

Article

Station-Free Bike Rebalancing Analysis: Scale, Modeling, and Computational Challenges

Xueting Jin and Daoqin Tong *

School of Geographical Sciences and Urban Planning, Arizona State University, 975 S Myrtle Ave, Tempe, AZ 85281, USA; xjin50@asu.edu

* Correspondence: daoqin.tong@asu.edu

Received: 17 September 2020; Accepted: 18 November 2020; Published: 19 November 2020



Abstract: In the past few years, station-free bike sharing systems (SFBSSs) have been adopted in many cities worldwide. Different from conventional station-based bike sharing systems (SBBSSs) that rely upon fixed bike stations, SFBSSs allow users the flexibility to locate a bike nearby and park it at any appropriate site after use. With no fixed bike stations, the spatial extent/scale used to evaluate bike shortage/surplus in an SFBSS has been rather arbitrary in existing studies. On the one hand, a balanced status using large areas may contain multiple local bike shortage/surplus sites, leading to a less effective rebalancing design. On the other hand, an imbalance evaluation conducted in small areas may not be meaningful or necessary, while significantly increasing the computational complexity. In this study, we examine the impacts of analysis scale on the SFBSS imbalance evaluation and the associated rebalancing design. In particular, we develop a spatial optimization model to strategically optimize bike rebalancing in an SFBSS. We also propose a region decomposition method to solve large-sized bike rebalancing problems that are constructed based on fine analysis scales. We apply the approach to study the SFBSS in downtown Beijing. The empirical study shows that imbalance evaluation results and optimal rebalancing design can vary substantially with analysis scale. According to the optimal rebalancing results, bike repositioning tends to take place among neighboring areas. Based on the empirical study, we would recommend 800 m and 100/200 m as the suitable scale for designing operator-based and user-based rebalancing plans, respectively. Computational results show that the region decomposition method can be used to solve problems that cannot be handled by existing commercial optimization software. This study provides important insights into effective bike-share rebalancing strategies and urban bike transportation planning.

Keywords: station-free bike sharing system; rebalance; scale; optimization

1. Introduction

Bike sharing systems (BSSs) are an important component in today's urban transportation system [1,2]. As a more sustainable mode of transportation, bike-sharing has the potential to reduce car usage, solve the first/last mile problem, and contribute to the local retail sales [3–5]. The history of bike sharing can be traced back to the 1960s when the concept was first introduced in Amsterdam, Netherlands [6]. Over the past decades, bike sharing systems have evolved through multiple generations. Early generations involved no prior registration and failed, as many bikes were vandalized or turned into private use [6,7]. Starting from 1995, prior registration has been required. A registered customer can rent and return a bicycle at a number of fixed bike stations. These systems are also known as station-based bike sharing systems (SBBSSs) [8] and have been successfully deployed in multiple cities around the world. Leveraged by the recent advanced technologies such as smart phones, GPS, and integrated payment systems, the newest generation of bike sharing involves having no fixed bike stations. Without docking stations, the station-free bike sharing systems (SFBSSs) provide users the

flexibility to use a smart phone app to locate a bike nearby and park it at any appropriate place after use. Due to the high flexibility and convenience, SFBSSs have gained considerable popularity over the past few years and have been widely adopted in many countries, including China, United Kingdom, Singapore, the United States, and the Netherlands [9]. It was estimated that in 2018, 16 to 18 million station-free bikes were in use worldwide, compared to 3.7 million station-based bikes [10].

A balanced status refers to the situation when bike supply meets bike demand. Both station-free and station-based bike sharing systems frequently suffer from the spatiotemporal fluctuations of demand, leading to the imbalance problem. Without timely rebalancing efforts, the imbalance problem may substantially degrade the system performance. For an SBBSS, the imbalance problem may lead to zero inventories at some stations and no space for parking at others. As a result, rentals and returns may be only possible at a limited number of stations, leaving many areas/demands un- or under-served [11]. For an SFBSS, the imbalance problem is more serious. While unmet demands remain, no restriction on where a user can pick up or park a bike may lead to bikes piling up and blocking city sidewalks.

Many studies have been conducted to assess the imbalance status of an SBBSS and design the associated rebalancing strategy [12–16]. While the imbalance evaluation in an SBBSS is rather straightforward [17,18], due to the lack of fixed stations in an SFBSS, there is no clearly defined spatial scale or geographical areas to evaluate bike inventories, assess the imbalance status, and perform the subsequent rebalancing analysis. Existing studies have explored various analysis scales, including traffic analysis zones and regular grids of various sizes [19–22]. The selection of these analysis scales has been rather arbitrary. There has been no discussion on the suitable scale needed to evaluate and address the imbalance issue in an SFBSS. On the one hand, a balanced status evaluated using large areas (e.g., 10×10 km grids) may contain multiple local bike shortage/surplus sites, which can significantly compromise the system performance. On the other hand, analysis conducted using small areas (e.g., 10×10 m grids) may greatly add to computational burden. Therefore, analysis scale, modeling/analysis, and computational requirements are related, and selection of the appropriate analysis scale is essential for the imbalance assessment of an SFBSS and the following rebalancing analysis.

2. Literature Review

Recently, a number of studies have examined bike sharing systems by focusing on two areas [23–26]. The first area concerns identifying the factors affecting bike sharing demand, and these factors include socio-economic status, built environment, traffic infrastructure, air quality, and weather [27–29]. The second area focuses on the imbalance issue of a bike sharing system and the associated rebalancing strategies. In this study, we are mainly interested in the latter.

To solve the imbalance issue in a bike-sharing system, various rebalancing strategies have been developed. Table 1 provides a summary of the recent literature on these rebalancing strategies. In particular, user-based and operator-based methods are the two major types of rebalancing strategies [8,12,30]. In a user-based rebalancing strategy, incentives in the form of a bonus voucher or discount are often used to encourage users to move bikes from surplus areas to shortage areas. For example, in 2008, the bike sharing system Vélib' in Paris launched a discount pricing strategy [31] to motivate users to return bikes to uphill stations. A pricing strategy was developed by Chemla et al. [32] to incentivize users to return bikes to the least loaded stations nearby, and a dynamic online pricing incentive strategy was proposed by Pfrommer et al. [33] to motivate users to choose alternative locations to pick up or return bikes. However, depending on users' participation, user-based rebalancing may not be sufficient to achieve the system level self-rebalancing [30]. Hence, user-based rebalancing strategies are often used as a supplement to operator-based ones. These two types of strategies can be combined to help reduce rebalancing costs [33]. Currently, most existing bike-sharing systems (e.g., Mobike and Ofo in China) employ operator-based rebalancing strategies [8].

Table 1. Summary of literature on bike rebalancing strategies.

Research Article	Bike Sharing System	Type of Rebalancing Strategy	Objective	Study Unit	Number of Units
Chemla et al. [12]	Station-based	Operator-based	Complete rebalance	Station	100
Raviv et al. [34]	Station-based	Operator-based	Minimize (a) total unmet demand, (b) total operating cost	Station	60
Ho and Szeto [35]	Station-based	Operator-based	Minimize total penalties	Station	400
Alvarez-Valdes et al. [36]	Station-based	Operator-based	Minimize unsatisfied demand	Station	28
Gaspero et al. [37]	Station-based	Operator-based	Maximize expected future bike demands	Station	92
Pal and Zhang [8]	Station-based	Operator-based	Minimize the operation time of rebalancing vehicle fleets	Station	476
Schuijbroek et al. [38]	Station-based	Operator-based	Minimize (a) the deviation from a given target inventory level, (b) the number of (un)loading operations, (c) total work time	Station	135
Dell’Amico et al. [39]	Station-based	Operator-based	Minimize total cost	Station	116
Chemla et al. [32]	Station-based	User-based	Minimize imbalance amount	Station	20 to 250
Pfrommer et al. [40]	Station-based	Operator-based and User-based	Minimize unsatisfied demand	Station	354
Pan et al. [19]	Station-free	User-based	Minimize cost	Grid/square	–
Ji et al. [41]	Station-free	User-based	Examine the influences of acceptable walking distance and user’s cooperative factor	Grid/hexagon with side length of 100 m	1185
Zhai et al. [21]	Station-free	Operator-based	Minimize travel cost	Grid/square	20 to 1100

A few studies have focused on developing efficient operator-based rebalancing methods. In an operator-based rebalancing operation, a fleet of vehicles are often sent to move bikes from site to site. Among others, spatial optimization methods [42] have been widely used to determine the number of relocation vehicles and the associated routes for performing the rebalancing operation [8,34,38,39]. All the studies included in Table 1 are based on optimization methods except for Ji et al. [41]. These problems have often been formulated using mixed integer programming (MIP) [34,38]. Some of these studies integrated GIS into classic location models, including the p -median problem and the coverage location problems, to achieve the optimal rebalancing design [8,18]. A range of objectives/goals have been examined in the bike rebalancing design. These objectives include minimizing costs, varying from travel costs to the overall rebalancing costs (including travel, trucks, loading and unloading costs), and maximizing the system level balance evaluated using different metrics (e.g., minimal absolute deviation from the target number of bikes) (also see Table 1). While some studies focused on a single objective [35,43], a few studies considered multiple goals simultaneously [44,45].

As shown in Table 1, existing bike rebalancing models are mainly station-based and have been constructed to address the imbalance issue of an SBBSS. In an SBBSS, stations are the analysis units used to perform the imbalance evaluation and the subsequent rebalancing analysis. In an SFBSS, if we divide a region into a set of sub-areas and treat each sub-area as a “station”, we can apply a station-based rebalancing model to solve the imbalance issue in an SFBSS. In fact, traffic analysis zones [20] and regular grids [19,21,41] have already been used to conduct the imbalance assessment for SFBSSs. Although traffic analysis zones (TAZs) are widely used in transportation planning, they are delineated largely based on motorized traffic and are generally too coarse to capture the variation of bike supply and demand given the short distance that bike users are normally willing to walk to access bikes. As for the grid-based approaches, the selection of grid size has been either arbitrary or unspecified [19,21,41]. To the best of our knowledge, there have been no studies exploring the impacts of analysis scale on the imbalance evaluation and the associated rebalancing design in an SFBSS. Scale can be critical in an SFBSS study: A balanced status at a large scale may contain multiple local bike shortage/surplus sites, and as a result, the rebalancing strategy designed based on a large scale

might not solve the imbalance issue completely. In contrast, analysis conducted using fine scales may encounter computational challenges, and a scale that is too fine (such as 1 m) might not be necessary or even meaningful.

Table 1 also summarizes the number of analysis units (i.e., sub-areas/grids in an SFBSS or stations in an SBBSS) used in current bike rebalancing studies, ranging from 28 to 1185. Since a rebalancing problem can be computationally intractable when a large number of analysis units are involved, many studies focused on small cities/areas or used a coarse scale [19,21]. For example, the model proposed by Chemla et al. [32] was unable to simulate the user-based rebalancing process for more than 250 stations because of the large problem size. Focusing on small areas or using coarse scale analysis units may provide limited insights into effective rebalancing strategies for a large urban area.

Therefore, identifying the appropriate analysis scale and designing efficient and effective bike rebalancing strategies are needed for SFBSSs in large cities. In this study, we aim to fill in the research gaps by focusing on two important questions: (1) how scale impacts the imbalance evaluation of an SFBSS and the associated rebalancing design, and (2) how to deal with computational challenges for large sized rebalancing problems. We develop a spatial optimization model to strategically optimize bike rebalancing efforts. We also propose a region decomposition method to solve large-sized problems that are constructed based on fine analysis scales. We apply the approach to study the SFBSS in downtown Beijing. Based on the empirical study, we discuss the strengths and weaknesses associated with alternative analysis scales and make recommendations on scale choice for different rebalancing strategies.

3. Methodology

3.1. Study Area and Data

Our study area consists of six districts in central Beijing: Dongcheng, Xicheng, Chaoyang, Haidian, Fengtai, and Shijingshan. Beijing, the capital of China, is among the few cities that first adopted SFBSSs in 2015. According to the Beijing Municipal Commission of Transport, in the second half of 2018, the average daily shared bike usage in Beijing was 1.3 million with the majority clustered in our study area. We use a trip level bike usage dataset in this study. The dataset was obtained from Mobike, a major station-free bike sharing company in China. The dataset contains the information on order ID, bike ID, user ID, start time, start location, and end location. We use the bike trip data on Wednesday, 10 May 2017 as a typical weekday to illustrate our method. We assume a daily rebalancing cycle by evaluating the bike pickups and returns in an area within a day.

3.2. Alternative Scales and Data Preprocessing

The whole study area is divided into regular grids. Each grid is regarded as a bike “station”, where bike pickups and drop-offs are summarized to evaluate the imbalance status. In this study, we examine seven alternative scales (2 km, 1.5 km, 1 km, 800 m, 400 m, 200 m, 100 m). In particular, the scale of 100 m was used in Ji et al. [41] for designing an user-based rebalancing strategy for an SFBSS; 400 m is the widely adopted walkable distance [46]; 800 m is the median distance of walking trips according to Yang and Diez-Roux’s survey [47]; 1 km² was the average area of the study units (traffic analysis zones) that Xu et al. [20] used to assess the imbalance status of an SFBSS; 2 km is the maximum distance people are willing to walk based on the general planning guidance in the US and UK [48]. Additionally, 200 m and 1.5 km are included in the set as intermediate scales to examine the scale impacts. Figure 1 shows an example of partitioning the study area into 1 × 1 km grids, while preserving the Jiedao boundary. Jiedao is the minimum census unit in China and is largely divided based on natural barriers or manmade structures, including rivers, railways, and main streets/highways. Therefore, Jiedao boundaries may have an impact on bike-share usage. In addition, by preserving the Jiedao boundaries, one can easily link the grids with census data. In this study, we also identified and excluded areas where no shared bikes are allowed, including mountain areas,

nature reserve areas, and parks. Within each grid, we count the bike pickups and returns based on the bike start and end locations in a day. Then, bike shortage/surplus is computed by comparing the pickup and return amounts: a bike shortage/surplus area is an area where bike pickups are more/less than bike returns. In this study, we also introduce a shortage/surplus lower bound and upper bound to allow an SFBSS to deviate slightly from the perfect balance status.

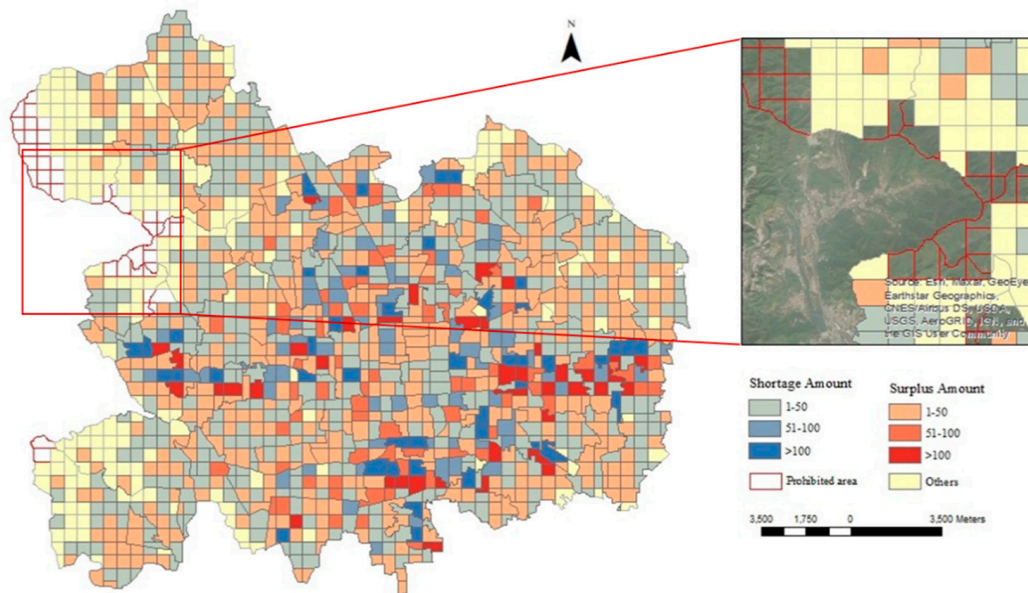


Figure 1. An example of partitioning the study area into regular grids (1 km \times 1 km).

3.3. Bike-Share Rebalancing Flow Model

In this study, we introduce a bike-share rebalancing flow model to strategically optimize bike rebalancing efforts. When addressing the imbalance issue, the model aims to identify the optimal bike repositioning with the minimal overall bike repositioning distance. In the model, we do not confine ourselves to a specific rebalancing strategy, user-based or operator-based. Therefore, the goal of this model is neither designing the optimal route for vehicles to redistribute bikes nor developing a pricing strategy to incentivize users to engage in the rebalancing efforts. This strategic analysis will help design the specific rebalancing strategies needed to solve the imbalance issue in an SFBSS.

In the model, we set a lower bound L and an upper bound U on the rebalancing ratio to allow an SFBSS to deviate slightly from the complete balance status. For an area with a certain amount of bike shortage/surplus, a complete rebalancing involves shipping in/out bikes that have the same amount of shortage/surplus. In this model, L and U specify the minimum and maximum percentage of imbalanced bikes that need to be shipped in/out to achieve a balance status. If we set $L = 0.9$ and $U = 1.1$, for an area with surplus imbalance of 100, the number of bikes that need to be shipped out during the rebalancing process should be more than or equal to 90 but cannot exceed 110. For a complete rebalancing operation, one can set L and U to be 1. Consider the following notation (Table 2):

Table 2. Notation summary.

	Notation	Description
Parameter	L	Lower bound of the rebalancing ratio
	U	Upper bound of the rebalancing ratio
	P	Set of areas with surplus imbalance
	N	Set of areas with shortage imbalance
	Z	Set of areas prohibiting shared bikes
	i, j	Index of areas
	A_i	Imbalance amount of area i
Variable	d_{ij}	Distance between area i and area j
	X_{ij}	Number of bikes relocated from area i to area j

The bike-share rebalancing flow model is formulated as:

$$\text{Min} \quad \sum_i \sum_j X_{ij} * d_{ij} \quad (1)$$

S.T.

$$\sum_j X_{ij} - \sum_j X_{ji} \geq L * A_i \quad \forall i \in P \quad (2)$$

$$\sum_j X_{ij} - \sum_j X_{ji} \leq U * A_i \quad \forall i \in P \quad (3)$$

$$\sum_j X_{ji} - \sum_j X_{ij} \geq L * A_i \quad \forall i \in N \quad (4)$$

$$\sum_j X_{ji} - \sum_j X_{ij} \leq U * A_i \quad \forall i \in N \quad (5)$$

$$X_{ij} = 0 \quad \forall i, j \in Z \quad (6)$$

$$X_{ij} \in \text{non-negative Integers} \quad \forall i, j \quad (7)$$

In the objective function (1), we aim to minimize the overall travel distance of repositioned bikes. Constraints (2) and (3) ensure that for all areas with bike surplus the net rebalancing outflow is more than the lower bound but does not exceed the upper bound. Constraints (4) and (5) specify that for all areas with bike shortage, the net rebalancing inflow is more than the lower bound but does not exceed the upper bound. Constraints (6) ensure no rebalancing flows for areas prohibiting shared bikes. Constraints (7) impose a positive integer restriction on the amount of bikes to be repositioned. In our experimental study, we use the rectilinear distance metric to evaluate the overall travel distance:

$$d_{ij} = |a_i - a_j| + |b_i - b_j| \quad (8)$$

where, a_i is the x-coordinate of the centroid of area i , b_i is the y-coordinate of the centroid of area i . In many urban areas, high-density roads and intersections make travel distance close to the rectilinear distance [49]. In addition, roads in many parts of downtown Beijing show the grid pattern, which makes rectilinear distance a reasonable metric for measuring travel distance [50].

3.4. Region Decomposition Approach

As discussed earlier, rebalancing models constructed using fine scales are generally challenging to solve. Consider that people often use shared bikes to travel a short distance. For example, bike-sharing trips in China are 1.5 km on average (<http://global.chinadaily.com.cn/>). For a large area, there may exist multiple sub-regions where bike-sharing trips mainly stay within a single sub-region. Given this, we develop a heuristic by decomposing a large region into multiple self-contained sub-regions to reduce the problem size. Figure 2 shows the workflow of the heuristic. First, the entire

region is divided into several sub-regions at a relatively coarse scale. Then the imbalance status for each sub-region is evaluated. If the sub-region meets the balance criterion, the sub-region is considered as a self-contained sub-region. If there exists a sub-region that is not self-contained, an alternative region division is needed until all the resultant sub-regions are self-contained. Then, each sub-region is further divided into finer grids (i.e., analysis scale), and based on the finer scale, the imbalance status is evaluated. Finally, the bike-sharing rebalancing flow model is implemented to obtain the optimal solution for each sub-region.

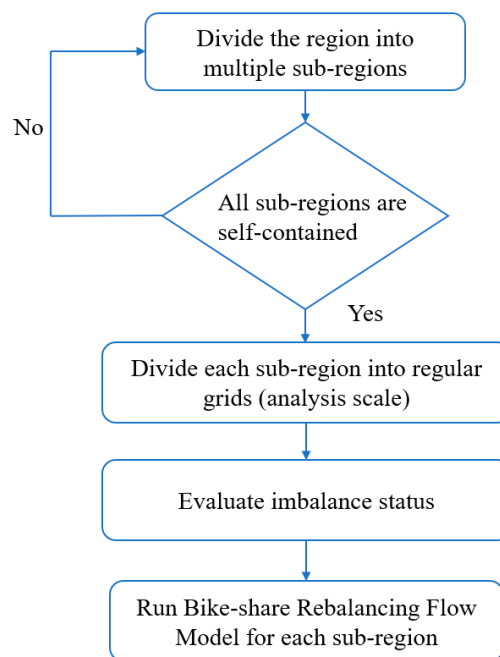


Figure 2. Workflow of the region decomposition approach.

In our case study, the computational experiments are carried out on a workstation powered by an Intel Core i7 CPU with a total RAM of 16 GB running a 64-bit operating system. Problem instances with analysis scales of 800 m and coarser are solved directly using the commercial optimization software CPLEX. Problems with finer scales are too large to be solved directly using CPLEX. For these problems, the region decomposition heuristic is applied. In the heuristic, the entire study area is divided into 14 sub-regions using the 10×10 km grids, while preserving the district boundaries (see Figure 3). These sub-regions have an average area of 98 km^2 . To examine whether these sub-regions are self-contained, we calculate the ratio of the shortage/surplus amount to the daily bike usage. If the absolute value of the ratio is less than 10%, the sub-region is considered self-contained. Then, each self-contained sub-region is further partitioned based on the analysis scale (e.g., 100, 200, and 400 m). Figure 3 shows the self-contained sub-regions derived in the case study along with the surplus/shortage evaluation conducted at the 200 m scale.

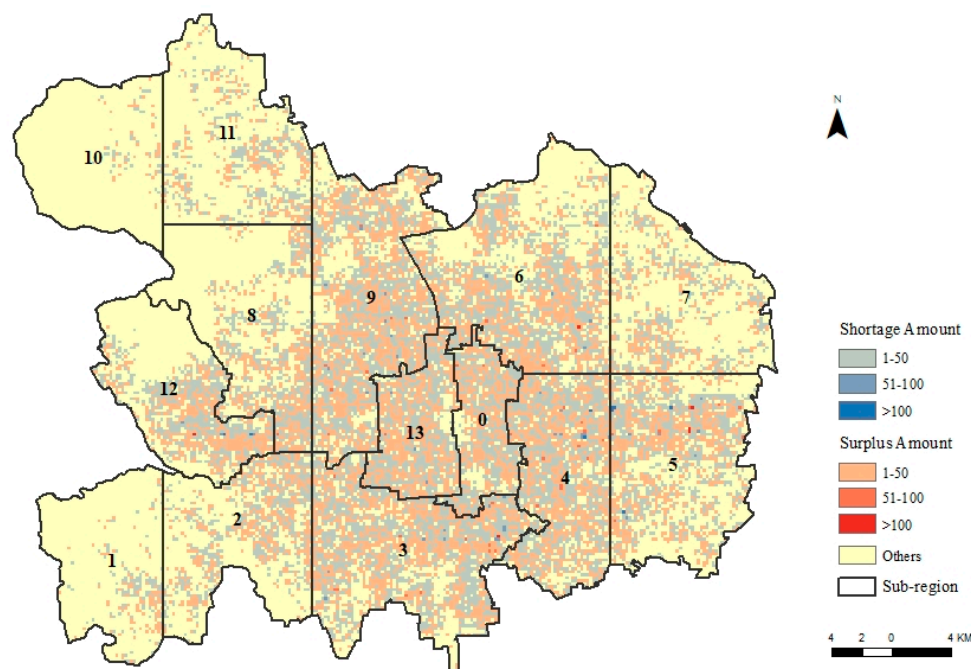


Figure 3. An example of sub-region division.

4. Results

4.1. Imbalance Assessment Across Scales

Table 3 provides a summary of the bike imbalance analysis in downtown Beijing. In addition to the seven scales discussed previously, we include two coarse scales, 10 km and 5 km, to provide the imbalance assessment at extreme large scales. With the grid size decreasing from 10 km to 100 m, the number of grids increases substantially. Columns “Maximum Surplus” and “Maximum Shortage” record the maximum bike surplus and shortage at each scale, respectively. Table 3 shows that while the maximum surplus and shortage appear to be relatively consistent across scale, the total imbalance increases substantially when fine grids are used. For example, when scale changes from 5 km to 200 m, the maximum surplus and shortage remain similar, but the total imbalance increases more than 17 times. This suggests that an imbalance assessment conducted at a large scale may dismiss many local imbalanced sites. Column “Imbalanced Grids” reports the total amount of grids with the surplus/shortage more than the 10% of the daily usage. Column “Percentage of Imbalanced Grids” computes the percentage of the imbalanced grids using the 10% imbalance threshold. In general, the percentage of imbalanced grids first increases, reaches the maximum at the 800 m scale, and then decreases when finer scales are used. This also suggests that a coarse scale analysis may ignore many local imbalance sites. When scale is extremely coarse (10 km in this case study), the entire system is balanced. We note that at a very fine scale, although more imbalanced local sites are identified, many grids are found to be sites with no bike inventory/usage. For example, our data show that there are 106,719 grids (about 75%) with no bike usage when the analysis is conducted at the 100 m scale.

Table 3. Imbalance assessment based on different scales.

Scale (m)	Number of Grids	Max Surplus	Max Shortage	Total Imbalance	Imbalanced Grids	Percentage of Imbalanced Grids
10,000	14	293	279	1364	0	0%
5000	63	363	227	5648	10	15.87%
2000	381	460	484	19,602	185	48.56%
1500	662	418	403	25,789	375	56.65%
1000	1470	367	473	35,974	856	58.23%
800	2297	367	316	42,996	1375	59.86%
400	9476	363	302	65,788	5004	52.81%
200	36,284	332	154	96,029	14,029	38.66%
100	138,475	304	148	113,618	24,446	17.65%

Figure 4 shows that imbalanced areas identified at a scale may change when an alternative scale is used. For example, at the 10 km scale, the northwestern grid (circled in Figure 4a) has bike surplus. At the 5 km scale, bike surplus is found to be concentrated in the northeastern corner of the area. If we continue decreasing the scale to 2 km, we start to find local bike shortage and surplus at sites that are considered to be balanced at the 5 km scale. A similar pattern is also observed when we further decrease the analysis scale. In general, when coarse grids are used to perform the imbalance assessment, local surplus and shortage tend to cancel out, leading to an overall more imbalanced status.

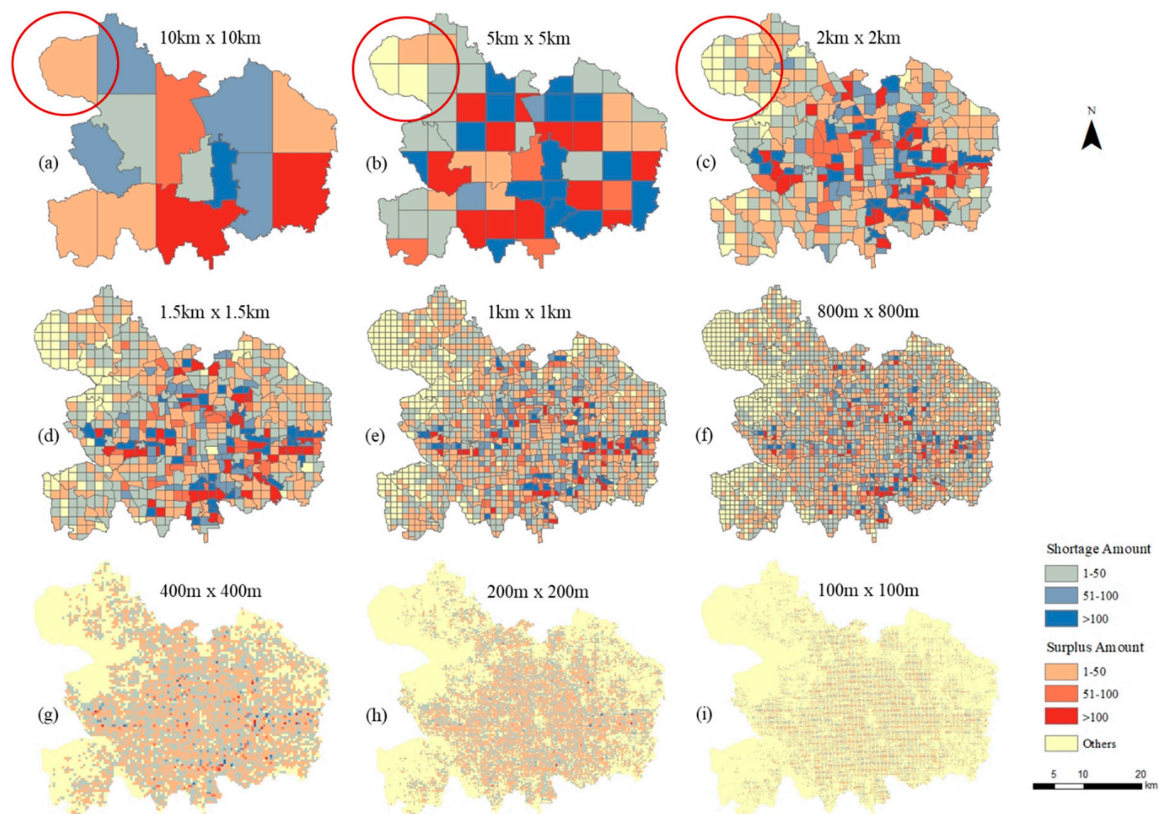


Figure 4. Imbalanced site identification at different scales: (a) 10 km × 10 km, (b) 5 km × 5 km, (c) 2 km × 2 km, (d) 1.5 km × 1.5 km, (e) 1 km × 1 km, (f) 800 m × 800 m, (g) 400 m × 400 m, (h) 200 m × 200 m, and (i) 100 m × 100 m.

4.2. Rebalancing Results at Different Scales

Table 4 provides a summary of the rebalancing results at different scales. For the rebalancing ratio, we set the lower bound $L = 0.9$ and upper bound $U = 1.1$. We exclude the scales of 5 km and 10 km when solving the rebalancing problem, as these two scales are too coarse. Table 4 shows that when finer scales are used, although the number of bikes repositioned increases, the travel distance per bike decreases with an overall average close to the scale used in the analysis. For example, at the 2 km scale, repositioned bikes need to travel slightly more than 2 km compared to the average bike travel distance of 100 m at the 100 m scale. Column “Objective” reports the total distance that repositioned bikes need to travel during the rebalancing process. When scale decreases from 2 km to 1 km, the total travel distance increases. This is because the rate of increase in the amount of bikes repositioned is slightly higher than the rate of decrease in the average bike travel distance. When the scale keeps decreasing from 1 km to 100 m, although more bikes need to be repositioned, the overall travel distance decreases. This is because the average bike travel distance decreases substantially at fine scales. For example, when scale decreases from 1 km to 100 m, the number of bikes repositioned increases about 4 times, whereas the average bike travel distance decrease by 13 times. Column “OD pairs” reports the number of origin–destination grid pairs involved in the bike repositioning process. Table 4 shows that the total amount of OD pairs increases when fine analysis scales are used. This is reasonable given that the overall amount of imbalanced grids increases with finer scales (also see Table 3). Column “Average bikes repositioned per OD pair” summarizes the average number of bikes repositioned between an OD grid pair. Results show that it decreases with finer scales; at the 100 m scale, on average only about three bikes are associated with a repositioning flow.

Table 4. Rebalancing results at different scales ($L = 0.9$; $U = 1.1$).

Scale (m)	Number of Grids	Objective (m)	Bikes Repositioned	Average Bike Travel Distance (m)	OD Pairs	Average Bikes Repositioned Per OD Pair
2000	381	23,264,185	9898	2350	322	31
1500	662	24,940,431	13,294	1876	533	25
1000	1469	25,504,588	19,146	1332	1135	17
800	2297	25,450,894	23,176	1098	1705	14
400	8715	25,338,955	52,358	484	5375	10
200	34,670	18,757,073	82,756	227	15,588	5
100	104,066	9,321,226	92,990	100	28,610	3

Figure 5 maps the rebalancing flows at the 200 m scale. As Figure 5 shows, most rebalancing flows occur among neighboring grids. Grids with surplus imbalance tend to be the origin of rebalancing flows, and grids with shortage imbalance are more likely to be the destination of rebalancing flows. We also notice that there are in/out flows in some grids that already meet the balance criterion. Such flows are often used to help neighboring grids to achieve the balance status. Some grids serve the “transfer” purpose. Rather than directly redistributing bikes from a bike surplus grid to a bike shortage grid, these grids receive bikes from bike surplus areas and redistribute some to one/multiple bike shortage areas. Since the rectilinear distance is used in the empirical study, the “transfer” does not increase the overall reposition distance but leads to a decrease in the average bike travel distance. Including these “transfers” may have the potential to reduce the overall cost by introducing “hubs” in the operator-based rebalancing operation and provide the flexibility for users to engage in one or multiple segments of the rebalancing flows in a user-based rebalancing process.

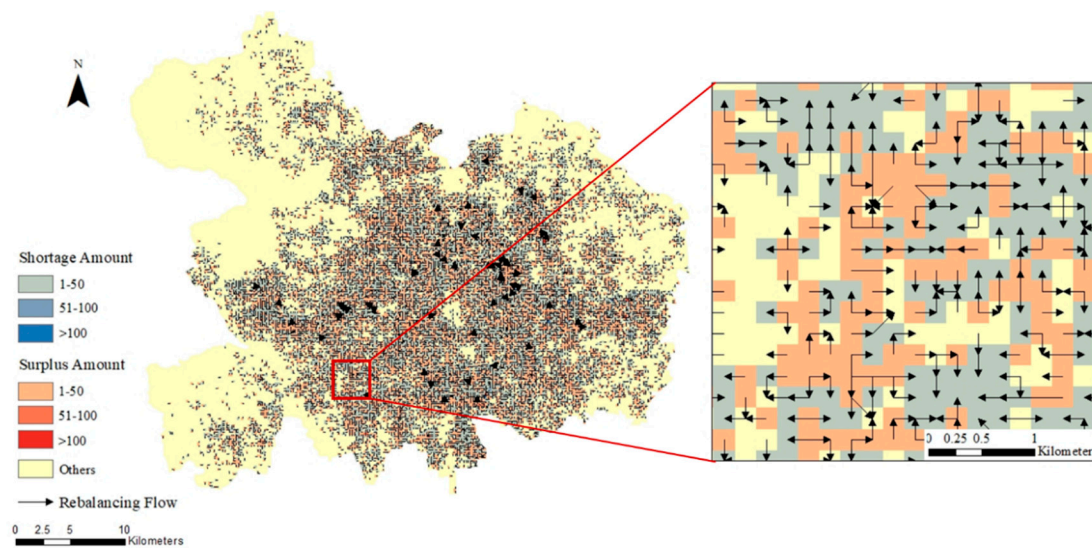


Figure 5. Rebalancing flows at the 200 m scale.

Figure 6 shows the frequency distribution of bike repositioning distances when the scale of 200 m is used in the analysis. According to the figure, around 71% of the repositioned bikes travel from 200 to 300 m, and 99% of the repositioned bikes travel less than 500 m. Figure 7 shows the frequency distribution of the amount of repositioned bikes associated with each rebalancing flow. According to Figure 7, 75% of the rebalancing flows contain one to five bikes, and 99% of the flows involve fewer than 35 bikes. A similar pattern can also be found at other scales: (1) rebalancing travel tends to concentrate on neighboring grids; (2) the average travel distance of repositioned bikes is similar to the analysis scale; (3) the number of repositioned bikes on each rebalancing flow tends to be small.

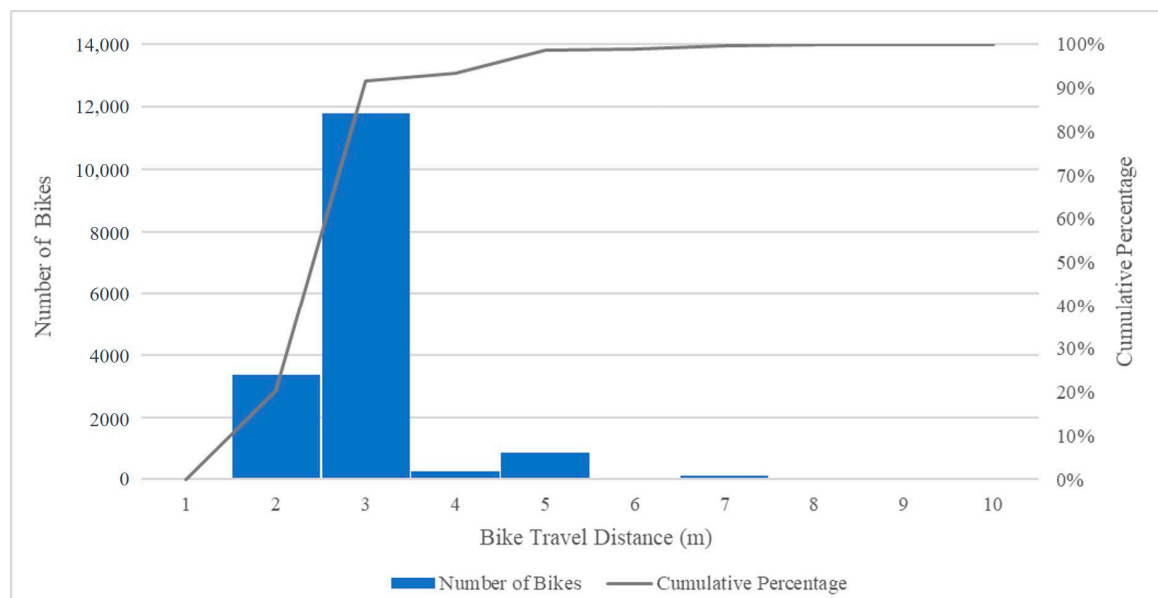


Figure 6. Frequency distribution of bike travel distance at the 200 m scale.

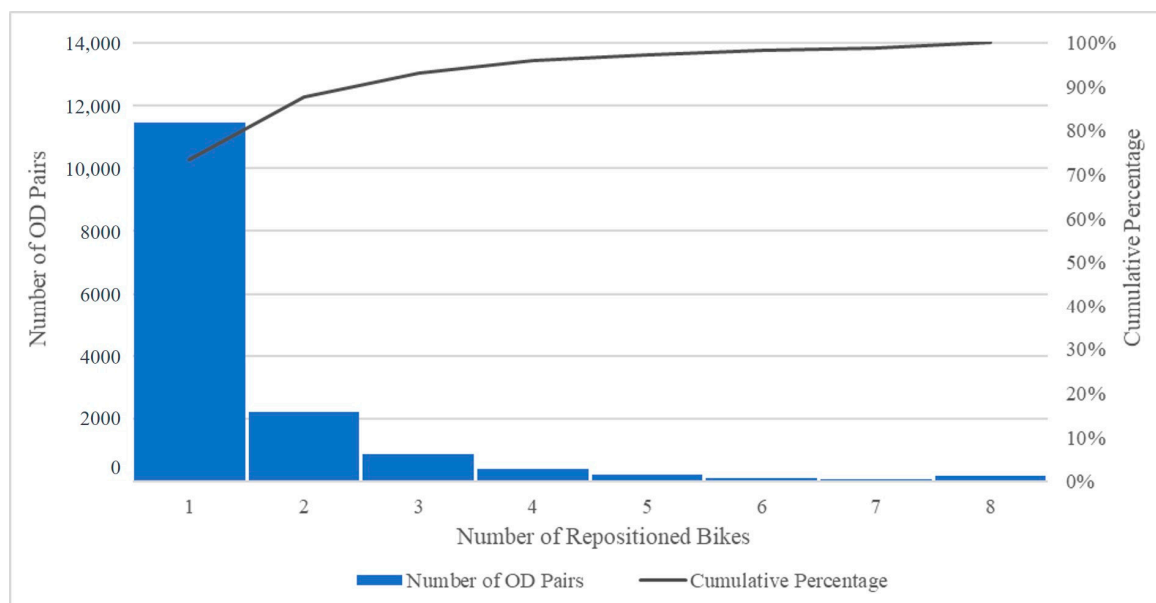


Figure 7. Frequency distribution of repositioned bikes on a rebalancing flow at the 200 m scale.

Table 5 presents the time needed for solving all the problem instances. Problem instances with the scale of 800 m or coarser are solved optimally using CPLEX. For these problems, the overall solution time increases when finer scales are used. For example, the time needed for solving the rebalancing model increases from 14 s to more than 3 min, when scale decreases from 2 km to 800 m. Problem instances with the scale of 400 m or finer are solved using the region decomposition heuristic. In general, problem solution time also increases when finer scales are used. For example, the overall computation time increases from 13 min to more than 2 h when scale decreases from 400 m to 100 m. To assess the quality of a solution obtained using the region decomposition heuristic, we also apply the region decomposition heuristic to solve the problem instance with the analysis scale of 800 m. The solution given by the region decomposition heuristic is about 4% worse than the optimal solution.

Table 5. Summary of problem solution time.

Scale (m)	Number of Grids	Solution Approach	Time (second)
2000	381	Exact	13
1500	662	Exact	17
1000	1469	Exact	88
800	2297	Exact	195
400	8715	Region Decomposition Heuristic	787
200	34,670	Region Decomposition Heuristic	3486
100	104,066	Region Decomposition Heuristic	8143

5. Discussion

Scale is critical in the imbalance evaluation and rebalancing design of an SFBSS. Analysis conducted based on large scales (e.g., 1 km or coarser) provides a regional picture of the imbalance status and is computational efficient. From the operator-based rebalancing point of view, a large-scale analysis helps a strategic rebalancing planning. A rebalancing strategy based on a large-scale analysis usually involves repositioning fewer bikes and a consequent smaller fleet size and less costs associated with bike loading and unloading. However, given that large-scale analysis may miss many local imbalance sites, the effectiveness of a rebalancing operation based on a large-scale analysis may be compromised. In addition, a large-scale analysis might be less helpful for designing a user-based rebalancing approach, because users are unlikely to walk long distances to participate in the rebalancing process.

Analysis using small scales (e.g., 100, 200 m) identifies a large amount of bikes that need to be repositioned. For an operator-based rebalancing approach, repositioning a large amount of bikes would require high costs for bike loading and unloading. However, given that imbalance sites can be more accurately identified when fine analysis scales are used, an operator-based rebalancing strategy designed based on fine scales will be more effective than that using a coarse scale. Small scale analysis also provides valuable insights into the design of effective user-based rebalancing strategies. Our empirical results suggest that many local bike surplus and shortage can be addressed by moving bikes between neighboring areas. For example, at the 200 m scale the majority of rebalancing trips would involve a bike repositioning of less than 300 m. In this case, incentives might be effective for users to help rebalance the system.

Based on the empirical study, we would recommend 800 m as the suitable scale for designing operator-based rebalancing strategies. At the 800 m scale, the percentage of imbalanced bikes reaches the maximum. It suggests that at this scale many local imbalance sites are identified without introducing too many non-relevant areas. In addition, at the scale of 800 m, the rebalancing problem size is reasonable and solving the rebalancing problem is efficient, making it possible to design real-time dynamic rebalancing strategies. Furthermore, even if local shortage/surplus sites exist within an 800 m grid, considering that 800 m is the median distance of walking trips [47], it is acceptable for many users to walk from a local shortage site to a local surplus site to pick up a bike.

As for user-based rebalancing strategies, we would recommend fine scales, such as 100 and 200 m, as the analysis unit. At the 100 m scale, the empirical study suggests that a lot of rebalancing could be achieved within a very short repositioning distance (100 m on average). In this case, consumers do not need to walk far and will be highly motivated to participate in the rebalancing process. Our study also shows that on average only a few bikes need to be redistributed between an OD pair, so in most cases, a small number of users are needed to help move bikes between two specific sites. At the fine scale, many sites could serve as the transfer “stations”, which allows users the flexibility to participate in one or multiple segments of rebalancing trips.

In general, large-sized rebalancing problems present problem–solution challenges. In this study, we introduce the region decomposition heuristic to solve large-sized problem instances. The empirical study shows that the heuristic can be used to solve problems that cannot be handled by existing commercial optimization software. We note that the heuristic solution quality highly relies upon the delineation of “self-contained” sub-regions. In the empirical study, we use a random approach to delineate these sub-regions, and solutions generated by the heuristic are slightly worse than the optimal solution. Future study can focus on developing strategies to identify effective self-contained sub-regions and the associated scale to improve solution quality.

It should be acknowledged that this study has some limitations. First, similar to many existing methods, in an SFBSS, we partition a region into a number of sub-areas and treat each sub-area as a “station” with all bikes inside the same sub-area aggregated at the center of the sub-area. Such an approach might be problematic when the analysis units/sub-areas are large, as bikes inside a sub-area are more likely to be far from the center of a sub-area. Second, in the rebalancing model, we allow an SFBSS to deviate slightly from the complete balance status by introducing a rebalancing upper bound and lower bound. In the empirical study, we examine a lower bound of 0.9 and an upper bound of 1.1. Future work should examine whether and how varying upper/lower bounds affect the appropriate scale selection in the SFBSS rebalancing analysis.

6. Conclusions

SFBSSs have been widely adopted in many cities around the world. However, allowing users to pick up/drop off bikes at numerous sites in an SFBSS causes the bike imbalance issue in an SFBSS. The imbalance issue cannot only suppress potential demand in areas of low supply but also causes bike mess in areas of excess supply. With no fixed bike stations, previous studies have adopted arbitrary analysis units for analyzing an SFBSS. This study examines the impacts of scale on the imbalance

evaluation of an SFBSS and the associated rebalancing strategy design. A spatial optimization model is developed to perform the rebalancing strategically along with a heuristic to solve large-sized problems. The empirical study in downtown Beijing shows that imbalance evaluation results can vary substantially with scale. Imbalance assessment at a large scale cannot only miss many local imbalance sites but also fail to accurately identify the location where imbalance occurs. As for the rebalancing efforts, bike repositioning tends to take place among neighboring areas. Recommendations on the scale choice are provided for the two major rebalancing strategies. Research results provide important insights into sustainable bike transportation planning. Insights can also be gained from this study to help solve the imbalance issue in other shared mobility systems, including car-sharing and scooter-sharing systems.

Author Contributions: Conceptualization, Daoqin Tong; data curation, Xueting Jin; formal analysis, Xueting Jin; methodology, Xueting Jin and Daoqin Tong; supervision, Daoqin Tong; visualization, Xueting Jin; writing—original draft, Xueting Jin; writing—review and editing, Daoqin Tong. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

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

References

1. Midgley, P. The role of smart bike-sharing systems in urban mobility. In *Journeys Sharing Urban Transport Solutions*; LTA Academy: Singapore, 2009; pp. 23–31.
2. Jiang, Q.; Ou, S.-J.; Wei, W. Why Shared Bikes of Free-Floating Systems Were Parked Out of Order? A Preliminary Study based on Factor Analysis. *Sustainability* **2019**, *11*, 3287. [\[CrossRef\]](#)
3. Büttner, J.; Mlasowsky, H.; Birkholz, T.; Gröper, D.; Fernández, A.C.; Emberger, G.; Petersen, T.; Robèrt, M.; Vila, S.S.; Reth, P.; et al. *Optimizing Bike Sharing in European Cities—A Handbook*; Obis Project; Intelligent Energy Europe: Brussels, Belgium, 2011.
4. Rojas-Rueda, D.; de Nazelle, A.; Teixido, O.; Nieuwenhuijsen, M.J. Replacing car trips by increasing bike and public transport in the greater Barcelona metropolitan area: A health impact assessment study. *Environ. Int.* **2012**, *49*, 100–109. [\[CrossRef\]](#)
5. Buehler, R.; Hamre, A. Business and Bikeshare User Perceptions of the Economic Benefits of Capital Bikeshare. *Transp. Res. Rec.* **2015**, *2520*, 100–111. [\[CrossRef\]](#)
6. DeMaio, P. Bike-sharing: History, Impacts, Models of Provision, and Future. *J. Public Transp.* **2009**, *12*, 3. [\[CrossRef\]](#)
7. Nielsen, B. *The Bicycle in Denmark: Present Use and Future Potential*; Ministry of Transport: Copenhagen, Denmark, 1993.
8. Pal, A.; Zhang, Y. Free-floating bike sharing: Solving real-life large-scale static rebalancing problems. *Transp. Res. Part C Emerg. Technol.* **2017**, *80*, 92–116. [\[CrossRef\]](#)
9. Shen, Y.; Zhang, X.; Zhao, J. Understanding the usage of dockless bike sharing in Singapore. *Int. J. Sustain. Transp.* **2018**, *12*, 686–700. [\[CrossRef\]](#)
10. Schmidt, C. Active travel for all? The surge in public bike-sharing programs. *Environ. Health Perspect.* **2018**, *126*, 082001. [\[CrossRef\]](#)
11. Nair, R.; Miller-Hooks, E. Fleet Management for Vehicle Sharing Operations. *Transp. Sci.* **2011**, *45*, 524–540. [\[CrossRef\]](#)
12. Chemla, D.; Meunier, F.; Calvo, R.W. Bike sharing systems: Solving the static rebalancing problem. *Discret. Opt.* **2013**, *10*, 120–146. [\[CrossRef\]](#)
13. O'Mahony, E.; Shmoys, D.B. Data Analysis and Optimization for (Citi) Bike Sharing. In Proceedings of the 29th AAAI Conference on Artificial Intelligence, Austin, TX, USA, 25–30 January 2015.
14. Liu, J.; Sun, L.; Chen, W.; Xiong, H. Rebalancing Bike Sharing Systems: A Multi-source Data Smart Optimization. In Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, San Francisco, CA, USA, 13–17 August 2016; Volume 10, pp. 1005–1014.
15. Fricker, C.; Gast, N. Incentives and redistribution in homogeneous bike-sharing systems with stations of finite capacity. *EURO J. Transp. Logist.* **2016**, *5*, 261–291. [\[CrossRef\]](#)

16. Li, Y.; Zheng, Y.; Yang, Q. Dynamic bike reposition: A spatio-temporal reinforcement learning approach. In Proceedings of the 24th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, London, UK, 19–23 August 2018; pp. 1724–1733.
17. Bulhões, T.; Subramanian, A.; Erdoğan, G.; Laporte, G. The static bike relocation problem with multiple vehicles and visits. *Eur. J. Oper. Res.* **2018**, *264*, 508–523. [\[CrossRef\]](#)
18. Chiariotti, F.; Pielli, C.; Zanella, A.; Zorzi, M. A Dynamic Approach to Rebalancing Bike-Sharing Systems. *Sensors* **2018**, *18*, 512. [\[CrossRef\]](#) [\[PubMed\]](#)
19. Pan, L.; Cai, Q.P.; Fang, Z.X.; Tang, P.Z.; Huang, L.B. A Deep Reinforcement Learning Framework for Rebalancing Dockless Bike Sharing Systems. In Proceedings of the AAAI Conference on Artificial Intelligence, 27 January–1 February 2019, Honolulu, HI, USA; Volume 33, pp. 1393–1400. [\[CrossRef\]](#)
20. Xu, C.; Ji, J.; Liu, P. The station-free sharing bike demand forecasting with a deep learning approach and large-scale datasets. *Transp. Res. Part C Emerg. Technol.* **2018**, *95*, 47–60. [\[CrossRef\]](#)
21. Zhai, Y.; Liu, J.; Du, J.; Wu, H. Fleet Size and Rebalancing Analysis of Dockless Bike-Sharing Stations Based on Markov Chain. *ISPRS Int. J. Geo Inf.* **2019**, *8*, 334. [\[CrossRef\]](#)
22. Xing, Y.; Wang, K.; Lu, J. Exploring travel patterns and trip purposes of dockless bike-sharing by analyzing massive bike-sharing data in Shanghai, China. *J. Transp. Geogr.* **2020**, *87*, 102787. [\[CrossRef\]](#)
23. Beecham, R.; Wood, J. Exploring gendered cycling behaviours within a large-scale behavioural data-set. *Transp. Plan. Technol.* **2014**, *37*, 83–97. [\[CrossRef\]](#)
24. Fishman, E.; Washington, S.; Haworth, N.; Watson, A. Factors influencing bike share membership: An analysis of Melbourne and Brisbane. *Transp. Res. Part A* **2015**, *71*, 17–30. [\[CrossRef\]](#)
25. Yang, H.; Xie, K.; Ozbay, K.; Ma, Y.; Wang, Z. Use of Deep Learning to Predict Daily Usage of Bike Sharing Systems. *Transp. Res. Rec. J. Transp. Res. Board* **2018**, *2672*, 92–102. [\[CrossRef\]](#)
26. Zhou, Y.; Wang, L.; Zhong, R.; Tan, Y. A Markov Chain Based Demand Prediction Model for Stations in Bike Sharing Systems. *Math. Probl. Eng.* **2018**, *2018*, 1–8. [\[CrossRef\]](#)
27. Li, H.; Wang, Q.; Shi, W.; Deng, Z.; Wang, H. Residential clustering and spatial access to public services in Shanghai. *Habitat Int.* **2015**, *46*, 119–129. [\[CrossRef\]](#)
28. Zhang, Y.; Thomas, T.; Brussel, M.; van Maarseveen, M. Exploring the impact of built environment factors on the use of public bikes at bike stations: Case study in Zhongshan, China. *J. Transp. Geogr.* **2017**, *58*, 59–70. [\[CrossRef\]](#)
29. El-Assi, W.; Salah Mahmoud, M.; Nurul Habib, K. Effects of built environment and weather on bike sharing demand: A station level analysis of commercial bike sharing in Toronto. *Transportation* **2017**, *44*, 589–613. [\[CrossRef\]](#)
30. Caggiani, L.; Ottomanelli, M. A Dynamic Simulation based Model for Optimal Fleet Repositioning in Bike-sharing Systems. *Proc. Soc. Behav. Sci.* **2013**, *87*, 203–210. [\[CrossRef\]](#)
31. Vélib'. Le Bonus V'+ Sera en Service dans une Centaine de Stations Vélib' dès le 14 Juin. 2008. Available online: <http://www.velib.paris.fr/> (accessed on 15 October 2020).
32. Chemla, D.; Meunier, F.; Pradeau, T.; Calvo, R.W. *Self-Service Bike Sharing Systems: Simulation, Repositioning, Pricing*; Technical Report hal-00824078; Centre d'Enseignement et de Recherche en Mathématiques et Calcul Scientifique–CERMICS, Laboratoire d'Informatique de Paris-Nord–LIPN, Parallélisme, Réseaux, Systèmes d'information, Modélisation–PRISM 2013: Aachen, Germany, 2013.
33. Müller, J.; Schmöller, S.; Giesel, F. Identifying Users and Use of (Electric-) Free-Floating Carsharing in Berlin and Munich. In Proceedings of the 2015 IEEE 18th International Conference, Las Palmas, Spain, 15–18 September 2015; pp. 2568–2573.
34. Raviv, T.; Tzur, M.; Forma, I.A. Static repositioning in a bike-sharing system: Models and solution approaches. *EURO J. Transp. Logist.* **2013**, *2*, 187–229. [\[CrossRef\]](#)
35. Ho, S.; Szeto, W. Solving a static repositioning problem in bike-sharing systems using iterated tabu search. *Transp. Res. Part E Logist. Transp. Rev.* **2014**, *69*, 180–198. [\[CrossRef\]](#)
36. Alvarez-Valdes, R.; Belenguer, J.M.; Benavent, E.; Bermudez, J.D.; Muoz, F.; Vercher, E.; Verdejo, F. Optimizing the level of service quality of a bikesharing system. *Omega* **2016**, *62*, 163–175. [\[CrossRef\]](#)
37. Gaspero, D.; Andrea, R.; Tommaso, U. Balancing bike sharing systems with constraint programming. *Constraints* **2015**, *21*, 318–348. [\[CrossRef\]](#)
38. Schuijbroek, J.; Hampshire, R.; van Hoes, W.J. Inventory rebalancing and vehicle routing in bike sharing systems. *Eur. J. Oper. Res.* **2017**, *257*, 992–1004. [\[CrossRef\]](#)

39. Dell’Amico, M.; Hadjicostantinou, E.; Iori, M.; Novellani, S. The Bike Sharing Rebalancing Problem: Mathematical Formulations and Benchmark Instances. *Omega* **2014**, *45*, 7–19.
40. 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](#)]
41. Ji, Y.; Jin, X.; Ma, X.; Zhang, S. How Does Dockless Bike-Sharing System Behave by Incentivizing Users to Participate in Rebalancing? *IEEE Access* **2020**, *8*, 58889–58897. [[CrossRef](#)]
42. Tong, D.; Murray, A. Spatial Optimization in Geography. *Ann. Assoc. Am. Geogr.* **2012**, *102*, 1290–1309. [[CrossRef](#)]
43. Contardo, C.; Morency, C.; Rousseau, L.-M. *Balancing a Dynamic Public Bike-Sharing System*; Technical Report; CIRRELT: Montreal, QC, Canada, 2012.
44. Caggiani, L.; Ottomanelli, M. A Modular Soft Computing Based Method for Vehicles Repositioning in Bike-sharing Systems. *Transp. Res. Proc.* **2012**, *10*, 364–373.
45. Szeto, W.Y.; Liu, Y.; Ho, S.C. Chemical Reaction Optimization for Solving a Static Multi-vehicle Bike Reposition Problem. *Transport. Res. Part B Methodol.* **2016**, *109*, 176–211. [[CrossRef](#)]
46. Atash, F. Redesigning suburbia for walking and transit: Emerging concepts. *J. Urban Plan. Dev.* **1994**, *120*, 48–57. [[CrossRef](#)]
47. Yang, Y.; Diez-Roux, A.V. Walking distance by trip purpose and population subgroups. *Am. J. Prev. Med.* **2012**, *43*, 11–19. [[CrossRef](#)]
48. Mátrai, T.; Tóth, J. Comparative assessment of public bike sharing systems. *Transp. Res. Proc.* **2016**, *14*, 2344–2351. [[CrossRef](#)]
49. O’Brien, O.; Cheshire, J.; Batty, M. Mining bicycle sharing data for generating insights into sustainable transport systems. *J. Transp. Geogr.* **2014**, *34*, 262–273. [[CrossRef](#)]
50. Miyagawa, M. Distribution of the difference between distances to the first and second nearest facilities (ISOLDE XII). *J. Oper. Res. Soc. Jpn.* **2013**, *56*, 167–176. [[CrossRef](#)]

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



© 2020 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/>).