Understanding the usage of dockless bike sharing in Singapore

Yu Shen[a,b], Xiaohu Zhang[b], and Jinhua Zhao[c]

[a]Key Laboratory of Road and Traffic Engineering of the Ministry of Education, Tongji University, Shanghai, China; [b]Singapore–MIT Alliance for Research and Technology Centre, Singapore; [c]Department of Urban Studies and Planning, MIT, Cambridge, MA, USA

ABSTRACT

A new generation of bike-sharing services without docking stations is currently revolutionizing the traditional bike-sharing market as it dramatically expands around the world. This study aims at understanding the usage of new dockless bike-sharing services through the lens of Singapore's prevalent service. We collected the GPS data of all dockless bikes from one of the largest bike sharing operators in Singapore for nine consecutive days, for a total of over 14 million records. We adopted spatial autoregressive models to analyze the spatiotemporal patterns of bike usage during the study period. The models explored the impact of bike fleet size, surrounding built environment, access to public transportation, bicycle infrastructure, and weather conditions on the usage of dockless bikes. Larger bike fleet is associated with higher usage but with diminishing marginal impact. In addition, high land use mixtures, easy access to public transportation, more supportive cycling facilities, and free-ride promotions positively impact the usage of dockless bikes. The negative influence of rainfall and high temperatures on bike utilization is also exhibited. The study also offered some guidance to urban planners, policy makers, and transportation practitioners who wish to promote bike-sharing service while ensuring its sustainability.

1. Introduction

The history of bike-sharing programs, as summarized in DeMaio (2009) and Shaheen, Guzman, and Zhang (2010), started with a dockless bike-sharing program—the White Bikes—over half a century ago. The White Bikes were unlocked, free-of-charge, and for public use. However, the availability of technology at that time was not able to support the operation of the system. The program survived for only a short time, ultimately succumbing to a series of problems in fleet management, e.g., theft and vandalism. To better manage shared bikes, docking stations were introduced in the 1990s for locking, payment, renting, and return of the bikes. This initiated the second generation of bike-sharing programs, based on coin-deposit systems. However, due to technological constraints, there was no time limit for bike use and the cash payment was anonymous. Thus, bike theft still failed to be fully addressed, which made the service quite unreliable. With the rapid development of information technology (IT), a renovated bike-sharing system, i.e., the third generation of bike sharing, emerged a few years later. As characterized in Shaheen et al. (2010), the new IT technology made available cashless payment, real-name registration, and dynamic pricing schemes, which enabled the IT-based systems with docking stations to quickly spread from Europe to Asia Pacific, North America, and South America.

But one of the barriers that still hindered bike-sharing services was the ease of access to docking stations (Fishman, Washington, Haworth, & Mazzei, 2014). To overcome this barrier, a system can either increase the number of docking stations or change the program to dockless. Often, limited space in a city constrains the number of docking stations that can be installed. In addition, docking stations not only store and lock the bikes, but they also provide an infrastructure for payment. Absorbing the advantage of a program with docks, a successful dockless bike-sharing program may integrate the functions of docking stations directly into the shared bikes. Two modern pilot programs along these lines started in Germany in the early 2000s, operated by Call-a-Bike and Nextbike. However, these programs are not fully dockless: There are still a number of docking stations installed in the service area.

In 2015, two start-up companies, Ofo and Mobike, initiated an innovative generation of fully dockless bike-sharing services in China. This innovation became possible only with the prevalence of smartphones and cashless mobile payment. The new service has expanded rapidly since 2016 and later spread to other countries such as Singapore, the United Kingdom, and the United States. The new dockless bike-sharing programs integrate mobile payment and GPS tracking into the system, which greatly increases the ease of use and management. Bikes report their locations to the central server through an embedded GPS sensor and communication module. A customer can easily locate bikes via a smartphone app. After finding a bike, the customer can unlock the bike by scanning its QR code or using near field communication (NFC) technology, and the ride starts. All the needed functionality is integrated into the app. To accommodate customers, a credit system can also be established.

The emergence of free-floating, dockless bike-sharing services revolutionizes the market. The new services make renting and returning bikes more convenient than ever. The recent dramatic increase in bike fleet is far beyond the expectations of transportation and urban planners. According to statistics summarized by Zhang, Shaheen, and Chen (2014), in February 2012 there were just over 180,000 third-generation shared bikes in China. In contrast, by March 2017, the fleet size of dockless bikes in China had reached over 4 million, and Shanghai alone had about 450,000 dockless shared bikes (iiMedia Research, 2017). The growth of dockless bike fleets in China is far beyond the shared bike growth projected of bike-sharing programs in any other country (Parkes, Marsden, Shaheen, & Cohen, 2013). In Europe, Nextbike—which operates dock-based, dockless, and hybrid systems—is one of the largest bike-sharing companies. But it operates only 5,000 bikes at 700 stations in Berlin, which is considered the largest German bike-sharing market (Nextbike GMBH, 2016).

Dockless bike-sharing services are significantly different from traditional ones in many aspects. For instance, since docking stations are not needed, the subscription and payment of service can be made via smartphone app. One can locate the bikes based on the GPS and subscriber identity modules installed in the bikes. The bikes can be parked in any places that allows bike parking. The fleet size is not constrained by the capacity of docking stations as well. Therefore, understanding the impact of new dockless bike-sharing systems is a vital need, currently absent from the literature. The typical fleet size of dockless services is much larger than that of conventional shared bikes, and free-floating shared bikes greatly improve ease of access. However, if there are too many bikes in the system while the utilization of the bikes remains at a low level, such services could be fiscally unsustainable or potentially harm the urban transport system. The unused bikes may occupy a lot of public space that could be used for other purposes. Thus, this study aims at understanding the usage of new dockless bike-sharing services through the lens of Singapore's prevalent service, based on spatiotemporal analysis and statistical modeling. We examine the impact of bike fleet size on the usage of dockless bike sharing service. In addition, the influence of surrounding built environment, access to public transportation, bicycle infrastructure, and weather conditions are investigated as well.

The paper is organized as follows. Section 2 elaborates the data collection and processing methods and the identification of valid bike trips. Section 3 presents a descriptive analysis, which includes the spatiotemporal distributions of the dockless bike fleet, the fleet's usage, and the impact of weather conditions. Spatial autoregressive models are built in Section 4 to explore potential influential factors, including bike fleet size, weather, transportation infrastructure, and the built environment. The modeling results are presented and analyzed in Section 5. The final section concludes this paper.

2. The dataset and preprocessing

The dockless bike-sharing system entered the Singaporean market in February 2017. To examine the usage of bike-sharing

service, previous studies have focused mainly on the capacity and occupancy of bike-sharing docking stations (O'Brien, Cheshire, & Batty, 2014). However, those studies' methods provide limited opportunity for analyzing the mobility patterns of dockless bikes. New data collection and analysis methods should be explored. GPS sensors are embedded in dockless bikes, which allow us to analyze the bikes' usage through their spatiotemporal distribution and moving trajectories.[1] We collected real-time bike-sharing data from one of the largest dockless bike operators in Singapore. The data are fully anonymous—no user information is associated. The collection of real-time dockless bike GPS data started on April 25th, 2017, when the service was still expanding. The data used for the study ranges from April 26th (one day after the initiation) to May 4th, 2017, consisting of six workdays, a two-day weekend, and the Labor Day holiday. The unique 9-digit bike ID and the real-time GPS location of every available dockless bike in Singapore were continually recorded with a frequency of 5 minutes on average. The raw outcome contains over 14 million records. Each record has GPS coordinates of one specific bike and an observed timestamp. During the time when a bike is rented, the bike location could not be detected until it is returned. By sorting the GPS locations of each bike chronologically, we can extract the movements of all bikes in Singapore. However, some redundant information and errors exist in the raw data, so we ran a series of preprocessing steps described below.

The first preprocessing step reduces redundant coordinate information about stationary bikes and cleans up some errors due to GPS drifting. For example, one kind of GPS drifting occurs from instabilities in civilian GPS sensors, which can cause a bike to seemingly teleport from one location to another before shifting back to exactly the same location. With these kinds of problems in mind, the preprocessing examines all sequences of three GPS locations for each bike; if the first location is identical to the third, we flag the middle location as being false data. Finally, we remove all flagged locations. As a result, if a bike is stationary for a period of time, we keep only the first and the last coordinates and remove all duplicates in between. Additionally, the process cleans some of the GPS drifting errors as well. The specific process is explained in Step 1 of the pseudocode in Appendix A.1.

We then remove some unrealistically short- or long-distance movement because such movements might not be associated with an actual cycling activity. For instance, the movement of a bike over a very short distance could result from a round trip, or it could be attributed to noncycling causes such as GPS instability, local bike relocation by bike-sharing operators, etc. On the other end of the spectrum, aerobic cycling could lead to a long bike trip of tens of kilometers, but such a movement could also be caused by maintenance and reallocation of bikes. Therefore, to avoid these inconclusive events, we further exclude extremely short- and long-distance trips by devising a strategy based on results from

---

[1]Romanillos et al. (2016) have summarized a series of studies using GPS data from smartphone apps to investigate cycling behaviors. Those studies collected their GPS data by either recruiting a group of participating cyclists or acquiring data directly from smartphone app companies (e.g., fitness apps). Nevertheless, these data are not suitable for studies of bike sharing and do not necessarily reflect the utilization of bike-sharing services.

the household interview travel survey (HITS) of 2012. In all, HITS provides the travel time and postal codes of the origin and destination (OD) of 961 valid cycling trips.

To single out overly long trips based on duration, we first noted that 99.7% (958) of the HITS trips were finished within one hour. There is a half dollar charge for every 15 minutes of dockless bike ride. The cost of a one-hour dockless bike rental (2 dollars) is sufficient to go anywhere in Singapore by public transit. Taking these two facts into account, we decided to select only bike trips shorter than one hour.

We also wanted to exclude some trips based on distance traveled. So, for the HITS data, we geocoded the origin and destination of all cycling trips and calculated the road-network travel distance using Google Maps Services (more specifically, the Geocoding and Direction services). The results are summarized in the box plot shown in Figure A1 of Appendix A.2. In the HITS data, almost all the trips are longer than 150 meters; the only exceptions are six cycling trips with the same origin and destination. And trips longer than 7.5 km are identified as outliers (see Figure A1 in Appendix A.2). Hence, we wanted to select only those trips between 150 meters and 7.5 km, although we needed a small modification to apply this rule to the dockless bike data. Since the distance calculated between each OD pair for dockless bikes is a great-circle distance, we scaled down the reference range from the Google-calculated road-network distance by about 1.47 (the average ratio between the network distance and great-circle distance in HITS). We thus selected only GPS movements with a great-circle distance longer than 100 meters and shorter than 5 kilometers as valid bike trips.

Finally, as dockless bikes are not designed for racing, we used 15 km/h as a cycling reference speed. If the travel speed between two locations is more than twice the reference, we believe that the trips are not valid bike trips.

The complete data filtering procedure is shown in Step 2 of the pseudocode in Appendix A.1. The bike trips considered in this paper are thus defined as the processed GPS movements of the dockless bikes. However, the inherent limitations of the dockless bike GPS data collected in this study must also be acknowledged. The bike-sharing company provided the GPS locations of a bike only when it is idle. No GPS information is known when the bike is being used. Therefore, the specific cycling route remains unknown to us. A potential group of cycling activities—round trips which have same or very similar origin and destination—may have been removed by the preprocessing. However, there are very few round trips (0.6%, or 6 out of 961) reported in the HITS, which suggests that excluding potential round trips may not affect the validity of our analysis.

3. Descriptive analysis

3.1. *Fleet size and utilization*

The everyday fleet size of available bikes during the study period and the total number of bikes that have ever occurred in the system (identified using the bikes' unique 9-digits IDs) are given in Table 1. The number of available bikes by hour in each day is shown in Figure A2(1.1) in Appendix A.3. In general, since the dockless bike-sharing service was newly launched, the total fleet size increased rapidly during the study period. The company was continually introducing new bikes into the city, and the total number of bikes increases by nearly 3,000 during the study period. Despite of the steady increase of fleet size, the availability of bikes each day fluctuated. The number of available bikes surged from 8,217 on April 26th to 9,690 on Labor Day (probably due to a special free-ride promotion); after the holiday, the number dropped slightly before recovering to over 9,800 at the end of the study period.

The average number of trips made per bike and the average usage time per bike are also listed in Table 1. The distribution of the average bike trips, travel distance, and travel time per bike per day can be found in Figure A2 in Appendix A.3. In general, the utilization of bikes in Singapore is quite low. About 14.3% of the bikes were not used at all during the study period. The average idle time per bike was about 23.5 hours per day, while the average number of rides per bike was less than one per day with under 30 minutes of riding. During the study period, special promotions were made for the long weekend. On Saturday (April 29th) and Sunday (April 30th), bike rides for the first half hour were free of charge and bike rides were totally gratis during the whole of Labor Day. Therefore, the utilization of bikes during the long weekend was higher than that in the workdays. On Labor Day, each bike was used about 1.64 times with an average usage time of around 40 minutes.

A detailed temporal distribution of bike trips by hour is plotted in Figure 1(1) below and Figure A2(1.1) in Appendix A.3. Peak usages are observed around 8 a.m. and 6 p.m. every day. Usage is in general higher during daytime on weekends and its trend keeps going up during the observation period. The highest usage is seen around 6–8 p.m. on Labor Day when there was no charge for riding.

Figure 1(2) shows the distribution of average cycling time during the study period for each hour of the day. It should be noted that, since the real-time data was collected with an average time interval of several minutes, there could be a small variation between the actual cycling time and the time computed by the difference of timestamps in the data. Therefore, the real-world travel time may be slightly shorter than the recorded time. As indicated in the figure, most trips ended in less than 30 minutes with an estimated cost of 1 Singapore dollar. During morning peak hours

Table 1. Fleet size and utilization of dockless bikes during the study period.

Date	Total fleet size of bikes	Available bikes	Average usage per bike per day	
			Number of trips	Usage time (min)
2017-04-26	8,519	8,217	0.65	16.5
2017-04-27	8,918	8,261	0.62	16.0
2017-04-28	9,436	8,639	0.73	20.5
2017-04-29	10,116	9,308	1.14	35.2
2017-04-30	10,323	9,274	1.21	33.6
2017-05-01	10,491	9,690	1.64	40.9
2017-05-02	10,728	9,235	0.93	22.9
2017-05-03	11,002	9,436	1.00	23.1
2017-05-04[†]	11,346	9,835	0.72	18.1

Note: The data after 6 p.m. on May 4th is incomplete and are thus excluded from the analysis. The statistics of average usage per bike during this day may not precisely reflect the situations for the full 24 hours.
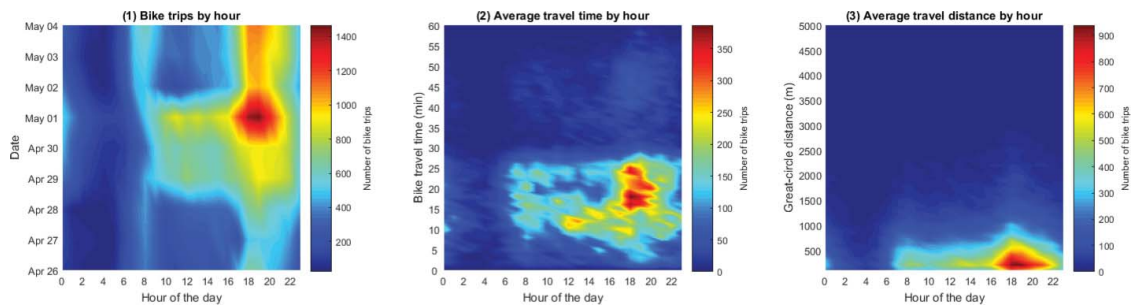
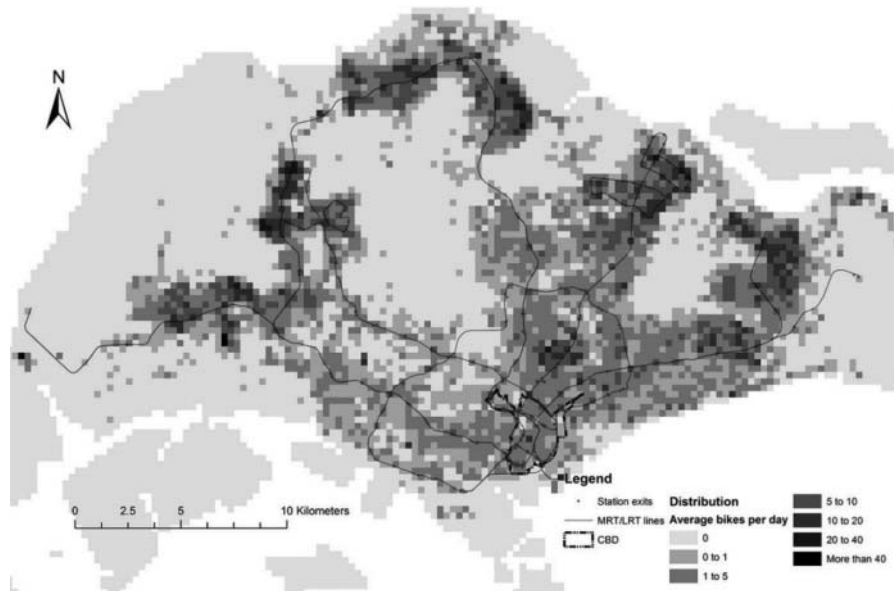Figure 1. Distribution of dockless bike utilization.

Figure 2. Spatial distribution of available bikes in Singapore.

from 7 to 9 a.m., trips are dominantly commuting journeys (or stages of commuting journeys), and most trips were 10 to 25 minutes. During the daytime, from 10 a.m. to 4 p.m., bike travel times became shorter, and most trips ended within 15 minutes. The peak usage ran from 4 to 10 p.m. every day.

Figure 1(3) shows the distribution of the average great-circle distance of bike usage for each hour. Before 4 p.m., most of the trips were less than 500 meters (equivalent to a road-network travel distance of approximately 730 meters, scaled by 1.47). After 4 p.m., the travel distance became longer. Some long-distance trips had a great-circle distance of more than 1 km. But even around 6 to 7 p.m., when there were more long-time and long-distance trips, most still had a great-circle distance of less than 3 km. In general, the number of trips quickly fell off for distances longer than 1 km. As previously discussed, we set a 5-km great-circle distance as the upper limit. The figures in Appendix A2 show that, there are almost no additional bike trips above this limit.

The spatial distribution of available bikes is shown in Figure 2. The map presents the average number of available dockless bikes per day during the study period in each 300 by 300-meter fishnet cell. Mass rapid transit (MRT) stations and the rail network layout are also plotted.[2] The bikes were mainly

concentrated in the peripheral residential areas with high population density and last-mile travel demand to access to the MRT. In the central business district (CBD) with dense MRT stations, the number of dockless bikes was lower than that in the residential areas mentioned above. The allocation of dockless bikes was in general consistent with the last-mile travel demand for accessing the MRT by bus. (The daily last-mile travel demand by bus is plotted in Figure A3 in Appendix A.4).

3.2. The impact of weather on bike utilization

The above analysis presents the variation of dockless bike hourly usage on workdays and on one weekend with a promotion. However, cycling is prone to influence by inclement weather (Corcoran, Li, Rohde, Charles-Edwards, & Mateo-Babiano, 2014; Meng, Zhang, Wong, & Au, 2016; Wojan & Hamrick, 2015). Therefore, to understand the impact of weather conditions, this work takes into account real-time weather information during the data collection period. We collected the readings of air temperature and precipitation from 52 meteorological stations in Singapore with a 5-minute frequency through an API provided by the Singapore open data sharing portal. The statistics are summarized in Figure 3. In each subfigure, the x-axis is the hour from 0 to 23; the left y-axis shows the number of bike trips per hour, plotted in bold black curves; and, the right y-axis illustrates the hourly

---

[2]In this paper, we use MRT to refer to both MRT and LRT (light-rail transit) LRT systems for simplification.

precipitation in centimeters (cm) and the average hourly temperature in degrees centigrade (°C). The precipitation is presented in gray bars, whereas the temperature is drawn in gray curves.

Singapore consists of a main island with a surrounding archipelago. It lies at the latitude of 1.35° N. With an equatorial climate, there is no summer or winter season in the year. The temperature stays high year-round (around 30°C). Even though the Singaporean residents are accustomed to hot weather, the fluctuation of temperature can still influence bike usage. The temperature increased during the midday, and the number of bike trips started to increase only when temperature began to drop during late afternoon.

Rainfall is frequent in Singapore and it sometimes rains heavily. During the study period, the most observable impact of rainfall on bike usage occurred during the 5 p.m. hour of April 27th. In a one-hour period, the accumulated precipitation reached 30 cm. The rainfall is associated with a large decline of bike usage in a time period that normally sees high usage. However, the influence of heavy rain during off-peak daytime hours (e.g., May 2nd from 10 a.m. to 12 p.m.; May 4th from 11 a.m. to 1 p.m.) was less significant than during peak hours. This implies that heavy rainfall during peak hours may reduce overall usage of bikes, but in off-peak periods, precipitation is less influential.

Precipitation in Singapore normally falls for a relatively short time over a small area, not island-wide. To examine the spatial impact of precipitation on bike usage, we constructed an isohyetal precipitation map for the study period based on rainfall readings from 52 meteorological stations in Singapore. The average rainfall in each land cell was interpolated based on an inverse distance weighted method. The isohyetal map is presented in Figure 4(1), with the GPS locations of all dockless bikes plotted as well. Dark gray dots in the map show bikes that were not used in the hour, whereas red dots indicate bikes that moved during the period. For comparative purposes, the utilization of bikes in the hour immediately after the rain is also drawn in Figure 4(2). During the hour of rain, a significant difference between the areas with precipitation below 4 mm and the areas with heavier rain can be observed. Areas with heavy rain showed sporadic bike usage, and areas with light rain had little bike usage as well. Comparing the two figures, the number of bike trips more than doubled from 228 to 500 after the rain. And just after the rain, we see many more bike trips in the area that had heavy rain. The patterns shown in the figures strengthen our findings above: During peak hours, the number of bike trips can be greatly influenced by heavy rainfall.

4. Models

4.1. Impacting factors of bike use

The descriptive analysis shows the potential impacts of the availability of dockless bikes, weather conditions such as air temperature and precipitation, and special promotions on the usage of dockless bikes. In addition, bike usage is also affected by the built environment—such as supportive cycling facilities, land use, etc.—and accessibility to transportation infrastructure. To facilitate further analysis, we divided the study area into regular grid cells so that different measures could be aggregated into each cell. Then, we modeled the impact of all the
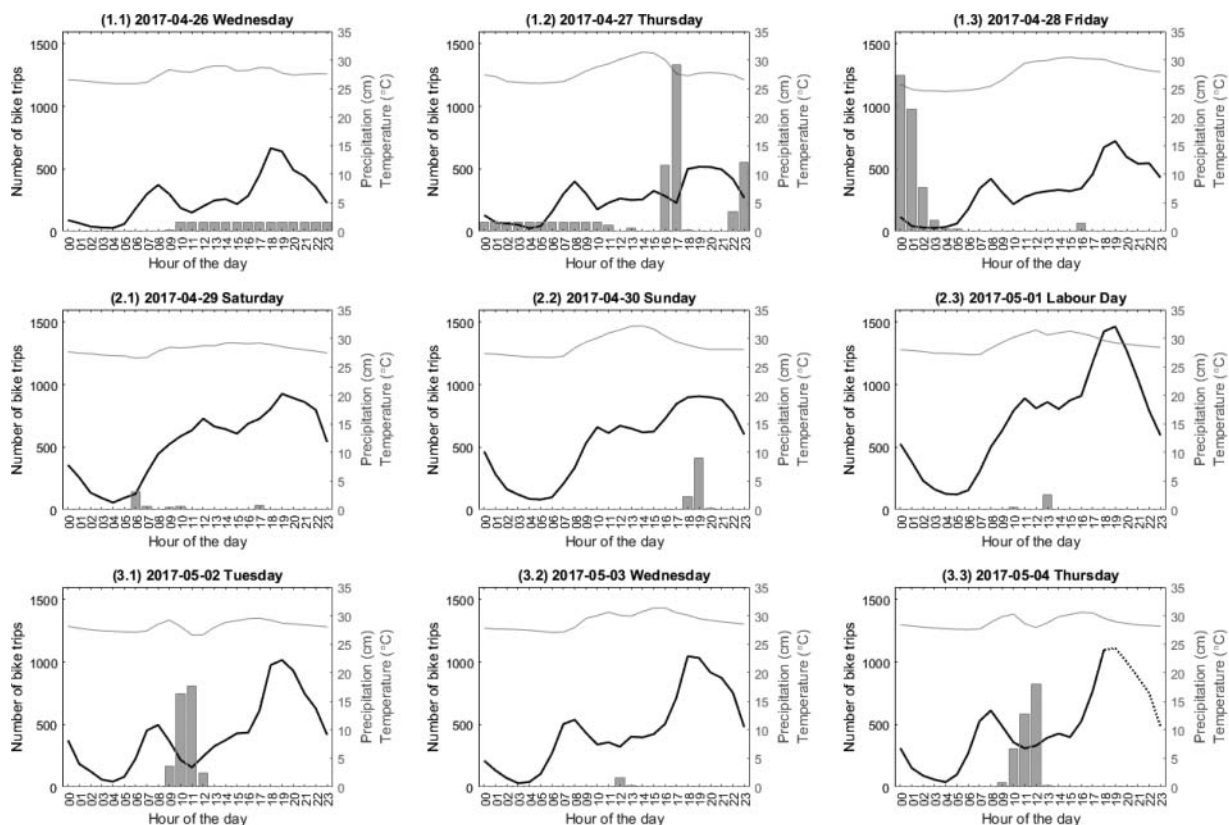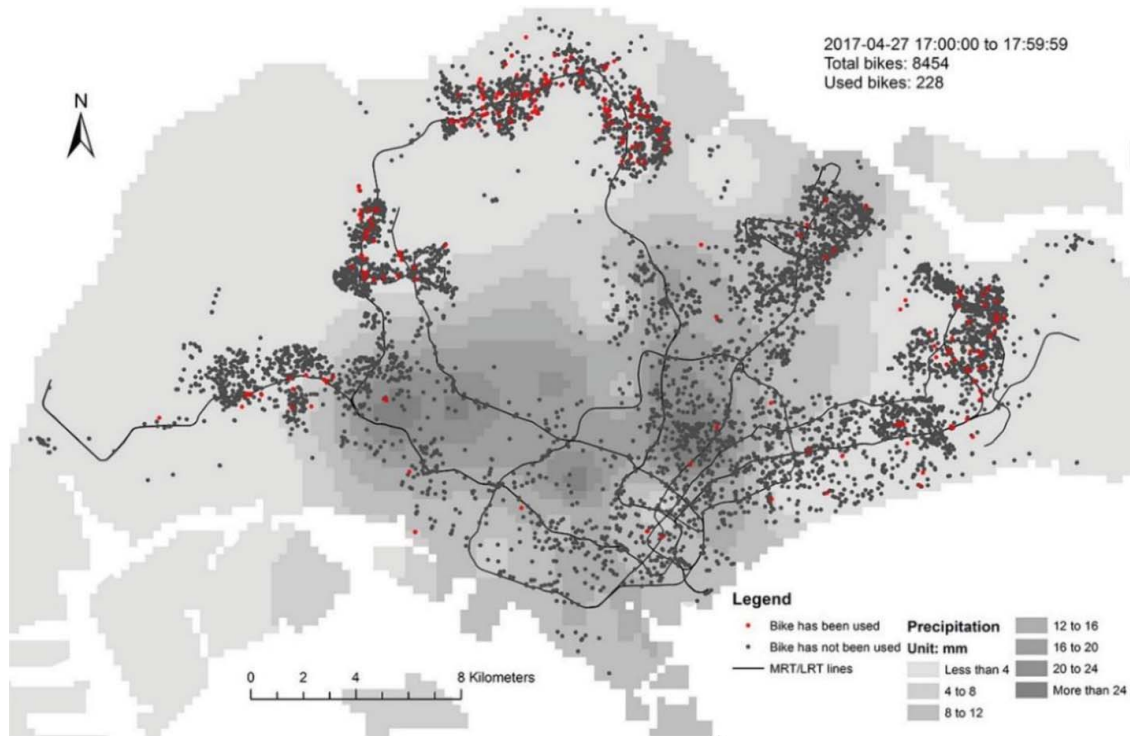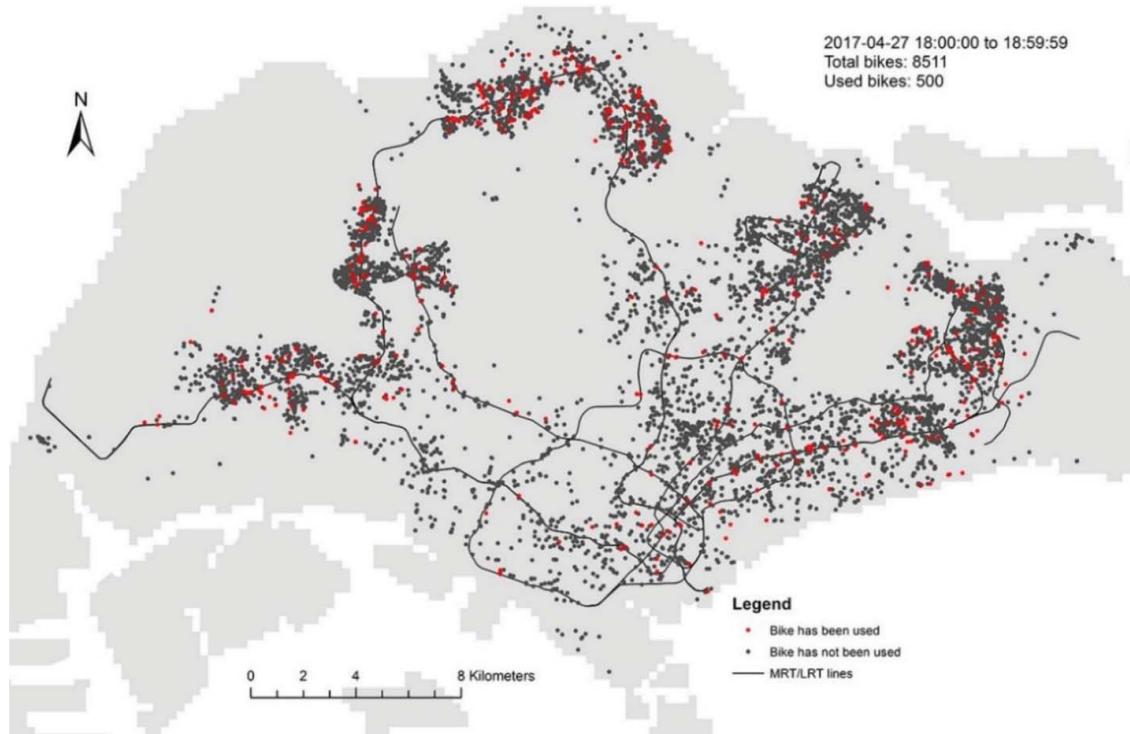

Figure 3. Bike utilization and weather conditions.

2017-04-27 17:00:00 to 17:59:59
Total bikes: 8454
Used bikes: 228

Legend

		Precipitation		12 to 16
•	Bike has been used	**Unit: mm**		16 to 20
•	Bike has not been used		Less than 4	20 to 24
—	MRT/LRT lines		4 to 8	More than 24
			8 to 12	

0 2 4 8 Kilometers

(1) Bike utilization during the rains (Apr. 27th, from 5 p.m. to 6 p.m.)

2017-04-27 18:00:00 to 18:59:59
Total bikes: 8511
Used bikes: 500

Legend

•	Bike has been used
•	Bike has not been used
—	MRT/LRT lines

0 2 4 8 Kilometers

(2) Bike utilization after the rains (Apr. 27th, from 6 p.m. to 7 p.m.)

Figure 4. Relationship between rainfall and bike utilization.

aforementioned factors on the usage of dockless bikes in each spatial unit i (i.e., each 300-by-300 m land cell) at time t as

$$B_{it} = f(S_{it} + BE_i + T_i + C_{it} + O_t) + \rho \Omega B_{it} + e \quad (1)$$

where the variables are designed to capture the impact of the following aspects:

B_{it}: the number of dockless bike trips started from land cell i at time t;

S_{it}: the supply of dockless bikes in land cell i at time t;

BE_i: the built environment factors at land cell i;

T_i: transportation-related factors at land cell i;

C_{it}: weather conditions (e.g., precipitation) in land cell i at time t;

O_t: other factors such as temporal factors at time t;
Ω: a spatial weights matrix;
ρ: a spatial lag coefficient;
e: an unobserved error term.

Alternatively, if there are omitted variables influencing the usage of dockless bikes across space, a spatial error model may fit, as

$$B_{it} = f(S_{it} + BE_i + T_i + C_{it} + O_t) + e, \text{ with } e = \lambda\Omega e + \mu \quad (2)$$

where the additional variables are:

λ: a spatial error coefficient;
μ: an unobserved error term.

The built environment and transportation-related variables were calculated as follows.

4.2. The impact of the built environment

A considerable body of literature investigates the linkages between the built environment and travel behavior (Ewing & Cervero, 2001, 2010). Land use and street configuration are commonly evaluated in cycling-related studies (Cervero, Sarmiento, Jacoby, Gomez, & Neiman, 2009; Faghih-Imani, Eluru, El-Geneidy, Rabbat, & Haq, 2014).

This study used the floor-area ratio (FAR) to indicate land use densities. Our FAR values were derived from a building database synthesizing a variety of datasets using the methodology proposed by Zhu and Ferreira (2015). We measured the land use density in four categories: public residential (i.e., HDB, named after the Housing and Development Board), private residential, commercial, and industrial areas. All the density measures were aggregated on grid level. The diversity (or mixture) of land use was also measured in our study through 40,782 point-of-interests (POIs) collected via the Google Place API. We reclassified the POIs into seven categories (see Table A1 in Appendix A.5). The Shannon entropy index was then calculated to represent the level of mixture of land use (Shannon, 1948) as

$$H = -\sum_i p_i \log_n p_i \quad (3)$$

where:

H: the value of entropy ranging between 0 and 1;
p_i: the percentage of the ith category of POI;
n: the number of categories.

4.3. The impact of transportation infrastructure

Transportation infrastructure and its accessibility also play a significant role in promoting cycling. Our model includes measures for the availability of designated cycling paths, accessibility to bike racks, connectivity of the road network, and distance to MRT stations.

To capture the availability of cycling paths, we obtained the length and location of the paths from the land transport authority (LTA) of Singapore, and those values were summed in each grid cell. Public bike racks in Singapore are usually located close to MRT stations. To quantify rack accessibility in each grid cell, we employed the classic accessibility measure proposed by Hansen (1959). In particular, we calculated the accessibility of bike racks in cell i as

$$A_i = \sum_j \frac{S_j}{D_{ij}} \quad (4)$$

where:

S_j: the supply of bike racks at land cell j;
D_{ij}: the distance from the centroid of land cell i to bike rack at cell j.

This formula uses road-network distance instead of Euclidean distance since it more precisely captures the effect of distance decay, especially when the landscape is extremely heterogeneous. Since our analysis was conducted on equal-area cells, we used the number of road intersections in each cell as a proxy for road connectivity and network design. More road intersections falling into a cell implies more connectivity and smaller street blocks. Figure 2 shows that dockless bikes were mostly distributed along the MRT network. We therefore assumed that cycling trips were more likely to occur around MRT stations. Thus, the distance to the nearest MRT station was also included in the model, calculated as the road-network distance from the centroid of each cell to nearest MRT station.

4.4. Descriptive statistics

For each hour from April 26th to May 4th, our data includes the dockless bike trips in each land cell that had available bikes. In all, there were over 460,000 observations, so to reduce the heavy computational burden in the statistic model, we took into account only the bike trips from 7 a.m. to 10 p.m. (the period when most trips occurred). We further aggregated data by partitioning the period into five three-hour groups. The data after 6 p.m. on May 4th are missing and therefore excluded from modeling. Descriptive statistics are shown in Table 2. The number of hourly bike trips, which is the dependent variable, is the average number of bike trips per hour in each land cell during the three-hour period. Similarly, the hourly available dockless bike fleet is the average number of dockless bikes in each land cell during the same period.

5. Results and analysis

A primary goal of this study was to investigate the impact of dockless bike fleet size on the usage of bikes. Toward this end, we tested whether the marginal utility of bike usage decreases when fleet size increases by regressing on two models: *Model (1)* uses the original bike fleet size as an explanatory variable, and *Model (2)* takes the square root of it.

5.1. Goodness-of-fit for the models

To test which spatial regression model has the better fit, we used the spatial dependence of our data based on the diagnostic process of (robust) Lagrange Multiplier tests (Anselin, 2005). The spatial weights matrix was defined as the inverse distance between 5-km neighborhoods. The modeling results are listed in Table 3, with

Table 2. Descriptive statistics.

Name of variables	Mean	Std. dev.	Minimum	Maximum
Hourly dockless bike trips	0.26	0.52	0.00	10.33
Hourly available dockless bike fleet	4.14	4.89	1.00	90.00
Length of cycling path (km)	0.02	0.09	0.00	0.93
Accessibility to bike racks	1.53	2.02	0.51	95.16
FAR of public residence	0.51	0.70	0.00	6.70
FAR of private residence	0.21	0.39	0.00	3.86
FAR of commercial building	0.10	0.32	0.00	4.56
FAR of industrial building	0.10	0.34	0.00	4.77
Shannon entropy	0.36	0.28	0.00	0.89
Number of road intersections	8.00	5.68	0.00	40.00
Distance to MRT station (km)	1.46	1.07	0.02	14.70
Number of bus stops	1.49	1.30	0.00	8.00
Distance to the CBD (km)	13.21	5.93	0.12	28.33
Accumulated precipitation (cm)	0.09	0.23	0.00	5.20
Average temperature (°C)	29.06	1.32	24.85	33.71

Note: Std. dev. = Standard deviation; FAR = Floor-area ratio.

different manipulations of the explanatory variables (i.e., the dockless bike fleet size or the square root of it) and spatial structures (i.e., a linear regression or spatial autoregressive model). The goodness-of-fit values of the four models are all presented.

The diagnostic results for spatial dependence suggest a spatial lag structure. The goodness-of-fit values of all models are over 0.2. Thus, considering the large sample size with over 100,000 observations, both models present a decent overall fit. The results also show that the inclusion of spatial structure captures more variations of the model (as seen in R^2 of Table 3). Additionally, in both models, the spatial pseudo R^2 values, which are a relatively less optimistic assessment of the model fit (Anselin & Rey, 2014), provide a slightly better overall fit than the corresponding linear regression models.

5.2. Fleet size

The coefficients of fleet size are positive and significant in all models. This implies that the usage of dockless bikes is strongly associated with the number of available bikes—i.e., on average, with more available bikes in the system, more trips will occur.

In the two models, *Model (1)* implies that the impact of dockless bike fleet size on the usage of bikes is linear—i.e., a given additional number of bikes put into the system at any point promotes a constant volume of new bike trips, while *Model (2)* assumes that the marginal return of additional bike trips declines with increasing fleet size, or more specifically, a larger fleet size is still associated with more bike trips, but for a given number of new bikes, the number of additional bike trips declines as the fleet size increases. As a result, *Model (2)*, which uses the square root of fleet size, fits better than *Model (1)*: In particular, when moving from *Model (1)* to *Model (2)*, the goodness-of-fit values (i.e., pseudo R^2) increase from 0.207 to 0.224 with the linear structure, and from 0.253 to 0.258 with the spatial lag structure. The superiority of *Model (2)* means that as more bikes are put into the city, the additional effect of each new bike on bike usage diminishes. Based on the specific results for *Model (2)* with spatial lag, increasing the number of available bikes per cell from none to 10 will induce an average of 0.7 trips per hour. However, when the fleet size grows from 100 to 110, the average may decrease to as low as 0.1.

In terms of the other variables, the modeling results are quite consistent. Thus, in the following analysis, we focus on *Model (2)*.

5.3. Built environment

Public residential density which is negatively associated with bike usage may imply that there are too many bikes allocated in HDB areas. Because of oversupply, the utilization rate is much lower than average. Some empirical evidence for this phenomenon is presented in Table A2 in Appendix A.6. In addition, public housing is usually located in high accessibility area that may counterbalance a potentially high demand of bike usage. This negative impact of public residential density is thus shown in our model.

The impact of private residential density on the usage of dockless bikes is positive but not significant in the model. The coefficients hint that, to some degree, the supply of bike-sharing services for residents living in private residential area is lower than its actual demand. In areas with high commercial land use density, the usage of dockless bikes is high as well, which might be due to the high occurrence of last-mile trips in commercial areas. In areas with high industrial land use density, the usage of dockless bikes is low. For land use mixture in our study, an entropy value of 1 means extreme diversity of land use—a perfectly even distribution of all seven categories of POIs in one land cell—whereas a value of 0 indicates the least diversity or extreme homogeneity of land use—only one category of POIs in the cell. Positive coefficients of entropy mean that a higher level of land use mixture is associated with more dockless bike usage, controlling for all other variables.

5.4. Other factors

The effect of road intersections is positive but only significant at 90% confident interval with spatial lag. Cycling facilities, such as longer cycling paths and more accessible bike racks for parking, positively influence the usage of dockless bikes. The usage of dockless bikes is most concentrated around MRT stations, which is validated by the coefficient of the distance to MRT station. This suggests that dockless bike-sharing programs may

Table 3. Modeling results.

	Model (1)		Model (2)	
	Linear	Spatial Lag	Linear	Spatial Lag
Constant	−0.005 (−0.79)	−0.072 (−11.40)*	−0.236 (−37.99)*	−0.272 (−42.99)*
Supply				
Available dockless bike fleet	0.039 (117.17)*	0.036 (110.39)*	n.a.	n.a.
Square root of dockless bike fleet	n.a.	n.a.	0.236 (127.22)*	0.225 (118.77)*
Cycling facilities				
Length of cycling path (in km)	0.310 (17.59)*	0.264 (15.37)*	0.295 (16.96)*	0.264 (15.42)*
Accessibility to bike racks ($\times 10^2$)	0.462 (6.30)*	0.383 (5.38)*	0.419 (5.78)*	0.366 (5.16)*
Density				
FAR of public residence	−0.005 (−1.86)•	−0.004 (−1.71)•	−0.033 (−12.48)*	−0.031 (−11.94)*
FAR of private residence	0.006 (1.48)	0.010 (2.59)†	0.001 (0.23)	0.004 (1.06)
FAR of commercial buildings	0.067 (13.88)*	0.062 (13.18)*	0.064 (13.23)*	0.060 (12.79)*
FAR of industry buildings	−0.036 (−8.15)*	−0.028 (−6.47)*	−0.027 (−6.07)*	−0.022 (−4.97)*
Diversity				
Shannon entropy	0.043 (7.59)*	0.049 (8.94)*	0.023 (4.14)*	0.029 (5.25)*
Design				
Number of road intersections ($\times 10^3$)	0.201 (0.66)	0.820 (2.79)†	0.046 (0.15)	0.493 (1.68)•
Access to public transportation				
Distance to MRT stations (in km)	−0.055 (−34.90)*	−0.037 (−22.32)*	−0.043 (−27.92)*	−0.031 (−19.25)*
Number of bus stops	0.009 (7.59)*	0.011 (9.29)*	0.005 (3.81)*	0.006 (5.17)*
Weather conditions				
Accumulated precipitation (in cm)	−0.122 (−18.71)*	−0.043 (−6.25)*	−0.121 (−18.73)*	−0.065 (− 9.47)*
Temperature over 31°C [(d)]	−0.022 (−3.57)*	−0.012 (−1.92)•	−0.024 (−3.90)*	−0.016 (−2.71)†
Temporal and other factors				
Unlimited free ride promotion [(d)]	0.100 (18.96)*	0.040 (7.89)*	0.099 (18.84)*	0.056 (12.00)*
Weekend and holiday [(d)]	0.081 (22.06)*	0.025 (5.42)*	0.080 (22.00)*	0.040 (6.35)*
Late afternoon (4 to 7 p.m.) [(d)]	0.086 (23.13)*	0.031 (4.47)*	0.086 (23.28)*	0.047 (49.05)*
Spatial factors				
Distance to the CBD (in km, $\times 10^2$)	0.708 (24.18)*	−0.024 (−0.62)	0.520 (17.79)*	0.013 (0.35)
Spatial lag coefficient	n.a.	0.617 (27.30)*	n.a.	0.438 (20.06)*
Summary of statistics				
Number of observations	100,122	100,122	100,122	100,122
(Pseudo) R^2	0.207	0.253	0.224	0.258
Adjusted (pseudo) R^2	0.207	n.a.	0.223	n.a.
Spatial pseudo R^2	n.a.	0.207	n.a.	0.224
Diagnostics for spatial dependence with p-value in parentheses				
Lagrange multiplier (lag)	20,575.84*	n.a.	17,033.79*	n.a.
Robust Lagrange multiplier (lag)	309.77*	n.a.	175.92*	n.a.
Lagrange multiplier (error)	29,128.17*	n.a.	27,341.36*	n.a.
Robust Lagrange multiplier (error)	8,862.11*	n.a.	10,483.49*	n.a.

Note: n.a. = not applicable; FAR = Floor-area ratio;
[(d)]dummy variable.
*< 0.001.
†< 0.01.
< 0.05.
•< 0.1.

facilitate last-mile connection to the MRT. In areas with more bus stops, the usage of bikes is also higher.

The coefficient of rainfall is consistent with the descriptive analysis presented above. Adverse weather conditions negatively influence the usage of bikes. The number of bike trips is greatly reduced in heavy rainfall. In addition, if the weather is too hot, i.e., with a temperature above 31°C, fewer people are likely to use bikes.

As for the other factors, the promotion of free rides on Labour Day led to more bike trips. More bike trips were also observed during the weekend[3] and in the late afternoon. The impact of the spatial related factor may be insignificant because the effects are already captured by the spatial lag structure, such as the distance to the CBD, in our model.

6. Discussion and conclusion

Dockless bike-sharing is an emerging shared mobility service. It revolutionizes traditional docked bike-sharing services. Docking stations are built earlier to regulate customers' behavior and facilitate payment. Now, the functions of docking stations have been integrated into a smartphone App with scanning QR code (or NFC) for payment and GPS sensors are embedded in bikes for fleet tracking and management.

Dockless bike service offers scholars unprecedented access to large-scale ridership data. It provides new opportunities for data analysis in investigating cycling behaviors. In docked bike-sharing, cycling trips are largely constrained by docking stations. Origins and destinations of cycling trips are biased toward locations around the stations. Thus, bike usage and movement pattern we observed from a docked system may not match real travel demand unless stations are deployed in a very high-density configuration, which however is not the case for

---

[3]There were also 30-minute free ride promotions during the weekend.

most of dock based systems around the world. Now, GPS tracking from dockless bikes can help capture more accurate trip records. And, as its fleet size is no longer limited by station capacity, it can be easily expanded into a large fleet size. This will serve more unmet travel demands and potentially improve our observation of human mobility. In addition, data availability of detailed GPS records can benefit many other parties. Operators can use these GPS records to assist in rebalancing. Users can benefit from knowing where bikes are available. Transportation planners, by aggregating cycling trips, can identify mismatches of cycling demand and infrastructure supply.

This work pioneers the study of dockless bike-sharing programs based on real-time GPS data. We investigate the usage of over 10,000 dockless bikes in Singapore during a nine-day period with more than 1.5 million observations. In the paper, we elaborate the two-step data mining methodology for GPS data, and we provide a visualization of the spatiotemporal distributions of the dockless bike usage in Singapore during the study period. We also relate bike usage to weather conditions, e.g. precipitation and temperature, during the same period.

The paper presents a spatial autoregressive model that we implemented to explore the impacts of fleet size, built environment, and weather conditions on the usage of dockless bike-sharing services. The results indicate that dockless bike fleet is positively associated with bike usage. However, as the fleet size grows, the marginal impact decreases, i.e., with larger fleet size, each new bike may induce fewer new trips. The built environment plays an import role as well. Higher land use density of commercial areas, more diverse economic activities, and smaller street blocks positively influence the usage of dockless bikes. Supportive cycling facilities and better transportation infrastructure also encourage bike usage. Rainfall and hot weather could, not surprisingly, reduce bike usage. Finally, the spatial lag coefficient of our model shows that dockless bike usage is spatially autocorrelated which means adjacent land cells of one cell with high bike usage also tend to exhibit high usage and vice versa.

Altogether, to effectively promote usage of dockless bikes, a straightforward solution is to make more dockless bikes available. And for dockless bike service, it is much easier to increase fleet size than traditional dock based system. This is what the bike-sharing operators have done in many cities, with a rapid fleet expansion aiding the emergence of the new service, although we suspect their main interests are to squeeze competitors out of the market rather than to promote usage. Nevertheless, such growth is not sustainable since public space and road resources are limited. We cannot keep introducing more bikes into the city. Besides, the utilization level per bike is low while the marginal benefit is decreasing as well.

Overall, this study advances our understanding of the new bike-sharing service, and the findings lead to practical guidance for transportation planning and urban management in the deployment of dockless bikes.

- The fleet size of dockless bike should be highly regulated, taking into the consideration of the capacity and management of urban public space. Oversupply of bike fleet may hurt its economic sustainability, occupy much public space, and cause visual pollution.
- The bicycle infrastructure needs to match with the demand of bike trips. More cycling facilities, including dedicated cycling paths and accessible bike racks, are needed. The potential locations of new cycling paths and bike racks can be based on an analysis of the spatial distribution of bike usage. Besides, owing to the tropical climate, Singapore rains frequently and sometimes heavily. The shelters, similar to the covered sidewalks, can also be built above the most heavily used cycling paths to facilitate cycling, which could also potentially benefit other active travel modes, e.g., walking, scooter (kick scooter or electronic version of it, not the motorcycle like Vespa).
- Rebalance strategy needs to consider temporal variation of bike usage, and the demand and supply for shared bikes in areas of different land uses, e.g., public/private residential and commercial areas. The system may likely reach equilibrium without rebalancing, but this equilibrium is likely to be far from optimal with lots of unmet demand.
- High bike usage is observed near MRT stations and bus stops, which might imply bike-sharing serves lots last-mile trips. Thus, integrating dockless bike service with public transportation, possibly through smart payment, shall be considered. It has a great potential to increase bike usage.
- The positive effect of a heterogeneous land use mixture indicates that planning urban areas with more diverse economic activities can also be beneficial.
- The impact of free ride promotions shows the "power of free." If a service improves the general social welfare, some financial support or subsidies from governing authorities can be considered, because it helps foster clean and energy-efficient travel mode and hence reduce the environmental impact of transportation.

The study can be further improved with more data from other sources. For instance, the filter of invalid GPS data is based on HITS-referenced thresholds due to the absence of GPS information during cycling. Some potential long-distance and long-time dockless bikes trips are thus filtered out. Cyclists' information is unknown in our study, which makes analyzing cycling behaviors of different demographics impossible. Solving these problems would require GPS and other data about cycling routes from bike-sharing operators or additional surveys of dockless bike riders. Also, it is the early stage of dockless bike-sharing service in Singapore, the customers' behavior might change as the service matures. And since the data used for this work only cover a 9-days period, it is possible that some extreme weather conditions are not captured in our observation. Future studies shall consider analyzing behavior changes with data collected over a longer period.

Past studies have looked into traditional bike-sharing programs by considering various topics from bike-sharing usage, barriers to bike-share programs, modal shifting between bike sharing and other transportation modes, bike fleet management, business models, product-service system design, etc. These areas and others provide several promising paths for future research on dockless bike sharing, particularly by using big data and information technologies, as in our study. Here, we list some potential future research directions as an extension of past work:

- *Comparison of bike-sharing system.* As dockless bike-sharing enters the market, it will co-exist with the old docking

systems for a while. Some studies have compared bike sharing systems in different cities (O'Brien et al., 2014; L. Zhang, Zhang, Duan, & Bryde, 2015). But these systems are mainly third-generation IT-based systems. More comparative studies could be made to compare dockless bike-sharing program with traditional dock based ones, and also private bike trips. The comparison can focus on usage rate, user motivation, user preferences, and demographics of bike users.

- *Barriers and facilitators to dockless bike-sharing.* Barriers and facilitators may change in the new type of bike-sharing service. Previous studies have identified barriers of docked bike-sharing (Fishman et al., 2014). Some of them, such as poor accessibility to docking stations and lengthy sign-up process, are considered as main barriers, but will no longer be evident in dockless bike-sharing. Re-identifying top barriers can help improve the usage of bike-sharing.

- *Synergy with public transportation.* Integrating bike sharing with public transportation has been shown to strengthen the benefits of both modes (Fishman, Washington, & Haworth, 2013). Ridership data of dockless bikes provide accurate origins and destinations of bike trips. Data-mining their collocation with metro stations or bus stops can deepen our understanding about to which extent it facilitates first- or last-mile connection to public transportation. This can be further enhanced with a comparative analysis of public transportation usage between the time periods before and after deployment of dockless bike service.

- *Mode substitution and impacts to other motorized travel modes.* Bike sharing is considered as a sustainable transport option that could help reduce car usage. Some studies find it improves transit connect and leads to more public transportation usage (Martin & Shaheen, 2014; S. Shaheen, Martin, & Cohen, 2013) while others suspect its efficiency of reducing motor vehicle usage if truck usage for rebalancing is counted (Fishman et al., 2014). How effective could the new dockless bike-sharing service improve transit usage and reduce car usage need future studies to evaluate.

- *Management and rebalancing.* There is also potential future research from a management point of view. Management issues include finding the optimal bike distribution, which is influenced by local supply and demand. But since docking stations are not required in new programs, the overall management, optimization, and rebalancing of free-floating bikes could be entirely different from the previous experience in bike fleet management with docking stations. For example, rebalancing of dock-based systems only need to consider pickup-and-delivery in stations (Dell'Amico, Hadjicostantinou, Iori, & Novellani, 2014). Dockless bikes can be parked anywhere that is legal to park, which potentially makes rebalancing a more complicated issue.

- *Business models.* Earlier studies (S. A. Shaheen et al., 2010; L. Zhang et al., 2015) have analyzed advantages and disadvantages of existing business models. As the new dockless bike-sharing services emerge, new models of provision may also continue to experience growth. Now, new bike-sharing companies are expanding very quickly and creating their own business models. The objectives of these dockless bike companies and urban transport authorities are not always aligned, so that whether or what regulation is necessary is a direction for future studies. For instance, when there are a number of bike-sharing operators competing in the market, tragedy of the commons could happen that will lead to an oversupply of dockless bikes. Hence, the behavior of operators should be regulated, and the design of appropriate regulation guidelines for dockless bike-sharing riders and operators should be discussed.

- *Safety and legislation.* Cycling also involves safety concerns that are revealed as a big barrier to people's decision-making in shifting to cycling trips (Fishman et al., 2014). How to improve road safety from a legislation perspective is, thus, worth investigation. In China and Singapore, helmets are not mandatory. In other areas, such as parts of the US, wearing helmet is compulsory when cycling, particularly for children. Such regulations may be a barrier to bike-sharing services as introduced in previous studies. Additional legislation to ensure the safety of dockless bike riders in other ways can also be explored along with the impact of these regulations on the usage of dockless bikes.

- *Urban design and planning.* Dockless bike service usually comes with large fleet size. As bikes could now be parked freely across the city, occupying much public space could potentially induce conflicts in urban management. Also, cyclists may need to compete for road space with motorized vehicles. These will trigger dialogues among users, operators, and legislators. The discussions about new urban design with better accommodation for bikes and more cycling paths are expected to be made. In addition, studies on regulation of fleet size and spatial configuration of parking space with dockless bike sharing should be conducted as well.

Acknowledgement

ORCID

Yu Shen http://orcid.org/0000-0002-0270-578X
Xiaohu Zhang http://orcid.org/0000-0001-5654-1445
Jinhua Zhao http://orcid.org/0000-0002-1929-7583

References

Anselin, L. (2005). *Exploring spatial data with GeoDa: A workbook. Spatial Analysis Laboratory.* Urbana-Campaign, Urbana, IL: University of Illinois.

Anselin, L., & Rey, S. J. (2014). *Modern spatial econometrics in practice: A guide to GeoDa, GeoDaSpace and PySAL.* Urbana, IL: GeoDa Press LLC.

Cervero, R., Sarmiento, O. L., Jacoby, E., Gomez, L. F., & Neiman, A. (2009). Influences of built environments on walking and cycling: Lessons from bogotá. *International Journal of Sustainable Transportation, 3*(4), 203–226. https://doi.org/10.1080/15568310802178314

Corcoran, J., Li, T., Rohde, D., Charles-Edwards, E., & Mateo-Babiano, D. (2014). Spatio-temporal patterns of a Public Bicycle Sharing Program: The effect of weather and calendar events. *Journal of Transport Geography*, *41*, 292–305. https://doi.org/10.1016/j.jtrangeo.2014.09.003

Dell'Amico, M., Hadjicostantinou, E., Iori, M., & Novellani, S. (2014). The bike sharing rebalancing problem: Mathematical formulations and benchmark instances. *Omega*, *45*, 7–19. https://doi.org/10.1016/j.omega.2013.12.001

DeMaio, P. (2009). Bike-sharing: History, impacts, models of provision, and future. *Journal of Public Transportation*, *12*(4), 41–56. https://doi.org/https://doi.org/10.5038/2375-0901.12.4.3

Ewing, R., & Cervero, R. (2001). Travel and the built environment: A synthesis. *Transportation Research Record: Journal of the Transportation Research Board*, *1780*(1), 87–114. https://doi.org/10.3141/1780-10

Ewing, R., & Cervero, R. (2010). Travel and the built environment. *Journal of the American Planning Association*, *76*(3), 265–294. https://doi.org/10.1080/01944361003766766

Faghih-Imani, A., Eluru, N., El-Geneidy, A. M., Rabbat, M., & Haq, U. (2014). How land-use and urban form impact bicycle flows: Evidence from the bicycle-sharing system (BIXI) in Montreal. *Journal of Transport Geography*, *41*, 306–314. https://doi.org/10.1016/j.jtrangeo.2014.01.013

Fishman, E., Washington, S., & Haworth, N. (2013). Bike Share: A Synthesis of the Literature. *Transport Reviews*, *33*(2), 148–165. https://doi.org/10.1080/01441647.2013.775612

Fishman, E., Washington, S., & Haworth, N. (2014). Bike share's impact on car use: Evidence from the United States, Great Britain, and Australia. *Transportation Research Part D: Transport and Environment*, *31*, 13–20. https://doi.org/10.101.6/j.trd.2014.05.013

Fishman, E., Washington, S., Haworth, N., & Mazzei, A. (2014). Barriers to bikesharing: An analysis from Melbourne and Brisbane. *Journal of Transport Geography*, *41*, 325–337. https://doi.org/10.1016/j.jtrangeo.2014.08.005

Hansen, W. G. (1959). How accessibility shapes land use. *Journal of the American Institute of Planners*, *25*(2), 73–76. https://doi.org/10.1080/01944365908978307

iiMedia Research. (2017). 2017Q1 China Renting Bicycle Market Research Report (in Chinese).

Martin, E. W., & Shaheen, S. A. (2014). Evaluating public transit modal shift dynamics in response to bikesharing: A tale of two U.S. cities. *Journal of Transport Geography*, *41*, 315–324. https://doi.org/10.1016/j.jtrangeo.2014.06.026

Meng, M., Zhang, J., Wong, Y. D., & Au, P. H. (2016). Effect of weather conditions and weather forecast on cycling travel behavior in Singapore. *International Journal of Sustainable Transportation*, *10*(9), 773–780. https://doi.org/10.1080/15568318.2016.1149646

Nextbike GMBH. (2016). Nextbike Company Profile.

O'Brien, O., Cheshire, J., & Batty, M. (2014). Mining bicycle sharing data for generating insights into sustainable transport systems. *Journal of Transport Geography*, *34*, 262–273. https://doi.org/10.1016/j.jtrangeo.2013.06.007

Parkes, S. D., Marsden, G., Shaheen, S. A., & Cohen, A. P. (2013). Understanding the diffusion of public bikesharing systems: Evidence from Europe and North America. *Journal of Transport Geography*, *31*, 94–103. https://doi.org/10.1016/j.jtrangeo.2013.06.003

Shaheen, S. A., Guzman, S., & Zhang, H. (2010). Bikesharing in Europe, the Americas, and Asia: Past, present, and future. *Transportation Research Record: Journal of the Transportation Research Board*, *2143*, 159–167. https://doi.org/10.3141/2143-20

Shaheen, S., Martin, E., & Cohen, A. (2013). Public bikesharing and modal shift behavior: A comparative study of early bikesharing systems in North America. *International Journal of Transportation*, *1*, https://doi.org/10.14257/ijt.2013.1.1.03

Shannon, C. E. (1948). A Mathematical Theory of Communication. *Bell System Technical Journal*, *27*(3), 379–423. https://doi.org/10.1002/j.1538-7305.1948.tb01338.x

Wojan, T. R., & Hamrick, K. S. (2015). Can walking or biking to work really make a difference? Compact development, observed commuter choice and body mass index. *PLoS ONE*, *10*(7), e0130903. https://doi.org/10.1371/journal.pone.0130903

Zhang, H., Shaheen, S. A., & Chen, X. (2014). Bicycle evolution in China: From the 1900s to the Present. *International Journal of Sustainable Transportation*, *8*(5), 317–335. https://doi.org/10.1080/15568318.2012.699999

Zhang, L., Zhang, J., Duan, Z. Y., & Bryde, D. (2015). Sustainable bike-sharing systems: Characteristics and commonalities across cases in urban China. *Journal of Cleaner Production*, *97*, 124–133. https://doi.org/10.1016/j.jclepro.2014.04.006

Zhu, Y., & Ferreira, J. (2015). Data integration to create Large-scale spatially detailed synthetic populations. In S. Geertman, J. Ferreira, Jr., R. Goodspeed, & J. Stillwell (Eds.), *Planning Support Systems and Smart Cities* (1st ed., pp. 121–141). Springer International Publishing. https://doi.org/10.1007/978-3-319-18368-8_7

Appendix A. Supplementary materials

A.1. Pseudocode

```
# Step 1. Data Cleaning of bike records (unmoved, GPS drifting)
input:
    bikes ← a list of (bike)
    bike ← a list of (i, timestamp, location)
    is_removable ← a list of flag
for each bike in all bikes:
    for each timestamp i:
        if location i == location i+2 (unmoved or GPS drifting):
            label i+1 in is_removable
    clear records labeled in is_removable
return:
    bikes after cleaning
# Step 2. Filter and calculation of bike utilization
input:
    bikes_od ← a list of (bike_od)
    bike_od ← a list of OD pairs (origin i, destination i)
    total_count ← the number of all bikes
for each bike_od in all bikes_od:
    total_time ← 0
    moved_time ← 0
    for each OD pair i:
        calculate great-circle_distance between origin and destination
        calculate time_elapse between origin and destination
        calculate speed based great-circle_distance and time_elapse
        total_time + = time_elapse
        if great-circle_distance is within a certain range …
            and time_elapse is within a certain range …
            and speed is smaller than a predefined threshold:
            moved_count ++
            moved_time + = time_elapse
    utilization + = moved_time / total_time
return:
    trips made per bike
    utilization per bike
```

A.2. Statistics of bike trips from HITS

Figure A1 shows the distance distribution of cycling trips from Household Interview Travel Survey (HITS) of 2012. HITS is a comprehensive transportation demand survey in Singapore that includes about 10,000 households. The trip distance was calculated based on Google Map Distance API. A box plot of all distances is presented in Figure A1 (1), revealing the majority of trips fall into a distance range between 150 meters to 7.5 kilometers. Further, Figure A1 (2) shows a distance distribution in this specific range.

A.3. Dockless bike usage

Figure A2 shows the usage of dockless bikes from different angles. Number of available bikes each hour and circadian usage pattern is presented in Figure A2 (1.1). The histogram of

number of average trips per bike per day is shown in Figure A2 (1.2), from which we can conclude that most bikes are used less than 2 times per day. Figure A2 (2.1) and (2.2) display the histograms of average travel distance and of average travel time respectively.

A.4. Last mile travel by bus in Singapore

Due to the restricted control on auto ownership, people in Singapore largely rely on public transit, i.e., bus and MRT. The last-mile travel demand is quantified as the number of bus trips connecting to the MRT stations, including that of bus trips before accessing to and after egressing from the MRT stations. The travel demand is calculated based on the public transit smart card data in August 2013 (the whole month) provided by the LTA, with over 175 million records. The results are aggregated from bus stops to the 300×300-meter land cells, plotted in Figure A3.

A.5. Classification of POIs

Table A1. Classification of POIs.

Number	Category	Examples
1	Company and small business	Accounting services, banks, health service, travel agency, laundry
2	Government, organizations, institutions	Church, city hall, embassy, museum, police, post office, university, school
3	Entertainments	Art gallery, bar, beauty salon, casino, gym, hair care, movie theater, spa
4	Hotels	Lodging, hotels
5	Retail	Bakery, book store, department store, gas station, supermarket, liquor store, shopping mall
6	Restaurant	Café, food, restaurant
7	Transportations	Airport, bus stations, subway stations, taxi stand, train stations

A.6. Usage of dockless bikes in different land use

Table A2. Usage of dockless bikes in different land use.

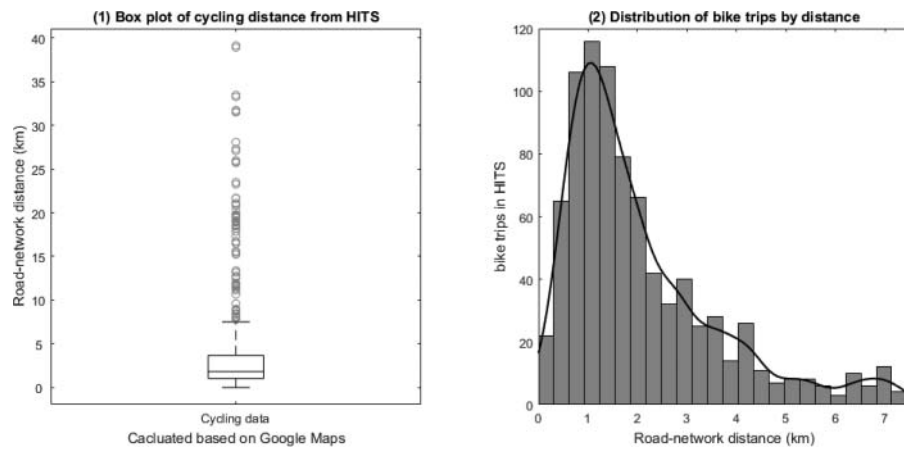|	Average bike trips	Average fleet size	Average usage ratio
Public residence land use	0.34	5.50	6.12%
Private residence land use	0.23	3.47	6.61%
Commercial land use	0.26	3.66	7.12%
Industrial land use	0.17	3.30	5.26%

Figure A1. Bike trips from HITS.

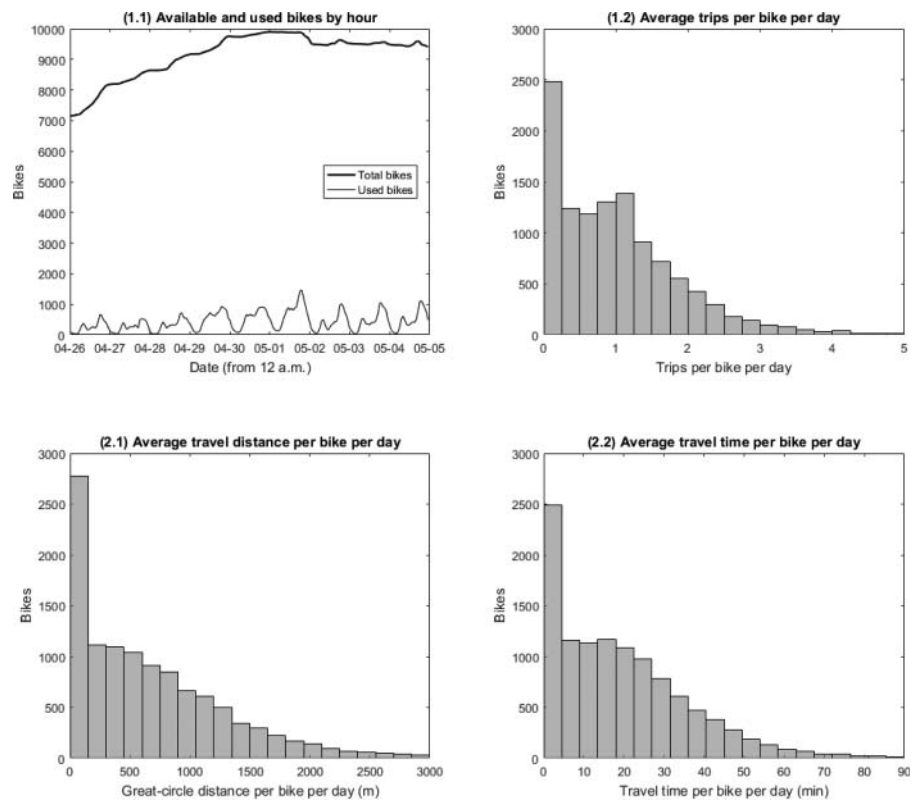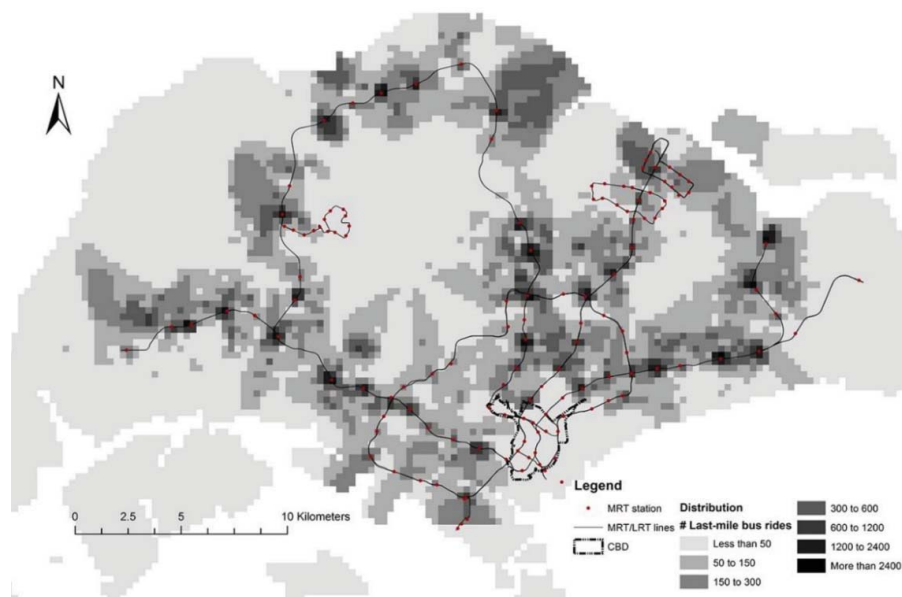**Figure A2.** Distribution of bike usage.

Figure A3. Distribution of daily last-mile bus travel demand.