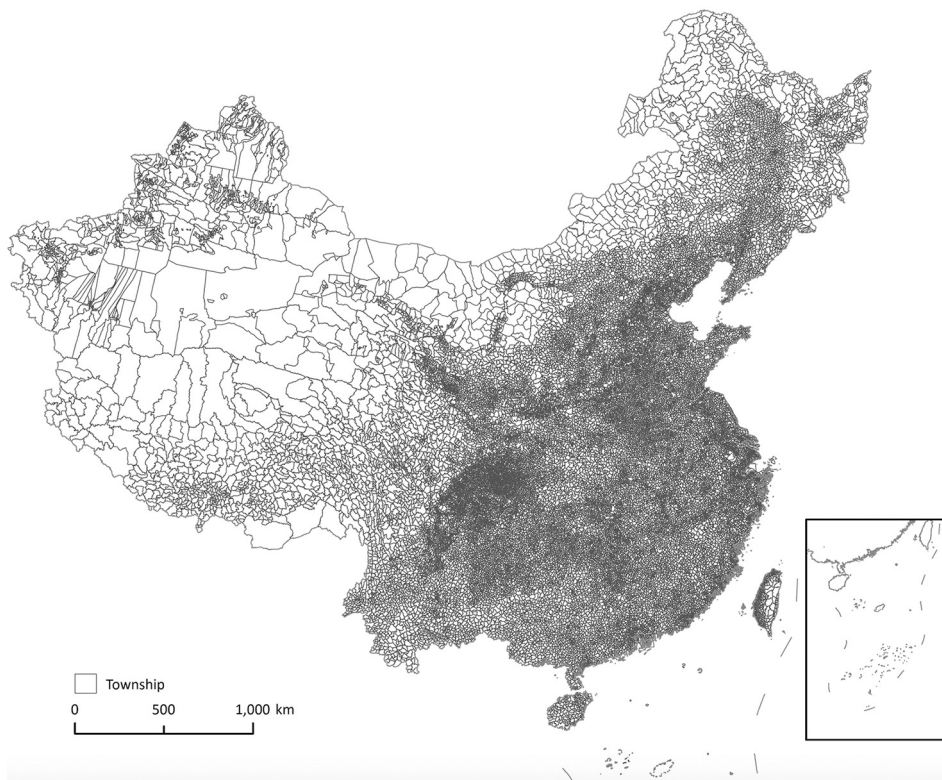

Fig. 2. Townships in China.

Fig. 3. The boundaries of the 23 officially recognized urban agglomerations in China.

the specific steps to delineate an FUA.

The potential cores of the FUAs were required to be "big enough" and "dense enough" (Duranton, 2015), supported by the idea that the core areas had to be townships that (a) had at least 40% urban built-up areas within their boundaries, or (b) had at least a 100/km² of flow density. We calculated flow density using the flow volume (flow in and flow out)/township area in each township (Fig. 5i). A parameter for 40% urban built-up areas within townships has been used in previous studies to define physical city systems in China as "big enough" (Ma & Long, 2019). In this study, we proposed a new approach for delineating FUAs that is performed at a fine-grained large scale that can fill in the blanks of current FUA delineations in China. The threshold—that is, the 100/km² flow density—did not have a theoretical basis in this method. This was determined by field studies in pilot cities and by the authors' personal experiences. The periphery areas were defined as those townships that had a high degree of socioeconomic integration with the core area that could be measured using commuting ties. In this study, we explicitly defined these areas as townships that (a) had commutes that were as long as a threshold of 10 commuting flow volume (only flow out) and (b) had at least 15% of the residents living in the townships that traveled to core areas to work. This percentage was evaluated using the flow volume (only flow out) to FUA cores/total flow volume (flow in and flow out) in townships during the commuting time period (Fig. 5ii). Previous studies widely used a 15% commuting figure as the threshold to incorporate a hinterland municipality into an FUA (OECD, 2002), but in some cases, this threshold has been used as an alternative. For example, Glaeser et al. (2001) stated that less than 10% of employment in a few U.S. cities such as Los Angeles is concentrated within 5 km of their centers. The threshold of 10 commuting flow volume was defined using observation and experience. In the future, commuting thresholds must be defined relative to the number of workers, and a relative threshold should be used whenever possible.

4. Results

In total, we delineated 308 FUAs that cover 4456 townships out of 39,007 townships in China (Fig. 6i). In general, the delineated FUAs exhibited a similar pattern as land urbanization in China (Fig. 4). Table 2 lists the top 10 FUAs in size to demonstrate the urban area hot spots in China from the FUA perspective. The FUAs of Guangzhou_Shenzhen, Beijing_Langfang, and Shanghai_Suzhou (only the two largest administrative city names were used to title these FUAs) are the top three largest FUAs located in the three most developed urban agglomerations: the Pearl River Delta Urban Agglomeration, the Beijing-Tianjin-Hebei Urban Agglomeration, and the Yangtze River Delta Urban Agglomeration, respectively. Fig. 6ii–iv illustrate these three FUAs and the FUAs patterns in their particular urban agglomerations.

The delineated FUAs in the Pearl River Delta Urban Agglomeration are the most integrated of the three urban agglomerations. The Guangzhou_Shenzhen FUAs include partial territories of eight administrative cities: Guangzhou, Shenzhen, Foshan, Huizhou, Zhongshan, Dongguan, Sihui, and Zengcheng. These FUAs occupy 80.7% of all of the FUAs in the Pearl River Delta Urban Agglomeration. Only two other independent FUAs surround the city of Zhuhai and Conghua in this urban agglomeration. In the Beijing-Tianjin-Hebei Urban Agglomeration, we identified 566 townships with a total area of
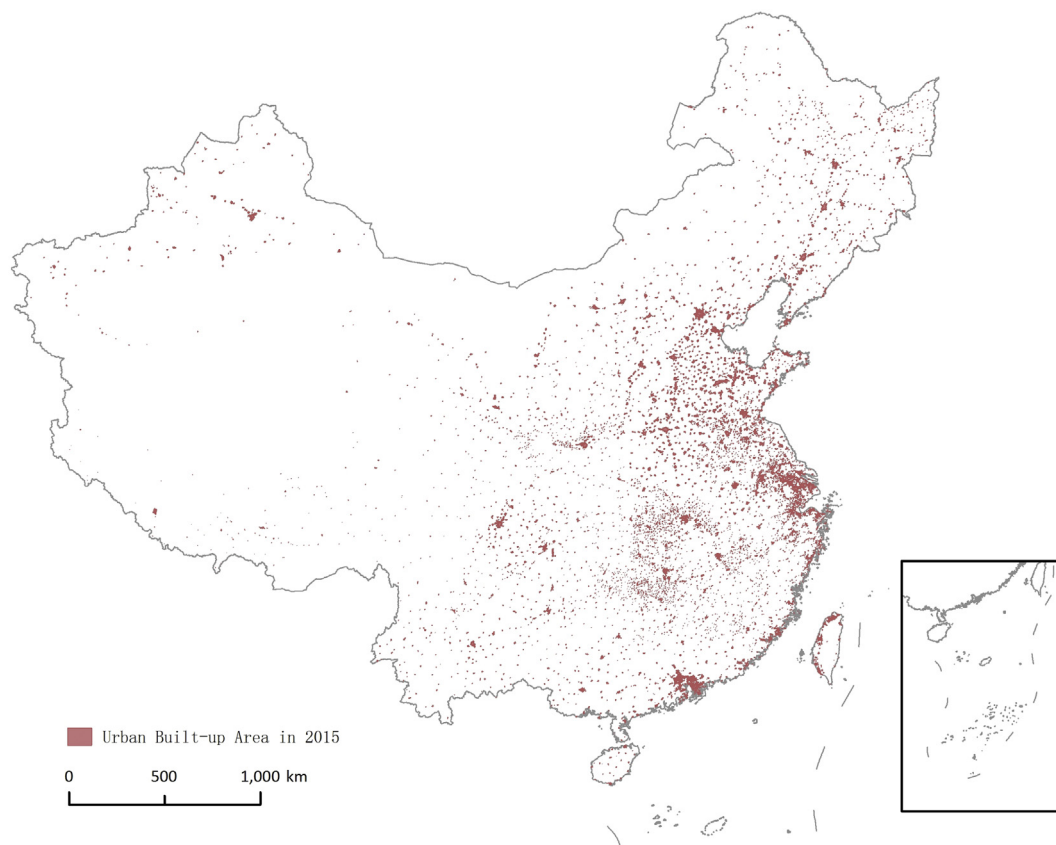
Fig. 4. The 2015 urban built-up areas.

21,580.9 km² as FUAs. The most continuous FUAs are the Beijing_Langfang FUAs, which consist of 264 townships covering 10,882.7 km² This size consists of more than half (50.4%) of the FUAs in this urban agglomeration. The other main FUAs are situated around the city of Tianjin, Tangshan, and Shijiazhuang. We identified 38,026.8 km² as FUAs, including 833 townships in the Yangtze River Delta Urban Agglomeration. In this urban agglomeration, the Shanghai_Suzhou FUAs occupy 25.3% of the total FUAs in the Yangtze River Delta. Another large FUA is the Hangzhou_Shaoxing, which is also one of the top 10 largest FUAs in China. This FUA occupies 25.3% of the total FUAs and is composed of Hangzhou, Shaoxing, and three county-level cities: Haining, Lin'an, and Fuyang. Other large FUAs in the Yangtze River Delta are situated around Nanjing, Ningbo, and Jiaxing, and these three large FUAs are 5052.3 km², 3137.1 km², and 1899.2 km,² respectively.

5. Potential applications

5.1. Comparison of delineated FUAs with the existing administrative boundaries of Chinese cities

Chinese cities are not each a "city" in the strictest sense of the term, but rather are an administrative unit composed of both urban areas and the surrounding rural areas. The delineation of FUAs creates a new lens for understanding Chinese cities and has applications to assess the potential for city administrative boundary adjustments that will accommodate appropriate urban planning and management schemes.

In this study, we evaluated the administrative boundaries of 36 Chinese cities, including municipality level cities, provincial capital cities, and subprovincial cities. The average difference between the administrative size and total FUA size of China's most important 36 (excluding Shanghai, Guangzhou, and Shenzhen) cities is 2572 km². In Chongqing, Urumchi, Chengdu, and Beijing, the administrative size is

more than 5000 km², which is a mismatch to the size of their FUAs. We classified these cities into four categories according to the spatial relationships between the administrative boundaries and the delineated FUAs (see Table 3 for a profile and Fig. 7 for examples). Type 1 includes the cities of Dalian and Nanjing. The administrative boundaries of these cities are good examples for which the administrative cities and their FUAs are consistent. Type 2 illustrates cities similar to Shijiazhuang, Chengdu, Changsha, Nanchang, Hefei, and Zhengzhou that should expand their administrative boundaries, which are quite smaller than the delineated FUAs. The FUAs in Type 3, such as in Tianjin, Haikou, and Shenyang, are smaller than their administrative boundaries, indicating that these cities are associated with very large hinterlands. This condition is not rare in China considering the frequent spatial adjustments of administrative cities in the recent two decades, which have resulted in the expansion of these Chinese cities from the administrative dimension. To some degree, the administrative boundaries of these cities should shrink. Type 4 is the most common type of the four types in this comparison. In this category, city boundaries and urban planning schemes can shrink in some directions and expand in other directions for the purpose of effective urban management.

5.2. Evaluation of the development of urban agglomerations in China using the delineated FUAs

A Chinese urban agglomeration is officially defined as a megalopolis, and government policies aim to make areas in every urban agglomeration connect tightly by enhancing transportation links, creating industrial cooperation, and defining more efficient divisions of labor (Wikipedia, 2019). Urban agglomerations need to contribute to the national population and economy, and they need to be treated as a core form in the next round of economic growth (Fang & Yu, 2017). In 2010, China's 10 largest urban agglomerations accounted for slightly more than one-fifth of total export originations from urban areas.
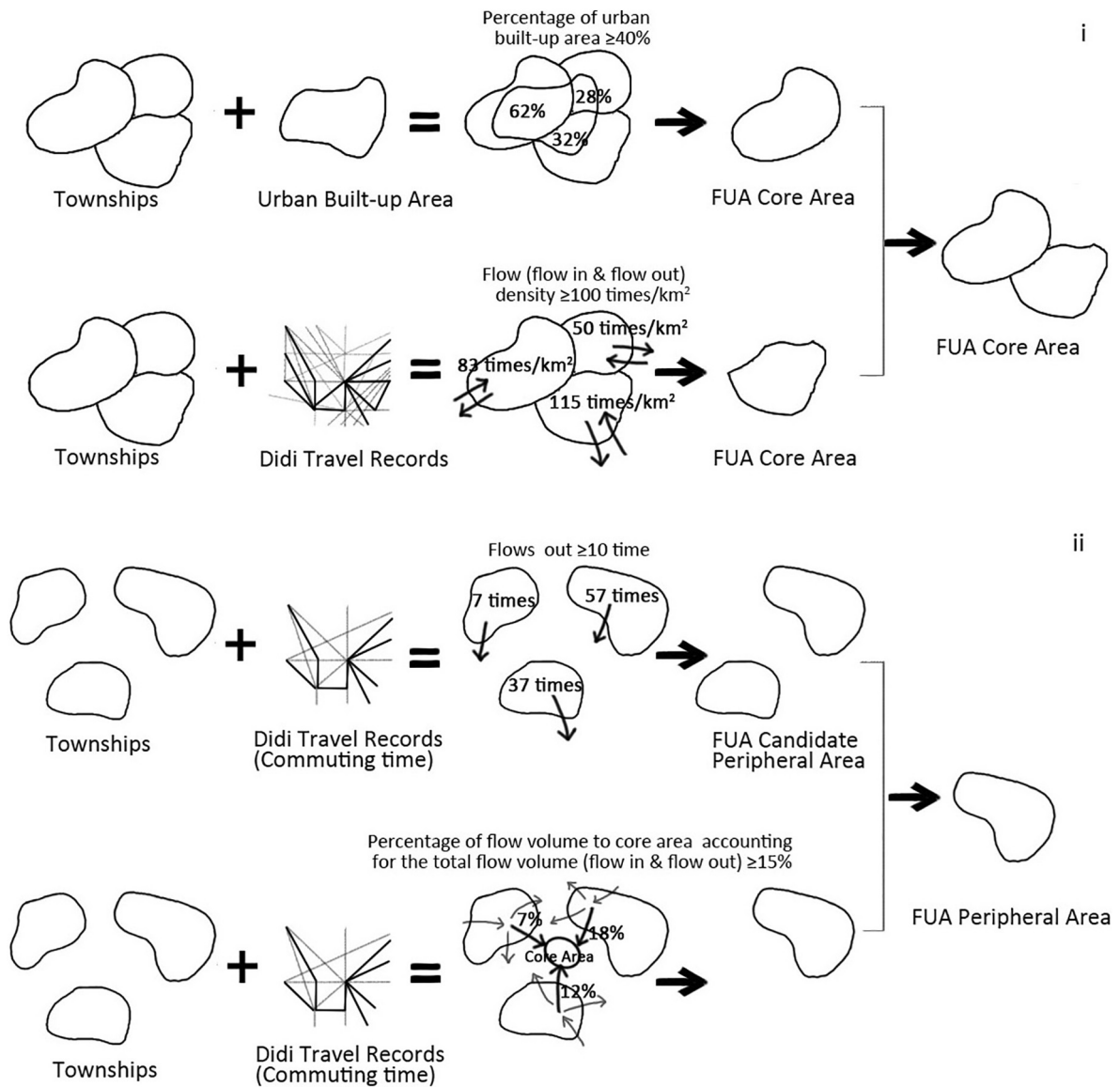
Fig. 5. The FUA delineation method using massive Didi ride-hailing records.

According to research by the National Bureau Statistics of China, the Yangtze River Delta Urban Agglomeration, the Pearl River Delta Urban Agglomeration, and the Beijing-Tianjin-Hebei Urban Agglomeration are the most imperative centers for Chinese finance, trade, and industry. In 2015, together they constituted only about 5.2% of China's land area, but 23% of the country's population and 39.4% of the country's GDP (Yuan & Ye, 2018).

An urban agglomeration is a dynamic system with complex capital interactions, information, populations, and cargo flows. The quantitative evaluation of urban agglomerations is not a well-defined task. FUAs are integrated in geographic space, and they are useful to measure the development quality of urban agglomerations using the FUA ratio, which indicates the extent and pattern of social and economic interactions in an urban agglomeration. An urban agglomeration that has a greater FUA ratio means that it has relatively high economic functions, productivity levels, and geospatial relationships. Fig. 8 shows the FUA ratios for the Chinese mainland officially recognized urban agglomerations. The advantages of the Yangtze Delta, Pearl River Delta, and Shandong Peninsula are obvious in terms of these ratios. The FUA ratios for the eastern coastal agglomerations are relatively high as well, such as in the Western Taiwan Straits Urban Agglomerations. The ratios for a few central regions and most of the western regions of the FUA areas are significantly lower. For instance, in the Northern Tianshan Mountains, Lanzhou-Xining-Baiyin, Central Guizhou, and Harbin-Daqing-Changchun, the FUA ratios are less than 8%. This trend is expected because of the long period of dual development in China in terms of the economy and society. Some urban agglomerations in central China, such as the Beijing-Tianjin-Hebei, Central Plains, Wuhan, and Central Shanxi urban agglomerations, have stable and medium FUA ratios. These urban agglomerations with very limited FUA ratios should be considered for boundary adjustments to better represent their actual development stage.

6. Discussion and conclusions

6.1. Discussion

Wang and Ye (2004) attempted to delineate FUAs in the absence of an official definition for FUAs in China. They followed the procedures used in a study performed by Hu, Zhou, and Gu (2000), and chose cities proper and counties as spatial units. In this study, we used non-agricultural indicators, as opposed to commuting patterns, to identify
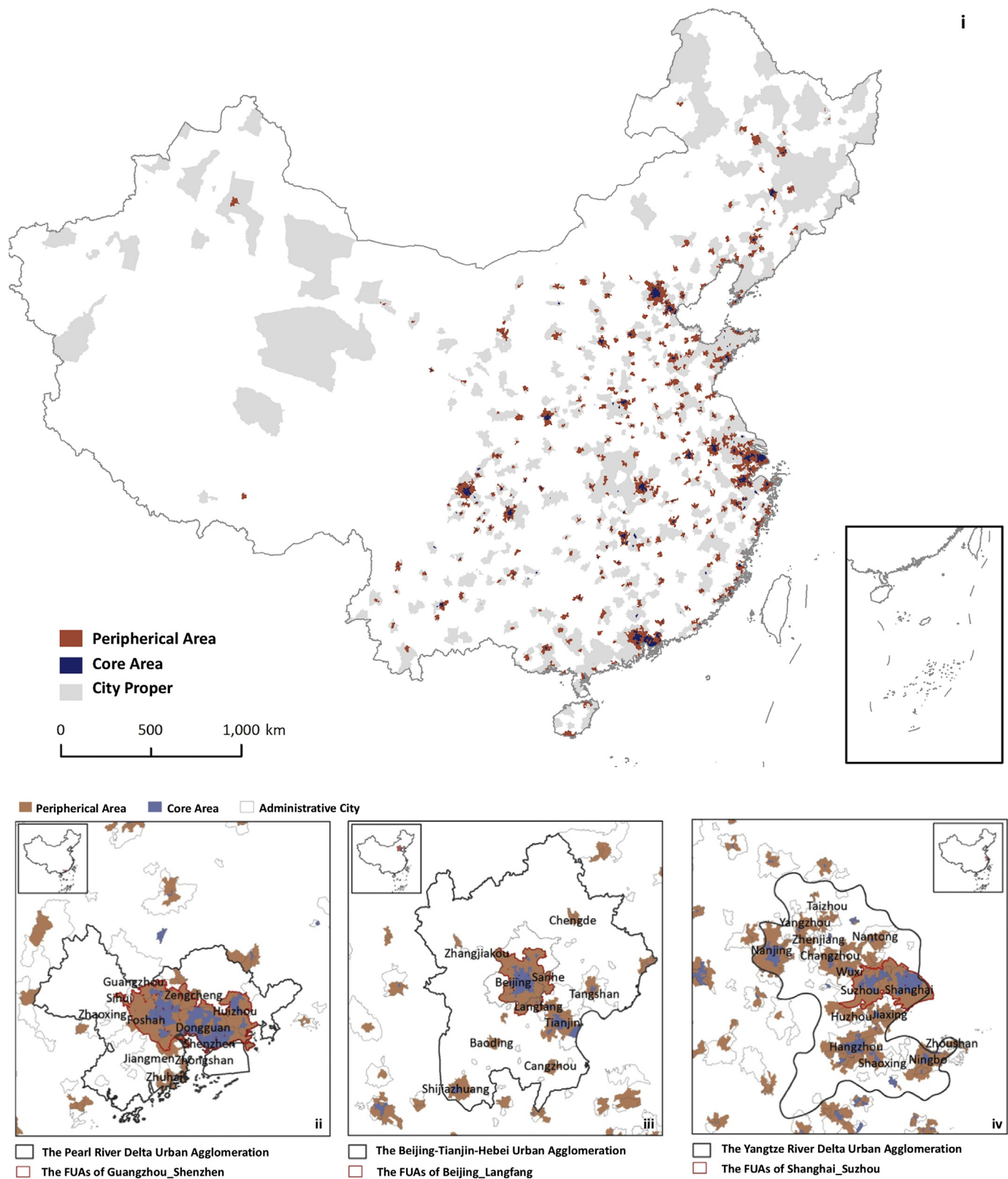
Fig. 6. Delineated FUAs in China. (i) FUAs of the entire territory of the Chinese mainland, (ii) FUAs of Guangzhou_Shenzhen and the FUA pattern in the Pearl River Delta Urban Agglomeration, (iii) FUAs of Beijing_Langfang and the FUA pattern in the Beijing-Tianjin-Hebei Urban Agglomeration, and (iv) FUA of Shanghai_Suzhou and the FUA pattern in the Yangtze River Delta Urban Agglomeration.

integrated and interconnected areas on the Chinese mainland.

They collected data from the Chinese Urban Statistics Yearbook (2001) and the Economic and Social Statistics Yearbook (2001) that

included the nonagricultural population ratio (P), the nonagricultural GDP ratio (A), the per capita GDP (G), and the population density (D). The former two data sets represented the economic connections

Table 2

Top 10 largest FUAs in China.

FUAs	Area in km^2	Townships amount
Guangzhou_Shenzhen	13,716.6	206
Beijing_Langfang	10,882.7	264
Shanghai_Suzhou	9626.8	265
Chengdu_Deyang	7323.8	214
Hangzhou_Shaoxing	6377.9	146
Wuhan	5045.1	39
Chongqing	4920.7	62
Nanjing	4481.7	88
Tianjin	4331.9	133
Xi'an_Xianyang	4005.9	91

between urban and rural areas, whereas the later data sets reflected the urbanization level of a spatial unit. F represented the value of a chosen spatial unit (cities proper or counties) used to identify FUAs. It was calculated using the following equation: $F = [(1 - P)/0.6)] \times [(1 - A)/0.75] \times (G/7000) \times (D/400)$. The thresholds for the nonagricultural population ratio, the nonagricultural GDP ratio, the per capital GDP, and the population density were 0.6, 0.75, 7000 CNY, and 400/km^2, respectively. When the F value in a chosen spatial unit was greater than 1, we identified this unit as an FUA. Specifically, once we found the identified FUA to be located in a city proper of a prefecture-level city with more than 200,000 in a nonagricultural population, we identified it as a core area of the FUA. Fig. 9 illustrates the outcome. The FUAs were primarily distributed in the eastern part of China or in the industrial cities in Inner Mongolia, Ningxia, Gansu, and Xinjiang in western China. The two largest FUAs were in the Yangtze Delta Urban Agglomeration and the Pearl River Delta Urban Agglomeration. The

Harbin-Daqing-Changchun Urban Agglomeration situated in the northern part of China also had a high percentage of FUAs.

A comparison of the delineation results of this study to that of Wang and Ye (2004) revealed that 30.1% of FUAs were identified by both researchers. Researchers in this field believe that FUAs across China have increased in the past 10 years. This growth is despite the fact that partial shrinkages in old industrial cities also have occurred, such as in the city of Yinchuan and Jingdezhen and in the Harbin-Daqing-Changchun Urban Agglomeration, which was the largest Chinese old industrial base that suffered from industrial structural transformation and resource exhaustion. The FUAs in the urban agglomerations of the Yangtze Delta, the Pearl River Delta, the Beijing-Tianjin-Hebei, Shandong Peninsula, Chengdu-Chongqing, and the Central Shanxi all have expanded as a result of accelerated economic centralization.

In a comparison of this approach with that of Wang and Ye (2004), we concluded that this approach was more straightforward and reliable in the following aspects. First, the study of Wang and Ye (2004) relied heavily on statistical yearbooks that were officially issued. The quality of the data among cities is inconsistent, however, because the indicators from the yearbooks were collected from local statistical departments using a bottom-up method and because professional staff training and levels of technological development vary across different cities, (Long, 2016). Second, as Long (2016) has mentioned, a large number of towns or downtown areas in counties should be regarded as cities according to the general definition of built-up areas. These areas are not addressed in the yearbooks, however, and they are termed as "invisible Chinese cities" or "neglected Chinese cities." Thus, using the administrative area to study FUAs is problematic. Moreover, county and cities proper are too large as spatial units for FUA delineations. The average size of an administrative unit at the county level is 3364 km^2 in China, which is

Table 3

Profiles of the four types of cities according to the spatial relationships between their administrative boundaries and FUAs.

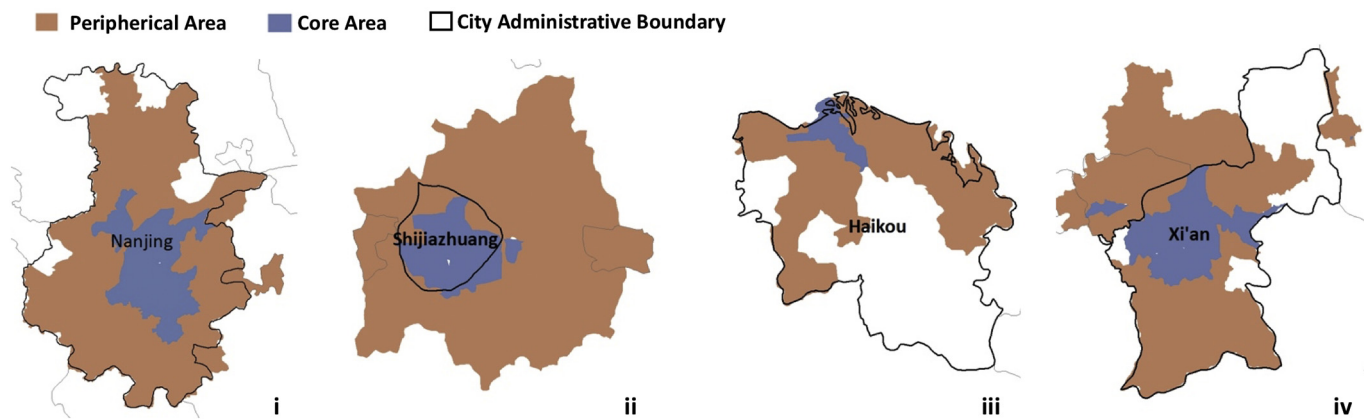| Type	City	Core Area (km^2)	Periphery Area (km^2)	Total FUAs (km^2)	Administrative Areas (km^2)
Type 1	Nanjing	825	4223	4482	4739
FUAs that agree quite well with the city administrative boundaries	Dalian	527	1385	1912	2358
Type 2	Guangzhou_Shenzhen	4715	9002	13,717	3545_1,966
City administrative boundaries that should expand	Chengdu	1485	5839	7324	2184
	Zhengzhou	476	2008	2484	1058
	Hefei	608	1695	2303	907
	Changsha	458	1834	2292	1790
	Nanchang	147	1692	1839	545
	Shijiazhuang	214	1618	1832	248
Type 3	Wuhan	1048	3997	5045	8548
City administrative boundaries that should shrink	Chongqing	1066	3855	4921	27,181
	Tianjin	1206	3126	4332	7191
	Harbin	195	2062	2257	7033
	Shenyang	221	1725	1946	3445
	Nanning	132	1703	1835	6439
	Urumchi	0	1491	1491	9427
	Haikou	96	1051	1147	2188
	Lanzhou	0	358	358	1667
Type 4	Beijing	2236	8647	10,883	16,404
City administrative boundaries that should shrink in some directions	Shanghai	2674	6953	9627	5132
and expand in other directions	Hangzhou	1222	5156	6378	3236
	Xi'an	777	3229	4006	3594
	Qingdao	838	2905	3743	3005
	Ningbo	370	2767	3137	2345
	Taiyuan	368	2017	2385	1445
	Jinan	215	2210	2425	3317
	Yinchuan	47	2232	2279	1632
	Changchun	756	1086	1842	3606
	Kunming	320	1424	1744	4616
	Fuzhou	233	1466	1699	949
	Guiyang	84	1318	1402	2464
	Xiamen	277	990	1267	1487
	Hohhot	0	1028	1028	2065
	Lhasa	0	710	710	523
	Xining	96	367	463	323

Fig. 7. Typical FUA patterns for each type: (i) the FUAs in Nanjing (Type 1); (ii) the FUAs in Shijiazhuang (Type 2); (iii) the FUAs in Haikou (Type 3); and (iv) the FUAs in Xi'an (Type 4).

larger than the size of New York City (789 km²) and Paris (105 km²) (Tao, 2018). Finally, the nonagricultural population ratio and the nonagricultural GDP ratio are not direct indicators that reflect interconnected areas as commuting patterns do. Therefore, the increasing availability of ride-hailing data creates new opportunities for FUA delineations throughout China, making FUA delineation a more thorough, straightforward, and efficient approach.

According to the OECD, industrialization and mobility improvements have enabled the movement of goods and labor between cities to

expand across far wider spatial areas in China. This has led to the emergence of functional areas on the Chinese mainland (Kamal-Chaoui, Leman, & Zhang, 2009). Another approach to FUA delineations in China was noted in the OECD report entitled "Urban Trends and Policy in China." The report identified 28 Regional Urban Systems (RUSs) at the subprovincial scale, and they included a far wider network of cities, towns, and villages with comparatively strong economic linkages. Because information on economic and social flows between cities, towns, and villages was not available in China at that time, four proxies were
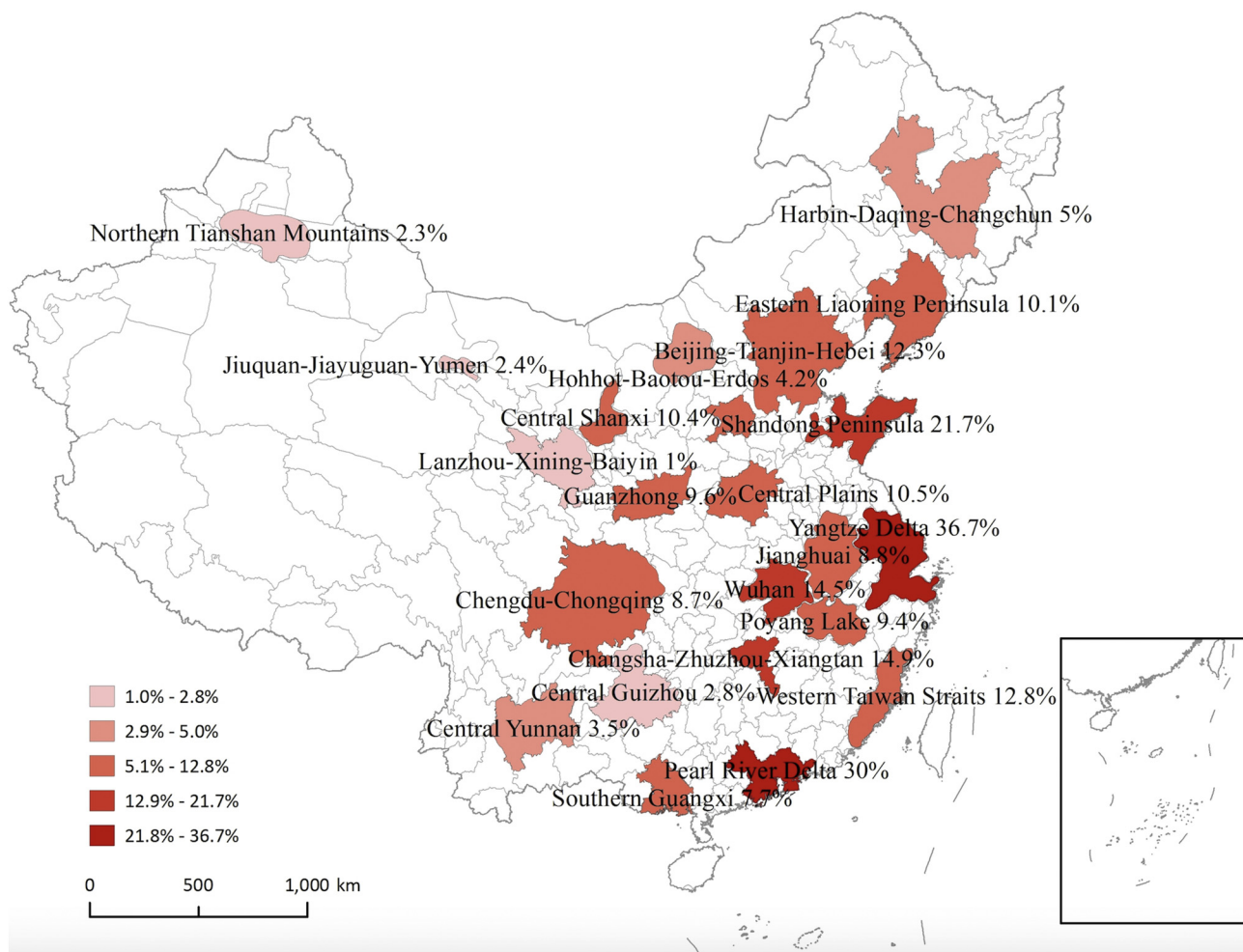

Fig. 8. The map of FUA ratio distributions in 23 urban agglomerations in China.
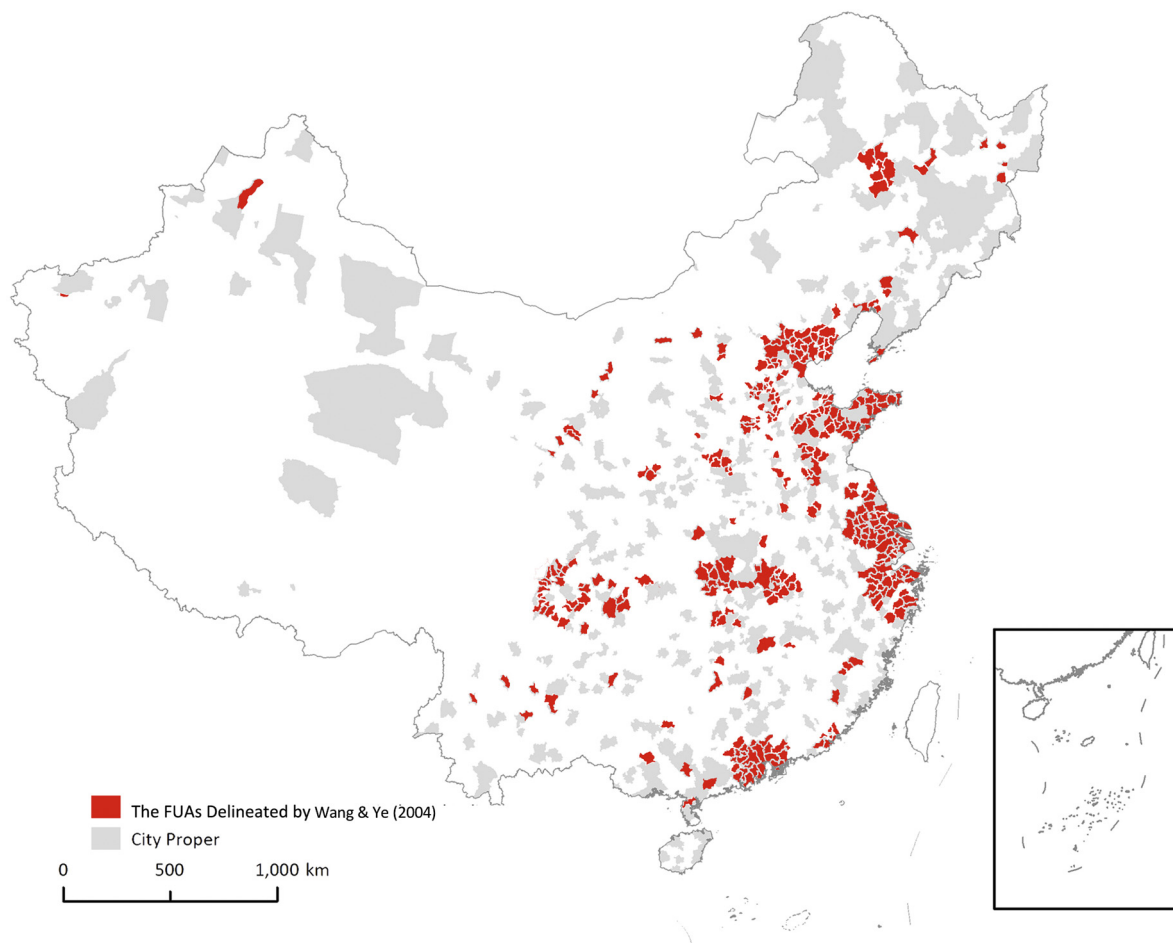
Fig. 9. FUA delineations of Wang and Ye (2004).

used to approximate linkages on the basis of "(1) population densities of all counties and cities in China, and the non-agricultural population size of statutory cities in 2000; (2) an analysis of average daily traffic volumes along more than 3000 segments of the national highway network; (3) an analysis of the location and capacities of the existing and planned National Trunk Highway System; and (4) an analysis of the location and capacities of the national railway network" (Kamal-Chaoui et al., 2009).

The results from the OECD study and those from this study have three main differences. First, the OECD study was based on population density at the county or city level, whereas this study used the massive ride-hailing records based on traffic flows at the township level, which resulted in a finer spatial resolution. Second, we used actual commuting flows instead of proxies, including the capacities of existing and planned national highway and railway networks, which are indirect measurements. Third, this study relied on increasingly available open-access large data sets that consist of higher spatial resolutions and result in more accurate FUA delineations.

6.2. Conclusion

In this article, we proposed a straightforward approach for FUA delineations for all of China that used massive Didi ride-hailing records. We collected a total of 43,846,160 records from Didi Chuxing, which has a high market penetration rate with a total of 654 cities. We delineated a total of 308 FUAs that cover 4456 townships out of a total of 39,007 townships. According to the results of this study, the top 10 FUAs in size were situated in Guangzhou_Shenzhen, Beijing_Langfang, Shanghai_Suzhou, Chengdu_Deyang, Hangzhou_Shaoxing, Wuhan,

Chongqing, Nanjing, Tianjin, and Xi'an_Xianyang. These results revealed hot spots of urbanization in China from the FUA perspective. The FUAs in the Pearl River Delta Urban Agglomeration were the most integrated among the Yangtze River Delta Urban Agglomeration, the Pearl River Delta Urban Agglomeration, and the Beijing-Tianjin-Hebei Urban Agglomeration, which are recognized as China's three most important centers of finance, trade, and industry. In the Beijing-Tianjin-Hebei Urban Agglomeration, the FUAs covered a total area of 21,580.9 km^2 in 566 townships. In the Yangtze River Delta Urban Agglomeration, the FUAs covered 38,026.8 km^2. This study also revealed two potential applications of delineating FUAs. First, through a comparison of the delineated FUAs with existing administrative boundaries, we found that the FUAs did not conform well with city administrative boundaries in most Chinese cities. The sizes of these administrative cities were smaller than their FUAs in cities such as Chengdu and Hangzhou. In these two cities, the administrative areas are 5140 km^2 and 3142 km^2 smaller than their FUAs. In contrast, in Chongqing and Urumchi, the administrative area FUAs are 22,260 km^2 and 7936 km^2 larger than their FUAs, and the area differences are the largest among Chinese cities. The next administrative division adjustment in China should be done according to FUA delineations. Second, we evaluated the development of urban agglomerations in China using FUA ratios. The FUA ratios of the eastern coastal agglomerations were relatively high compared with the west. The FUA ratios of the Northern Tianshan Mountains, Jiuquan-Jiayuguan_Yumen, and Central Guizhou were 2.3%, 2.4%, and 2.8% respectively, with a lower development than all of the urban agglomerations.

According to this study's outcomes, the authors suggest that urban planning in China should recognize the FUAs as basic urban planning

boundaries rather than administrative boundaries to (1) prepare for future urban development; (2) rationally organize the urban land; and (3) coordinate urban spatial layout, properties, scale, and development direction. The government should conduct a construction plan of the energy, communication, information network, and landscape based on these FUAs. In addition, in urban policy making, the role of administrative cities should be diminished and public services, such as education, medical treatment, housing purchase, and social security should be promoted in every FUA. From the perspective of urban agglomerations, the clear urban and planning policy recommendations are as follows. First, for urban agglomerations with FUA ratios less than 5% (e.g., the Northern Tianshan Mountains, the Jiuquan-Jiayuguan-Yumen, Lanzhou-Xining-Baiyin, Lanzhou-Xining-Baiyin, Hohhot-Baotou-Erdos, Central Guizhou, and Central Yunnan), the central government should cancel the delineation of these urban agglomerations. Instead, they should develop individual cities (e.g., Lanzhou, Xining, and Hohhot) to enhance the agglomeration effect of central cities. The government should support the advantages of the Yangtze Delta, Pearl River Delta, and Shandong Peninsula, which are obvious in terms of FUA ratios. The development of Chengdu-Chongqing, Central Shanxi, and Guanzhong urban agglomerations, which are three urban agglomerations in West China with relatively high FUA ratios, should be highlighted to ensure the balanced development of western and eastern China. From the perspective of potential administrative division adjustments, in the past few decades, the administrative boundaries of China have changed frequently. As a result of these findings, the authors suggest that in the next round of urban administrative division adjustment, the government should consider these FUAs. More precisely, for current administrative cities with small FUAs, the administrative boundaries should shrink and vice versa. The administrative boundaries of Chongqing, Chengdu, and Chengdu should be the first to be adjusted. In addition, county-level cities with no FUAs, also should be changed to counties.

In this study, we delineated the FUAs in China at the township scale, and this filled a gap in common approaches that primarily have focused on a small number of cities. Commuting data are foundational for reflecting the socioeconomic interactions between cities and their surrounding cities, villages, and rural areas. With the development of the global ride-hailing industry, some companies, such as Didi, Uber, and Lyft, have generated a large amount of commuting data worldwide. Thus, this study based on massive ride-hailing records is likely a better reference for the delineation of FUAs from a global view. In summary, this is the first FUA delineation for China that has used large-scale and fine-grained big data, along with increasingly popular resident ride-hailing data. This straightforward and new approach can be applied to delineate FUAs across wide geographic regions.

This research has some potential limitations. In this study, we used online ride-hailing data (three working days in August 2016) to generate the FUAs for the Chinese mainland. Although the Didi Chuxing represents China's largest car-hailing database, people who use offline methods (such as phone calls) or other offline avenues (such as driving) or people who use public transportation were not recorded. Second, individual situational biases of Didi Chuxing usage in different cities may introduce bias into FUA identification. The only way to deal with these limitations is to recognize them fully and to attempt to access the most accurate data possible. It is expected that travel data will only advance in the future and will provide researchers with higher spatial and temporal resolutions. Additionally, further research should be improved by checking other thresholds and using more detailed comparisons of FUA delineation outcomes.

Funding

Declaration of competing interest

The authors declare that they have no conflict of interest.

References

Agatz, N. A., Erera, A. L., Savelsbergh, M. W., & Wang, X. (2011). Dynamic ride-sharing: A simulation study in metro Atlanta. *Transportation Research Part B: Methodological, 45*(9), 1450–1464.

Arcaute, E., Hatna, E., Ferguson, P., Youn, H., Johansson, A., & Batty, M. (2015). Constructing cities, deconstructing scaling laws. *Journal of the Royal Society Interface,* 1–8.

Bode, E. (2008). Delineating metropolitan areas using land prices. *Journal of Regional Science, 48*(1), 131–163.

Bosker, M., Park, J., & Roberts, M. (2018). Definition matters: Metropolitan areas and agglomeration economies in a large developing country. *Policy research working paper series*.

Chaudhari, H. A., Byers, J. W., & Terzi, E. (2018). Putting data in the driver's seat: Optimizing earnings for on-demand ride-hailing. *Proceedings of the eleventh ACM international conference on web search and data mining* (pp. 90–98). .

Chen, Y., Liu, X., Li, X., Liu, X., Yao, Y., Hu, G., ... Pei, F. S. (2016). Delineating urban functional areas with building-level social media data: A dynamic time warping (dtw) distance based k-medoids method. *Landscape and Urban Planning, 160*, 48–60.

Chi, J., Jiao, L., Dong, T., Gu, Y., & Ma, Y. (2016). Quantitative identification and visualization of urban functional area based on poi data. *Journal of Geomatics, 41*(2), 68–73.

Cörvers, F., Hensen, M., & Bongaerts, D. (2009). Delimitation and coherence of functional and administrative regions. *Regional Studies, 43*(1), 19–31.

Danko, D. M. (1992). The digital chart of the world project. *Photogrammetric Engineering & Remote Sensing, 58*(8), 1125–1128.

Demsar, U., Reades, J., Manley, E., & Batty, J. (2014). Edge-based communities for identification of functional regions in a taxi flow network. *American Journal of Ophthalmology, 83*(2), 267–271.

DidiChuxing (2018). Didi Chuxing corporate citizenship report 2017. Available at https://www.didiglobal.com/news/newsDetail?id=384&type=news.

Dijkstra, L., & Poelman, H. (2014). A harmonised definition of cities and rural areas: The new degree of urbanization. *Regional working paper, directorate-general for regional and urban policy*. Brussels: European Commission.

Dong, Y., Wang, S., Li, L., & Zhang, Z. (2018). An empirical study on travel patterns of internet based ride-sharing. *Transportation Research Part C: Emerging Technologies, 86*, 1–22.

Duranton, G. (2015). Delineating metropolitan areas: Measuring spatial labour market networks through commuting patterns. *The economics of interfirm networks. Vol. 4. The economics of interfirm networks* (pp. 107–133). Springer Japan.

Elvidge, C. D., Kihn, K. E., & Davis, E. R. (1996). Mapping City lights with nighttime data from the DMSP-OLS operational linescan system. *Photogrammetric Engineering & Remote Sensing, 63*(6), 727–734.

ESPON (2007). Study on urban functions. *ESPON Project, 1*(4), 3.

Fang, C., & Yu, D. (2017). Urban agglomeration: An evolving concept of an emerging phenomenon. *Landscape and Urban Planning, 162*, 126–136.

Farmer, C. J. Q., & Fotheringham, A. S. (2011). Network-based functional regions. *Environment & Planning A, 43*(11), 1–37.

Fotheringham, A. S., & O'Kelly, M. E. (1989). *Spatial interaction models: Formulations and applications.* Dordrecht: Kluwer Academic Publishers.

Gajović, V. (2013). Comparative analysis of different methods and obtained results of delineation of functional urban areas. *SPATIUM International Review, 29*(7), 8–15.

Glaeser, E. L., Kahn, M., & Chu, C. (2001). *Job Sprawl: Employment location in U.S. metropolitan areas.* Washington, D.C.: The Brookings Institution.

Goddard, J. B. (1973). Office linkages and location: A study of communications and spatial patterns in Central London. *Progress in Planning, 1*, 109–232.

Houssou, N. L. J., Guillaume, J., & Prigent, A. (2019). A graph based approach for functional urban areas delineation. *Proceedings of the 34th ACM/SIGAPP symposium on applied computing*https://doi.org/10.1145/3297280.3297341.

Hu, X. W., Zhou, Y. X., & Gu, C. L. (2000). *On the spatial agglomeration and dispersion in coastal regions.* Beijing: Science Press.

Kamal-Chaoui, L., Leman, E., & Zhang, R. (2009). *Urban trends and policy in China.* OECD Regional Development Working Papershttps://doi.org/10.1787/225205036417.

Karlsson, C., & Olsson, M. (2006). The identification of functional regions: Theory, methods, and applications. *The Annals of Regional Science, 40*(1), 1–18.

Korcelli, P. (2008). Review of typologies of European rural-urban regions. *PLURAL*. Available at http://www.plurel.org/images/D211.pdf.

Long, Y. (2016). Redefining Chinese city system with emerging new data. *Applied Geography, 75*, 36–48.

Ma, S., & Long, Y. (2019). Identifying spatial cities in China at the community scale. *Journal of Urban and Regional Planning, 11*(1), 9–16.

Marshall, A. (1890). *Principles of economics.* London: Macmillan.

MOHURD (2014). *Chinese city construction statistics yearbook 2014.* Beijing: China Planning Press.

Mulíček, O. (2013). Vymezení funkčního území Brněnské metropolitní oblasti a Jihlavské sídelní aglomerace. Available at http://www.brno.cz/fileadmin/user_upload/sprava_

mesta/Strategie_pro_Brno/doplnujici_studi e/BMO_JSA_vymezeni_2013_web.pdf.

Newman, M. E. J. (2004). Detecting community structure in networks. *European Physical Journal B, 38*(2), 321–330.

OECD (2002). Definition of functional regions. Available at www.oecd.org/cfe/regional-policy/15181241.doc.

OMB (2010). Standards for delineating metropolitan and micropolitan statistical areas. *Federal Register, 75*(123), 37246–37252.

Pham, A., Dacosta, I., Jacot-Guillarmod, J., Huguenin, K., Hajar, T., Tramer, F., ... Hubaux, J. P. (2017). *Proceedings on Privacy Enhancing Technologies, 201*(2), 35–53.

Ratti, C., Sobolevsky, S., Calabrese, F., Andris, C., Reades, J., Martino, M., ... Martino, M. (2010). Redrawing the map of great Britain from a network of human interactions. *PLoS One, 5*(12), e14248.

Reba, M., Reitsma, F., & Seto, K. C. (2016). Spatializing 6,000 years of global urbanization from 3700 BC to AD 2000. *Scientific Data, 3*, 160034.

Schaller, B. (2017). The growth of app-based ride services and traffic. *Travel and the future of new York City*. Available at http://schallerconsult.com/rideservices/unsustainable.htm.

Shen, Y., & Batty, M. (2019). Delineating the perceived functional regions of London from commuting flows. *Environment and Planning A: Economy and Space, 51*(3), 547–550.

Shen, Y. N., Sun, F. Y., & Che, Y. (2017). Public green spaces and human wellbeing: Mapping the spatial inequity and mismatching status of public green space in the Central City of Shanghai. *Urban Forestry & Urban Greening, 27*, 59–68.

Shih, G. (2015). China taxi apps DiDi Dache and Kuaidi Dache announce $6 billion tie-up. *Reuters February 14*. Available at https://www.yahoo.com/tech/china-taxi-apps-didi-dache-kuaidi-dache-announce-023235253-finance.html.

Takhteyev, Y., Gruzd, A., & Wellman, B. (2012). Geography of twitter networks. *Social Networks, 34*(1), 73–81.

Tao, D. Y. (2018). Development history and evolutionary characteristics of the statistical system of metropolitan area abroad and its implication for China. *Shanghai Urban Planning Review, 1*, 46–52.

USCB (2014). *History: Metropolitan areas.* U.S. Census Bureau.

Wang, J., & Ye, T. (2004). Preliminary research on division of metropolitan area and metropolitan interlocking region in China. *Areal Research and Development, 23*(3), 13–17.

Wang, Q., Huo, H., He, K., Yao, Z., & Zhang, Q. (2008). Characterization of vehicle driving patterns and development of driving cycles in Chinese cities. *Transportation Research Part D Transport & Environment, 13*(5), 289–297.

Wikipedia (2019). Article: Megalopolis of China. Available at https://en.wikipedia.org/wiki/Megalopolis_of_China.

Wöhrnschimmel, H., Zuk, M., Martínez-Villa, G., Cerón, J., Cárdenas, B., & Rojas-Bracho, L. (2008). The impact of a bus rapid transit system on commuters' exposure to benzene, co, pm in Mexico city. *Atmospheric Environment, 42*(35), 8194–8203.

Xie, S. H., & Ning, Y. M. (2005). *Study on the development of big cities and the formation of metropolitan. Vol. 1*, Urban Problems11–15.

Yuan, N. J., Zheng, Y., Xie, X., Wang, Y., Zheng, K., & Xiong, H. (2015). Discovering urban functional zones using latent activity trajectories. *IEEE Transactions on Knowledge and Data Engineering, 27*(3), 712–725.

Yuan, Q. Q., & Ye, T. L. (2018). *Development quality evaluation of three metropolitan agglomerations in China. Qianxian. Vol. 7*, 73–75.

Zhang, J. G. (2010). New metropolitan regionalism: The new governance of the metropolitan areas in the USA. *Journal of Sun Yatsen University, 50*(1), 131–141.

Zhang, Z. X., Wang, Y. H., Liu, Q. J., Li, L. L., & Pang, W. (2016). A CNN based functional zone classification method for aerial images. *2016 IEEE international geoscience and remote sensing symposium (IGARSS)* (pp. 5449–5452). .

Zhu, Z., Yang, J., Zhong, C., Seiter, J., & Tröster, G. (2015). Learning functional compositions of urban spaces with crowd-augmented travel survey data. *Proceedings of the 23rd SIGSPATIAL international conference on advances in geographic information systems*New York, NY, USA: ACM22. https://doi.org/10.1145/2820783.2820832.

Zong, F., Juan, Z., & Liu, B. (2007). Accessibility of the staggered shifts policy. *International conference on transportation engineering* (pp. 3076–3081). .