

Article

Detecting and Analyzing Urban Centers Based on the Localized Contour Tree Method Using Taxi Trajectory Data: A Case Study of Shanghai

Mengqi Sun ¹  and Hongchao Fan ^{2,*}

¹ School of Remote Sensing and Information Engineering, Wuhan University, Wuhan 430079, China; Mengqi_Sun@whu.edu.cn

² Department of Civil and Transport Engineering, Norwegian University of Science and Technology, 7491 Trondheim, Norway

* Correspondence: hongchao.fan@ntnu.no

Abstract: Urban structure is of vital importance to urban planning, transportation, economics and other applications. Since detecting and analyzing urban centers is crucial for understanding urban structure, a large number of studies on urban center extraction have been performed. In this paper, we propose an analysis framework to identify urban centers by using taxi trajectory data. The proposed approach differs from previous methods by employing a novel way to simulate taxi trajectory data with the topographic surface. We extracted pick-up and drop-off spots from taxi trajectory data and employed the localized contour tree method to delineate the boundaries and hierarchies of urban centers. The experiments show that the proposed method can successfully detect urban centers and analyze their temporal patterns in different periods in Shanghai, China.

Keywords: urban structure; urban center; taxi trajectory data



Citation: Sun, M.; Fan, H. Detecting and Analyzing Urban Centers Based on the Localized Contour Tree Method Using Taxi Trajectory Data: A Case Study of Shanghai. *ISPRS Int. J. Geo-Inf.* **2021**, *10*, 220.
<https://doi.org/10.3390/ijgi10040220>

Academic Editors: Giuseppe Borruso and Wolfgang Kainz

Received: 20 February 2021
Accepted: 29 March 2021
Published: 2 April 2021

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Copyright: © 2021 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

Urban structure is defined as the spatial arrangement of different land uses in urban areas [1]. It is an important part of urban geography and has major effects on urban planning, economics, transportation and many other aspects of cities. Given the acceleration of urbanization in recent years, the study of urban structure is of great importance to government decision-making and policy implementation.

In 1925, Burgess proposed the first monocentric model to describe the city structure [2]. Since then, many researchers have focused on studying urban structure, and most of them have concentrated on structure models, urban centers, land use, functional areas, urban density, etc. (detailed in Section 2.1).

In many early works, researchers referred to employment and population census data to identify employment centers [3–17]. However, census data are always based on a preset census tract. The preset shape of the census tract may prevent the identification of a center. Therefore, the analysis results may be affected by the modifiable areal unit problem [18].

Since remote sensing can obtain a wide range of urban data, it has been used to measure urban density [19] and polycentricity [20]. In 2017, Chen et al. employed nighttime light (NTL) data to analyze urban structure using a localized contour tree method [18]. They successfully identified most of the centers specified in the master plan of Shanghai. In addition, they inventively drew an analogy between the NTL intensity and topographic surface. Then, they constructed a hierarchy of the detected urban centers and calculated the attributes of each center. Although this is an innovative and effective approach, there are still some limitations of using NTL data. Firstly, the spatial and temporal resolutions of NTL data are not high. The NPP-VIIRS NTL data used by Chen et al. [18] only has a resolution of 500 m. Although the spatial resolution of the latest Luojia-1 data is 130 m,

the regression cycle takes 3–5 days and does not reflect the details of finer spatial and temporal scales [21]. Secondly, NTL satellites only capture nighttime data, which means that it can only reflect activities at night. Therefore, daytime activities cannot be inferred from NTL data. Thirdly, the proximity of night light to human activities is rather weak in certain places in the urban area; for instance, streetlights and many urban illuminations are independent of human flow. In sum, NTL data are unable to completely reveal the urban structure dynamic.

Urban centers and their functions are always connected to sociological and economical activities in a city. As humans are the main body of social and economic activities, employing data that are more related to human mobility, e.g., taxi trajectory data, to identify urban centers is more reasonable and reliable. Taxis are a large proportion of urban traffic in most large cities and can even account for as much as 50–60% of the traffic in certain key areas [22]. Taxi trajectory data provide precise information on citizens' favorite departure sites and destinations, and they are an important type of digital footprint for extracting human activities [23]. Furthermore, taxi trajectory data have the advantages of wide coverage, high sampling density, high position accuracy, large volume and information richness [24,25]. Compared with NTL data, taxi data have flexible spatial and temporal resolutions, are more related to human activities and can reflect these activities throughout a 24-h period.

Previous studies using taxi trajectory data related to the urban structure have mostly focused on human mobility [26–31], land use and functional regions [32–37] (detailed in Section 2.2). Notably, by creatively using kernel density estimation (KDE) to transfer taxi data to a continuous surface, Hu et al. proposed an algorithm to extract geometric features from the surface, and they successfully detected and analyzed hotspots in Shanghai [38]. However, similar to many previous studies, the specific boundaries of the detected hotspots were not delineated. Although researchers have delineated the boundaries of regions by using a clustering algorithm [30], census tract [28] or pre-shaped partition [29,31,35–37], we did not find any research that used taxi data to analyze the hierarchy and attributes of urban centers.

Pick-up and drop-off spots extracted from taxi trajectory data have proven to be related to urban land use [32,35]. Therefore, we assumed that the use of taxi trajectory data could improve the identification of urban centers. Inspired by the work of Chen et al. [18], in this study, we simulated taxi data with the topographic surface and employed the localized contour tree method to identify urban centers with boundaries in Shanghai. In addition, the hierarchy and attributes of the detected centers were analyzed and calculated.

In this work, the following research questions (RQ) are addressed:

RQ1: Can taxi trajectory data replace NTL data for the purpose of identifying urban centers?

RQ2: What are the advantages of using taxi trajectory data in urban center identification?

The remainder of this paper is organized as follows. A short review of previous research on urban structure is provided in the next section. We introduce our proposed methodology for identifying and analyzing urban centers in Section 3. In Section 4, we present our analysis results, followed by a discussion of our results and methods in Section 5. The conclusion of our work is given in Section 6.

2. Related Work

This section provides an overview of existing research approaches related to this work: urban structure research (Section 2.1) and taxi trajectory data research related to urban structure (Section 2.2).

2.1. Urban Structure Research

A city is a spatial combination of different land uses and functional areas. Initially, researchers proposed monocentric models to describe the spatial distribution in a city. In 1925, Burgess introduced the concentric zone model [2]. This is the first model that explains the urban structure. Five concentric functional zones, namely, the central business district

(CBD), zone of transition, zone of workers' homes, the residential area and the commuter zone, were created to describe the urban structure of Chicago. This concentric model matched the city pattern when America was influenced by the Industrial Revolution [39]. However, some urban areas have become more active, and they continuously grow and expand. Therefore, in 1939, Hoyt proposed the sector model [40]. The sector model also has a single center, but some functional areas can extend from the CBD to the urban edge. Harris and Ullman developed the multiple nuclei model in 1945 [1]. According to the multiple nuclei model, a city can contain several centers. The CBD is no longer in the dominant position as it was before, and several sub-centers have been derived. Nevertheless, with urban development and population growth in recent decades, the urban structure has become complex, and it cannot be simply described by any traditional model.

Scholars employ different kinds of data to discover structures within cities. Census data, such as employment data, population data and commuting flows, are the most popular and are often used in urban structure analysis. In 1987, McDonald suggested that gross employment density and the employment–population ratio were the best measures of employment sub-centers [3]. Then, these measures were applied to identify sub-centers in Chicago [3,4]. Giuliano and Small developed a method for identifying employment sub-centers in Los Angeles using journey-to-work data [5]. They defined a sub-center as a continuous set of zones with an employment density and total employment above preset thresholds. Then, they analyzed the functions, distribution and commuting flows associated with the detected centers. Following in their steps, numerous works based on thresholds have been reported [6,7,9,10,12]. As noted by McMillen, the regression approach is less sensitive to the analysis unit [11]; therefore, he used a two-stage regression procedure to identify employment sub-centers in six cities. In contrast to the parametric regression model proposed by McDonald and Prather [8], McMillen's procedure is non-parametric, and its identification of employment sub-centers can better adapt to the reality of the urban structure [15]. Another work of McMillen proposed an algorithm using the standard properties of contiguity matrices to identify sub-centers [13]. This algorithm combines regression and threshold methods. Further work by Redfearn used the topography of the employment density obtained by non-parametric regression methods to detect both the sub-centers and their areas of influence, respectively [14,15]. Census data have still been commonly used in recent studies. Hu et al. detected the urban spatial structure using employment data and characterized the dynamic commuting patterns in Beijing, China [16]. Shen and Batty detected functional regions perceived by people with different occupations using disaggregated flow data and thus showed differences across occupations [17].

Instead of census data, 3D building models derived from remote sensing data were applied to measure the urban density [19] and analyze polycentricity [20] in four German city regions. Since nighttime light (NTL) data are receiving increasing attention, Chen et al. applied NTL data to delimit urban sub-centers and determine the topological relationships between them by adopting a graph-theory-based localized contour tree method [18]. In addition, as GPS is widely used, location-based data have become a research hotspot. Location-based social network data were used to identify city centers with precise boundaries [41] and distinguish urban functions [42]. Taxi trajectories are also popular location-based data for studying urban structure (detailed in Section 2.2).

2.2. Taxi Trajectory Data Research Related to Urban Structure

Taxi trajectory data record routes based on GPS devices. Such data have significant spatial and temporal characteristics and are closely related to human mobility, social-economic activities and urban layout. Therefore, a large number of studies have used taxi trajectory data to study human mobility, functional regions and land use.

In their work, Peng et al. analyzed taxi trips in Shanghai and found that people traveled for three main purposes on workdays: commuting between home and the workplace, traveling from workplace to workplace and traveling for other activities such as leisure [26]. Zhang et al. estimated origin–destination (OD) flows and mined their semantics based on

periodic patterns [27]. In 2014, Scholz and Lu defined pick-up or drop-off spots as activity locations, and they proposed a method to detect the dynamic patterns of activity hotspots in a city [28]. After that, Tang et al. divided Harbin city into 400 transportation districts [29]. They clustered pick-up and drop-off locations using the DBSCAN algorithm and analyzed human mobility on weekdays and weekends. Employing road nodes to generate Thiessen polygons, Shen et al. analyzed the spatial distribution of pick-up and drop-off spots and used Moran's I index to explore the spatial autocorrelation of these regions, and hotspots were found by clustering [30]. Moreover, Kong et al. proposed a time–location relationship to find human mobility patterns in functional regions by analyzing pick-up and drop-off locations [31].

Since Zheng et al. detected inadequacy in urban planning using taxi trajectory data in 2011 [43], many taxi data researchers have focused on studying urban structure. Combined with points of interest (POIs), trajectory data were used to identify the different functions of regions partitioned by major roads [34]. The association between traffic patterns extracted from pick-up and drop-off spots and urban land uses in Shanghai has been revealed [35]. In addition, Qi et al. demonstrated that the social function of an urban region could be characterized by temporal variation in the number of pick-ups/drop-offs [32]. Pan et al. presented an improved DBSCAN algorithm for clustering pick-up and drop-off spots [33]. Then, they designed six features for regions extracted from clustering and classified these regions into different land uses. Using three large taxi datasets from three cities, Zhao et al. provided a new quad-tree region division method and then used the association rule to identify four kinds of functional regions based on the human mobility patterns revealed by taxi trips [36]. In 2016, Tong et al. used Voronoi decomposition to divide the city into regions, which were then classified into different functional areas by incorporating pick-up and drop-off information [37].

Thus far, in previous research work, pick-up and drop-off spots (origins and destinations) have been used for the analysis of human mobility and functional areas since these locations can reflect people's travel preferences [26–29,31–33,36,37]. From this point of view, it is logical to expect that pick-up and drop-off hotspots extracted from taxi trajectory data can directly indicate the urban structure and can therefore be used to detect urban centers.

3. Detecting Urban Centers by Using Taxi Trajectory Data

The main innovation of our proposed method is the adoption of a remote sensing method to analyze taxi trajectory data and extract results with finer details. In comparison with NTL data, taxi trajectory data are more related to human mobility and activities, as mentioned in Section 1, so they can better reflect human activity intensity in urban areas. In our analysis framework, we used a statistical grid of pick-up/drop-off spots and regarded it as a topographic surface of human activity intensity. Thus, an analogy was drawn between this surface and the statistical grid. The center of human activity is analogous to a peak on the topographic surface. Since a peak is represented by concentric contour lines on the topographic surface, an urban center can be detected by generating contour lines from the statistical surface (statistical grid). With this analogy, a localized contour tree was applied to our data to identify urban centers, analyze the hierarchy and calculate attributes. Our approach consists of five steps: taxi trajectory data preprocessing, contour map generation, localized contour tree construction and simplification, urban center identification and attribute calculation. The complete analysis framework is presented in Figure 1.

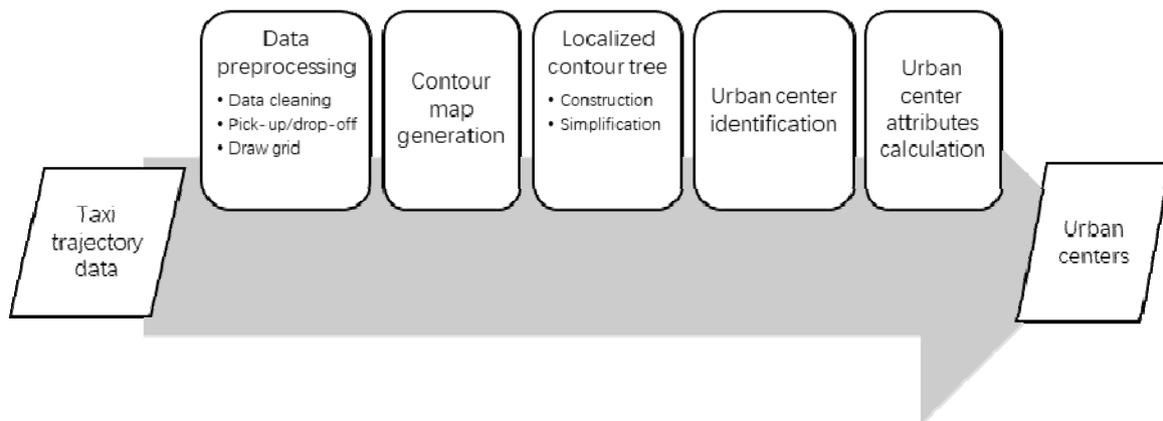


Figure 1. Analysis framework.

3.1. Taxi Trajectory Data Preprocessing

Taxi trajectory data are a set of points recorded by GPS devices. Each point has several features, including GPSID, timestamp, coordinates, speed, direction, passenger status, etc. After basic data cleaning, null and out-of-range values were eliminated. The remaining data values were projected onto locations in a planar coordinate system.

As the pick-up and drop-off locations reflect people's departure places and destinations, we chose them to represent human activities. The passenger status of the taxi trajectory data is 0 (vacant) or 1 (occupied). In order to extract pick-up and drop-off spots, we established the following premise:

For taxi data with the same GPSID in chronological order, if the passenger status of two adjacent records changes from 0 to 1, then the record with a passenger status of 0 is set as a pick-up spot. On the other hand, if the passenger status of two adjacent records changes from 1 to 0, then the record with a passenger status of 1 is set as a drop-off spot.

Following this premise, we extracted pick-up and drop-off spots from taxi trajectory data. Then, as an analogy to spatial resolution, we drew square grids and counted the total number of pick-up and drop-off spots in each grid. We calculated the hourly average of the total number in each grid. Thus, each grid has an attribute that represents the human activity intensity in that grid. The entire grid is similar to topography with peaks.

In order to compare this study with the work of Chen et al. [18], we chose a grid size of 500 m, which is equivalent to the pixel size of their NTL data. Pick-up and drop-off spot statistics in different periods were also computed to analyze the temporal patterns of urban centers.

3.2. Contour Map Generation

As we treated the statistical grid as topography, we used it to generate contour lines. To smooth noisy data, we applied a 5×5 Gaussian filter with a sigma value of 1 to our statistical grid. The initial contour value was set to 1, and the contour interval was set to 1.

Since taxis mainly pick up and drop off passengers near entrances and exits, the centers can be fragmented into small areas. However, areas that are too small cannot be considered as a center. Given these considerations, we defined the minimum area of a detected center as 1 square kilometer. Therefore, we removed all closed contours with areas less than 1 square kilometer. We believe that all major urban centers can be detected with these preset parameters.

3.3. Localized Contour Tree Construction and Simplification

The contour tree was first used to store contour maps [44,45]. Then, in 1994, Kweon and Kanade built a contour map from a digital elevation model (DEM) and then created a connectivity tree to extract topographic features [46]. They called the resulting tree a topographic change tree. The localized contour tree method was first described by Wu et al.

to detect and delineate surface depressions [47,48]. They generated a contour tree from a LiDAR DEM. Then, a localized fast priority search algorithm was applied over the contour tree to identify depressions.

A contour tree is composed of nodes and links. Nodes represent contour lines, while links represent topological inclusion relationships between adjacent contour lines. According to Wu et al. [47], the localized contour tree is constructed in a bottom-up manner. Sample contour lines are shown in Figure 2, and sample contour trees based on Figure 2 are shown in Figure 3a. First, we identified seed contour lines. A seed contour line contains no other contour lines and is a leaf node in a contour tree. In Figure 2, contours *g*, *h*, *i*, *j* are seed contour lines. In Figure 3a, nodes *g*, *h*, *i*, *j* are leaf nodes. Starting from seed contour lines, the rest of the contour trees can be constructed. If the external contour line contains only one contour line or branch, then the former is considered to be at the same level as the latter. Otherwise, the external contour line is considered to be at a higher level. As shown in Figures 2 and 3a, contours *g*, *h*, *i*, *j* are at level 1. Contours *d*, *e*, *f* contain only one contour line each, so they are also at level 1. However, contour *c* contains more than one branch, so it is considered as level 2. Similarly, contour *a* contains two branches, and contour *b* contains only one, so they are identified as level 3 and level 1, respectively. For each branch, the outermost contour line is the root node. Following this rule, contour trees were constructed as shown in Figure 3a.

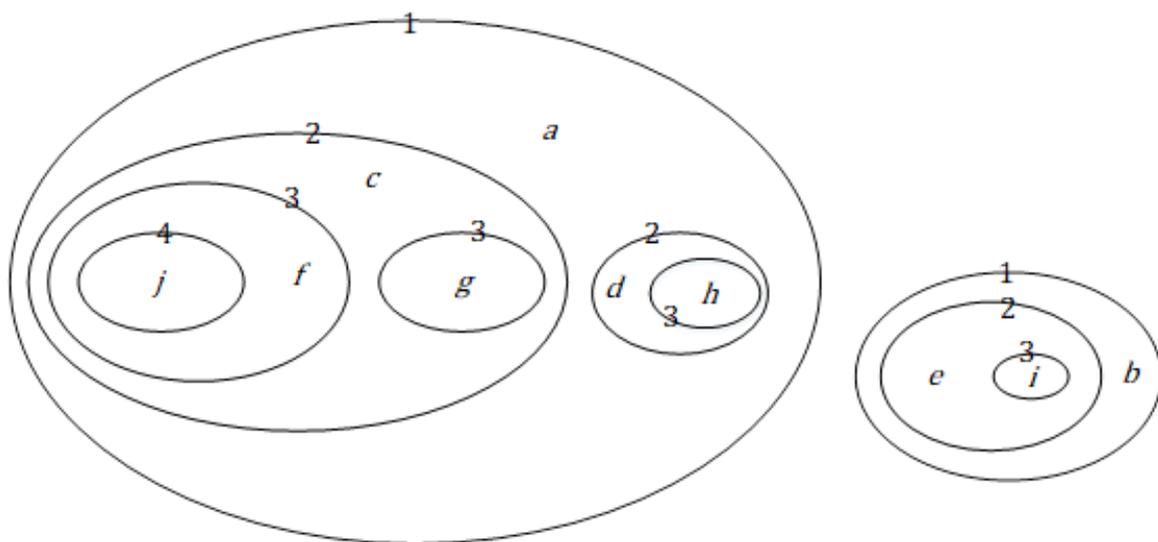


Figure 2. Sample contour lines.

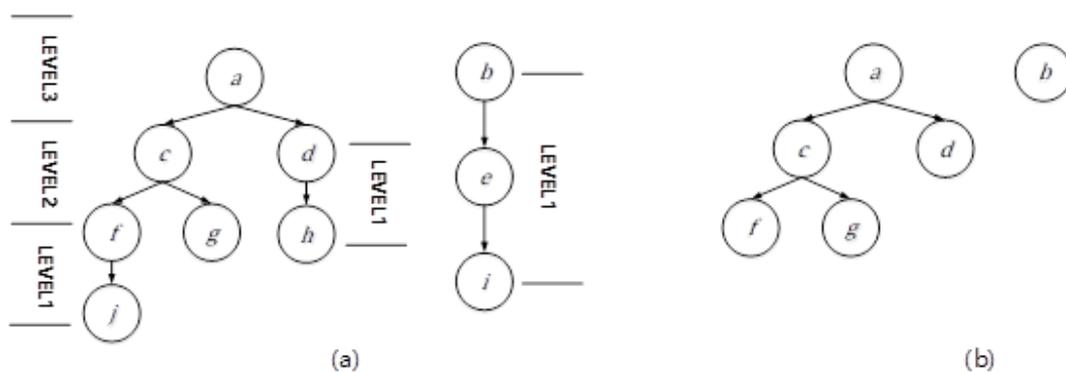


Figure 3. (a) Sample contour trees; (b) simplified sample contour trees.

In order to find the hierarchy of contour lines and detect the urban center, the contour trees were simplified. Adjacent contour lines at the same level were considered to be

different locations of one peak, and we only kept the outermost contour line, i.e., the highest node with the same level in the contour tree. Therefore, as shown in Figure 3b, contours a, b, c, d, f, g are retained as the simplified contour tree.

3.4. Urban Center Identification

As mentioned above, we calculated the hourly average number of total pick-up and drop-off spots to reflect human activity intensity. Urban centers are usually areas of high human activity intensity. Since the contour lines define ranges with the same activity intensity, they can be used to identify urban centers.

The simplified contour tree represents the entire hierarchy of all contour lines. In Figure 3b, there are two sub-trees in the simplified contour tree. The root nodes are the outermost contour lines of each sub-tree. Some sub-trees can have several branches, while others may contain only one node, as shown in Figure 3b. In this paper, we regard each leaf node, including one-node sub-trees, as a monocentric urban center (MUC). In addition, sub-trees that contain more than one leaf node or branch are considered to be polycentric urban centers (PUCs). The MUCs in a PUC are sub-centers, and the largest sub-center is the main center of the PUC. For example, in Figure 3, there is one MUC, and one PUC contains three sub-centers. We consider the main center in the largest contour tree to be the main center of a city, and the number of levels can be regarded as a reference to the complexity of PUCs.

3.5. Calculation of Urban Center Attributes

In order to better understand the characteristics of urban centers, we calculated several attributes for each one. Since we used the hourly average number of pick-up and drop-off spots to represent human activity intensity, for each MUC, the minimum intensity ($MinI$), the maximum intensity ($MaxI$), the total intensity (TI), the average intensity ($AveI$) and the standard deviation of intensity ($StdI$) were calculated as statistical attributes. In addition, the area (S), urban development orientation (Φ), compactness index (C) and elongatedness (E) [18] were calculated for MUCs as morphometric attributes to analyze geometric characteristics [49,50]. Moreover, we used the slope of topography to calculate the gradient (G) of activity intensity to identify intensity changes from the center to the periphery of a center. For each PUC, we calculated a polycentric indicator (PI) to quantify polycentricity [51,52]. The detailed definitions of the attributes are listed in Table 1.

In Table 1, n is the number of grids in a center, and x_i is the hourly average number of pick-up and drop-off spots in the i th grid. The compactness index (C) is designed to quantify the compactness of the shape, and p is the perimeter of the center boundary. A circle is the most compact shape in Euclidean space, and its C value is 1 [17]. Elongatedness (E) also indicates the compactness of a shape. Len and Wid are the length and width of the MBR (Minimum Bounding Rectangle) of the center. If a center has a high C and a low E , then the center is regarded as compact. Otherwise, the center is incompact. When calculating the polycentric indicator (PI), we consider all of the sub-centers in a PUC. n_c is the number of sub-centers in a PUC. Std_c is the standard deviation, and \overline{AveI} is the average of all $AveI$ of those sub-centers. $AveI_i$ is the $AveI$ of the i_{th} sub-center. $AveI_{max}$ is the maximum $AveI$ of all sub-centers, and Std_{max} is the standard deviation between $AveI_{max}$ and an ideal sub-center with an $AveI$ of 0. PI ranges from 0 to 1. If the PI of a PUC is 0, then the PUC has a dominant main center, while if PI is 1, then the sub-centers equally impact the PUC.

Table 1. Definitions of urban center attributes.

Attribute	Definition
Monocentric Urban Center (MUC)	
Minimum Intensity (<i>MinI</i>)	$MinI = \min\{x_i i = 1, 2, \dots, n\}$
Maximum Intensity (<i>MaxI</i>)	$MaxI = \max\{x_i i = 1, 2, \dots, n\}$
Total Intensity (<i>TI</i>)	$TI = \sum_{i=1}^n x_i$
Average Intensity (<i>AveI</i>)	$AveI = TI/n$
Standard Deviation of Intensity (<i>StdI</i>)	$StdI = \sqrt{\frac{1}{n} \sum_{i=1}^n (x_i - AveI)^2}$
Area (<i>S</i>)	Area within the boundary of a center
Urban Development Orientation (Φ)	Angle between the <i>x</i> -axis and the long side of MBR (Minimum Bounding Rectangle) of the center
Compactness Index (<i>C</i>)	$C = 4\pi S/p^2$
Elongatedness (<i>E</i>)	$E = Len/Wid$
Intensity Gradient (<i>G</i>)	Average slope within a center
Polycentric Urban Center (PUC)	
Polycentric Indicator (<i>PI</i>)	$PI = 1 - \frac{Std_c}{Std_{max}}$ $Std_c = \sqrt{\frac{1}{n_c} \sum_{i=1}^{n_c} (AveI_i - AveI)^2}$ $Std_{max} = \sqrt{\frac{1}{2} ((AveI_{max} - AveI_{mean})^2 + (0 - AveI_{mean})^2)}$ $AveI_{mean} = \frac{AveI_{max} + 0}{2}$

4. Case Study and Results

This section presents the results of our case study in Shanghai, China. We first introduce our study area and data, and then we present the results of the localized contour tree and urban center detection. Finally, we analyze the attributes and temporal patterns of urban centers in Shanghai.

4.1. Study Area and Data

Researchers have verified that the polycentric planning strategy is beneficial to urban economics in China [53]. Among the metropolitans in China, Shanghai has been demonstrated to be a polycentric city, which is consistent with its master plan [18,54]. Shanghai is located in the east of China. It is an international center of economics, trade, science and technology. As shown in Figure 4, Shanghai has 16 districts (Huangpu, Xuhui, Changning, Jing'an, Putuo, Hongkou, Yangpu, Minhang, Baoshan, Jiading, Pudong, Jinshan, Songjiang, Qingpu, Fengxian and Chongming) with a total area of 6340.5 square kilometers, and its population was over 24 million in 2019 according to national statistical data [55]. Shanghai has three rings, which are connected by several roads and expressways. According to the Shanghai master plan, the central area of Shanghai is within the outer ring road [56].

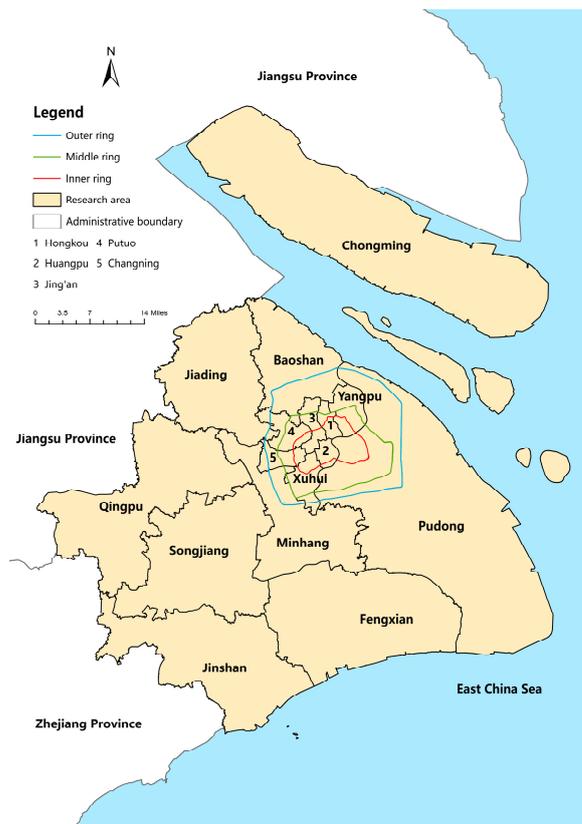


Figure 4. Study area.

According to the Shanghai Annual Report on Urban Traffic Operation in 2018, the daily average number of passenger trips of cruising taxis is around 1.75 million, and that of online car-hailing is around 1.05 million. Considering the daily average number of passenger trips via underground transportation (around 10.16 million) and public buses (around 5.76 million), taxis accounted for more than one-tenth of public transportation in Shanghai in 2018.

In this case study, we used the taxi trajectory data in Shanghai provided by a local company to verify our analysis framework. Our data include around 5500 taxis and cover an ordinary week in 2018 (from Monday 4th of June to Sunday 10th of June). One of the largest taxi companies, Qiangsheng, claimed that they had about 12,000 taxis in 2018, which accounted for 25 percent of taxis in Shanghai. According to this information, our taxis account for about one-tenth of taxis in Shanghai. The data include information on each taxi, and the GPS sampling frequency is 10 s. Table 2 shows sample records from our taxi trajectory data. The value of the field ‘Time’ ranges from 0 to 235,959, which represents the time in 24-h format. The field ‘GPSID’ is the unique identification of each taxi. The fields ‘Lon’ and ‘Lat’ are the longitude and latitude of the taxi. The fields ‘S’ and ‘D’ are the instantaneous speed and direction of the taxi. The field ‘State’ indicates whether there are passengers in the taxi and has a value of 0 (vacant) or 1 (occupied).

Table 2. Sample records of taxi trajectory data.

Date	Time	GPSID	Lon	Lat	S	D	State
20180604	6	10068	121.73326	31.035209	0	294	1
...
20180604	122149	10142	121.801537	31.149965	10	97	0

We removed missing and outlier values from the raw data. The coordinates were converted to locations in a projected coordinate system, which is WGS 1984 UTM Zone 51N in this case study. The pick-up and drop-off spots were extracted following the method described in Section 3.1. Each record in the pick-up and drop-off data was transferred to a point in shapefile format. Square grids with sides of 500 m covering the whole study area were drawn as shown in Figure 5. We calculated the total number of pick-up and drop-off spots in each grid. We repeated the steps above for seven days of data, which were separated into workday (Monday to Friday) and weekend (Saturday and Sunday) data. The hourly average of the total number in each of the two data groups was calculated in each grid. Then, we obtained the workday grids and weekend grids, which represent the human activity intensity on workdays and weekends in Shanghai, respectively. As an example, the statistical results for weekdays are shown in Figure 5 using graduated color.

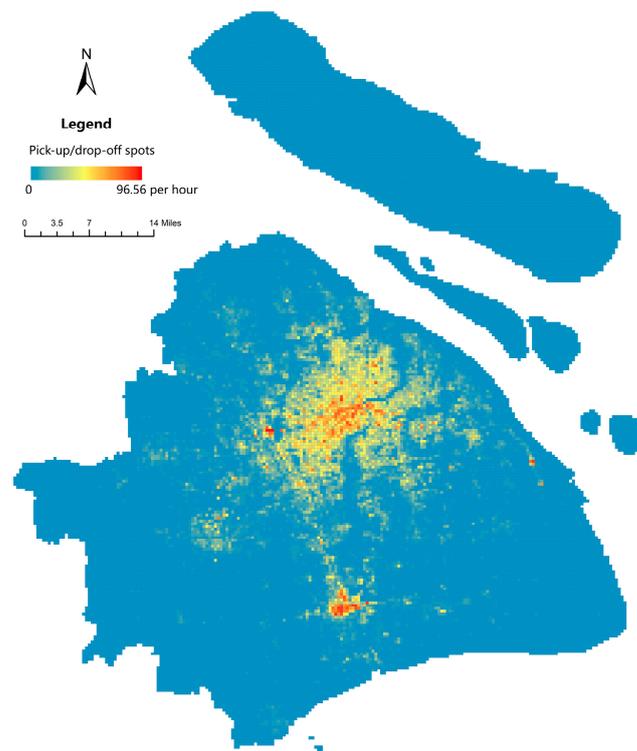


Figure 5. Grids and statistical results on weekdays.

4.2. Identification and Analysis of Urban Centers

We generated contour lines from the resulting statistical grids. In this case study, we used a 5×5 Gaussian filter with a sigma value of 1 to smooth the noisy data and set both the initial contour value and contour interval to 1. A contour level of 1 means that the grid has an average of 1 pick-up/drop-off spot within an hour. Similarly, a contour value of 60 means that the grid has an average of 60 pick-up/drop-off spots within an hour. The contour map of the weekday grids is shown as an example in Figure 6. The outlined area in the figure is enlarged in the lower-left corner to show the detailed contour lines since they are dense in some areas. The contour values increase as the color changes from yellow to blue. The contour lines are found to be much denser when the contour value is high.

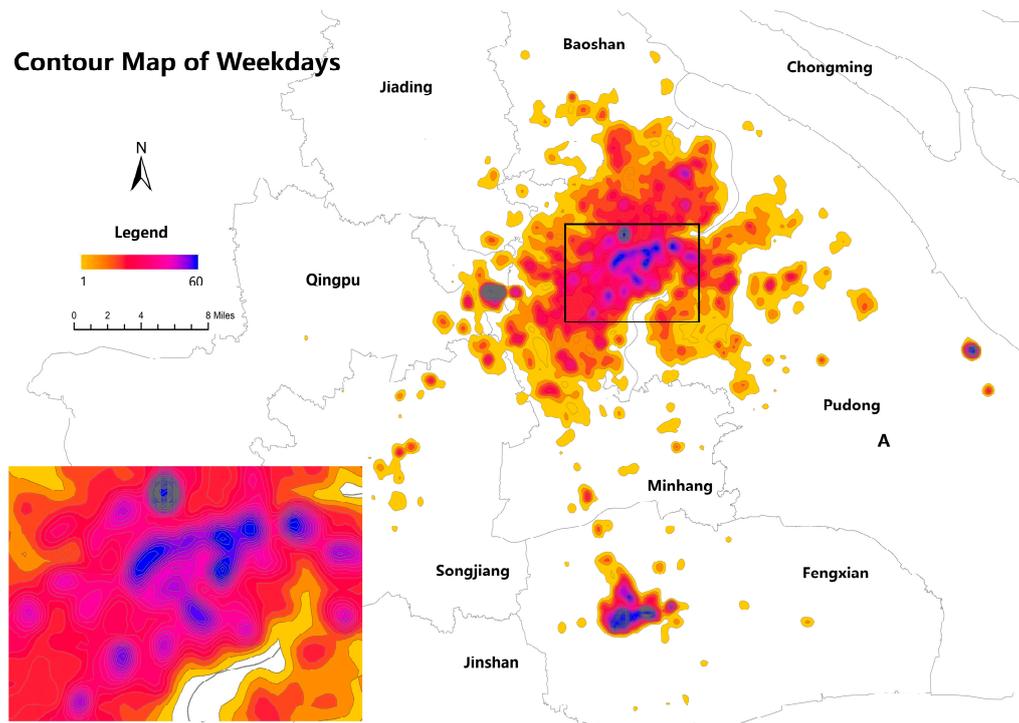


Figure 6. Weekday contour map.

As mentioned in Section 3.2, we removed the closed contour lines with areas less than 1 square kilometer to eliminate potential centers that were too small. After that, we applied the localized contour tree method to generate contour trees. Taking the simplified contour tree for weekdays as an example, we obtained 30 sub-trees, as shown in Figure 7a. There were 76 leaf nodes in total. Among these leaf nodes, 27 were one-node sub-trees (yellow features in Figure 7a), and the others belonged to three complex sub-trees (blue features in Figure 7a). One complex sub-tree had two levels, and another had three levels. The largest sub-tree had 11 levels and covered most of Shanghai's central area (within the outer ring road), as shown in Figure 7b. As the largest sub-tree is too complex (Figure 7b) and the space is limited, Fengxian District is used as an example to illustrate our tree construction (Figure 7c). In Figure 7c, the sub-tree is a 3-level contour tree. Leaf nodes 2, 3 and 4 at level 1 are three MUCs. Node 1 consists of nodes 3 and 4 and represents a PUC with two sub-centers at level 2. Then, node 0 is a root node of this small sub-tree and represents a more complex PUC than node 1. Even though nodes 2, 3 and 4 are all at level 1, Figure 7c indicates that nodes 3 and 4 have closer relationships since they belong to node 1. When verifying the map, we found a canal between nodes 1 and 2. This shows that the extracted hierarchy is in line with the actual situation. Following these rules, the hierarchy within the detected PUCs was analyzed.

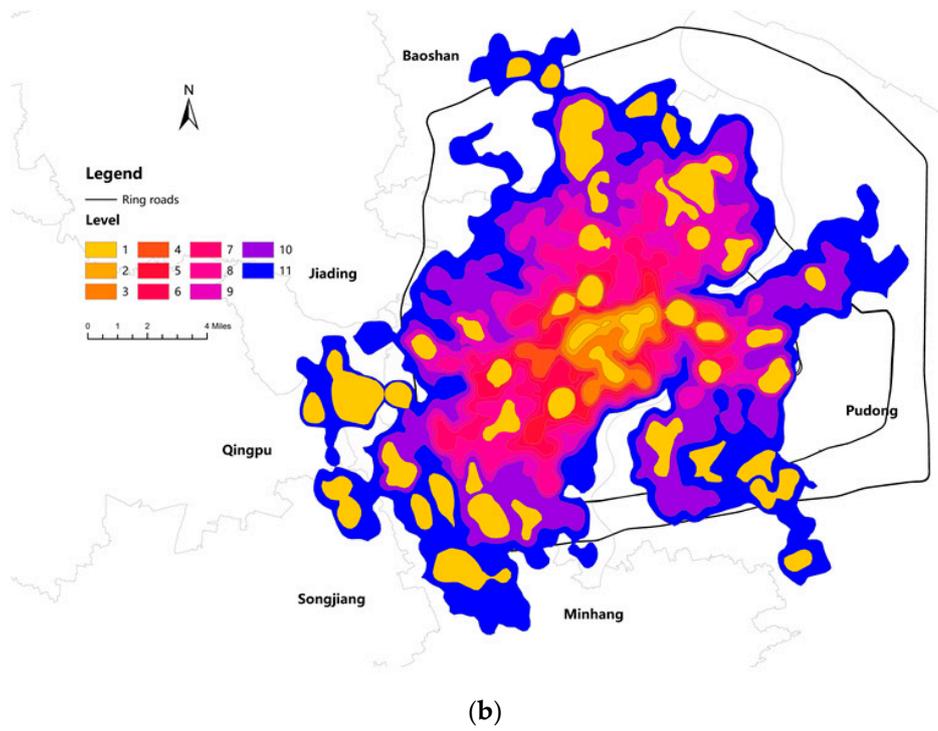
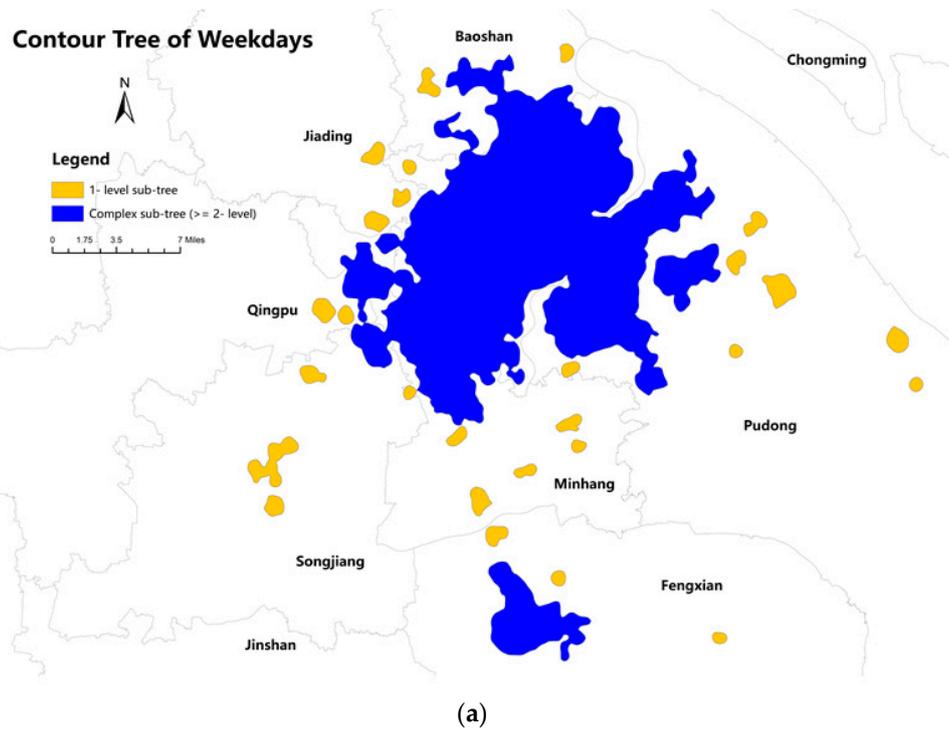


Figure 7. Cont.

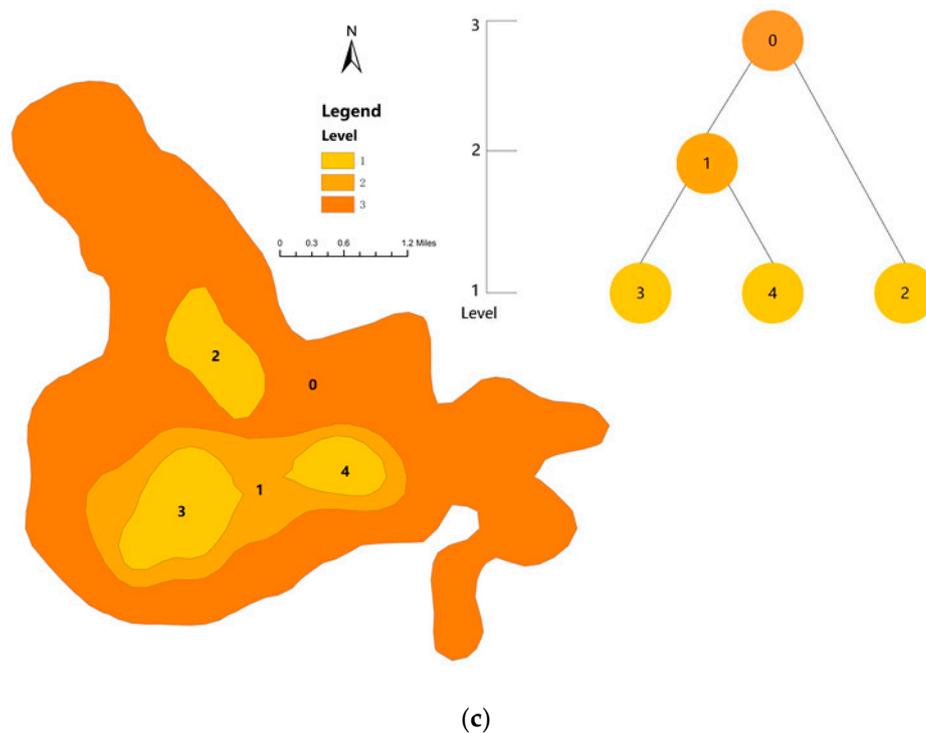


Figure 7. Contour tree for weekdays. (a) Thirty sub-trees for weekdays; (b) the largest sub-tree for weekdays; (c) the sub-tree in Fengxian District and its construction.

Continuing with weekday urban centers as an example, we detected 76 MUCs in total, as illustrated in Figure 8. We compared our detected centers with points of interest (POIs) and land use according to OpenStreetMap (<https://www.openstreetmap.org/>, accessed on 15 October 2020). PUC A is located in Fengxian District and consists of three MUCs. MUCs 71, 75 and 76 cover five commercial markets in Fengxian District, and PUC A covers almost the entire commercial and industrial area of Fengxian New City, which is included in the Shanghai City Master Plan (2017–2035) [56]. PUC B is located in Pudong New District and is within the outer ring road of Shanghai. It includes MUCs 41, 42 and 43. They cover most of Shanghai Zhangjiang Hi-Tech Park and several universities. PUC C is the largest center and contains 43 MUCs. It covers most hotspots and landmarks in Shanghai’s central area. MUCs 67 and 68 cover Lujiazui Financial Center, and MUC 70 contains Shanghai Railway Station and Sleepless City. MUC 72 is close to Ruijin Hospital and Huaihai Middle Road. MUC 74 includes People’s Square, Nanjing Road and the Shanghai Government. MUC 66 comprises Xujiahui. Hongqiao Airport and Hongqiao Railway Station are covered by MUCs 39 and 40 together. MUCs 53 and 61 contain Jiangwan Stadium, Wujiaochang, Tongji University and Fudan University. In addition to the MUCs that belong to the three PUCs, there are 27 individual MUCs. MUCs 12 and 17 are at Pudong International Airport. MUC 16 is Shanghai Disney Resort, and MUCs 4 and 6 are near Songjiang University Town and Songjiang Technological and Industrial Zone.

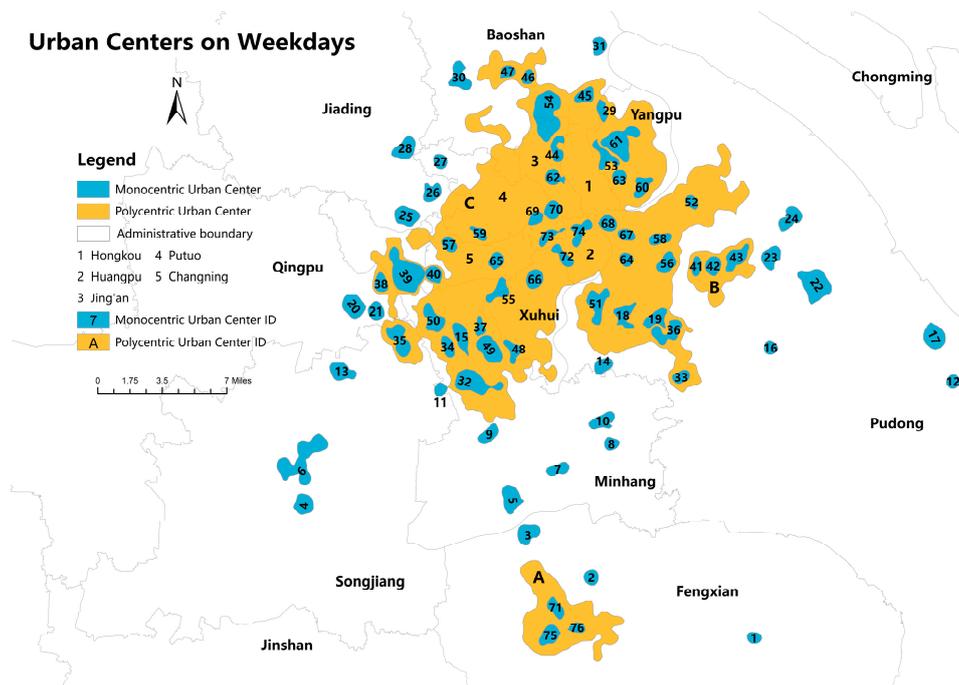


Figure 8. Urban centers on weekdays.

To verify the accuracy of the proposed method, in addition to comparing our results with those of Chen et al. [18], we also referred to the Shanghai City Master Plan (2017–2035) [56]. The plan describes Shanghai’s four-level system of public activity centers (a. the main center (central activity area); b. the sub-center; c. the district center; and d. the community center). In Table 3, we list all 74 MUCs on weekdays, indicate the locations covered by these MUCs and compare them with the planned activity centers in Shanghai.

As can be seen in Table 3, most of the 33 centers identified by the NTL data of Chen et al. [18] are included in our results. Since the centers obtained from taxi data are more detailed, multiple MUCs commonly correspond to one NTL center. For example, the main center detected in NTL data was defined as the center of Lujiazui, People’s Square and Shanghai Railway Station, whereas we had six MUCs covering this area, which includes these landmarks.

Compared with Shanghai’s four-level system of public activity centers, we cover most of the key locations in the planned central activity area (marked “a” in Table 3) and include 9 sub-centers out of 16 planned ones (marked “b” in Table 3) and 19 district centers (marked “c” in Table 3).

Most of the undetected centers that are included in the two references are in suburban areas such as Jiading District, Qingpu District, the northern part of Baoshan District, Jinshan District, the southern part of Pudong New District and Chongming District. These are areas outside the outer ring road of Shanghai, which are economically underdeveloped, industrial or new planned development zones. The intensity of human activity is low in these areas, and taxis are not the most economical mode of transportation for remote areas. Since taxi drivers can choose their activity range independently, they may prefer to avoid these areas.

Table 3. Validation of urban centers.

Location(s)	MUCID (PUCID) ¹	Identified by Chen et al. ²	Planned Activity Centers ³	Location(s)	MUCID (PUCID)	Identified by Chen et al.	Planned Activity Centers
People's Square, Shanghai Government, The Bund, Nanjing Road	74(C)	✓	a	Jiangqiao Town	25, 26		c
Jing'an Temple, Jing'an Park	73(C)		a	Taopu Town	27		c
Ruijin Hospital, Huaihai Middle Road	72(C)		a	Nanxiang Town	28	✓	c
Shanghai Railway Station, Sleepless City	70(C)	✓	a	Kongjiang Road, Xinhua Hospital	63(C)		c
Lujiazui	67(C), 68(C)	✓	a	Gaoqing Road	18(C)		c
Xujiahui	66(C)	✓	a	Gubei, Caohejing Development Zone	55(C)		c
Yangpu Riverside	60(C)		a	Gucun Town	30, 46(C), 47(C)	✓	c
Hongqiao Development Zone	65(C)		a	Zhoupu Town	33(C)	✓	c
Fengxian	71(A), 75(A), 76(A)		b	Qibao Town	34(C), 50(C)		c
Zhangjiang Town, Zhangjiang Hi-tech Park	41(B), 42(B), 43(B)	✓	b	Kangqiao, Yuqiao	36(C), 19(C)		c
Songjiang University Town, Songjiang New City	6	✓	b	Gongkang	54(C)		c
Chuansha New Town	22		b	Daning	62(C)		c
Wusong	31		b	Wujing Town	7		
Jiangwan Stadium, Wujiaochang, Tongji University, Fudan University	53(C), 61(C)	✓	b	Pudong International Airport	12, 17	✓	

Table 3. Cont.

Huamu-Longyang Road	56(C)		b	Shanghai Disney Resort	16	
Xinzhuang	11, 32(C)	✓	b	Xujing Town	20, 21	
Jinqiao	52(C)	✓	b	Hongqiao Railway Station, Hongqiao International Airport	39(C), 40(C)	✓
Fengcheng Town	1		c	Renji Hospital	64(C)	
Songjiang Old City	4		c	Changfeng Ecological Business District	59(C)	✓
Jiangchuan, Zizhu Science Park	5	✓	c	Government of Pudong New District, Shanghai Science and Technology Museum	58(C)	
Pujiang Town, Renji Hospital	8, 10	✓	c	Songnan Town	45(C)	
Zhuanqiao Town	9		c	National Convention and Exhibition Centre	38(C)	
Sijing Town	13		c	Other (residential area, major road, etc.)	2, 3, 14, 15(C), 24, 29(C), 35(C), 37(C), 44(C), 48(C), 49(C), 51(C), 57(C), 69(C)	
Tang Town	23		c			

¹ For example, 74(C) represents MUC 74, which is included in PUC C. ² ✓ means that the center is detected in Chen et al. [18]. ³ a, b and c mean that the center is located in one of the first three levels of planned activity centers (a. the main center (central activity area); b. the sub-center; c. the district center) mentioned in the Shanghai City Master Plan (2017–2035) [56].

Furthermore, we detected some centers that are not mentioned in one or both references, which are listed in Table 3. For example, Shanghai Disney Resort (MUC 16), Jing'an Temple (MUC 73), Yangpu Riverside (MUC 60), Fengxian (MUCs 71, 75 and 76), some towns, etc., are not included in the NTL results. Shanghai Disney Resort is an amusement park that closes around 8:30 pm. Such areas are not illuminated for long at night, so they cannot be captured by NTL data. As a result, NTL data cannot detect such centers, whereas taxi data, which can capture 24-h human activity, can identify them.

Some small towns, as well as Shanghai Disney Resort (MUC 16), Hongqiao Railway Station and Hongqiao International Airport (MUC 39, 40), Pudong International Airport (MUC 12, 17), etc., are not mentioned by name in the first three levels of Shanghai's public activity center system. Although Shanghai Disney Resort is not explicitly named in the Shanghai Master Plan, it is located near Chuansha New Town, which is planned as an activity center. As railway stations and airports are traffic hubs in the city, we consider them to be urban centers with transportation functions. Moreover, some MUCs do not cover any distinct landmark, but they either include residential areas and small hospitals, such as Pediatric Hospital of Fudan University (MUC 49), or encircle major roads, such as the outer ring highway (MUC 15). They are all preferred destinations or routes used by people in their daily activities.

Overall, except for suburban areas, our results are reliable, and most urban centers can be successfully detected in more detail.

4.3. Attributes of Urban Centers

The five statistical attributes and areas of MUCs on weekdays are illustrated by box plots in Figure 9a. The statistical attributes of 76 MUCs on weekdays show that the intensity is not equal between MUCs, which means that the MUCs do not attract an equal degree of activity due to their different locations. The three MUCs within PUC A have relatively high intensity, as they are located in the important development center of Fengxian District. Other MUCs with relatively high intensity are mostly transportation and commercial centers such as Hongqiao Railway Station (MUC 39), Pudong International Airport (MUC 12), Lujiazui Financial Center (MUC 68), etc.

The total area of the MUCs detected on weekdays is 172.02 km². The average area of 76 MUCs is 2.26 km². Most MUCs have areas less than 4 km², and only six MUCs (MUCs 6, 54, 39, 22, 61, 32) are larger than 4 km². The largest MUC is around Songjiang University Town (MUC 6) and is 8 km². The centers of Hongqiao Railway Station (MUC 39), Wujiaochang (MUC 61) and Chuansha New Town (MUC 22) are larger than 5 km². Additionally, MUC 32 covers a large region that encompasses residential areas in Xinzhuang Town, and MUC 54 covers residential areas and a section of the north–south viaduct.

Among the three PUCs on weekdays, PUC C has an area of 511.04 km² and covers most parts inside the outer ring. According to the statistical data of Shanghai, the central area within its outer ring is around 660 km² [56]. Therefore, PUC C covers around 77% of the central area of Shanghai.

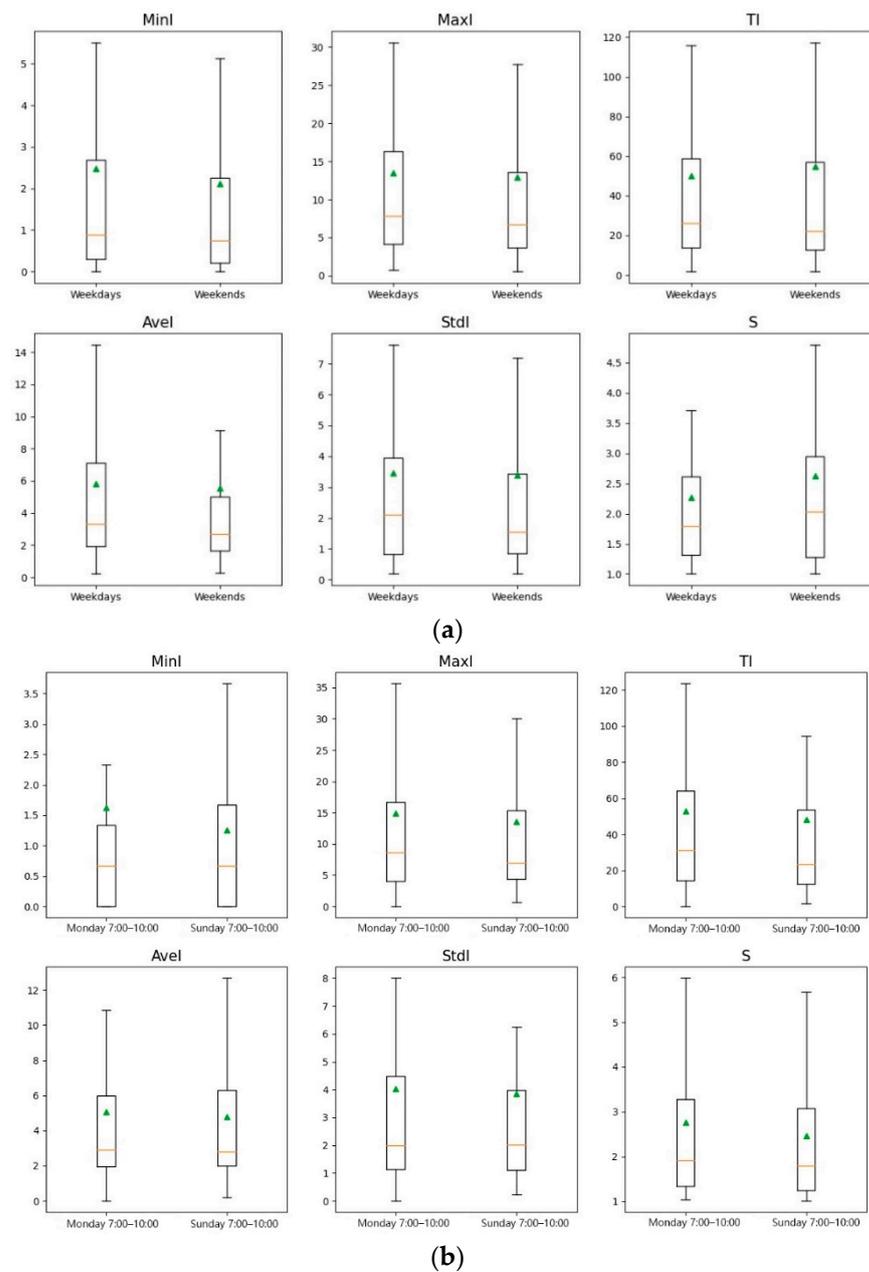


Figure 9. The statistical attributes and area of MUCs. (a) A comparison between weekdays and weekends; (b) a comparison between Monday and Sunday. The green triangle represents the mean, and the orange line represents the median.

For the other four morphometric attributes, the Compactness Index (C) and Elongatedness (E) together reflect the compactness of MUCs. If the C value is large and the E value is small, then the MUC's shape is relatively compact. As shown in Figure 10a, most MUCs on weekdays are relatively incompact, and the most compact MUC on weekdays is MUC 17 (Pudong International Airport). The Urban Development Orientation (Φ) reflects the development orientation of MUCs. We find that, in Figure 11a, most urban centers develop along the road or around intersections, such as the typical MUCs 15 and 48. Furthermore, the shape and development direction of the urban center is related to the shape of the building (e.g., MUC 39, Hongqiao Railway Station) and the shape of the town or residential area (e.g., MUC 22, Shachuan New Town). The Intensity Gradient (G) reflects the changes in activity intensity within each urban center. Figure 11a also shows that the G value is relatively low in most urban centers, indicating that the intensity change is not very large

inside the center. The three transportation centers (MUC 70 (Shanghai Railway Station), MUCs 39 and 40 (Hongqiao Airport and Hongqiao Railway Station) and MUC 17 (Pudong International Airport)) and some commercial (e.g., MUCs 71, 75 and 76 (Fengxian)) and financial centers (e.g., MUC 68 (Lujiazui)) have high G values, which means that the activity intensity changes sharply within these centers.

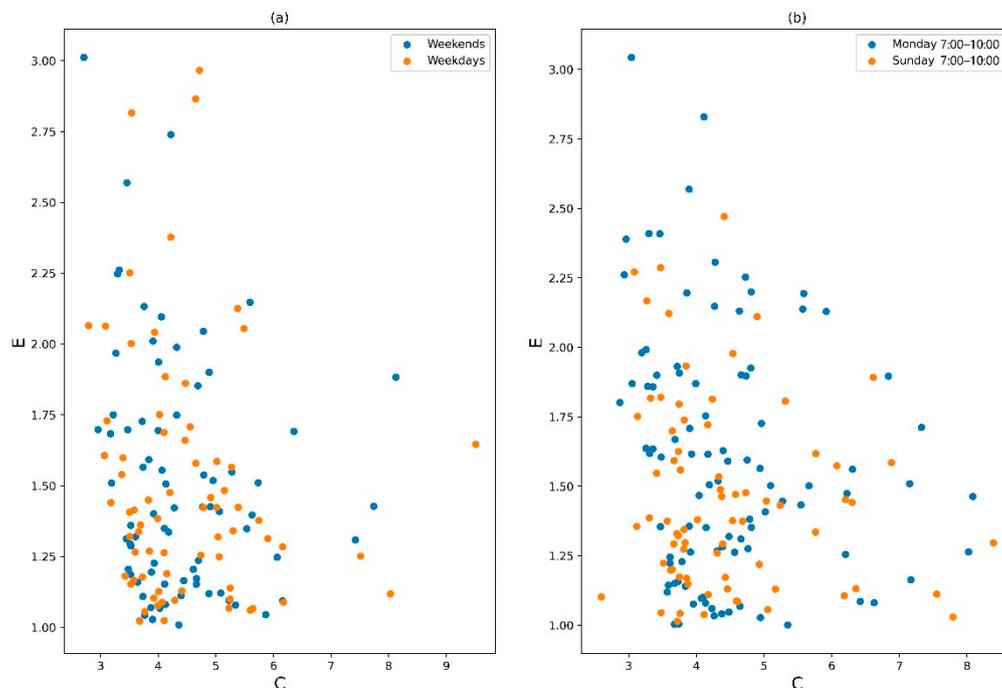


Figure 10. The compactness of MUCs. (a) A comparison between weekdays and weekends; (b) a comparison between Monday and Sunday.

The Polycentric Indicator (PI) was calculated for the three PUCs on weekdays. PUC A has a PI value of 0.66. The largest PUC C has a PI value of 0.47, which indicates its polycentricity. We believe that the polycentricity of PUC C is affected by certain MUCs with high intensity. PUC B has a high PI value of 0.85, which means that the three MUCs (MUCs 41, 42 and 43) within PUC B have a nearly equal impact.

4.4. Temporal Patterns of Urban Centers

We analyzed the urban centers detected on weekends to compare them with those on weekdays. Figure 12a illustrates the urban centers on weekends. Compared with urban centers on weekdays (Figure 12a), the overall range does not change significantly. Table 4 shows the basic information on weekday and weekend urban centers. The number of contour trees on weekends is 34 (more than that on weekdays), but the highest level of the largest contour tree is 10 (less than that on weekdays). The decrease in level means that the largest contour tree on weekends is less complex than that on weekdays. There are fewer MUCs on weekends than on weekdays, but the number of PUCs remains the same.

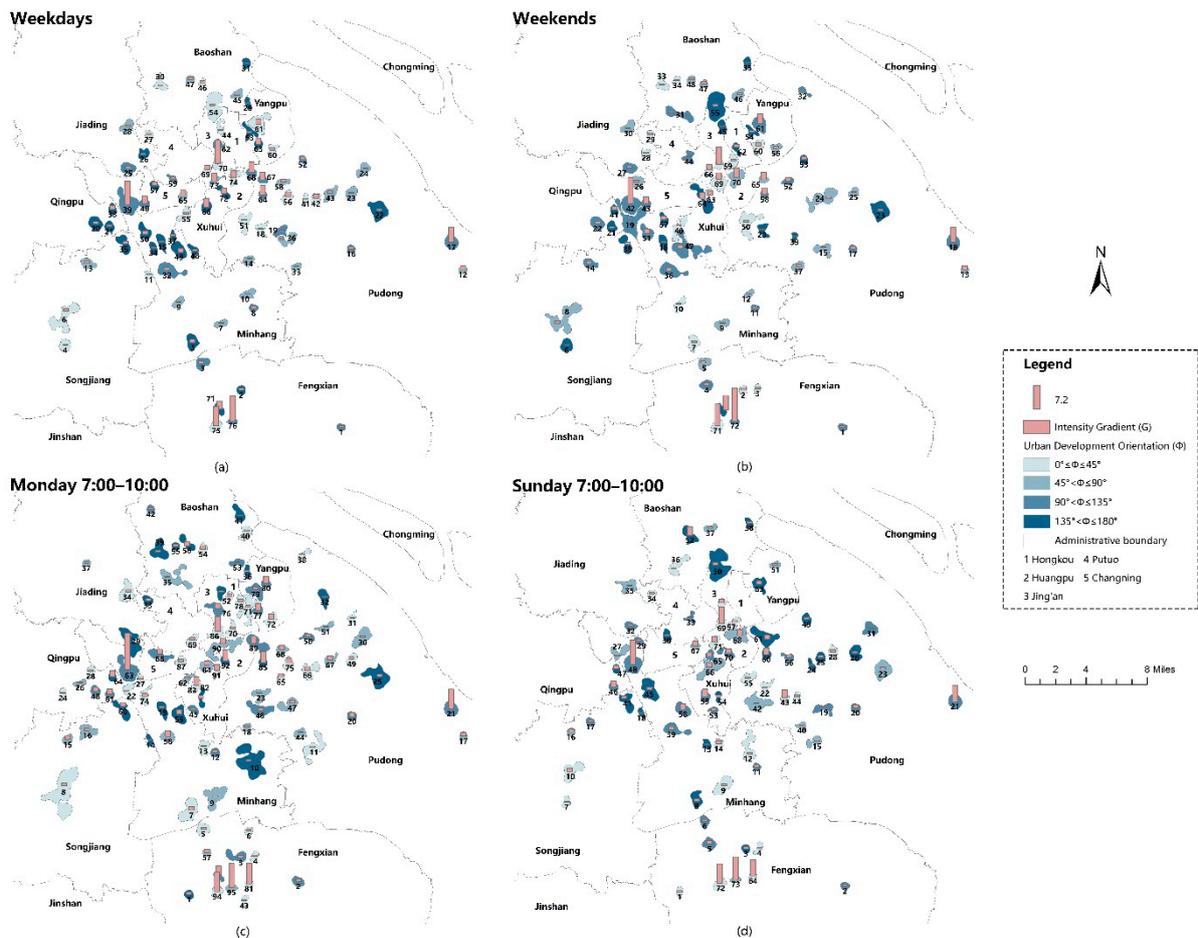


Figure 11. The urban development orientation (Φ) and intensity gradient (G) of MUCs. (a) Weekdays; (b) weekends; (c) Monday 7:00–10:00; (d) Sunday 7:00–10:00.

Table 4. Basic information on urban centers on weekdays and weekends.

	Highest Level	Contour Tree	MUC Number	MUC Area (km ²)	PUC Number	PUC Area (km ²)
Weekdays	11	30	76	172.02	3	511.04
Weekends	10	34	72	188.83	3	529.52

However, PUCs A' and C' on weekends shrink compared with PUC A and C on weekdays. In addition, the area of PUC B' on weekends is a part of PUC C on weekdays, and PUC B on weekdays becomes a smaller MUC (MUC 24 (Zhangjiang Hi-tech Park) on weekends in Figure 12a) on weekends. We find that a large number of MUCs on weekends have almost the same positions and boundaries as those on weekdays, for example, MUCs 18 and 13 (Pudong International Airport) on weekends (corresponding to MUCs 17 and 12 on weekdays). These MUCs represent urban centers that do not change between weekdays and weekends, such as transportation centers. There are still some MUCs on weekdays that disappear on weekends (e.g., MUCs 14, 15, 19, 24, 29, 56, 57, 59, 62 and 68 on weekdays), and other MUCs emerge on weekends but do not exist on weekdays (e.g., MUCs 3, 15, 19, 26, 31, 32, 34, 44, 57, 59 and 62 on weekends). Lujiazui, Changfeng Ecological Business District and Hongqiao Development Zone (MUCs 68, 59 and 65 on weekdays) all disappear on weekends. Some residential areas, such as Qibao Town (MUC 34 on weekdays), also disappear. The centers that disappear on weekends are mostly employment centers (e.g., Lujiazui) and some residential areas. Three urban centers that are planned in the Shanghai

City Master Plan (2017–2035) but are not detected on weekdays are delineated on weekends: these centers are Zhenru, Luxun Park and Sichuan North Road, and Gaoqiao Town (MUCs 44, 62 and 32 on weekends). The new centers on weekends are mostly residential areas (e.g., Gaoqiao Town) and parks (e.g., Luxun Park). Some urban centers with commercial and leisure functions, for instance, People’s Square, Shanghai Government, The Bund, Nanjing Road, Jing’an Temple and Jing’an Park (MUCs 69 and 70 on weekdays), are expanded compared with their sizes on weekdays. The most prominent one, MUC 64 on weekends, extends from Xujiahui to Zhongshan Park. This is because people’s commuting needs decrease significantly on the weekend, but leisure activities increase.

As shown in Figure 9a, the statistical attributes of MUCs on weekends are mostly distributed in a lower range than those on weekdays. However, the total intensity of all MUCs on weekends is 3953.94 and is higher than that on weekdays (3797.13). Even though some MUCs on weekends have the same locations and boundaries as they do on weekdays, their activity intensity is higher (e.g., People’s Square). Moreover, the total area of MUCs and PUCs on weekends is larger than on weekdays. For morphometric attributes, the most compact center on weekends is still Pudong International Airport. MUCs that have high *G* values on weekdays have even higher *G* values on weekends. Overall, in terms of the distributions of statistical attributes, many detected urban centers are less active on weekends, but there are several vital centers, such as transportation and commercial centers, that are even more active than on weekdays. This further reflects the increased demand for leisure and entertainment on weekends.

In order to extract finer temporal patterns, we analyzed urban centers on Monday and Sunday between 7:00 and 10:00, which is when people are on their way to work. The detected urban centers on Monday and Sunday are shown in Figure 12b. The centers clearly shrink on Sunday morning compared with Monday. The number and total area of MUCs significantly decrease on Sunday morning as shown in Table 5. Many MUCs on Monday morning disappear on Sunday morning, and those that still exist on Sunday morning mostly shrink. The MUCs that disappear on Sunday morning are mostly residential areas, towns, some major roads, intersections and viaducts. There are also some new MUCs on Sunday morning compared with Monday, for instance, Zhenru (MUC 33), Zhongshan Park (MUC 67), Shanghai Botanical Garden (MUC 54), Shanghai World Expo Park (MUC 55), etc., which shows that people tend to travel to leisure and entertainment destinations on Sunday.

Table 5. Basic information on urban centers on Monday and Sunday.

	Highest Level	Contour Tree	MUC Number	MUC Area (km ²)	PUC Number	PUC Area (km ²)
Mon. 7:00–10:00	8	40	95	264.13	3	602.12
Sun. 7:00–10:00	7	38	73	180.28	3	490.72

The statistical and morphometric attributes of MUCs on Monday and Sunday morning are illustrated in Figures 9b, 10b and 11c,d. The statistical attributes on Monday have a higher mean value than on Sunday. The total activity intensity on Monday morning is 5061.33 and is much higher than that on Sunday morning (3526), which means that Shanghai is much more active on Monday morning than on Sunday morning. The compactness, urban development orientation and intensity gradient of the centers do not vary much between Monday and Sunday. Overall, there are more centers on Monday morning due to commuting demands, but vital urban centers, such as transportation centers, are not influenced by temporal factors.

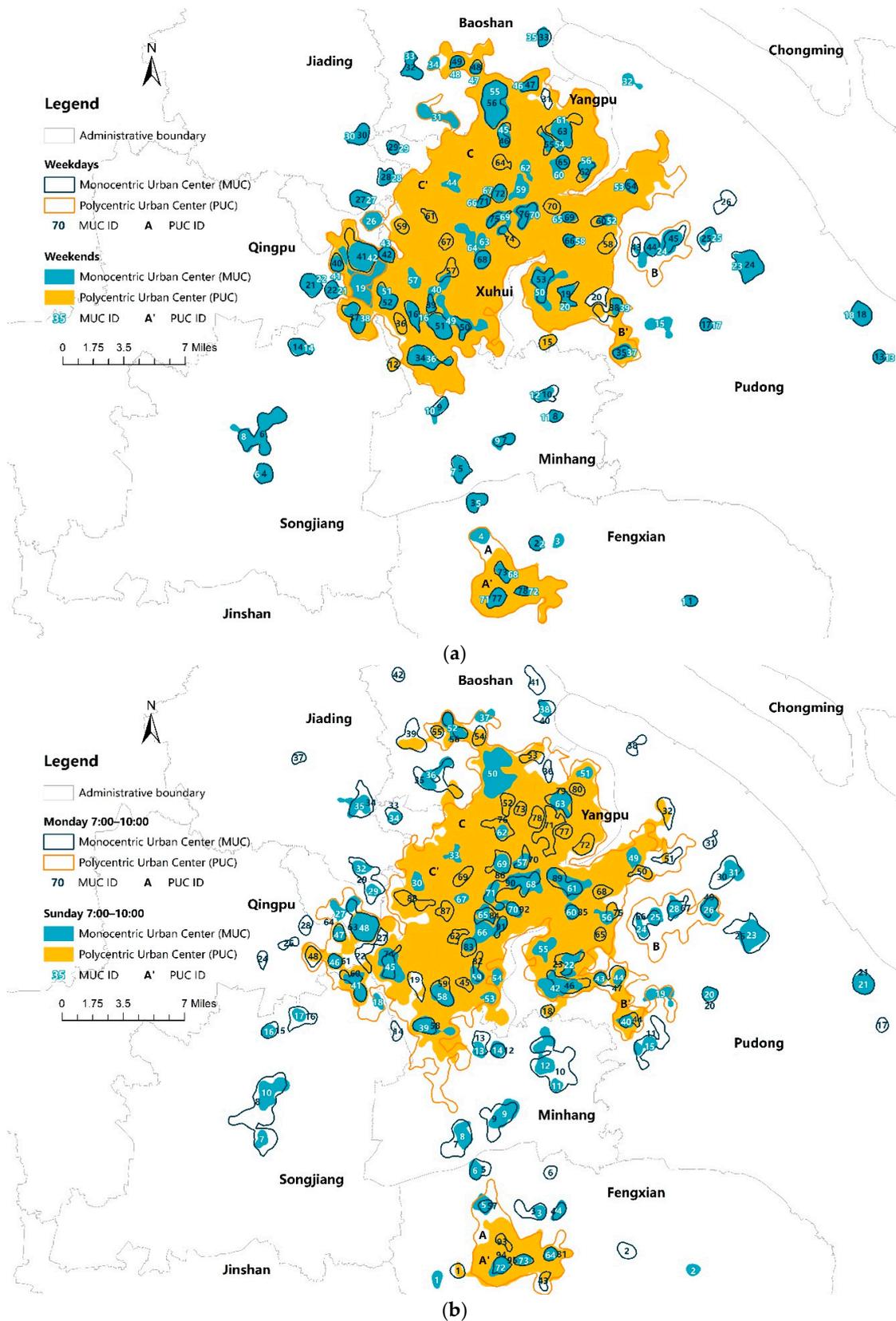


Figure 12. (a) A comparison between urban centers on weekends and weekdays; (b) a comparison between urban centers on Monday 7:00–10:00 and Sunday 7:00–10:00.

5. Discussion

Summarizing the results of our case study (Section 4), the urban centers of Shanghai were successfully detected and analyzed by applying our analysis framework (Section 3). We focused on the attributes and locations of monocentric urban centers (MUCs). The hierarchy and polycentricity of polycentric urban centers (PUCs) were also revealed. There were fewer urban centers on weekends than on weekdays, but the total area and activity intensity of the centers were larger and higher. In addition, we detected urban centers during the commuting period (7:00–10:00) on Monday and Sunday. The number and area of centers both increased on Monday relative to those on Sunday, which fits the typical lifestyle. Furthermore, the results also suggest that major urban centers are not affected by temporal factors.

We used our proposed analysis framework to apply taxi trajectory data to detect urban centers. The detected urban centers were mostly small in range but large enough to cover key hotspots, which is suitable for analyzing urban centers on small scales. Our results covered more than half of the 33 urban centers detected by nighttime light (NTL) data [18], as listed in Table 3. We identified a greater number of centers; for instance, 76 MUCs were detected on weekdays. Most of the delineated centers were relatively small compared with the NTL results. To be more specific, several detected centers corresponded to one center detected from NTL data. Therefore, our results were more detailed. Among our detected centers, some were not included in the NTL results, such as Fengxian New City in Fengxian District and several other residential areas, as explained in Section 4.2 and listed in Table 3.

Those centers that we did not identify were all located in the suburban areas of Shanghai or along the coastline, such as Jiading District, Qingpu District, the northern part of Baoshan District, etc. This omission is due to the lack of taxis in suburban areas. Since taxi drivers can choose their activity ranges independently, they naturally choose areas with high human activity intensity so that they can serve more passengers. This, in turn, will result in high taxi numbers in these areas, making them more likely to be urban centers.

Our results demonstrate the advantages of flexible resolution, high sampling frequency and rich semantic information for taxi trajectory data. First, we used the average number of pick-up/drop-off spots in a week to extract urban centers to eliminate deviations caused by a special event on a certain day. The average data in a long period can produce more generalizable results. In this study, some urban centers were found to be time-dependent. For example, Zhenru (MUC 44 on weekends in Figure 12a and MUC 33 on Sunday in Figure 12b) was only detected on weekends and Sunday morning. In addition, the results for Monday and Sunday morning show that most urban centers, with the exception of some major centers such as transportation centers, are strongly affected by time. Therefore, dynamic urban centers can be analyzed by dividing taxi trajectory data into more detailed time ranges, which cannot be done with NTL data.

However, there are still limitations in our method. This paper is based on the premise that our taxi data can represent all taxis in Shanghai and that taxis travel to every corner of the city. Unfortunately, as shown in our results, taxis are distributed unevenly and cannot cover the city completely. The taxis tend to be sparse in suburban areas. This is why we detected few centers in suburban areas. In addition, taxis are limited by road networks and buildings, so some places are inaccessible to them. Thus, further work should be conducted to address these limitations.

6. Conclusions and Future Work

In this paper, we propose an analysis framework to detect urban centers with boundaries from taxi trajectory data. We simulated taxi trajectory data with the topographic surface and generated contour maps. Then, we employed the localized contour tree method to delineate the boundaries of urban centers and analyzed the hierarchy within polycentric urban centers. We calculated the attributes of the detected urban centers and compared the results between different periods. We integrated a remote sensing method and taxi trajectory data to provide a new perspective for analyzing urban centers.

The work in this paper provides an affirmative answer to RQ1 raised in the Introduction. Taxi trajectory data can be used as a replacement for NTL data when detecting urban centers. In the presented study, we identified 76 monocentric urban centers (MUCs) on weekdays, and most of them agree with Shanghai's actual situation and the master plan, as illustrated in Section 4.2. The results are more detailed than those of NTL data, and we performed our method for different time periods to analyze temporal patterns of urban centers. Although urban centers in suburban areas were rarely detected due to the limitations of taxi trajectory data, as discussed in Section 5, the results are still reliable and representative.

In answer to RQ2, the advantages of using taxi trajectory data for urban center detection are apparent. The topographic surface simulated by taxi trajectory data has an unrestricted spatial and temporal resolution. We used grids with sides of 500 m in this analysis, but theoretically speaking, our analysis framework is suitable for grids with any side length. In addition, more detailed urban centers can be analyzed by using taxi trajectory data, as detected centers with small areas are more likely to correspond to actual landmarks. Moreover, the detection of urban centers is not limited by time period since taxi trajectory data can reflect human activity throughout a 24-h period. Therefore, urban centers can be extracted at finer temporal scales. Finally, due to the rich semantic information of taxi trajectory data, different statistical parameters can be designed according to different purposes. Then, different simulative topographic surfaces can be obtained, and centers with different meanings can be extracted.

There is much work to do in the future. The effects of the limitations of taxi trajectory data can be reduced by adding more taxis, using data from a longer period and including data on other forms of transportation. In addition to addressing the limitations of taxi trajectory data, as one of the future research directions, dynamic urban centers can be analyzed by using our analysis framework over shorter periods. Moreover, urban functional areas can be detected directly through spatiotemporal variation in urban centers. We can also verify the degree of completion of the city master plan by analyzing urban centers in different years. Overall, the analysis framework that we propose can be applied to study urban structure, urban transportation and many other aspects.

Author Contributions: Hongchao Fan and Mengqi Sun designed the experiments. Mengqi Sun performed the experiments under the supervision of Hongchao Fan. Mengqi Sun wrote the paper, and Hongchao Fan revised it. Both authors have read and agreed to the published version of the manuscript.

Funding: This research was supported by the NSFC (National Natural Science Foundation of China) project no. 41771484.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: No new data were created or analyzed in this study. Data sharing is not applicable to this article.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Harris, C.D.; Ullman, E.L. The nature of cities. *Ann. Am. Acad. Political Soc. Sci.* **1945**, *242*, 7–17. [[CrossRef](#)]
2. Burgess, E.W. The growth of the city: An introduction to a research project. In *Urban Ecology: An International Perspective on the Interaction Between Humans and Nature*; Marzluff, J.M., Shulenberg, E., Endlicher, W., Alberti, M., Bradley, G., Ryan, C., Simon, U., ZumBrunnen, C., Eds.; Springer: Boston, MA, USA, 2008; pp. 71–78.
3. McDonald, J.F. The identification of urban employment subcenters. *J. Urban Econ.* **1987**, *21*, 242–258. [[CrossRef](#)]
4. McDonald, J.F.; McMillen, D.P. Employment subcenters and land values in a polycentric urban area: The case of Chicago. *Environ. Plan. A* **1990**, *22*, 1561–1574. [[CrossRef](#)]
5. Giuliano, G.; Small, K.A. Subcenters in the Los Angeles Region. *Reg. Sci. Urban Econ.* **1991**, *21*, 163–182. [[CrossRef](#)]
6. Song, S. Modelling worker residence distribution in the Los Angeles region. *Urban Stud.* **1994**, *31*, 1533–1544. [[CrossRef](#)]
7. Small, K.A.; Song, S. Population and employment densities: Structure and change. *J. Urban Econ.* **1994**, *36*, 292–313. [[CrossRef](#)]

8. McDonald, J.F.; Prather, P.J. Suburban employment centres: The case of Chicago. *Urban Stud.* **1994**, *31*, 201–218. [[CrossRef](#)]
9. Cervero, R.; Wu, K.-L. Polycentrism, commuting, and residential location in the San Francisco Bay area. *Environ. Plan. A* **1997**, *29*, 865–886. [[CrossRef](#)]
10. McMillen, D.P.; McDonald, J.F. Suburban subcenters and employment density in metropolitan Chicago. *J. Urban Econ.* **1998**, *43*, 157–180. [[CrossRef](#)]
11. McMillen, D.P. Nonparametric employment subcenter identification. *J. Urban Econ.* **2001**, *50*, 448–473. [[CrossRef](#)]
12. Shearmur, R.; Coffey, W.J. A tale of four cities: Intrametropolitan employment distribution in Toronto, Montreal, Vancouver, and Ottawa–Hull, 1981–1996. *Environ. Plan. A* **2002**, *34*, 575–598. [[CrossRef](#)]
13. McMillen, D.P. Identifying sub-centres using contiguity matrices. *Urban Stud.* **2003**, *40*, 57–69. [[CrossRef](#)]
14. Redfearn, C.L. The topography of metropolitan employment: Identifying centers of employment in a polycentric urban area. *J. Urban Econ.* **2007**, *61*, 519–541. [[CrossRef](#)]
15. Roca Cladera, J.; Marmolejo Duarte, C.R.; Moix, M. Urban structure and polycentrism: Towards a redefinition of the sub-centre concept. *Urban Stud.* **2009**, *46*, 2841–2868. [[CrossRef](#)]
16. Hu, L.; Sun, T.; Wang, L. Evolving urban spatial structure and commuting patterns: A case study of Beijing, China. *Transp. Res. Part D Transp. Environ.* **2018**, *59*, 11–22. [[CrossRef](#)]
17. Shen, Y.; Batty, M. Delineating the perceived functional regions of London from commuting flows. *Environ. Plan. A Econ. Space* **2019**, *51*, 547–550. [[CrossRef](#)]
18. Chen, Z.; Yu, B.; Song, W.; Liu, H.; Wu, Q.; Shi, K.; Wu, J. A new approach for detecting urban centers and their spatial structure with nighttime light remote sensing. *IEEE Trans. Geosci. Remote Sens.* **2017**, *55*, 6305–6319. [[CrossRef](#)]
19. Krehl, A.; Siedentop, S.; Taubenböck, H.; Wurm, M. A comprehensive view on urban spatial structure: Urban density patterns of German city regions. *ISPRS Int. J. Geo Inf.* **2016**, *5*, 76. [[CrossRef](#)]
20. Taubenböck, H.; Standfuß, I.; Wurm, M.; Krehl, A.; Siedentop, S. Measuring morphological polycentricity—A comparative analysis of urban mass concentrations using remote sensing data. *Comput. Environ. Urban Syst.* **2017**, *64*, 42–56. [[CrossRef](#)]
21. Zhang, G.; Wang, J.; Jiang, Y.; Zhou, P.; Zhao, Y.; Xu, Y. On-Orbit Geometric Calibration and Validation of LuoJia 1-01 Night-Light Satellite. *Remote Sens.* **2019**, *11*, 264. [[CrossRef](#)]
22. Yang, H.; Ye, M.; Tang, W.H.; Wong, S.C. Regulating taxi services in the presence of congestion externality. *Transp. Res. Part A Policy Pract.* **2005**, *39*, 17–40. [[CrossRef](#)]
23. Castro, P.S.; Zhang, D.; Chen, C.; Li, S.; Pan, G. From taxi GPS traces to social and community dynamics: A survey. *ACM Comput. Surv.* **2013**, *46*, 1–34. [[CrossRef](#)]
24. Pan, G.; Qi, G.; Zhang, W.; Li, S.; Wu, Z.; Yang, L.T. Trace analysis and mining for smart cities: Issues, methods, and applications. *IEEE Commun. Mag.* **2013**, *51*, 120–126. [[CrossRef](#)]
25. Wu, H.; Huang, R.; You, L.; Xiang, L. Recent progress in taxi trajectory data mining. *Acta Geod. Sin.* **2019**, *48*, 1341–1356.
26. Peng, C.; Jin, X.; Wong, K.-C.; Shi, M.; Liò, P. Collective human mobility pattern from taxi trips in urban area. *PLoS ONE* **2012**, *7*, e34487. [[CrossRef](#)]
27. Zhang, W.; Li, S.; Pan, G. Mining the semantics of origin-destination flows using taxi traces. *UbiComp* **2012**, 522, 943.
28. Scholz, R.W.; Lu, Y. Detection of dynamic activity patterns at a collective level from large-volume trajectory data. *Int. J. Geogr. Inf. Sci.* **2014**, *28*, 946–963. [[CrossRef](#)]
29. Tang, J.; Liu, F.; Wang, Y.; Wang, H. Uncovering urban human mobility from large scale taxi GPS data. *Phys. A Stat. Mech. Appl.* **2015**, *438*, 140–153. [[CrossRef](#)]
30. Shen, J.; Liu, X.; Chen, M. Discovering spatial and temporal patterns from taxi-based Floating Car Data: A case study from Nanjing. *Gisci. Remote Sens.* **2017**, *54*, 617–638. [[CrossRef](#)]
31. Kong, X.; Xia, F.; Wang, J.; Rahim, A.; Das, S.K. Time-location-relationship combined service recommendation based on taxi trajectory data. *IEEE Trans. Ind. Inform.* **2017**, *13*, 1202–1212. [[CrossRef](#)]
32. Qi, G.; Li, X.; Li, S.; Pan, G.; Wang, Z.; Zhang, D. Measuring social functions of city regions from large-scale taxi behaviors. In Proceedings of the 2011 IEEE International Conference on Pervasive Computing and Communications Workshops (PERCOM Workshops), Seattle, WA, USA, 21–25 March 2011; pp. 384–388. [[CrossRef](#)]
33. Pan, G.; Qi, G.; Wu, Z.; Zhang, D.; Li, S. Land-use classification using taxi GPS traces. *IEEE Trans. Intell. Transp. Syst.* **2012**, *14*, 113–123. [[CrossRef](#)]
34. Yuan, J.; Zheng, Y.; Xie, X. Discovering regions of different functions in a city using human mobility and POIs. In Proceedings of the 18th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, Beijing, China, 12–16 August 2012; pp. 186–194. [[CrossRef](#)]
35. Liu, Y.; Wang, F.; Xiao, Y.; Gao, S. Urban land uses and traffic ‘source-sink areas’: Evidence from GPS-enabled taxi data in Shanghai. *Landsc. Urban Plan.* **2012**, *106*, 73–87. [[CrossRef](#)]
36. Zhao, K.; Chinnasamy, M.P.; Tarkoma, S. Automatic city region analysis for urban routing. In Proceedings of the 2015 IEEE International Conference on Data Mining Workshop (ICDMW), Atlantic City, NJ, USA, 14–17 November 2015; pp. 1136–1142. [[CrossRef](#)]
37. Zhang, D.; Wan, J.; He, Z.; Zhao, S.; Fan, K.; Park, S.O.; Jiang, Z. Identifying region-wide functions using urban taxicab trajectories. *ACM Trans. Embed. Comput. Syst.* **2016**, *15*, 1–19. [[CrossRef](#)]

38. Hu, Y.; Miller, H.J.; Li, X. Detecting and analyzing mobility hotspots using surface networks. *Trans. GIS* **2014**, *18*, 911–935. [[CrossRef](#)]
39. Models of Urban Structure. Available online: <https://www.owen.k12.ky.us/userfiles/257/Classes/17727/APHGModelsofUrbanStructure.pdf> (accessed on 20 July 2020).
40. Adams, J.S.; Hoyt, H. 1939: The structure and growth of residential neighborhoods in American cities. Washington, DC: Federal Housing Administration. *Prog. Hum. Geogr.* **2005**, *29*, 321–325. [[CrossRef](#)]
41. Sun, Y.; Fan, H.; Li, M.; Zipf, A. Identifying the city center using human travel flows generated from location-based social networking data. *Environ. Plan. B Plan. Des.* **2016**, *43*, 480–498. [[CrossRef](#)]
42. Ye, C.; Zhang, F.; Mu, L.; Gao, Y.; Liu, Y. Urban function recognition by integrating social media and street-level imagery. *Environ. Plan. B Urban Anal. City Sci.* **2020**. [[CrossRef](#)]
43. Zheng, Y.; Liu, Y.; Yuan, J.; Xie, X. Urban computing with taxicabs. In Proceedings of the 13th International Conference on Ubiquitous Computing, Beijing, China, 17–21 September 2011; pp. 89–98. [[CrossRef](#)]
44. Boyell, R.L.; Ruston, H. Hybrid techniques for real-time radar simulation. In Proceedings of the Fall Joint Computer Conference, New York, NY, USA, 12–14 November 1963; pp. 445–458. [[CrossRef](#)]
45. Freeman, H.; Morse, S. On searching a contour map for a given terrain elevation profile. *J. Frankl. Inst.* **1967**, *284*, 1–25. [[CrossRef](#)]
46. Kweon, I.S.; Kanade, T. Extracting topographic terrain features from elevation maps. *CVGIP Image Underst.* **1994**, *59*, 171–182. [[CrossRef](#)]
47. Wu, Q.; Liu, H.; Wang, S.; Yu, B.; Beck, R.; Hinkel, K. A localized contour tree method for deriving geometric and topological properties of complex surface depressions based on high-resolution topographical data. *Int. J. Geogr. Inf. Sci.* **2015**, *29*, 2041–2060. [[CrossRef](#)]
48. Wu, Q.; Lane, C.R. Delineation and quantification of wetland depressions in the Prairie Pothole Region of North Dakota. *Wetlands* **2016**, *36*, 215–227. [[CrossRef](#)]
49. Liu, H.; Wang, L.; Sherman, D.; Gao, Y.; Wu, Q. An object-based conceptual framework and computational method for representing and analyzing coastal morphological changes. *Int. J. Geogr. Inf. Sci.* **2010**, *24*, 1015–1041. [[CrossRef](#)]
50. Yu, B.; Shu, S.; Liu, H.; Song, W.; Wu, J.; Wang, L.; Chen, Z. Object-based spatial cluster analysis of urban landscape pattern using nighttime light satellite images: A case study of China. *Int. J. Geogr. Inf. Sci.* **2014**, *28*, 2328–2355. [[CrossRef](#)]
51. Green, N. Functional polycentricity: A formal definition in terms of social network analysis. *Urban Stud.* **2007**, *44*, 2077–2103. [[CrossRef](#)]
52. Liu, X.; Derudder, B.; Wu, K. Measuring polycentric urban development in China: An intercity transportation network perspective. *Reg. Stud.* **2016**, *50*, 1302–1315. [[CrossRef](#)]
53. Zhang, T.; Sun, B.; Li, W. The economic performance of urban structure: From the perspective of Polycentricity and Monocentricity. *Cities* **2017**, *68*, 18–24. [[CrossRef](#)]
54. Liu, X.; Wang, M. How polycentric is urban China and why? A case study of 318 cities. *Landsc. Urban Plan.* **2016**, *151*, 10–20. [[CrossRef](#)]
55. Statistical Communique of Shanghai on National Economic and Social Development in 2019. Available online: <http://tjj.sh.gov.cn/tjgb/20200329/05f0f4abb2d448a69e4517f6a6448819.html> (accessed on 31 March 2021).
56. Shanghai Master Plan 2017–2035. Available online: <http://www.shanghai.gov.cn/newshanghai/xgkfj/2035001.pdf> (accessed on 31 March 2021).