International Journal of
Geo-Information

MDPI

Article

Revealing the Correlation between Population Density and the Spatial Distribution of Urban Public Service Facilities with Mobile Phone Data

Yi Shi [1], Junyan Yang [1,*] and Peiyu Shen [2]

[1] School of Architecture, Southeast University, Nanjing 210096, China; shiyi@seu.edu.cn
[2] School of Landscape Architecture, Nanjing Forestry University, Nanjing 210037, China; spy@njfu.edu.cn
* Correspondence: yangjunyan@seu.edu.cn; Tel.: +86-139-5190-9977

Received: 15 November 2019; Accepted: 9 January 2020; Published: 13 January 2020

check for updates

Abstract: Some studies have confirmed the association between urban public services and population density; however, other studies using census data, for example, have arrived at the opposite conclusion. Mobile signaling data provide new technological tools to investigate the subject. Based on the data of 20 million 2G mobile phone users in downtown Shanghai and the land use data of urban public service facilities, this study explores the spatiotemporal correlation between population density and public service facilities' locations in downtown Shanghai and its variation laws. The correlation between individual population density at day vs. night and urban public service facilities distribution was also examined from a dynamic perspective. The results show a correlation between service facilities' locations and urban population density at different times of the day. As a result, the average population density observed over a long period of time (day-time periodicity or longer) with census data or remote sensing data does not directly correlation with the distribution of public service facilities despite its correlation with public service facilities distribution. Among them, there is a significant spatial correlation between public service facilities and daytime population density and a significant spatial correlation between non-public service facilities and night-time population density. The spatial and temporal changes in the relationship between urban population density and service facilities is due to changing crowd behavior; however, the density of specific types of behavior is the real factor that affects the layout of urban public service facilities. The results show that mobile signaling data and land use data of service facilities are of great value for studying the spatiotemporal correlations between urban population density and service facilities.

Keywords: spatial distribution of urban public service facilities; population density; spatiotemporal distribution; mobile phone data; floor area ratio

1. Introduction

Urban public service facilities can be defined as components of a city of which the primary function is to provide public goods and services, either wholly or partly supported by government [1,2]. Urban public service facilities play a crucial role in the public's everyday quality of life. Public facilities can be classified into historical amenities and relatively modern amenities, which are generated mainly by past and present government decisions regarding investment in education, medical care, transportation, and other infrastructure [3–5]. There are various types of public service facilities with different functions to support community activities [6]. The equitable distribution of public facilities is one of the major concerns of urban planners. To address the equity aspect of sustainable development, urban facilities should not only be accessible at, for example, transportation modes, but should be equally accessible among all population groups [7]. A large volume of research has been done in this respect and on related issues, such as transportation [8–15]. The goal of equity considering public services is ensuring equal distribution of resources, services, and benefits among individuals and fair and appropriate distribution in society [16]. A spatial equity assessment of urban facilities helps policy-makers and urban authorities to evaluate the effectiveness of the current urban services and facilities [17].

As cities develop and their inner structures become more complex, the study of population distribution within city boundaries becomes increasingly important [18]. It is also of great significance for urban planning—such as the division of urban functional areas, the allocation of public service facilities, and urban traffic planning [19–21]—and plays a critical role in the protection of urban public safety and crisis management [22].

Considerable research has been carried out on the spatial effect of population density and the correlation between population density and the distribution of public service facilities [23,24]. Li reported that the distribution of commercial centers is influenced by habitant distribution density [25]. Thus, commercial centers are often found in areas with high habitant distribution density. In recent years, additional studies have validated the correlation between population density and public service facilities distribution using quantitative analysis. Shen analyzed the level of correlation between population distribution and the spatial distribution of service facilities in Shanghai by combining street demographic data with night-time light remote sensing and land use data [26]. The author found a strong positive correlation between population distribution and the distribution of service facilities in Shanghai. In a study on the relationship between population and public transportation system of urban buses in Daejeon and Gwangju, Kwon found a strong positive correlation between the service load of the public transportation system and population distribution [27]. Reigadinha et al. demonstrated a positive correlation between the distribution of retail stores and population density using geographic information system data along with linear regression and correlative analysis [28].

However, other studies have drawn the opposite conclusion while identifying areas in cities with high population density but low density of public service facilities. This phenomenon, which is not hard to find in the city, is referred to as incongruity between population distribution and public service facilities [29]. The above studies illustrate that it is too early to make an conclusion about how population density effect on the distribution of urban public service facilities, although population density distribution is undoubtedly an important factor in the planning of urban public service facilities. The data on which population density distribution is measured may be the cause of this inconsistency. Population distribution data obtained via means such as household registration do not reflect the actual distribution and behavior of people; such data only reflect the geographical spatial distribution of the population. In contrast, data such as street demographic data can truly reflect the overall population distribution over a long time period. Although night-time light remote sensing data reflect the distribution of human behavior at a particular point, they cannot recognize and classify the types of behaviors. Due to the above constraints, the correlation between population distribution and service facilities distribution may be affected by various factors, making it difficult to explain the contradiction and incongruity observed in past studies.

Recently, the rapid development of big data technology and mobile information has greatly increased the available methods for urban studies, especially in the area concerning the dynamic spatial and temporal relationships of human behavior and environmental studies [30–32]. Mobile phone data are of the most popular datasets among newly available data, as it can provide precise spatial and temporal information, which can be used with precision in urban studies, such as urban structure research [33–35], research on accessibility to service facilities [36], and identification of urban functional areas [37]. Due to the popularity of modern communication facilities in developing countries, mobile phone data have become a data source that can be used to dynamically measure population density. With mobile phone data of 20 million 2G users of Shanghai, this article investigated the relationship between urban public service facilities distribution and the spatiotemporal distribution of population density. The correlation between individual population density at day vs. night and urban public service facilities distribution from a dynamic perspective was also examined from a dynamic perspective. This paper aimed to answer the following questions. (1) Are the effects of high population density on the development of urban public service facilities, compare with other kind of land use really differ? (2) How is the correlation between the distribution of public service facilities and the dynamic population density from day to night? (3) How is the correlation between non-public service facilities and the population activity density from day to night as a contrast? (4) What can be applied in urban planning and urban sustainable development after understanding the relationship between dynamic population density and public service distribution?

The remainder of this study is structured as follows. Section 2 introduces the methodology and data, while the correlation results are shown in Section 3. Section 4 reveals the mechanism and Section 5 presents the main conclusions.

2. Data and Methodology

2.1. Research Area

This study utilizes mobile phone data and land use data on Shanghai, China. Shanghai is one of the most economically developed cities in the country and serves as the economic, financial, trade, and shipping center of mainland China. It functions as the hub of China's economic interaction with the world and is home to the head offices of many multinational companies in China. It is also the leading urban area of the Yangtze River Delta economic zone. At the end of 2013, there were about 24.1515 million permanent urban residents living in the city. In the same year, the GDP generated by the city was approximately 2.13982 trillion Yuan, with a per capita GDP of USD 14,442.33, reaching a similar GDP per capita level to that of developed countries. With the booming consumer demand, there is an urgent need to provide the city with rich, diversified, and specialized public service facilities.

This study focuses on the central urban area of Shanghai, an area of ~66,000 hectares. (Figure 1) The spatial unit is of great significance when analyzing the dynamic distribution characteristics of urban population density. Land use (lot) is a commonly used data unit in the field of urban planning, and it is relatively simple to convert data units such as blocks, traffic districts, and administrative districts to this unit. Therefore, this study utilizes 12,584 plots in the central area of Shanghai as the analysis unit, which is of great value for urban planning and management.

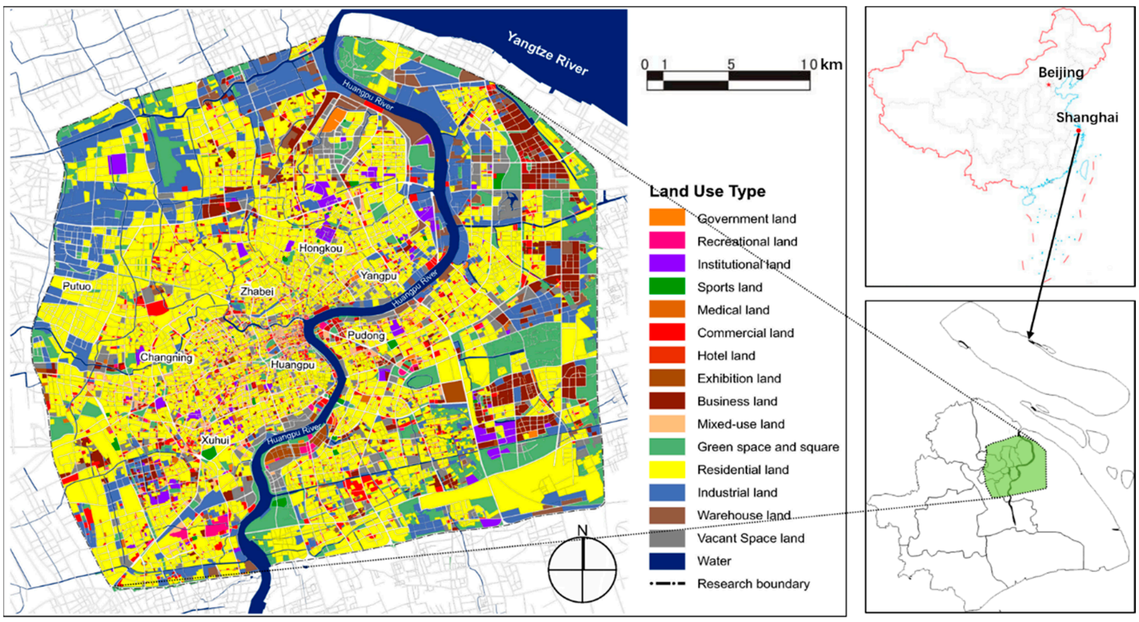

Figure 1. Study area. Source: Plotted by the author.

2.2. Data

2.2.1. Mobile Phone Data

The study "Mobile phone landscape: urban analysis using mobile phone location data", published in 2006, can be seen as the first work on the intervention of mobile phone signaling data in urban planning. Moreover, the relationship between individual behavior density and urban location is one of the earliest concerns of planning scholars. Previous works present a series of interviews concerning global cities such as Milan [38,39], Rome [40,41], and New York [42]. There have also been studies on mobile signaling data in several cities around the world, such as the work of Isaacman [43] and other studies on Boston [44] and Los Angeles [43]. As further validation, Isaacman [43] investigated mobile phone data in New York City, such as mobile phone call records (CDR) and text messages (SMG), and summarized and compared it with census data. Deville [45] used Portuguese and French mobile phone data (CDR and SMG) and compared it with remote sensing data of population distribution. Janzen [46] used mobile phone data to make up for the limitations of traditional survey methods in analyzing long-distance travel demand. All of these studies prove that a population distribution density calculation based on mobile phone data obtained from mobile base stations is very accurate. Recently, the use of mobile phone signal data was introduced. As CDR and SMG data does not have a high sampling rate (can record all on mobile devices positioning signal) or a high update rate (once every 5 to 10 min for positioning), mobile signaling data have become the preferred data type for urban studies of behavior density calculation, because of its accuracy.

The data used in this study consist of the 2G user data of a Chinese telecom operator in Shanghai and include base station data, mobile signaling static data, and mobile signaling dynamic data. The base station data include base station numbers and geographic coordinates. The mobile signaling static data refer to the number of cell phone services per hour of each base station on a given day. The mobile signaling dynamic data provide the location information recorded by the operator when the mobile phone user is active in the communication network, which is involuntarily provided. When a user's mobile phone starts up, shuts down, makes or receives calls, and receives an SMS, the action is recorded at the base station or mobile switching center. The location, mobile phone identification number, signaling success time, and cell station number at that time are all saved in the mobile phone signaling data. This signaling data use the encrypted mobile phone identification number after removing the user's attributes and do not involve the personal information of the mobile phone user. The mobile phone signaling data used in this study were recorded on eight designated days in 2013, including four working days and four holidays. On the designated days, approximately 20 million different mobile phone identification numbers were recorded by 9578 base stations in Shanghai per day, representing ~20 million 2G mobile phone users. The population sampling rate within the city area was approximately 70%. The data show that the average number of mobile phone users in the central area is ~1.5 million users per hour. The mobile phone data involved in this research are all anonymous data processed by communication service providers, and do not involve any mobile phone user privacy.

This study used the Thiessen polygon algorithm to calculate the range of each base station in Shanghai, so that there is only one base station within the range of each base station cell, and the distance from any point in the base station area to the station is smaller than the distance to other base stations. The average area of the base stations within Shanghai is 88.1 hectares, the average area of the base stations within the central urban area is 37.7 hectares.

The average population density per hour period from 0:00 a.m. to 12:00 p.m. were used from all data collected on weekdays in central Shanghai. Considering special events like major festival activities that may distort population density at certain times of the day in some areas, it is not possible to reflect the daily population density in these areas. Therefore, the variance coefficient approach was used to detect the variations in population density at the same time points among different days. With the exceptions of a few regions and fringe areas, the variance coefficients were all less than 20%.

2.2.2. Service Facilities Data

This study obtained data concerning the construction land area and building area of Shanghai central area. The statistical results show that the total area of the Shanghai central area is 66,000 hectares. The land types are divided into 21 categories, such as administration, retail, business, culture and art, and parks and green space. The building area of public service facilities physically built in the central area of Shanghai is 1.39 million square meters. (Table 1) Each land use data block (a total of 12,584 plots) provides information such as the land area and floor area of service facilities, address, and latitude and longitude. The building area data comes from baidu map open platform. The land use data come from the land use status chart published on the website of Shanghai planning bureau.

Table 1. Scale of public service facilities in central area of Shanghai.

Types of Public Service Facilities	Building Area (10,000 m^2)
Retail	3858.37
Business	5674.56
Financial	735.68
Hotel	749.48
Conference and Exhibition	286.82
Wholesale	293.80
Administration	396.52
Culture and Art	290.91
Hospital	384.08
Sports	138.78
Research Institution	1105.72
Total	13,914.73

Floor area ratio (FAR) was chosen as the metric for investigating the distribution of public service facilities in this study. Murphy and Vance proposed the quantitative definition method in 1954 [47]. Since then, the FAR of public service facilities has been used to evaluate public service facilities, and this ratio has been fully tested and extensively used in studies on Chinese downtowns [48]. The FAR of public service facilities can be used to evaluate the development degree of urban areas ranging from low- to high-grade commercial networks. All commercial centers grow from low-grade commercial networks together with the gradual development of public service facilities. Therefore, the FAR of public service facilities can be used as an index to represent the development level of public service facilities in various urban areas.

We processed the plot data as follows. First, we used the ArcGIS 10.2 spatial analysis platform to extract the land use data of the study area and 12,584 plots were obtained and numbered. Second, the floor areas of public service facilities and non-public service facilities in each plot were incorporated into the plot data. Finally, the plot ratio of public service facilities and non-public service facilities in each plot was calculated.

2.3. Methodology

2.3.1. Population Density Index

First, we calculated individual population density based on a three-dimensional activity space area. This is the core step, and the main difficulty of this algorithm is converting the mobile signaling data based on cell phone base stations into user spatial distribution data based on land use (lot). Deville et al. proposed an algorithm to calculate the dynamic population density of different administrative regions by taking the land area as the distribution weight in his population density study in southern Europe using mobile phone data [45]. This algorithm overlaps the boundaries of administrative regions with the boundaries of base station districts. Assuming that the population in each base station district is evenly distributed on a two-dimensional plane, the population of each superimposed block is calculated. Then, the population of superimposed blocks belonging to the same administrative region is summarized to obtain the population of each administrative region. The algorithm proposed in this study is further improved using Deville's method and uses a three-dimensional active space area instead of a two-dimensional land area as the allocation weight. To make the distinction, this algorithm is called the algorithm based on the three-dimensional active space area, whereas the algorithm based on the plane area used by Pierre and others as the allocation weight is called the algorithm based on the two-dimensional area.

The three-dimensional activity space defined in this study refers to the activity space of the main population in the city. This space is composed of architectural space and outdoor public space (but does not include public space that is difficult for people to use, such as open water or vegetation). For specific urban areas, the three-dimensional active space area can be expressed as follows,

$$A = A_0' + A_0$$

or

$$A = A' + A_0 - A_g'$$

where, A is the three-dimensional activity space area, $A_0{}'$ is the outdoor activity area of the area, A_0 is the total construction area, A' is the area, and $A_g{}'$ is the total floor area.

This study defines individual behavior density as the number of individuals moving on the land per unit area at a certain time. The following mobile phone user behavior density algorithm based on three-dimensional activity space area is proposed, and the formula is as follows.

$$\rho_{c_i} = \frac{1}{A'_{c_i}} \sum_{c_i} \frac{A_{(c_i \cap v_j)} D_{v_j}}{A_{v_j}}$$

Here, v_j represents the base station cell number j, c_i represents the land number i, ρ_{c_i} is the density of mobile phone users in area c_i, A'_{c_i} represents the area of land in area c_i, D_{v_j} represents the number of mobile phone users at a certain time in the base station cell v_j, A_{v_j} represents the area of the three-dimensional moving space of the base station cell v_j, $A_{(c_i \cap v_j)}$ represents the area of the three-dimensional moving space in the overlapping area formed by the area c_i, and the base station is v_j.

This algorithm is applied to calculate the signaling data in the n period of the mobile phone base station with a quantity of j within the scope of the study. The data format with the base station as the unit is effectively transformed into the data format with the land as the unit. The format of the obtained behavior density data is shown in Table 2 below.

Table 2. Population density data format based on land use.

	t_1 Time Density	t_2 Time Density	t_3 Time Density	……	t_n Time Density
c_1 plot	$\rho_{c_1 t_1}$	$\rho_{c_1 t_2}$	$\rho_{c_1 t_3}$	……	$\rho_{c_1 t_n}$
c_2 plot	$\rho_{c_2 t_1}$	$\rho_{c_2 t_2}$	$\rho_{c_2 t_3}$	……	$\rho_{c_2 t_n}$
c_3 plot	$\rho_{c_3 t_1}$	$\rho_{c_3 t_2}$	$\rho_{c_3 t_3}$	……	$\rho_{c_3 t_n}$
……	……	……	……	……	……
c_m plot	$\rho_{c_m t_1}$	$\rho_{c_m t_2}$	$\rho_{c_m t_3}$	……	$\rho_{c_m t_n}$

Source: authors.

2.3.2. Linear Regression Analysis

Linear regression was employed to examine the correlation between public service facilities distribution and population density. Linear regression based on data from a large number of experiments is widely used as a metering method used to determine interactions between variables, the levels of influence of variables, and the static rules underlying numerical distributions. A geographically weighted regression (GWR) is normally an appropriate method for processing spatial nonstationary data. However, the correlation between the block plot ratio and population dynamic density in the different urban areas in this study is not caused by their different spatial locations, but by their different land use properties (land for public service facilities and land for non-public service facilities). As the distribution of land used for public service facilities and that for non-public service facilities is not a function significantly related to space, it is not necessary to discuss the nonstationarity of the space. Therefore, GWR or other spatial analysis models are not used in this study. Establishing a model for linear regression requires the following conditions.

1. The independent variables refer to nonrandom variables that are not interrelated, namely, $\mathrm{Cov}\left(x_i, x_j\right) = 0$;
2. Random error terms are independent of each other and follow a normal distribution with the expectation being zero and the standard deviation σ; namely, $\varepsilon_i \sim \mathrm{N}\left(0, \sigma^2\right)$; and
3. sample number is more than the number of parameters, namely, $n > p + 1$ $n > p + 1$.

Based on the theoretical hypothesis above, a model is created as follows,

$$y = \alpha + \beta x + \mu,$$

where y is the dependent variable, x is the independent variable, α interceptis the regression intercept, β is the regression coefficient, and μ is the random error.

The linear regression model adopts the significance test based on the regression coefficient R^2, which reflects the reasonableness of the independent variables. The regression does not pass the significance test if the test statistic t is less than the critical value, and vice versa. Those variables with coefficients that do not pass the test should be eliminated based on actual conditions, which is a widely used method in the choice of independent variables.

In this study, the linear regression equations were assessed based on R^2 and standardized coefficients. R^2 stands for the proportion of explicable part of sample data in regression equation. The larger the proportion is, the closer R^2 is to 1, which refers that more samples in the regression equation can be explicable, and the model will be more accurate. When the multiple regression method R^2 is between 0.8 and 1, it means that the goodness of fit of the model is relatively high. Given the large differences in complexity and accuracy between micro and macro data, appropriate changes in the evaluation criteria should be considered. If R^2 is between 0.5 and 0.8, the goodness of fit of the model is considered reasonable. The regression coefficient is the coefficient after eliminating the effects of the units of dependent variable y and independent variable x. The value of regression coefficient directly reflects the effect of x on y. Thus, the larger regression coefficient a is, the greater influence of x has on y. If regression coefficient is positive, y increases with increasing x; if it is negative, y decreases with increasing x.

The relationship between urban population density distribution and the FAR of (*non-*)public service facilities is analyzed by adopting an ordinary least square (OLS) model. The FAR P, the FAR of public service facilities P_a, and the FAR of non-public service facilities P_b were selected as the main dependent variables in our regression model. Daily population density ρ, the day-time (7 a.m.–6 p.m.) population density ρ_d, and the night-time (7 p.m.–6 a.m.) population density ρ_n were chosen as the main independent variables. The day-time population density ρ_d and night-time population density ρ_n represent the population densities at 4 a.m. and 2 p.m., respectively.

Table 3 defines all variables involved in the regression analysis, including the meanings these values represent and their measurement units.

As shown in Table 2, P represents the FAR (i.e., the ratio of total construction area to land use area) of each area. For public service and management lands, land for commercial and service facilities, and mixed-use land (commercial and residential mixed areas are not included here), P_a is equal to that of the FAR, while P_b is 0. For land for non-public service facilities (e.g., residential land), industrial land,

and land for warehouses, P_b is equal to that of the FAR P, while P_a is 0. For commercial and residential land, P_a and P_b, respectively, represent the ratio of commercial land to residential land in mixed-use land. The independent variable ρ stands for daily population density; ρ_d is the day-time population density from 10 a.m. to 11 a.m. and ρ_n is the night-time population density from 4 a.m. to 5 a.m.

Table 3. Definition of regression variables.

Dependent Variables	Corresponding Evaluation Factor	Measurement Units
P	FAR	Ratio of total construction area to land use area (%)
P_a	FAR of public service facilities	Ratio of construction area of public service facilities to land use area (%)
P_b	FAR of non-public service facilities	Ratio of construction area of non-public service facilities to land use area (%)
Independent Variables	**Corresponding Evaluation Factor**	**Measurement Units**
ρ	Daily population density	Per capita/per square meter
ρ_d	Day-time population density (10 a.m.–11 a.m.)	Per capita/per square meter
ρ_n	Night-time population density (4 a.m.–5 a.m.)	Per capita/per square meter

Data source: plotted by the author.

3. Characteristics and Results

3.1. Characteristics of the Temporal Evolution of Urban Population Density

The results of the behavior density calculation provided an intuitionistic and accurate cognition of the spatial and temporal law of citizen activities on a macro-level. For example, on an autumn day in Shanghai's central city, the variations in the density difference is larger within each block, with the highest density being 150,000 people per square kilometers and the lowest density being less than 2000 people per square kilometer. Moreover, the density of space in a single day differed significantly. The kernel density method was used to draw a contour distribution map of behavior density in each time period. This map shows that the behavior density showed a dark-colored concentration core in the central area during the day, which shifted to other areas outside the city center at night (Figure 2).

According to the behavior density calculation results, the temporal changes of mobile phone user density can be measured on different scales. Within the central area of Shanghai, the average numbers of mobile phone users on each working day and each holiday period are shown in Figure 3. The average number of users is 9.037 million on working days and 8.971 million on holidays. The maximum number of users on weekdays is 9.608 million, whereas the maximum number of users on holidays is 9.534 million. The minimum number of weekday users was 8.446 million, whereas the minimum number of holiday users was 8.411 million. The minimum users' numbers were recorded at 5 am. The fluctuations between the maximum and minimum values were mainly caused by mobile phone users entering or leaving the research area at different times.

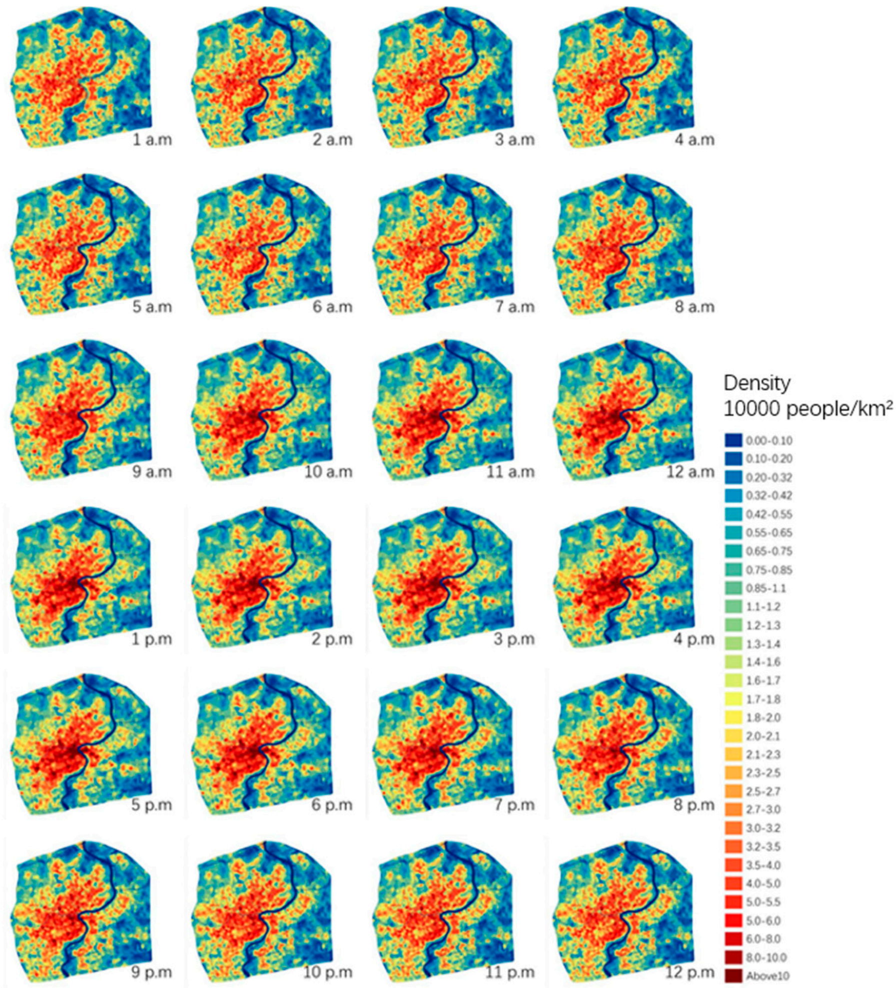

Figure 2. Mobile phone users in central city during weekdays and weekends. Data source: plotted by the author.

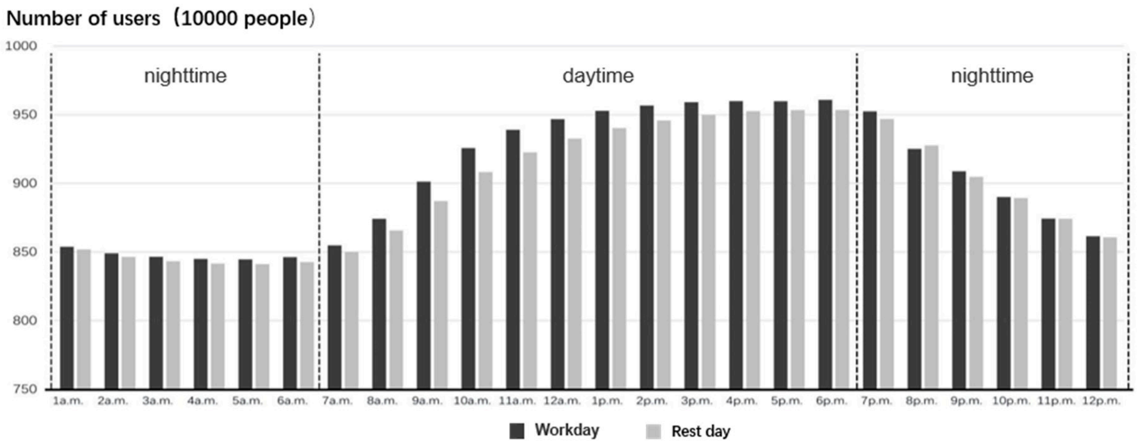

Figure 3. Average mobile phone users in central city during working days and weekends. Data source: plotted by the author.

3.2. Correlation between the FAR and Average Population Density by Day

The simple linear regression on dependent variables (FAR of public service facilities P_a and FAR of non-public service facilities P_b) and the independent variable (daily population density ρ) was performed to test the relationship between population density and the distribution of public service facilities. The linear regression model is shown in Table 4.

Table 4. Regression results for daily population density.

Independent Variable	Dependent Variable	R2	Standardized Coefficient	Sig.
daily population density (ρ)	FAR (P)	0.644	0.802	0.000
daily population density (ρ)	FAR of public service facilities (P_a)	0.653	0.808	0.000
Daily population density (ρ)	FAR of non-public service facilities (P_b)	0.637	0.798	0.000

Data source: plotted by the author.

According to the regression results shown in Table 4, the R^2 value of the linear regression equation between the FAR of public service facilities P_a and daily population density ρ is 0.653, and the regression coefficient is 0.808. However, the daily population density is also highly related to land FAR along with the FAR of non-public service facilities. The R^2 value of the linear regression equation between daily population density and land FAR is 0.644, and the regression coefficient is 0.802. The R^2 value of the linear regression equation between daily population density and the FAR of non-public service facilities is 0.637, and the regression coefficient is 0.808. Thus, daily population density is highly related to the FAR of both public and non-public service facilities; both FAR increase with increasing daily population density. In addition, the R^2 values and regression coefficients of these three linear regression equations are similar. This indicates that although daily population density is strongly related to FAR, there is no significant difference between its effects on public and non-public service facilities.

The above conclusions indicate that in the overall distribution of population density in days or even longer periods, although population density did have an effect on the FAR of that area (areas with high population density tend to have a high FAR), this effect also happen on non-public service facilities similarly. The difference of effect of high population density on the development of urban public service facilities compared with other kind of land use does not exist. As a result, the average population density observed over a long period of time (day-time periodicity or longer) does not directly affect the distribution of public service facilities despite its correlation with public service facilities distribution.

3.3. Correlation between the FAR of Public Service Facilities and Population Density from Day to Night

The above Figure 4 shows the distribution of population density at typical time of day (3 p.m.–4 p.m.), and Figure 5 shows the distribution of population density at typical time of night (2 a.m.–3 a.m.). The regression results between day-time population density ρ_d and FAR P along with those between night-time population density ρ_n and FAR P are shown in Table 5. The R^2 value between FAR P and day-time population density is 0.613, and the regression coefficient is 0.783. The R^2 value between P and night-time population density is 0.625, and the regression coefficient is 0.791. These results indicate that changes in daytime and night-time population density have a negligible effect on FAR. However, the regression results also reveal that the relativity effect of day-time population density on the FAR of public service facilities is greater than that of night-time population density. The R^2 value between day-time population density and the FAR of public service facilities is 0.706, with a regression coefficient of 0.840. However, that between night-time population density and the FAR of public service facilities is only 0.441, with a regression coefficient of 0.664. This indicates that although changes in daytime and night-time population density have a negligible impact on FAR, their relative effects on the FAR of public service facilities are significantly different.

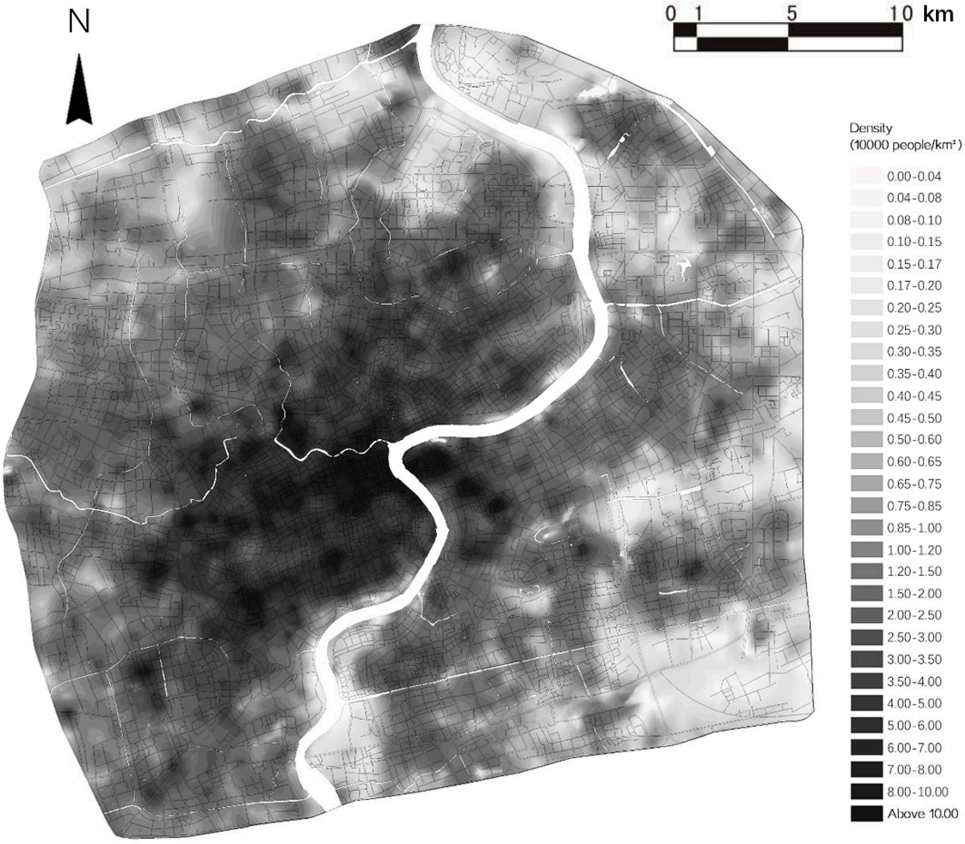

Figure 4. Distribution of daytime (3 p.m.–4 p.m.) population density. data source: plotted by the author.

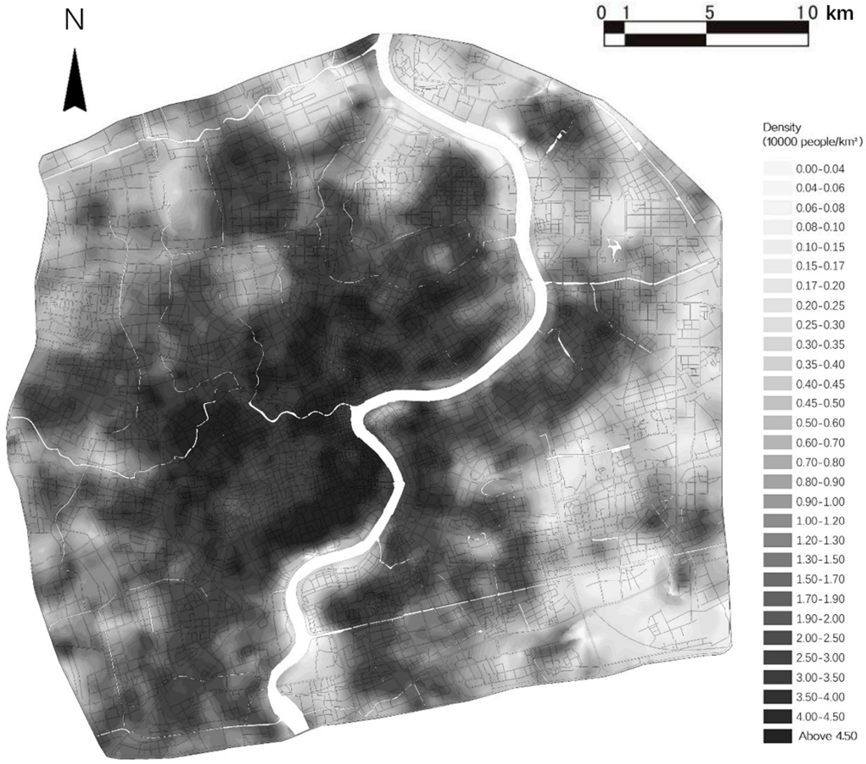

Figure 5. Distribution of night-time (2 a.m. to 3 a.m.) population density. data source: plotted by the author.

Table 5. Regression results for daytime and night-time population density.

Independent Variable	Dependent Variable	R2	Standardized Coefficients	Sig.
Day-time population density (ρ_d)	FAR (P)	0.613	0.783	0.000
Night-time population density (ρ_n)	FAR (P)	0.625	0.791	0.000
Day-time population density (ρ_d)	FAR of public service facilities (P_a)	0.706	0.840	0.000
Night-time population density (ρ_n)	FAR of public service facilities (P_a)	0.441	0.664	0.000

Data source: plotted by the author.

The regression results revealed a temporal heterogeneity between dynamic population density and public service facilities. Areas with high population density in the daytime tend to have more public service facilities. The higher the day-time population density, the more likely public service facilities are to aggregate. As a contrast, there was no obvious correlation with the distribution of public service facilities and the night-time population density.

3.4. Correlation between the FAR of Non-Public Service Facilities and Population Density from Day to Night

The above regression analysis established a significant correlation between day-time population density and public service facilities distribution. As another part of the investigation, linear regression analysis was applied to determine if relationships exist between night-time/day-time population density and the distribution of non-public service facilities.

The regression results between day-time population density ρ_d and FAR P along with those between night-time population density ρ_n and FAR P are shown in Table 6. Compared to day-time population density, there was a stronger correlation between night-time population density and non-public service facilities. The R^2 value between night-time population density and the FAR of public service facilities is 0.639, and the regression coefficient is 0.799. The R^2 value between day-time population density and the FAR of public service facilities is 0.585, with a regression coefficient of 0.765.

The results indicate that the relative impact of night-time population density on the FAR of non-public service facilities is greater than its impact on the FAR of public service facilities. The higher the night-time population density, the less likely public service facilities are to aggregate with a higher FAR of non-public service facilities such as residential land. This indicates that the areas with high night-time population density are more likely to accumulate non-public service facilities: the higher the night-time population density, the larger the scale of non-public service facilities such as residential land. And there was no obvious correlation with the distribution of non-public service facilities and the daytime population density.

Table 6. Regression results for daytime and night-time population density.

Independent Variable	Dependent Variable	R^2	Standardized Coefficients	Sig.
Day-time population density (ρ_d)	FAR (P)	0.613	0.783	0.000
Night-time population density (ρ_n)	FAR (P)	0.625	0.791	0.000
Day-time population density (ρ_d)	FAR of non-public service facilities (P_a)	0.585	0.765	0.000
Night-time population density (ρ_n)	FAR of non-public service facilities (P_a)	0.639	0.799	0.000

Data source: plotted by the author.

4. Discussion

Our results reveal a clear correlation between public service facilities and daytime behavior density, as well as a correlation between non-public service facilities and night-time behavior density, which denies daily or even longer correlation. This also explains why public service facilities have been found to be irrelevant to population density in previous studies [29], because the correlation between public service facilities and population density cannot be suitably analyzed with data of days or even longer time periods. In recent years, other scholars have also found the spatial and temporal distribution law of urban population density to be similar to this paper by using a Baidu heat map and mobile signaling data [22,49].

The results of this study reveal that the distribution of public service facilities is highly related to day-time population density. Our results show that the population density at a certain time—or more specifically, the density of specific behavior types—is the real influencing factor on the distribution of urban public service facilities. In western behavioral geography, the concept of "Action Space" has been proposed [50,51]. This includes not only people's behavior in daily life, but also their perception of their environment. Behavioral space is an analytical carrier used to reveal the "human–environment" relationship. Golledge proposed the concept of "activity space", which specifically refers to the observable movement and activities of individuals within their living space [52]. This means that personal activity space is the sum of all behavior systems in real space imprinted on space (including the place of origin, destination, mode of transportation, activity content, and time). Chapin believes that the daily activities of individuals living in a familiar spatial environment are composed of habitual behaviors such as going to work, going home, and going shopping [53]. According to Japanese geographer Yoshio Arai [54,55], living activity space refers to "the expansion of people's lives in space", and "the spatial range constituted by many activities of people in order to maintain daily life." The basic elements of living space include shopping space, leisure space, employment space, and other private space. The concept of living space also emphasizes that it is a kind of "phase moving space", that is, the relative activity space centering on oneself. Yoshio Arai [56] divides human behavior into four types from the perspective of behavior space: employment behavior, consumption behavior, recreation behavior, and habitation behavior. These behaviors can also be classified into public space behaviors and non-public space behaviors based on the places and nature of these behaviors.

The spatiotemporal distribution of urban population density is a process in which residents choose an activity space to meet their own needs. The interaction between the purpose of urban residents' activities and the spatial functional differences leads to the spatiotemporal evolution of urban population density. The time difference related to urban residents' activity purposes is one of the internal factors that cause the spatiotemporal evolution of urban population density.

Public space behaviors can be mainly classified into employment behavior, consumption behavior and recreation behavior. Employment behavior refers to those activities that people engage in to obtain remuneration or operating income. Employment behavior is often mandatory in terms of frequency and time of the activity; most employment behavior occurs on weekdays from 9 a.m. to 5 p.m. Consumer behavior refers to activities in which people obtain goods and services (e.g., window shopping, choosing, purchasing and using goods). Consumer behaviors in cities include obtaining staples like food along with other activities such as purchasing luxury goods and attending concerts. Purchasing staples has a certain time regularity, whereas the latter activities are related the proclivities of individual consumers. Recreation behavior refers to all kinds of activities that people do in their leisure time, usually during nonworking periods. Because most of these public space behaviors occur in the daytime and are closely related to specific urban public service facilities, it can be understood that the significant relationship between the daytime population density and the distribution of public service facilities is caused by the interaction between public space behaviors and public service facilities during the daytime.

Non-public space is where people conduct daily private activities. Behaviors in non-public space are mainly the activities of residents that occur in non-public space like residential areas. Residential behavior is how people use residential land and has strong universality: people often return to their house at the end of the day to enjoy rest and sleep, which is the reason for the highly association between night population density and distribution of non-public services.

5. Conclusions

The recent development of big data provides new opportunities for studies on population dynamics. This study explores the spatiotemporal characteristics and mechanism of urban population density by using mobile phone data and public service facility land use data.

Studying the evolution process of urban population density with time provides an initial insight into city dynamics, revealing the differences in population density in different urban districts as well as the differences between working days and holidays. According to our behavior density calculation results, the temporal changes of mobile phone user density can be measured on different scales.

Subsequently, using OLS models, we explored the relationship between population density in daytime and night-time and (non-)public service facilities. The results show that the average population density observed during a long period of time (day-time periodicity or longer) is not directly related to the distribution of public service facilities, despite its correlation with public service facilities distribution. The results also reveal a temporal heterogeneity between dynamic population density and public service facilities. Areas with high population density in the daytime tend to have more public service facilities. The higher the day-time population density, the more likely public service facilities are to aggregate. In contrast, there was no obvious correlation between the distribution of public service facilities and night-time population density. Moreover, compared with day-time population density, there was a stronger correlation between night-time population density and non-public service facilities. This indicates that the higher the night-time population density, the larger the scale of non-public service facilities, such as residential land. The relationship between the overall distribution of population density and public service facilities essentially answers the important question of how population distribution is related to the layout of public facilities.

Finally, we proposed a mechanism of spatiotemporal correlation between urban population density and public service facilities. Residents' daily activities during different times in a day lead to a spatiotemporal dynamic correlation between public service facilities and population density. In future, we plan to apply the results of this research to traffic planning and facilities allocation, while considering other spatial analysis models.

Public facilities planning is normally executed in relation to a planning standard, such as how many hectares of open space are required for a certain number of people in a district. However, planning standards only specify the area required and seldom specify where the public facility should be located. A location–allocation model aimed at finding the best sites for facilities would be a more useful tool for public facilities planning [57]. This study reveals the strong correlation between population density and public service facilities during daytime, which can guide urban decision-makers to make better decisions on public facilities' location using big data. We will continue to explore the potential of spatiotemporal data, such as using mobile phone signaling data to study the operation and development of urban systems.

Author Contributions: Conceptualization, Yi Shi; Data curation, Yi Shi; Formal analysis, Yi Shi; Funding acquisition, Yi Shi; Investigation, Yi Shi; Methodology, Junyan Yang; Project administration, Junyan Yang; Resources, Yi Shi; Software, Yi Shi; Supervision, Yi Shi; Validation, Yi Shi; Visualization, Yi Shi Writing—original draft, Yi Shi; Writing—review and editing, Junyan Yang and Peiyu Shen. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by the Youth Science Fund Project of National Natural Science Foundation of China, grant number 51708103.

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

References

1. De Verteuil, G. Reconsidering the legacy of urban public facility location theory in human geography. *Prog. Hum. Geogr.* **2000**, *24*, 47–69. [CrossRef]
2. Teitz, M.B. Toward a theory of urban public facility location. *Pap. Reg. Sci. Assoc.* **1968**, *21*, 35–51. [CrossRef]

3. Brueckner, J.K.; Thisse, J.F.; Zenou, Y. Why is central Paris rich and downtown Detroit poor? *Eur. Econ. Rev.* **1999**, *43*, 91–107. [CrossRef]
4. Li, H.; Wei, Y.H.D.; Yu, Z.; Tian, G. Amenity, accessibility and housing values in metropolitan USA: A study of Salt Lake County, Utah. *Cities* **2016**, *59*, 113–125. [CrossRef]
5. Yuan, F.; Wei, Y.D.; Wu, J.W. Amenity effects of urban facilities on housing prices in China: Accessibility, scarcity, and urban spaces. *Cities* **2020**, *96*, 102433. [CrossRef]
6. Setyono, D.A.; Cahyono, D.D.; Helmy, M. Measuring Service Capacity of Public Facilities Based on Supply Aspect (Case Study: Elementary School in Malang City). *Procedia Soc. Behav. Sci.* **2016**, *227*, 45–51. [CrossRef]
7. Tahmasbi, B.; Mansourianfar, M.H.; Haghshenas, H.; Kim, I. Multimodal accessibility-based equity assessment of urban public facilities distribution. *Sustain. Cities Soc.* **2019**, *49*, 101633. [CrossRef]
8. Delbosc, A.; Currie, G. Using Lorenz curves to assess public transport equity. *J. Transp. Geogr.* **2011**, *19*, 1252–1259. [CrossRef]
9. Welch, T.F.; Mishra, S. A measure of equity for public transit connectivity. *J. Transp. Geogr.* **2013**, *33*, 29–41. [CrossRef]
10. Manaugh, K.; Badami, M.G.; El-Geneidy, A.M. Integrating social equity into urban transportation planning: A critical evaluation of equity objectives and measures in transportation plans in North America. *Transp. Policy* **2015**, *37*, 167–176. [CrossRef]
11. Goddard, M.; Smith, P. Equity of access to health care services: Theory and evidence from the UK. *Soc. Sci. Med.* **2001**, *53*, 1149–1162. [CrossRef]
12. Marsh, M.T.; Schilling, D.A. Equity measurement in facility location analysis: A review and framework. *Eur. J. Oper. Res.* **1994**, *74*, 1–17. [CrossRef]
13. McLay, L.A.; Mayorga, M.E. A dispatching model for server-to-customer systems that balances efficiency and equity. *Manuf. Serv. Oper. Manag.* **2013**, *15*, 205–220. [CrossRef]
14. Chang, Z.; Chen, J.Y.; Li, W.F.; Li, X. Public transportation and the spatial inequality of urban park accessibility: New evidence from Hong Kong. *Transp. Res. Part D Transp. Environ.* **2019**, *76*, 111–122. [CrossRef]
15. Dadashpoor, H.; Rostami, F.; Alizadeh, B. Is inequality in the distribution of urban facilities inequitable? Exploring a method for identifying spatial inequity in an Iranian city. *Cities* **2016**, *52*, 159–172. [CrossRef]
16. Chen, A.; Bouferguene, Y.; Shen, M. Assessing accessibility-based service effectiveness (ABSEV) and social equity for urban bus transit: A sustainability perspective. *Sustain. Cities Soc.* **2019**, *44*, 499–510. [CrossRef]
17. Taleai, M.; Sliuzas, R.; Flacke, J. An integrated framework to evaluate the equity of urban public facilities using spatial multi-criteria analysis. *Cities* **2014**, *40*, 56–69. [CrossRef]
18. Jia, P.; Qiu, Y.; Gaughan, A.E. A fine-scale spatial population distribution on the High-resolution Gridded Population Surface and application in Alachua County, Florida. *Appl. Geogr.* **2014**, *50*, 99–107. [CrossRef]
19. Deng, C.; Wu, C.; Wang, L. Improving the housing-unit method for small-area population estimation using remote-sensing and GIS information. *Int. J. Remote Sens.* **2010**, *31*, 5673–5688. [CrossRef]
20. Murray, A.T.; Davis, R.; Stimson, R.J.; Ferreira, L. Public transportation access. *Transp. Res. Part D Transp. Environ.* **1998**, *3*, 319–328. [CrossRef]
21. Pattnaik, S.B.; Mohan, S.; Tom, V.M. Urban bus transit route network design using genetic algorithm. *J. Transp. Eng.* **1998**, *124*, 368–375. [CrossRef]
22. Zhong, W.; Wang, D.; Xie, D.; Yan, L. Dynamic characteristics of Shanghai's population distribution using cell phone signaling data. *Geogr. Res.* **2017**, *36*, 972–984.
23. Zhang, W. An analysis on the coordination of population distribution and service infrastructure distribution in Beijing. *Soc. Sci. Beijing* **2004**, *1*, 78–84.
24. Yu, L.; Zhang, S.; Han, G. Spatial Feature Analysis of Alternation of Population Distribution in Shanghai. *China Popul. Resour. Environ.* **2006**, *16*, 83–87.
25. Li, Z. Development Research of Tianjin Urban Commercial System and Planning Arrangement Structure. Ph.D. Thesis, Tianjin University, Tianjin, China, 2006. Available online: http://www.cnki.net/ (accessed on 9 January 2006).
26. Shen, Q. Spatial Coordination of Population Distribution and Service Infrastructure Distribution in Shanghai. *J. Gansu Sci.* **2014**, *26*, 139–142.
27. Kwon, O. Scaling laws between population and a public transportation system of urban buses. *Phys. A* **2018**, *503*, 209–214. [CrossRef]

28. Reigadinha, T.; Godinho, P.; Dias, J. Portuguese food retailers—Exploring three classic theories of retail location. *J. Retail. Consum. Serv.* **2017**, *34*, 102–116. [CrossRef]
29. Zhao, Z. The Study of Commercial Pattern Development in Changchun since 19th Century. 2014. Available online: http://www.cnki.net/ (accessed on 14 November 2019).
30. Tu, W.; Zhu, T.T.; Xia, J.Z.; Zhou, Y.L.; Lai, Y.N.; Jiang, J.C.; Li, Q.Q. Portraying the spatial dynamics of urban vibrancy using multisource urban big data. *Comput. Environ. Urban Syst.* **2019**, 101428. [CrossRef]
31. He, Q.S.; He, W.S.; Song, Y.; Wu, J.Y.; Yin, C.H.; Mou, Y.C. The impact of urban growth patterns on urban vitality in newly built-up areas based on an association rules analysis using geographical 'big data'. *Land Use Policy* **2018**, *78*, 726–738. [CrossRef]
32. Rathore, M.M.; Ahmad, A.; Paul, A.; Rho, S.M. Urban planning and building smart cities based on the Internet of Things using Big Data analytics. *Comput. Netw.* **2016**, *101*, 63–80. [CrossRef]
33. Niu, X.Y.; Ding, L.; Song, X.D. Understanding Spatial Structure of Shanghai Central City Based on Mobile Phone data. *Urban Plan. Forum* **2014**, *6*, 61–67.
34. Lee, K.S.; You, S.Y.; Eom, J.K.; Song, J.Y.; Min, J.H. Urban spatiotemporal analysis using mobile phone data: Case study of medium- and large-sized Korean cities. *Habitat Int.* **2018**, *73*, 6–15. [CrossRef]
35. Liu, Y.L.; Fang, F.G.; Jing, Y. How urban land use influences commuting flows in Wuhan, Central China: A mobile phone signaling data perspective. *Sustain. Cities Soc.* **2020**, *53*, 101914. [CrossRef]
36. Guo, S.H.; Song, C.; Pei, T.; Liu, Y.X.; Ma, T.; Du, Y.Y.; Chen, J.; Fan, Z.D.; Tang, X.L.; Peng, Y.; et al. Accessibility to urban parks for elderly residents: Perspectives from mobile phone data. *Landsc. Urban Plan.* **2019**, *191*, 103642. [CrossRef]
37. Anda, C.; Medina, S.A.O.; Fourie, P. Multi-agent urban transport simulations using OD matrices from mobile phone data. *Procedia Comput. Sci.* **2018**, *130*, 803–809. [CrossRef]
38. Calabrese, F.; Ratti, C. Real time Rome. *Netw. Commun. Stud.* **2006**, *20*, 247–258.
39. Pulselli, R.; Ramono, P.; Ratti, C.; Tiezzi, E. Computing urban mobile landscapes through monitoring population density based on cellphone chatting. *Int. J. Des. Nat. Ecodynamics* **2008**, *3*, 121–134.
40. Reades, J.; Calabrese, F.; Sevtsuk, A.; Ratti, C. Cellular census: Explorations in urban data collection. *IEEE Pervasive Comput.* **2007**, *6*, 30–38. [CrossRef]
41. Girardin, F.; Calabrese, F.; Dal, F.F.; Ratti, C.; Blat, J. Digital footprinting: Uncovering tourists with user-generated content. *IEEE Pervasive Comput.* **2008**, *7*, 36–43. [CrossRef]
42. Girardin, F.; Vaccari, A.; Gerber, A.; Biderman, A.; Ratti, C. Towards Estimating the Presence of Visitors from the Aggregate Mobile Phone Network Activity They Generate. In Proceedings of the International Conference on Computers in Urban Planning and Urban Management, Hong Kong, China, 16–18 June 2009.
43. Isaacman, S.; Becker, R.; Cáceres, R.; Kobourov, S.; Rowland, J.; Varshavsky, A. A Tale of Two Cities. In Proceedings of the Eleventh Workshop on Mobile Computing Systems Applications, Annapolis, MD, USA, 22–23 February 2010; ACM: New York, NY, USA, 2010; pp. 19–24.
44. Calabrese, F.; Di, L.G.; Liu, L.; Ratti, C. Estimating Origin-Destination flows using opportunistically collected mobile phone location data from one million users in Boston Metropolitan Area. *IEEE Pervasive Comput.* **2011**, *10*, 36–44. [CrossRef]
45. Deville, P.; Linard, C.; Martin, S.; Gilbert, M.; Stevens, F.R.; Gaughan, A.E.; Blondel, V.D.; Tatem, A.J. Dynamic population mapping using mobile phone data. *Proc. Natl. Acad. Sci. USA* **2014**, *111*, 15888–15893. [CrossRef] [PubMed]
46. Janzen, M.; Vanhoof, M.; Smoreda, Z.; Axhausen, K.W. Closer to the total? Long-distance travel of French mobile phone users. *Travel Behav. Soc.* **2018**, *11*, 31–42. [CrossRef]
47. Murphy, R.E.; Vance, J.E. Delimiting the CBD. *Econ. Geogr.* **1954**, *30*, 34. [CrossRef]
48. Yang, J.; Shi, B. Research on the Quantitative Definition Method of Urban Center Boundaries. *J. Hum. Settl. West China* **2014**, *29*, 17–21.
49. Li, J.G.; Li, J.W.; Yuan, Y.G.; Li, G.F. Spatiotemporal distribution characteristics and mechanism analysis of urban population density: A case of Xi'an, Shaanxi, China. *Cities* **2019**, *86*, 62–70. [CrossRef]
50. Kirk, W. Problems of geography. *Geography* **1963**, *48*, 357–371.
51. Horton, F.E.; Reynolds, D.R. Effects of urban spatial structure on individual behavior. *Econ. Geogr.* **1971**, *47*, 36–48. [CrossRef]
52. Golledge, R.G.; Stimson, R.J. *Spatial Behavior: A Geographic Perspective*; The Guilford Press: New York, NY, USA, 1997.

53. Chapin, F.S. *Human Activity Patterns in the City: Things People Do in Time and in Space*; John Wiley Sons: New York, NY, USA, 1974.
54. Arai, Y. Topological space of life and daily activities. *Reg. Dev.* **1985**, *10*, 45–56.
55. Arai, Y. Basic structure and problem of living activity space in city. *J. Econ. Shinshu Univ.* **1992**, *29*, 27–67.
56. Arai, Y.; Okamoto, K.; Kamiya, H. *Taro Kawaguchi: Urban Space and Time-Time Geography of Living Activities*; Kokon Shoin: Tokyo, Japan, 1996.
57. Yeh, A.G.O.; Chow, M.H. An integrated GIS and location-allocation approach to public facilities planning—An example of open space planning. *Comput. Environ. Urban Syst.* **1996**, *20*, 339–350. [CrossRef]