

Article

# Operating Characteristics of Dockless Bike-Sharing Systems near Metro Stations: Case Study in Nanjing City, China

Yuan Li, Zhenjun Zhu and Xiucheng Guo \*

School of Transportation, Southeast University, No.2 Dongnandaxue Road, Nanjing 211189, China; 220162614@seu.edu.cn (Y.L.); zhuzhenjun@seu.edu.cn (Z.Z.)

\* Correspondence: 101002320@seu.edu.cn

Received: 26 January 2019; Accepted: 14 April 2019; Published: 15 April 2019



**Abstract:** With the growth of dockless bike-sharing (DLBS) systems, the first-and-last mile connection to public transport, such as metro and light railway stations, could be improved. DLBS systems complete the trip chain by connecting metro stations with points of interest and enhance the sustainability of urban transportation. Therefore, it is necessary to understand the trans-shipment characteristics of DLBS systems for metro stations. In this study, we collected data from the Mobike DLBS system in Nanjing City, China and applied K-means clustering to analyse the activity patterns of DLBS systems near local metro stations. Metro stations were categorised into five types on workdays and three types on weekends. An analysis of the relationships between activity patterns and spatial distribution characteristics demonstrated that the distribution of clusters possesses a strong connection with the surrounding environment. Low land development rates and a sparse distribution of metro stations cause a large range of influences. This research has direct implications for understanding the operating state of DLBS systems near metro stations and promoting the proper management of DLBS systems.

**Keywords:** dockless bike-sharing system; metro station; activity pattern; cluster analysis; range of influences; data mining

## 1. Introduction

With rapid economic development and high-level urbanisation, large-scale cities in China experience traffic congestion [1,2]. Twenty-six percent of China's urban commuting peaks were in a state of congestion in 2017 [3]. Transit-oriented development (TOD) is a method for ensuring the sustainability of transportation and urbanisation. As a rapid, efficient and large-capacity method of transportation, the metro system is a priority in TOD strategies [4]. However, the availability of metro stations in high-traffic areas is often deficient because the infrastructure in these areas does not allow for the construction of metro stations. Patrons of the metro system usually access the stations by other modes, such as walking, cycling and taking the bus [5]; this transit process is described as the first-and-last mile problem [6]. Improvements in accessibility and the enhanced integration of other feeding modes and metro stations would definitely boost the ridership of metro systems [7].

Dockless bike-sharing (DLBS) systems have been developed rapidly all over the world within the last few years. These systems offer an environmentally-friendly and sustainable solution to the first-and-last mile connection, which completes the trip chain for existing public transportation modes, such as metros and bus systems [8]. DLBS systems have replaced traditional bike sharing systems to a large extent in most metropolitan areas in China owing to their flexibility and convenience. With traditional bike-sharing systems, commuters rent bikes from their closest station and return

them to a station near their destination [9]. It is likely that traditional bike-sharing stations are not located near commuters' origins and destinations. Therefore, the traditional bike-sharing systems do not significantly improve the first-and-last mile problem owing to the system's structure. Conversely, commuters can theoretically rent and return bikes anywhere in DLBS systems. Despite the popularity of DLBS systems in metropolitan areas in China, there is still little research on their operating characteristics. Therefore, for city managers, there is a lack of theoretical guidance for the operation and management of DLBS systems, especially when using DLBS systems as a transfer mode to metro stations. This study sought to analyse the operating characteristics of DLBS systems around metro stations and their influence ranges to provide theoretical guidance for the operation and management of DLBS systems near metro stations.

The remainder of this paper is organised as follows. Section 2 contains a literature review. Section 3 describes the data collection and pre-processing approaches employed. Section 4 details the feature extraction from the initial data, the application of K-means clustering to analyse activity patterns and the examination of the range of influence of DLBS systems for clusters of metro stations. Conclusions are presented in Section 5.

## 2. Literature Review

There is abundant research about the traditional bike-sharing system since it arose in the late 1990s. For bike-sharing systems in different cities, Pfrommer et al. [10] determined that weekday usage peaks from 7 am to 9 am and from 4 pm to 6 pm, while weekend usage is highest in the middle of the day. Ahmed et al. [11] determined that bike-sharing systems are busier during warmer months, which generally confirms the relationship between weather and the propensity for private bike riding. A study on the duration of bike-sharing trips based on data from Melbourne, Brisbane, Washington D.C., Minnesota and London determined that durations are within a tight band between 16 and 22 min [12]. Another study determined that casual users of a specific bike-sharing service typically take longer trips than annual members [13]. Tao et al. [14] analysed the global and spatio-temporal operating patterns of the traditional Public Bicycle Sharing system in Nanning City, China and studied the impact of urban morphology on these patterns. Froehlich et al. [15] provided a spatio-temporal analysis of thirteen weeks of bicycle station usage from Barcelona's shared bicycling system, applying clustering techniques to identify shared behaviour across stations and comparing experimental results from four predictive models of nearby station usage. Some researches focused on the sustainability of bike-sharing systems.

In terms of user preferences, multiple studies concluded that convenience is the major perceived benefit identified by bike-sharing users [16]; other investigations have demonstrated the importance of proximity of docking stations to users' homes [17]. These results point to the advantage of DLBS systems over traditional bike-sharing systems. Moreover, some studies determined that a significant proportion of users do not use bike-sharing systems frequently [13,18]. For trip purposes, some research demonstrated that the primary usage for long-term members was work-related, while the primary purpose for short-term members was leisure or sightseeing [13]. Bike-sharing systems are perceived as completing public transit. Some studies demonstrate that the majority of traditional bike-sharing trips are replacing trips formerly accomplished by public transport and walking [18,19]. Shaheen et al. found that bike-sharing competes with public transport in areas with more robust or congested transit networks. However, in areas with smaller public transit systems, bike-sharing serves a greater role as a first-and-last mile connector [20]. Fishman et al. [21] calculated bike-sharing's overall impact on kilometres travelled by vehicles and concluded bike-sharing reduces car use.

Owing to the tidal operating characteristics of bike-sharing systems, the number of vehicles in some bike-sharing stations does not match the demand during morning and evening peak hours. In such cases, vehicle rebalancing is required. Faghih-Imani et al. [22] have examined the factors associated with higher and lower levels of docking station activity and determined that weather and the presence of restaurants have a predictable impact. Parkes et al. [23] suggested that altering the

price to achieve rebalancing objectives could be employed as an option to resolve fleet distribution issues. Pfrommer et al. [10] used historical data of a bike-sharing system in London to model both the effectiveness of using lorries for rebalancing and the impact of introducing price incentives to mitigate fleet imbalance.

There are few studies on DLBS systems. Bao et al. [24] proposed a data-driven approach to develop bike lane plans. Dakshak Keerthi Chandra et al. [9] proposed a multidimensional tensor model to address the mismatching problem for supply and demand of DLBS systems. Wu et al. [25] discussed investigating the roles of this new bike-sharing system in urban mobility in China, especially in Shanghai, along with its influences in the society. Du et al. established a multinomial logit model to explore the influential factors associated with three patterns, Origin to Destination Pattern, Travel Cycle Pattern and Transfer Pattern, based on a survey of 4939 valid questionnaires in Nanjing, China [26]. Shi et al. employed the social network analysis method to recognise the critical factors and links in DBSPs' (dockless bike-sharing programs) sustainability [27].

While studies regarding the use of DLBS near metro stations are imperative, studies exploring the use of bicycles as a transfer mode to metro station areas in cities remain to be enriched. Karel Martens discussed the use of bike-and-ride in three countries, the Netherlands, Germany and the UK [28]. The research concluded that the majority of bike-and-ride users travel between 2 and 5 km to a public transport stop, with longer access distances reported for faster modes of public transport. Moreover, work and education are the primary travel motives, while car availability is not a strong influential factor. Zhao et al. determined that travel distance is the most important influence on rates of cycling for transfer trips between metro stations and home or the workplace [29]. Additionally, the presence of bicycle-sharing programs, mixed land use and green parks in metro station areas were associated with greater rates of cycling transfer. Lin et al. analysed the mode choices of passengers for connecting travels between trip origins/destinations and metro stations, and determined that collecting local empirical knowledge on travel behaviour is critical for developing bike-friendly environments for a city [30]. Ma et al. analysed the general characteristics of metro-bike-sharing transfer trips based on smart card data [31]. Cheng et al. conducted a cost-benefit analysis of public bicycle sharing system incorporation into the metro system to determine its cost-effectiveness [32]. Zhang et al. mapped the bicycle traffic on an equal population cartogram of Shanghai to distinguish overall patterns within the centre of Shanghai and determined that the usage frequency of bike-sharing systems from metro stations to outlying areas is gradually declining [33].

Generally, most existing researches on bike-sharing systems focused on the traditional models, which are restrained by the locations of dock station. However, the operating characteristics and management of DLBS systems are significantly different. Understanding the characteristics of DLBS systems is critical to optimizing their operation and management, so as to allocate corresponding parking facilities and vehicle rebalancing. Regarding the use of bicycles as a transfer mode to metro stations, existing literature [28–33] primarily focused on the influential factors of bicycles as a transfer mode to metro stations, such as travel distances, travel motivations and car availability. These studies provide great insight into the promotion of the use of bicycles as a transfer mode. However, there are still some research gaps on theoretical exploration for the daily operation and management of DLBS systems. There remains a necessity to understand the operating characteristics of DLBS systems near metro stations, which could provide guidance for their operation and management.

To fill in these knowledge gaps, this study was aimed at analysing the operating characteristics of a DLBS system near metro stations using data from Mobike based on the example of Nanjing, which is the capital of Jiangsu province. 'Operating characteristics' refers to the temporal usage of bikes near metro stations and their relationship with points of interest (POI). We use metro stations and POI as anchor points to collect the real-time Mobike data through its open mobile application-programming interface (API). Moreover, K-means clustering and spatio-temporal analysis was carried out to cluster the activity patterns and determine the range of influence of DLBS systems near metro stations. We consider that this study will contribute to the understanding of the operating characteristics of DLBS systems

and their effect on first-and-last mile connections for metro stations and promote the operation and management of DLBS systems in similar cities.

### 3. Data Collection and Pre-Processing

#### 3.1. Data Collection

For this study, Mobike location data, metro stations data and Nanjing POI data were analysed.

##### 3.1.1. Metro Stations and POI Data

The Nanjing metro system was established in 2005 and has developed 10 metro lines and 174 stations. Metro station data were obtained through a web API of Amap, which is a major electronic mapping service provider in China. The method to obtain data consists in sending a data request to Amap's server using the specified form of URL (Uniform Resource Locator), to which the server returns the corresponding data. The data contain information including name, ID, address, latitude and longitude, as shown in Table 1.

**Table 1.** Structure of metro data.

Field	Content
name	Xin Jie Kou
ID	BV10057753
address	Line 1; Line 2
latitude	32.041806
longitude	118.784136

The POI is a specific point location that someone could find useful or interesting, such as a residence, fuel station, public service point, etc. These points are usually considered origin or destination locations for trips in cities. These data were also obtained from Amap through its web API. The data include information including name, type, address, latitude and longitude, as shown in Table 2. A total of 225,555 POI data items across Nanjing City were collected. Considering there is a coverage for Mobike data collected later from every available point, this study selected the POI which are not in the same coverage. A total of 21,621 POI remained.

**Table 2.** Structure of points of interest (POI) data.

Field	Content
name	Fang Ting Pan Yuan
types	residence
address	No.1 Fang Ting road
latitude	32.219022
longitude	118.728905

##### 3.1.2. Mobike Location Data

Mobike location data comprised the dataset used to analyse the activity patterns and range of influence. There is an open mobile API for Mobike called by a WeChat applet, which requires latitude and longitude parameters, and returns a dataset of nearby Mobike locations in JavaScript Object Notation (JSON) format. The method to obtain data consists in posting a data request to Mobike's server using the specified form of data set, to which the server returns the corresponding data. A script written with Python 3.6 was used to collect Mobike location data near metro stations and POI. Considering that the DLBS system around the metro stations is used relatively frequently during the morning and evening commuting peaks, the number of bikes near stations also changes relatively rapidly. Therefore, in order to obtain more details and to better understand the operating

characteristics of DLBS systems during the morning and evening peak hours, we adopted different collection time intervals. The collection time interval was 10 min during the timeframes of 6–9 am and 5–10 pm; it was 30 min otherwise. The collection time interval for POI locations was 30 minutes at all times.

The data included collection time, acquired latitude, acquired longitude, bike ID, bike latitude, bike longitude and the distance between the acquired location and bike (in metres). An example of this data is shown in Table 3. Real-time data were collected from 12 June 2018 to 26 June 2018. A total of 12,640,899 observations were collected. The size of the total original data exceeded 3 GB.

Table 3. Structure of Mobike data.

Time	Acquired Latitude	Acquired Longitude	ID of Bike	Latitude of Bike	Longitude of Bike	Distance between Acquired Location and Bike
2018-06-12 00:00:00	32.049556	118.894656	8640215020	32.04994343	118.8948145	45
			0256507954	32.04988437	118.8949574	46
			0256539136	32.0494920	118.893895	72

### 3.2. Data Cleansing

It was necessary to cleanse the data because some of the data collected through the API was invalid or incomplete. This was due to multiple factors, such as Mobikes being prohibited in some metro stations, lack of Mobike availability in some remote stations, system failure, bad weather and other unexpected problems. There was a final total of 8,532,827 observations from 146 metro stations, and 17,365 POI remained for further analysis. The spatial distribution of the remaining metro stations is shown in Figure 1.

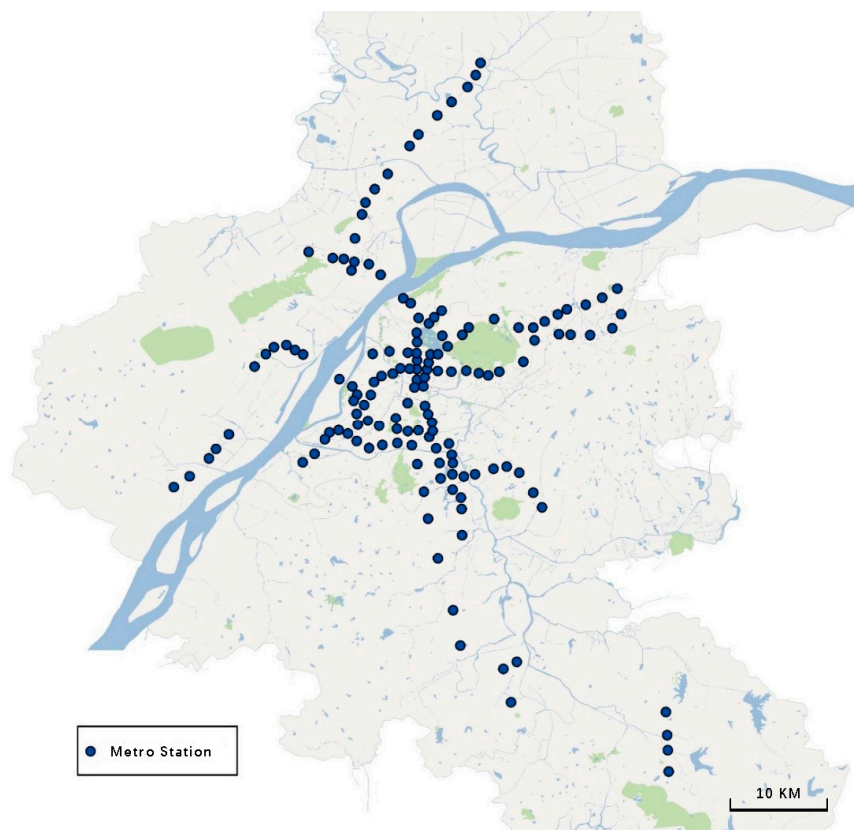


Figure 1. Spatial distribution of Nanjing metro stations.



### 3.3. Data Pre-Processing

As mentioned earlier, Mobike data were collected by time interval and saved as CSV files. Before analysing the data, they were integrated and reformatted with a short script with the NumPy and pandas packages for Python 3.6. Two types of data were integrated and reformatted: metro stations and POI locations. The structure of these data is shown in Table 4.

**Table 4.** Structure of integrated data collected in metro stations and POI.

Time	ID of Metro Station/POI	Latitude of Metro Station/POI	Longitude of Metro Station/POI	ID of Bike	Latitude of Bike	Longitude of Bike	Distance between Metro Station/POI and Bike
2018-06-12 00:00:00	1	32.049556	118.894656	256547535	32.04966271	118.89438071	28

## 4. Data Analysis

This study analysed the spatial and temporal distribution of Mobike locations near metro stations and POI. K-means clustering was applied to categorise all metro stations according to the temporal and spatial variation of the number of Mobikes nearby in order to illustrate the activity patterns of DLBS systems near metro stations in terms of workday and weekend data. In the last portion, this study searched Mobike data at every POI for any bike that appeared near any metro station within an hour to illustrate the coverage of DLBS systems for every type of metro station.

### 4.1. Feature Extraction

After data pre-processing, an initial data format for feature extraction was developed. The data extracted were from Mobikes whose user destination or origin was a metro station. Owing to the flexibility of DLBS systems, people usually return bikes near metro stations when they use a metro system; we tested distance ranges for these scenarios. A threshold of 100 metres was set such that any bike located within this distance of a metro station was considered to be related to that station. The number of related Mobikes was counted for every metro station for every time period along with the average value for both workday and weekend. The feature extracted data structure is shown in Table 5. For example, the value “1.428571429” in the cell with ID of Metro Station = 1 means that, during the data collection period, there was an average of 1.428571429 bikes within 100 meters of the metro station No. 1 at 00:00 on workdays/weekends.

**Table 5.** Structure of feature extracted data.

ID of Metro Station	00:00	00:30	...	23:00	23:30
1	1.428571429	1.142857143	...	2.000000000	1.857142857
2	10.57142857	9.000000000	...	11.28571429	9.714285714
...	...	...	...	...	...
145	10.71428571	12.00000000	...	9.714285714	9.428571429
146	3.285714286	3.285714286	...	5.714285714	4.857142857

### 4.2. Cluster Analysis

After the feature extraction was developed, the K-means clustering was applied to analyse the specific activity patterns of the DLBS systems using the sklearn package for Python 3.6.

K-means clustering is a method of vector quantisation which is popular for cluster analysis in data mining. It aims to partition  $n$  observations into  $k$  clusters, represented by their centres or means. The centre of each cluster is calculated as the mean of all the instances belonging to that cluster [34]. The K-means clustering algorithm is extremely efficient and concise for the classification of equivalent multidimensional data, which is consistent with our data type.

The algorithm begins with an initial set of cluster centres, chosen at random. The number of cluster centres is defined by  $k$ , which is provided in advance. In each iteration, each instance is assigned to its nearest cluster centre according to the Euclidean distance between the two. Then the cluster centres are re-calculated to reduce the partitioning error (defined by Equation (1)). The iteration would terminate when the partitioning error is no longer reduced by the relocation of the centres.

$$E_p = \sum_{i=1}^k \sum_{p \in C_i} |p - m_i|^2 \quad (1)$$

where  $E_p$  is the partitioning error,  $k$  is the number of clusters,  $p$  is the instance,  $C_i$  is the number  $i$  clusters and  $m_i$  is the cluster centre of number  $i$  clusters.

The determination of the  $k$  value is critical to the classification effect when using k-means clustering. Generally, the lower the partitioning error, the better the classification performance is. However, the partitioning error monotonically decreases as  $k$  increases. When the  $k$  value is extremely large, the classification becomes meaningless. We used the Elbow Method to define the  $k$  value to ensure that the value of  $k$  is balanced in both the classification performance and meaning. As the  $k$  value continues to increase, the improvement in the partitioning error continues to decrease. There is a relatively clear demarcation point. When the  $k$  value exceeds this point, the improvement in the partitioning error sharply decreases. Five is the inflection point for data of workdays and three is that of weekends, as shown in Figure 2. Therefore, we define  $k$  as five for workdays and three for weekends.

Five types of metro stations were clustered for workdays and three for weekends as shown in Figure 3.

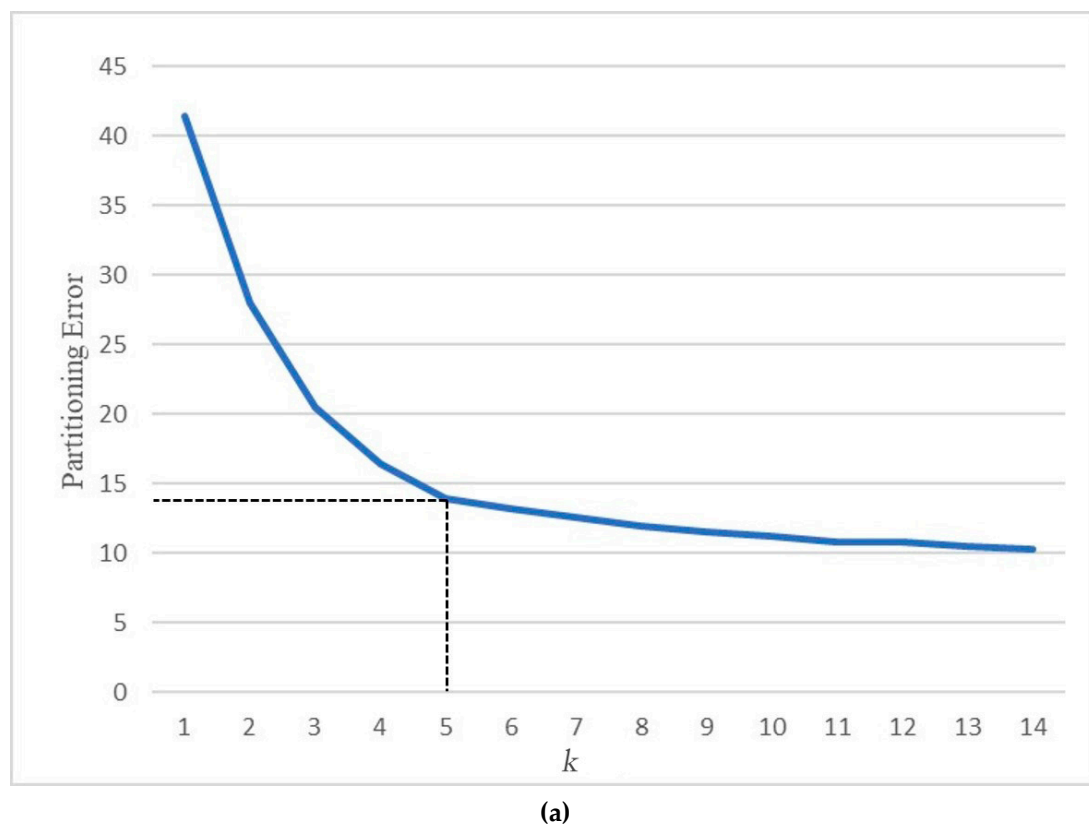
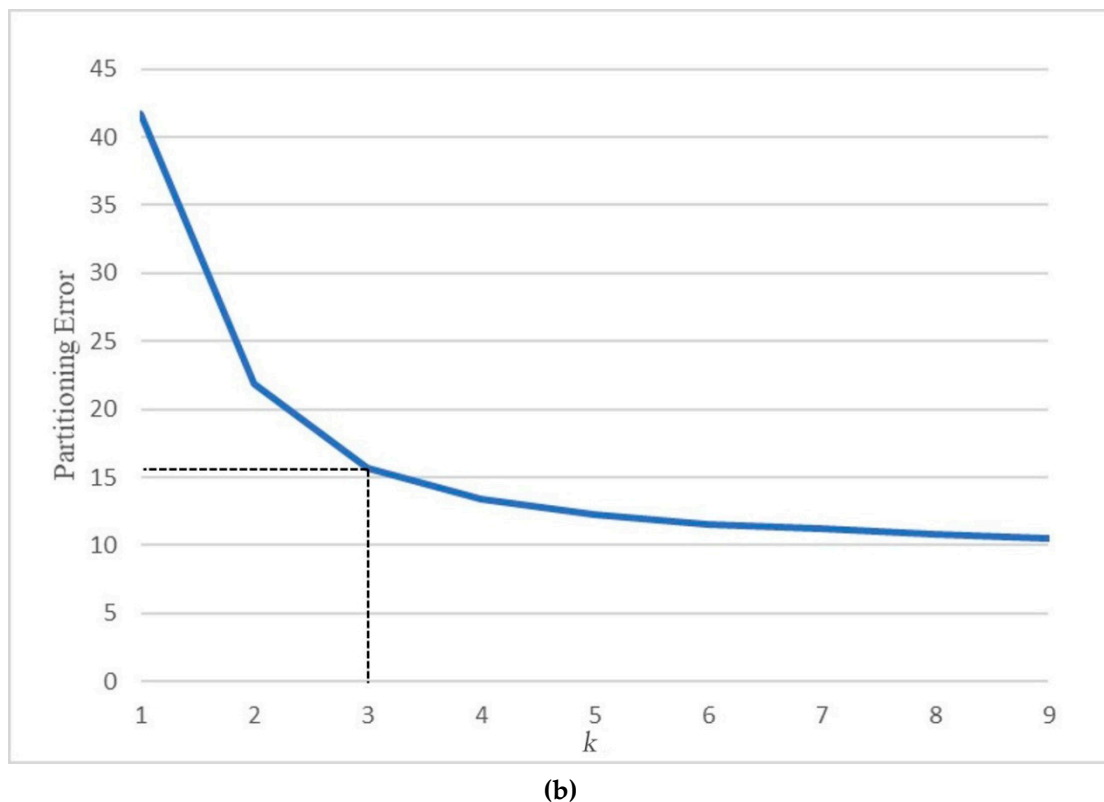


Figure 2. Cont.



**Figure 2.** (a) Variation trend of partitioning error with k value on workdays; (b) variation trend of partitioning error with k value on weekends.

The activity patterns of DLBS near metro stations could be indicated from the variation trend of the average number of Mobike. As shown in Figure 3a, the variation trend on workdays can be categorised into five types based on the shapes of the curves.

Cluster 1 reflects an inactive activity pattern. The Mobike count near those metro stations was stable, between 0 and 2, which indicates that the visitor flow rate of those systems is relatively low; moreover, the time-variant characteristic was not significant.

Cluster 2 represents a tidal characteristic activity pattern. The number of Mobikes near stations rapidly increases between 6 am and 8 am, slowly decreases between 8 am and 9 am, slowly increases until 6 pm and decreases significantly between 6 pm and midnight.

Cluster 3 possesses a more distinctive characteristic than the others. The number of Mobikes near stations is relatively stable between midnight and 7:30 am, rapidly decreases between 7:30 am and 9 am, remains relatively low from 9 am to 4 pm, rapidly increases between 4 pm and 6 pm and then stays relatively stable.

Cluster 4 displays an opposite characteristic to Cluster 3. The number of Mobikes near stations increases rapidly from 6 am to 8 am, remains stable between 8 am and 6 pm and significantly declines from 6 pm to midnight. Moreover, its integral level is above Cluster 2.

Ostensibly, Cluster 5 possesses a similar time-variant characteristic to Cluster 1. However, the integral level of the number of Mobikes near stations is significantly higher than Cluster 1, remaining stable at between 12 and 14.

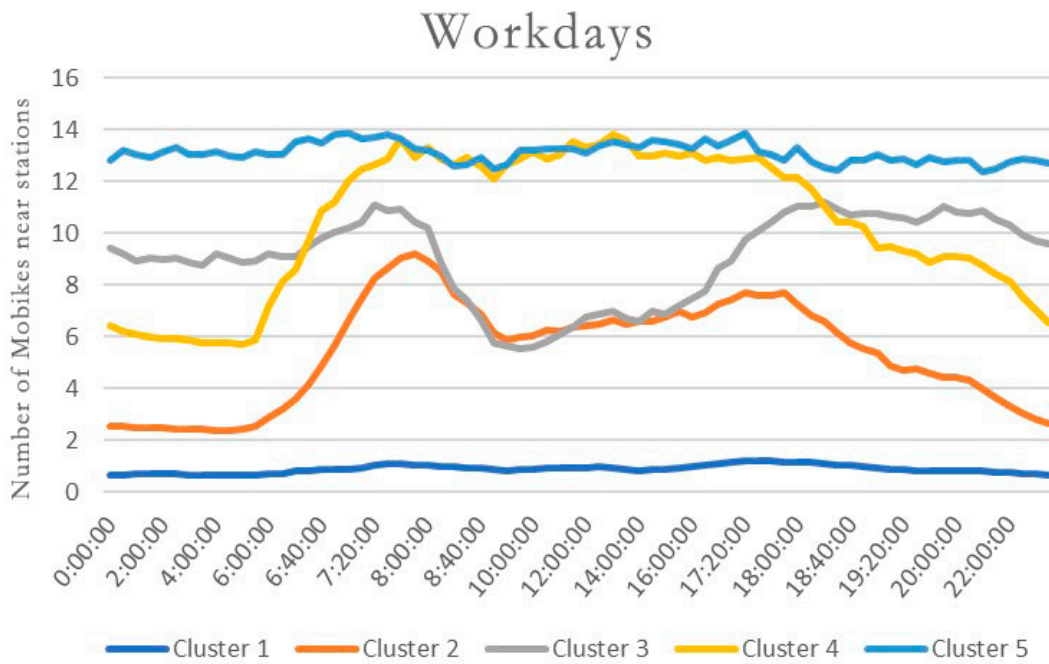
The variation trend on weekends can be categorised into three types based on the shapes of the curves as seen in Figure 3b.

Cluster 1 reflects an inactive activity pattern similar to Cluster 1 on workdays.

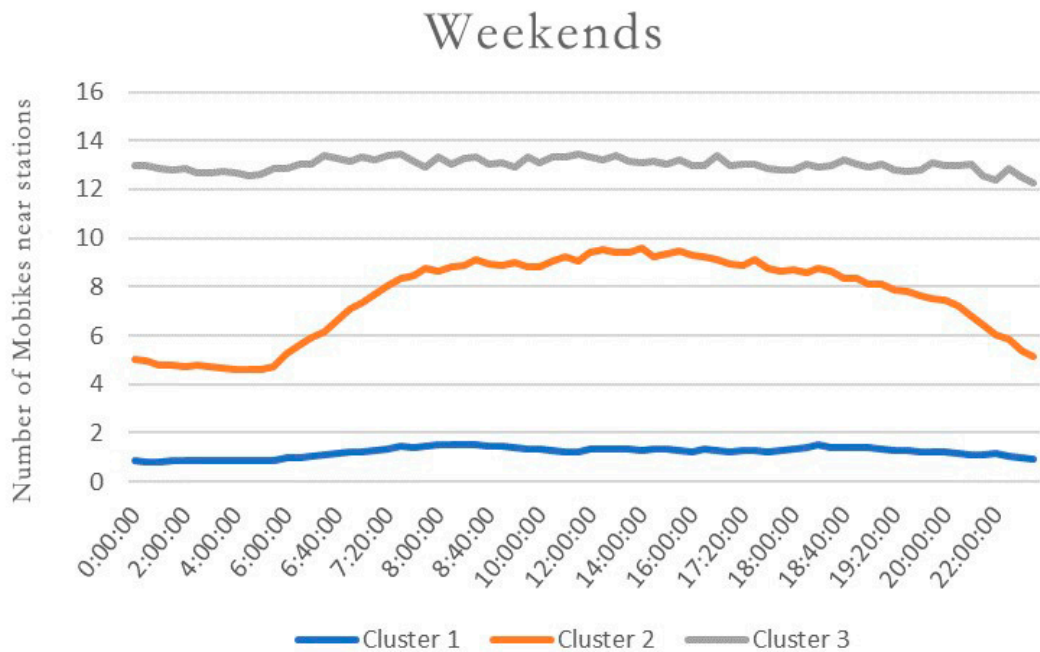
Cluster 2 significantly increases between 6 am and 8 am, remains relatively stable until 4 pm and then decreases slowly until midnight.



Cluster 3 has a similar time-variant characteristic to Cluster 5 on workdays, and its integral level is stable at between 12 and 14.



(a)



(b)

**Figure 3.** (a) Activity patterns of each cluster for metro stations on workdays; (b) activity patterns of each cluster for metro stations on weekends.

Figure 4a displays all types of metro stations on workdays, reflecting the relationship between spatial distribution and type of activity pattern. It can be observed that Cluster 1 stations commonly distribute around peripheral zones of the city where the land is undeveloped and economic activity is low. These conditions can reasonably explain the inactivity pattern. Cluster 2 distributes in new

development districts and suburban areas that usually contain both residential and business areas. Based on the curve of Cluster 2, it can be inferred that there is a certain percentage of residents that live in these areas but work in other districts. Residents who work in other districts are required to go to work using the metro in the early morning, which causes the increase in the number of Mobikes near metro stations between 6 am and 8 am. Workers residing in other areas arrive at the metro stations and ride Mobikes to work, which accounts for the decrease in Mobikes near the stations between 8 am and 9 am. A slow increase is due to workers who live in other areas who leave work at different times between 4 pm and 6 pm. Finally, a significant decrease occurs due to residents of these areas arriving home between 6 pm and midnight. Cluster 3 usually distributes in high-tech areas and industrial parks where jobs are concentrated. This explains the rapid decrease from 7:30 am to 9 am and the rapid increase from 4 pm to 6 pm. Cluster 4 commonly distributes in residential areas of the city, which explains the rapid increase between 6 am and 8 am due to residents going to work and the significant decrease from 6 pm to midnight due to residents arriving home after work. Cluster 5 commonly distributes in the downtown area and near tourist attractions. While the curve of Cluster 5 is flat, the integral level of the number of Mobikes is significantly higher than any other cluster, which indicates the activity pattern is not tidal, and the system maintains a high turnover rate at all times.

The types of metro stations used on weekends are shown in Figure 4b. Cluster 1 exhibits a similar characteristic with Cluster 1 on workdays. The stations distribute primarily in peripheral zones of the city. Cluster 2 distributes primarily in residential areas. It can be inferred that activities of residents on weekends causes the significant increase from 6 am to 8 am and slow decrease from 4 pm to midnight. Cluster 3 distributes primarily in the downtown area, near tourist attractions and the areas where metro lines intersect. It can be inferred that a high frequency passenger flow contributes to the high turnover rate of Mobikes and the large number of Mobikes near stations.

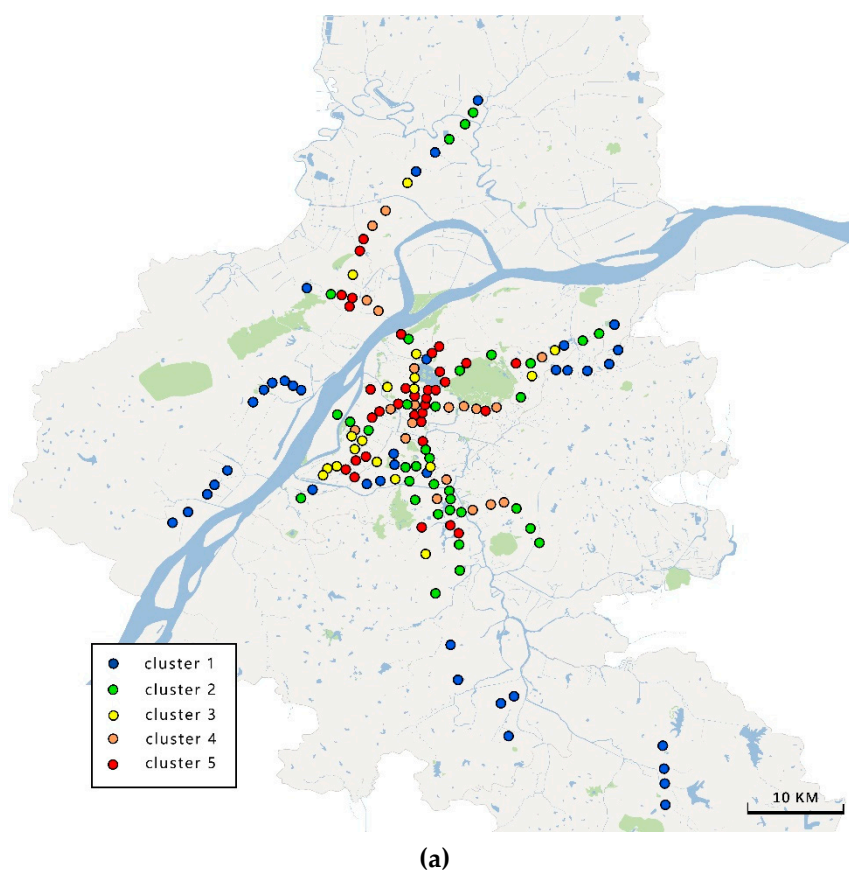
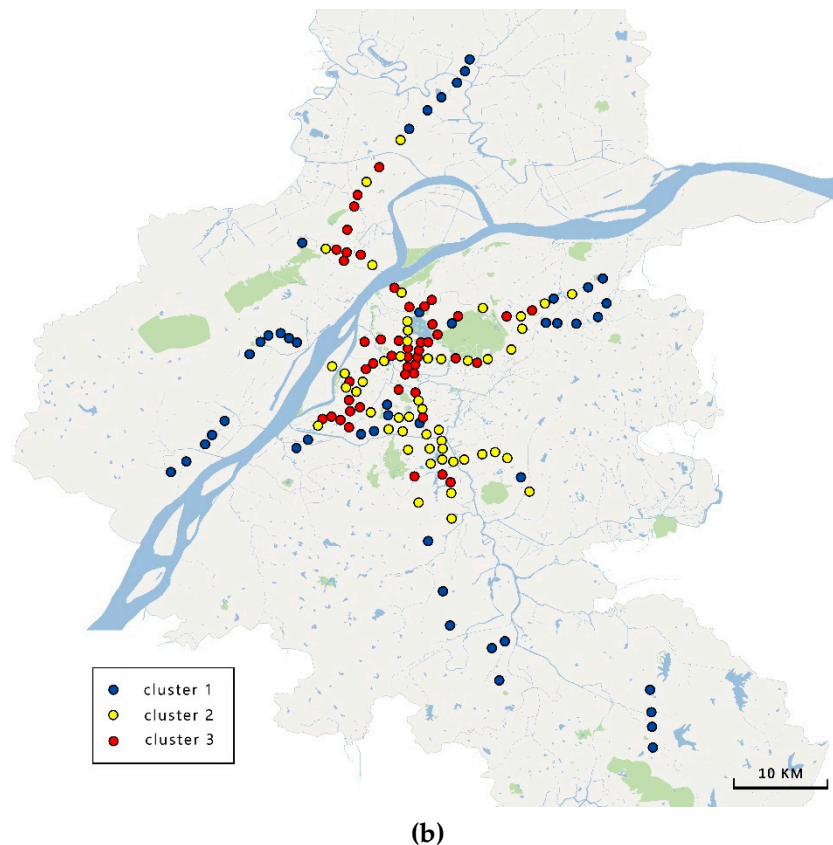
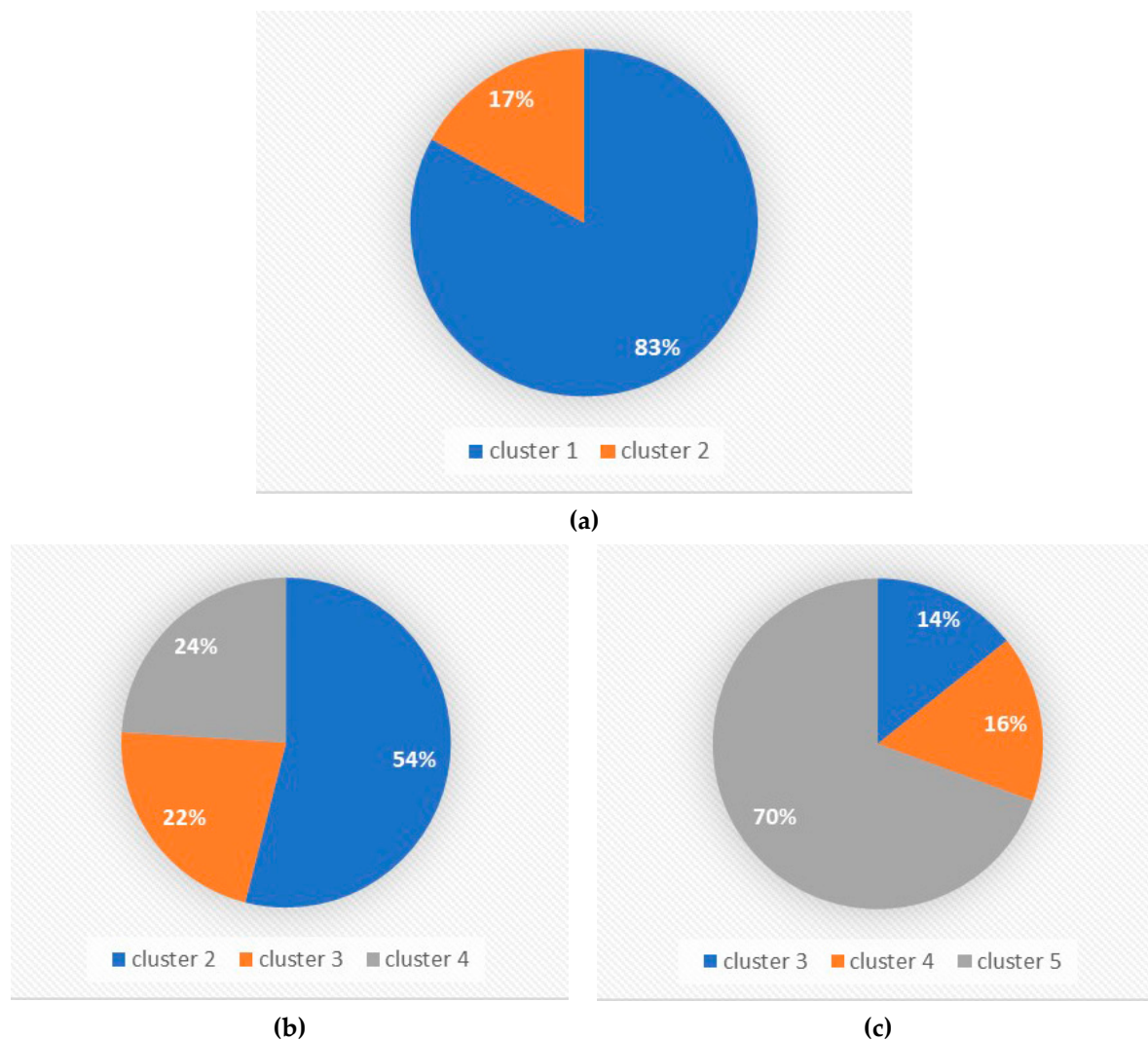


Figure 4. Cont.



**Figure 4.** (a) Geographical distribution of each cluster for metro stations on workdays; (b) geographical distribution of each cluster for metro stations on weekends.

The relationship between clusters on workdays and weekends is shown in Figure 5. It indicates that eighty-three percent of the metro stations in Cluster 1 on weekends are also in Cluster 1 on workdays, and seventeen percent are in Cluster 2 on workdays. It can be inferred that residents live around the stations of Cluster 2 for workdays in relatively remote areas that have less travel activities on weekends. Combined with Figure 4a,b, Figure 5b shows the percentage of weekend Cluster 2 stations that are also in a particular workday cluster. Specifically, fifty-four percent are in workday Cluster 2, twenty-two percent are in workday Cluster 3 and twenty-four percent are in workday Cluster 4. It can be inferred that residents who live around the stations in Clusters 2, 3 and 4 disperse into new development districts and suburban areas that are at a moderate distance from downtown; additionally, these residents have a tidal travel activity pattern on weekends. Combined with Figure 4a,b, Figure 5c shows that seventy percent of the metro stations in Cluster 3 on weekends are in Cluster 5 on workdays, fourteen percent are in Cluster 3 on workdays and sixteen percent of stations are in Cluster 4 on workdays. It can be inferred that some stations distribute in tourist attraction areas and recreational areas that have a higher visitor flow rate and turnover rate on weekends than workdays.



**Figure 5.** (a) Corresponding workday clusters of the stations that are in Cluster 1 on weekends; (b) corresponding workday clusters of the stations that are in Cluster 2 on weekends; (c) corresponding workday clusters of the stations that are in Cluster 3 on weekends.

#### 4.3. Analysis of Range of Influence of DLBS Systems Near Metro Stations

In the previous analysis, metro stations were classified into five clusters on workdays and three clusters on weekends. It was determined that different stations possess different activity patterns for nearby DLBS systems. Moreover, the relationship between the cluster types on workdays and on weekends was examined. The relationships between POI and metro stations of different clusters have significant differences, which has a direct impact on the range of vehicle rebalancing and the scale of parking facilities for operation and management of DLBS systems. The ranges of influence of DLBS systems in different clusters reflect the differences. This section will focus on the analysis of range of influences of DLBS systems near metro stations in terms of their cluster types.

We examined every POI and nearby Mobike for every metro station in every time point to filter the POI which contain the same Mobike that appeared in a nearby station within the previous or subsequent hour. The filter rules are listed as follows:

1. The Mobike whose distance is within 100 metres of a metro station is considered related to the station.
2. At every time point, the ID of every related Mobike is stored for every station.
3. Examine every POI within the previous and subsequent hours for every time point for every station to select those POI which contain a related Mobike and store in the initial list of influenced POI.
4. Any POI that appears at least three times in the initial list of influenced POI for every station is considered to be in the range of influence of DLBS systems near the station.

After filtering the data, distances between the influenced POI and corresponding metro stations were acquired using the web API for Amap. The average range of influence of DLBS systems near every metro station was calculated using Equation (2).

$$R_f = \frac{d_1 + d_2 + d_3 + \dots + d_n - d_{max} - d_{min}}{n - 2} \quad (2)$$

where  $R_f$  is the average range of influence,  $d_i$  is the distance between number  $i$  of POI in the range of influence and the station,  $d_{max}$  is the maximum among  $d_i$ ,  $d_{min}$  is the minimum among  $d_i$  and  $n$  is the sum of the POI that are in the range of influence.

Here is an example for calculating the average range of influence using Equation (2). Station A was determined to have 5 POI in its range of influence after being filtered by the rules above. The distances between those POI and station A are listed as follows: 1000 metres, 1100 metres, 1200 metres, 1300 metres and 1400 metres.  $R_f$  is calculated using Equation (2):

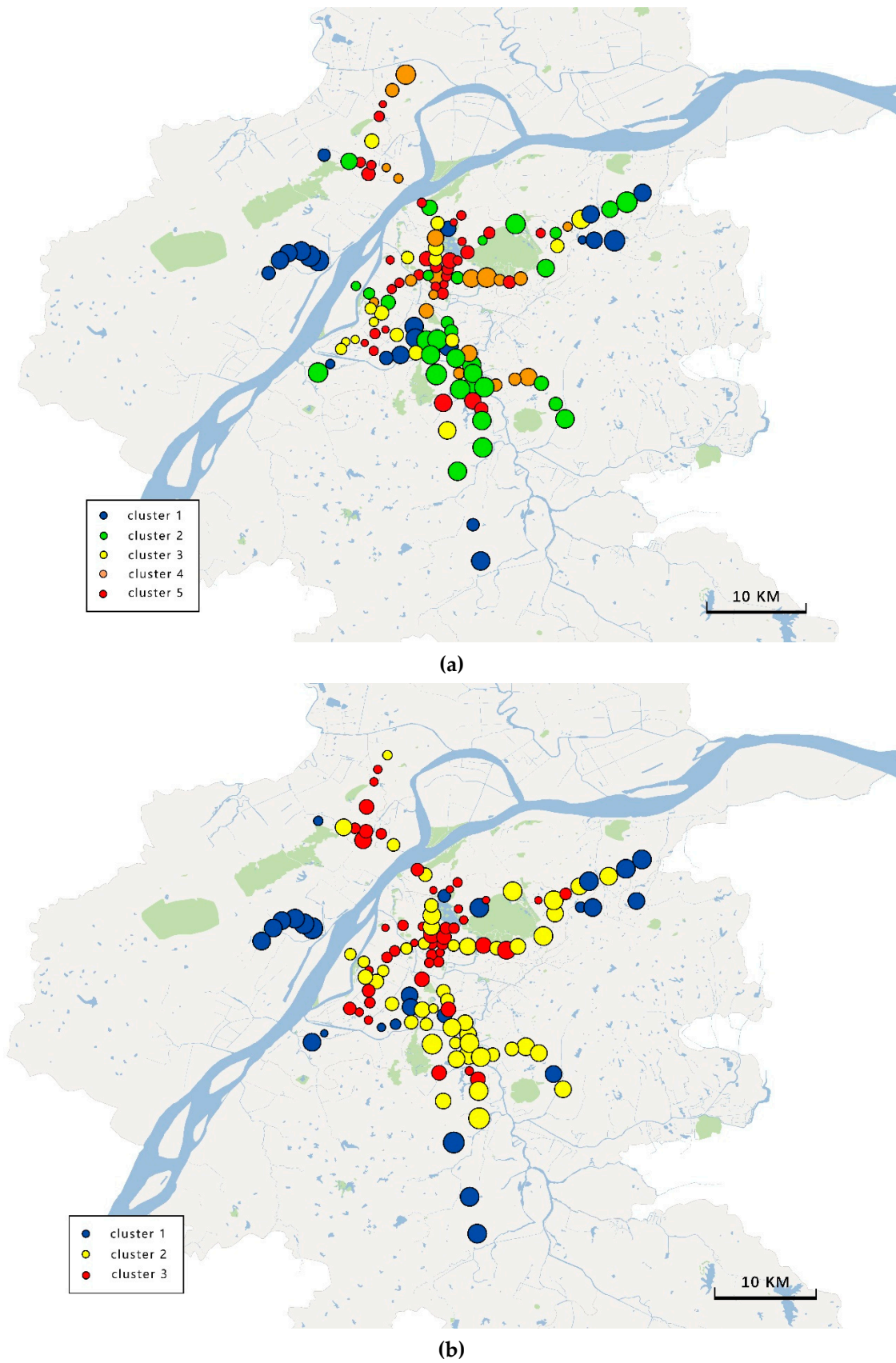
$$R_f = \frac{1000 + 1100 + 1200 + 1300 + 1400 - 1400 - 1000}{5 - 2} = 1200. \quad (3)$$

In Figure 6, the average range of influence of DLBS systems near every metro station is displayed relatively by the size of icons and coloured based on the type of cluster (there were some stations filtered out due to incomplete data). The specific average range of influence for clusters on workdays and weekends is shown in Figure 7.

Figure 6 shows a general tendency that the closer the metro station is to the city centre, the smaller the range of influence is for stations in the same cluster. For metro stations distributed in remote areas, the ranges of influence on weekends are commonly larger than those on workdays. Conversely, for stations distributed around downtown, the ranges of influence on weekends are smaller than those on workdays.

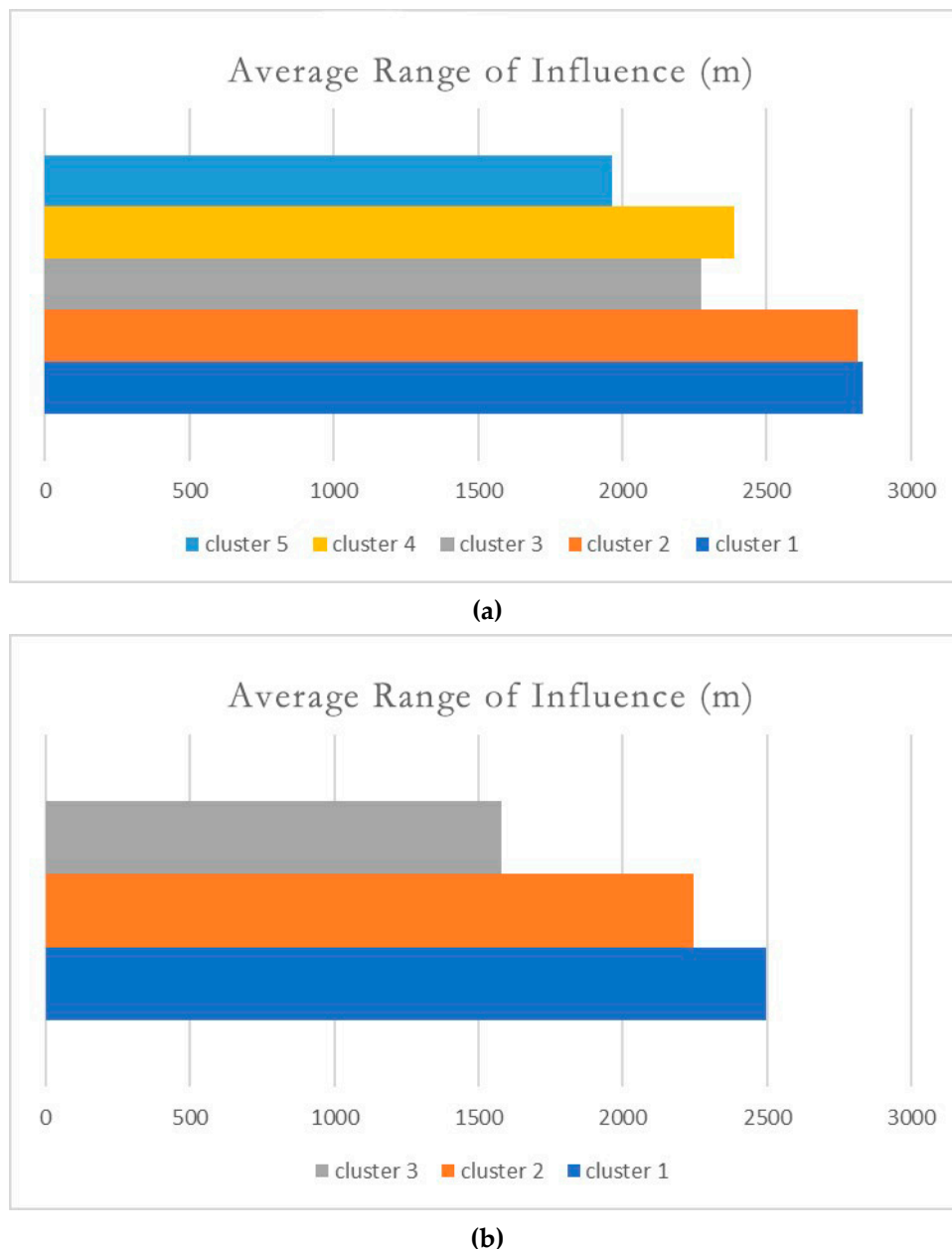
Figure 7a shows that metro stations in Clusters 1 and 2 have the largest average range of influence on workdays, followed by Clusters 4, 3 and 5, in that order. Thus, the low land development rate and sparse distribution of metro stations causes a large average range of influence and vice versa combined with the analysis. Figure 7b shows that the average ranges of influence in Clusters 1, 2 and 3 decrease in that order on weekends for the same reason.





**Figure 6.** (a) Geographical distribution and average range of influence of each cluster on workdays; (b) geographical distribution and average range of influence of each cluster on weekends.





**Figure 7.** (a) Average range of influence of metro stations from each cluster on workdays; (b) average range of influence of metro stations from each cluster on weekends.

## 5. Conclusions

The work presented in this paper focused on DLBS systems near metro stations in the city of Nanjing. This study included workday and weekend data and focused on analysing activity patterns of DLBS systems, examining the relationship between spatial distribution and activity patterns and the determination of the range of influence for nearby stations. The primary conclusions are as follows:

First, the metro stations of Nanjing can be clustered into five types on workdays and three types on weekends based on activity patterns of the DLBS system nearby. For workdays, Cluster 1 reflects an inactivity pattern and commonly distributes around peripheral zones of the city where the land is undeveloped and economic activity is low. Cluster 2 displays a tidal characteristic activity pattern that has two distinct peaks, and commonly distributes in new development districts and suburban areas containing both residential and business areas. Cluster 3 exhibits a concave characteristic activity pattern and usually distributes in high-tech and industrial parks where jobs are concentrated. Cluster 4

displays a convex characteristic activity pattern and commonly distributes in residential areas. Cluster 5 exhibits a flat but high turnover activity pattern and commonly distributes in the downtown area and near tourist attractions.

Second, for weekends: Cluster 1 exhibits an inactivity pattern and distributes primarily in peripheral zones of the city. Cluster 2 displays a convex characteristic activity pattern and distributes primarily in residential areas. Cluster 3 reflects a flat but high turnover activity pattern and distributes primarily in the downtown area, near tourist attractions and the area where metro lines connect.

Third, the majority of metro stations share similar activity patterns both on workdays and weekends. However, the stations distributed in areas where jobs are concentrated display clear differences in activity patterns between workdays and weekends.

Fourth, there is a general tendency that the closer the metro station is to the city centre, the smaller the range of influence of nearby DLBS systems in the same cluster is. The ranges of influence on weekends are usually larger than those on workdays for metro stations distributed in remote areas. The opposite is true for stations distributed around downtown.

Fifth, low land development rate and sparse distribution of metro stations cause large average ranges of influence.

Based on the conclusions above, some suggestions about operating DLBS systems to address the first-and-last mile connections for metro stations are proposed as follows:

For the stations in Cluster 1 on workdays and weekends, passengers usually travel a longer distance to metro stations when using the DLBS system. The demand for passengers to use the DLBS system to connect to metro stations is not strong. Therefore, the operator of DLBS systems should focus on the vehicle maintenance.

For the stations in Clusters 2, 3 and 4 on workdays and Cluster 2 on weekends, the demand for passengers using the DLBS system to connect to metro stations possesses a tidal characteristic. It is possible that an imbalance between demand and vehicle configuration occurs during the peak period. In addition, these stations generally distribute in residence areas and business districts. The operator of the DLBS should focus on the rebalancing of vehicle distribution. For stations in residential areas, the vehicles tend to gather around stations and are required to be moved back to surrounding areas between 6 am and 8 am; additionally, it is necessary to transport the vehicles to stations after 6 pm. For stations distributed in business districts, the rebalancing operation should be carried out in reverse. The stations in Cluster 3 on workdays are distinct in that the rebalancing operation is unnecessary, because the passengers themselves would do the rebalancing.

For the stations in Cluster 5 on workdays and Cluster 3 on weekends, the distance for which passengers use the DLBS system to connect to metro stations is the shortest. Additionally, the demand is constantly high. Moreover, these stations generally distribute in downtown areas and near tourist attractions. The operator of the DLBS should focus on providing enough parking spaces and services for the DLBS system and passengers near metro stations.

This study analyses the activity characteristics of DLBS systems near metro stations in Nanjing City, examining their temporal and spatial distribution features and their range of influence. This study contributes to the literature on the operation and management of DLBS systems in China and has implications for understanding the operating state of DLBS systems near metro stations, understanding the DLBS systems' effect on improving the first-and-last mile connection for metro stations and promoting the proper management of DLBS systems in similar cities. Owing to the limitations of time and space, this paper mainly focuses on the operating characteristics of DLBS systems in Nanjing City, but lacks the analysis of the relationship between DLBS systems and passenger flow characteristics of metro stations. The operating characteristics of DLBS systems in other types of cities still need to be studied. Further research should concentrate on the operating characteristics of the DLBS system in other types of cities and the relationship between DLBS bike use and metro use.

**Author Contributions:** Conceptualization, Y.L. and Z.Z.; methodology, Y.L.; software, Y.L.; validation, Y.L., Z.Z. and X.G.; formal analysis, Y.L.; investigation, Y.L. and Z.Z.; resources, X.G.; data curation, Y.L.; writing—original draft preparation, Y.L.; writing—review and editing, Z.Z.; visualization, Y.L.; supervision, Z.Z. and X.G.; project administration, X.G.; funding acquisition, X.G.

**Funding:** This research received no external funding.

**Acknowledgments:** We thank Tao Tao (University of Minnesota), Xiaolin Gong (Nanjing Forestry University) and Dewen Kong (Beijing University of Technology) for their valuable comments on early drafts.

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

## References

1. Wang, Y.; Lei, L.; Wang, Z.; Lv, T.; Wang, L. Mode shift behavior impacts from the introduction of metro service: Case study of Xi' an, China. *J. Urban Plan. Dev.* **2013**, *139*, 216–225. [CrossRef]
2. Zhu, Z.; Li, Z.; Chen, H.; Liu, Y.; Zeng, J. Subjective well-being in China: How much does commuting matter? *Transportation* **2017**, 1–20. [CrossRef]
3. Beijing Gaode Software Co. China Major City Traffic Analysis Report. 2017. Available online: <https://report.amap.com/share.do?id=8a38bb8660f9109101610835e79701bf> (accessed on 10 April 2019).
4. Li, Y.; Guo, H.; Li, H.; Xu, G.; Wang, Z.; Kong, C. Transit-oriented land planning model considering sustainability of mass rail transit. *J. Urban Plan. Dev.* **2010**, *136*, 243–248. [CrossRef]
5. Wang, Z.; Chen, F.; Xu, T. Interchange between metro and other modes: Access distance and catchment area. *J. Urban Plan. Dev.* **2016**, *142*, 1–9. [CrossRef]
6. Yang, M.; Liu, X.; Wang, W.; Li, Z.; Zhao, J. Empirical analysis of a mode shift to using public bicycles to access the suburban metro: Survey of Nanjing, China. *J. Urban Plan. Dev.* **2016**, *142*, 05015011. [CrossRef]
7. Cervero, R. *Transit-Oriented Development in the United States: Experiences, Challenges, and Prospects*; Transportation Research Board: Washington, DC, USA, 2004.
8. Demaio, P. Bike-sharing: history, impacts, models of provision, and future. *J. Public Transp.* **2009**, *12*, 41–56. [CrossRef]
9. Chandra, D.K.; Liu, K.; Fu, Y. How Unbalanced are Bicycle Dynamics? Demand-Supply Shortage Detection with Spatiotemporal Tensor Factorization in Station-Less Bike Systems. 2018. Available online: [https://illidanlab.github.io/big\\_traffic/2018/papers/chandra2018show.pdf](https://illidanlab.github.io/big_traffic/2018/papers/chandra2018show.pdf) (accessed on 10 April 2019).
10. Pfrommer, J.; Warrington, J.; Schildbach, G.; Morari, M. Dynamic vehicle redistribution and online price incentives in shared mobility systems. *Ieee Trans. Intell. Transp. Syst.* **2014**, *15*, 1567–1578. [CrossRef]
11. Ahmed, F.; Rose, G.; Jacob, C. Impact of weather on commuter cyclist behaviour and implications for climate change adaptation. In Proceedings of the ATRF 2010: 33rd Australasian Transport Research Forum, Canberra, Australia, 29 September–1 October 2010.
12. Fishman, E.; Washington, S.; Haworth, N. Bike share's impact on car use: Evidence from the United States, Great Britain, and Australia. *Transp. Res. Part D Transp. Environ.* **2014**, *31*, 13–20. [CrossRef]
13. Buck, D.; Buehler, R.; Happ, P.; Rawls, B.; Chung, P.; Borecki, N. Are bikeshare users different from regular cyclists? A first look at short-term users, annual members, and area cyclists in the Washington, D.C., Region. *Transp. Res. Rec. J. Transp. Res. Board* **2013**, *2387*, 112–119. [CrossRef]
14. Tao, T.; Guo, X.; Li, J.; Huang, Y. Operating characteristics of a public bicycle-sharing system based on the status of stations: Case study in Nanning City, China. In Proceedings of the Transportation Research Board 96th Annual Meeting, Washington, DC, USA, 8–12 January 2017.
15. Froehlich, J.; Neumann, J.; Oliver, N. Sensing and Predicting the Pulse of the City through Shared Bicycling. *Ijcai Int. Jt. Conf. Artif. Intell.* **2009**, *3*, 1420–1426.
16. Fishman, E.; Washington, S.; Haworth, N. Bike share: A synthesis of the literature. *Transp. Rev.* **2013**, *33*, 148–165. [CrossRef]
17. Fishman, E. Bikeshare: A review of recent literature. *Transp. Rev.* **2016**, *36*, 92–113. [CrossRef]
18. Fishman, E.; Washington, S.; Haworth, N.; Watson, A. Factors influencing bike share membership: An analysis of Melbourne and Brisbane. *Transp. Res. Part A Policy Pract.* **2014**, *71*, 17–30. [CrossRef]
19. Zhu, W.; Pang, Y. Travel behavior change after the introduction of public bicycle systems: Case study in Minhang District, Shanghai. In Proceedings of the Transportation Research Board 92nd Annual Meeting, Washington, DC, USA, 13–17 January 2013.

20. Shaheen, S.; Martin, E. Unraveling the Modal Impacts of Bikesharing. *ACCESS Magazine* **2015**, *1*, 8–15.
21. Fishman, E. Bikeshare: Barriers, Facilitators and Impacts on Car Use. Ph.D. Thesis, Queensland University of Technology, Brisbane, Australia, 2014.
22. Faghih-Imani, A.; Eluru, N.; El-Geneidy, A.M.; Rabbat, M.; Haq, U. How land-use and urban form impact bicycle flows: Evidence from the bicycle-sharing system (BIXI) in Montreal. *J. Transp. Geogr.* **2014**, *41*, 306–314. [[CrossRef](#)]
23. Parkes, S.D.; Marsden, G.; Shaheen, S.A.; Cohen, A.P. Understanding the diffusion of public bikesharing systems: Evidence from Europe and North America. *J. Transp. Geogr.* **2013**, *31*, 94–103. [[CrossRef](#)]
24. Bao, J.; He, T.; Ruan, S.; Li, Y.; Zheng, Y. Planning Bike Lanes Based on Sharing-Bikes' Trajectories. Available online: <https://www.microsoft.com/en-us/research/wp-content/uploads/2017/06/main.pdf> (accessed on 10 April 2019).
25. Wu, F.; Xue, Y. Innovations of Bike Sharing Industry in China—A Case Study of Mobike's Station-Less Bike Sharing System. Ph.D. Thesis, KTH Royal Institute of Technology, Stockholm, Sweden, 2017.
26. Du, M.; Cheng, L. Better understanding the characteristics and influential factors of different travel patterns in free-floating bike sharing: Evidence from Nanjing, China. *Sustainability* **2018**, *10*, 1244. [[CrossRef](#)]
27. Shi, J.; Si, H.; Wu, G.; Su, Y.; Lan, J. Critical factors to achieve dockless bike-sharing sustainability in China: A stakeholder-oriented network perspective. *Sustainability* **2018**, *10*, 2090. [[CrossRef](#)]
28. Martens, K. The bicycle as a feeding mode: Experiences from three European countries. *Transp. Res. Part D-Transp. Environ.* **2004**, 281–294. [[CrossRef](#)]
29. Zhao, P.; Li, S. Bicycle-Metro integration in a growing city: The determinants of cycling as a transfer mode in metro station areas in Beijing. *Transp. Res. Part A Policy Pract.* **2017**, *99*, 46–60. [[CrossRef](#)]
30. Lin, J.; Zhao, P.; Takada, K.; Li, S.; Yai, T.; Chen, C. Built environment and public bike usage for metro access: A Comparison of Neighborhoods in Beijing, Taipei, and Tokyo. *Transp. Res. Part D Transp. Environ.* **2018**, *63*, 209–221. [[CrossRef](#)]
31. Ma, X.; Ji, Y.; Yang, M.; Jin, Y.; Tan, X. Understanding bikeshare mode as a feeder to metro by isolating metro-bikeshare transfers from smart card data. *Transp. Policy* **2018**, *71*, 57–69. [[CrossRef](#)]
32. Cheng, Y.; Lin, Y. Expanding the effect of metro station service coverage by incorporating a public bicycle sharing system. *Int. J. Sustain. Transp.* **2018**, *12*, 241–252. [[CrossRef](#)]
33. Zhang, Z.; Qian, C.; Bian, Y. Bicycle-metro integration for the 'Last Mile': Visualizing cycling in Shanghai. *Environ. Plan. A Econ. Space* **2018**, *0*, 1–4. [[CrossRef](#)]
34. Kanungo, T.; Mount, D.M.; Netanyahu, N.S.; Piatko, C.D.; Silverman, R.; Wu, A.Y. An efficient k-means clustering algorithm: Analysis and implementation. *Ieee Trans. Pattern Anal. Mach. Intell.* **2002**, *24*, 881–892. [[CrossRef](#)]



© 2019 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 (<http://creativecommons.org/licenses/by/4.0/>).