T-ITS-23-05-1448.Final

Recovering Real Demand for Free-Floating Bike-Sharing System Considering Demand Truncation, Migration, and Spatial Correlation

Jianbiao Wang, Tomio Miwa, Xinwei Ma, Dawei Li, and Takayuki Morikawa

***Abstract*—Owing to the shareability and spatial-temporal imbalance of free-floating bike-sharing (FFBS), the users may fail to pick up the bike at the desired location (i.e., the demand is truncated). Consequently, the demand partially migrates to nearby areas or is lost. Thus, the observed demand recorded in the system is not the underlying real demand. To address this issue, a framework for Demand Recovery considering demand Truncation, Migration, and spatial Correlation (DRTMC) was proposed. In detail, the real demand recovery is first modeled as a maximum-a-posteriori (MAP) problem. Then, the prior term and conditional term in MAP are formulated to jointly consider the demand truncation, migration, and spatial correlation. Finally, a tailored simulated annealing approach is applied to estimate the real demand. We then present the results using simulation data for validation and real-world data for empirical analysis. The validation results indicate that the demand recovered by the DRTMC model is much closer to the real demand value than the observed demand, and it performs better than all the benchmarks. For the case study, the FFBS data from Shanghai City in a prosperous area is chosen. It shows the distribution of observed demand is significantly different from that of real demand, emphasizing the importance of real demand recovery. The proposed DRTMC model enables researchers to capture the underlying real demand for FFBS and develop more effective rebalance strategies to improve FFBS service levels.**

***Index Terms*—demand correlation; demand migration; demand truncation; free-floating bike-sharing; real demand recovery.**

I. Introduction

Bike-sharing systems provide a shared, economical, flexible, convenient, and sustainable travel alternative for travelers, which solves the last mile problems and mitigates traffic congestion in cities [1], [2], [3], [4], [5]. In general, the current bike-sharing systems can be categorized into two types: station-based bike-sharing (SBBS) and free-floating bike-sharing (FFBS) [6]. In SBBS, users pick up and drop off the bikes at a designated station. In contrast, FFBS eliminates the station constraint and is characterized by its "Internet plus" property, enabling users to pick up and drop off bikes at any allowable location with their smartphones [6], [7]. The introduction of bike-sharing benefits the travelers and transport system. To improve the service level of bike-sharing, an increasing amount of research has been conducted, which commonly includes analyzing the usage pattern [8], [9], investigating the influential factors [7], [10], [11] predicting future demand [5], [12], [13], and relocating the bike-sharing services [14], [15], [16], [17].

Accurately capturing the actual real-world demand of FFBS is vital for ensuring the effectiveness of the aforementioned studies. However, due to the feature of sharing, the supply would be limited among users. Consequently, the bike usage demand may not be totally satisfied [18]. Thus, some users fail to acquire bikes at their desired locations. Moreover, the failure to use the shared travel mode is a common issue not only in FFBS but also in other shared travel modes (e.g., taxi or ride-hailing). It is vital to consider the unmet or unassigned demand when analyzing or designing the shared travel system [19],[20]. Thus, in this paper, the FFBS as a typical shared travel mode is selected for the study. The example of usage details for FFBS and corresponding data are illustrated in Fig. 1.

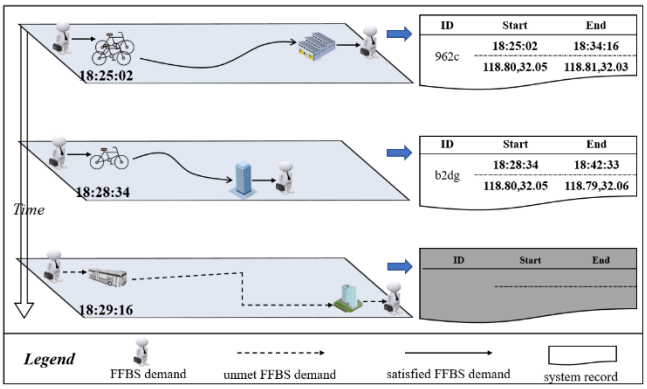


Fig. 1. The usage details and data samples of FFBS.

In this example, three users want to pick up the bike at the location (118.80, 32.05). For the first user arriving at 18:25:02, two bikes are available and then one bike is picked up. After the trip is finished, the FFBS system will record the information.

This work was supported in part by the Nagoya University Interdisciplinary Frontier Fellowship under Grant JPMJFS2120, and in part by the National Natural Science Foundation of China under Grants 52202387, 71971056, and 51608115. The authors gratefully acknowledge their support. *(Corresponding author: Jianbiao Wang.)*

Jianbiao Wang is with the Department of Civil and Environmental Engineering, Nagoya University, Nagoya 464-8603, Japan (e-mail: wang.jianbiao98@gmail.com)

Tomio Miwa is with the Institute of Materials and Systems for Sustainability, Nagoya University, Nagoya 464-8603, Japan (e-mail: miwa@nagoya-u.jp)

Xinwei Ma is with School of Civil and Transportation Engineering, Hebei University of Technology, Tianjin 300401, China (e-mail: xinweima@hebut.edu.cn)

Dawei Li is with School of Transportation, Southeast University, Nanjing, 211189, China (e-mail: lidawei@seu.edu.cn)

Takayuki Morikawa is with the Institute of Innovation for Future Society, Nagoya University, Nagoya 464-8603, Japan (e-mail: morikawa@nagoya-u.jp)

which includes the bike ID, the trip starting time and origin location, the trip ending time and the destination location. Similarly, another bike is picked up by the second user arriving at 18:28:34, and the usage information is recorded as well. However, when the third user wants to use the bike at 18:29:16, the bike is not available. Then, she/he chooses the secondary option and goes to the destination by bus. In this case, the FFBS demand is unmet. Unlike the previous two successful pickups, the information of the third user can't be recorded. As a consequence, the pickups recorded in the system cannot reflect the real bike usage demand [21].

In this study, we define the number of users planning to use bikes when the bike supply is sufficient as real demand (i.e. actual real-world demand without truncation), successfully using bikes as observed demand, and planning but failing to use bikes as the unmet demand [22], [23]. According to [21], which used the FFBS dataset in Beijing from May 14 to May 31, 2017, there were more than 80% of areas experienced demand truncation. Thus, the real demand and observed demand should be distinguished. However, most previous studies ignore the unmet demand and directly treat historical observed demand as the real demand [5], [7]. Thus, estimating the underlying real demand of bike-sharing receives little attention. But estimating the FFBS real demand is difficult due to the data limitation. Commonly, as presented in Fig.1, the dataset of FFBS contains only the bike ID, pick-up/drop-off time and location [4], [9]. In other words, only the successful usage can be recorded. The researchers have no idea about how many users search for the bikes in the APP, and how the users behave when no bike is available. Moreover, detailed GPS movement is rarely obtained in the FFBS system. Thus, there is no direct reference for the underlying real FFBS demand. All those data limitations make it a challenge to estimate the FFBS real demand.

In addition to the real demand, observed demand, and unmet demand, another kind of demand related to demand truncation analysis especially in the free-floating bike-sharing context is the "migration demand" [24]. It is the number of demands captured by the nearby area when bikes in the first-choice area are not available [23] [25]. It is worth mentioning that the concept of migration demand is closely related to substitution demand in market management. In detail, the migration demand can be viewed as the FFBS demand substituting among different areas. For the demand substituting to other travel modes, it is counted as the unmet demand. Here, to emphasize the context of FFBS and to avoid misleading, the migration demand instead of substitution demand is used for the paper exposition. An important concept related to the migration demand is the migration (or substitution) rate $\pi_{i\to j}$ [26], which is the probability of shifting to area j when bikes in area i are not available [25], [26]. When truncation occurs, the truncated demand spills over to the migration area, which captures extra demand that wasn't supposed to be there. In other words, the demand in different areas would be coupled together [23]. Therefore, the migration process must be considered and the spillover and captured demand must be decoupled when recovering the real demand for FFBS in each area [27], [28].

Further, it is proved that the demand for shared travel modes like car-sharing and public buses show an obvious spatial correlation, and the FFBS is also not the exception. Especially, the real demand for FFBS in the nearby area should be spatially correlated and tend to be similar [7], [23]. Therefore, in this paper, such an implicit positive correlation of FFBS will also be considered.

The aim of this study is to develop a model to recover the underlying real demand based on the observed information in Fig.1. Especially, what we can see is not the truth, and we need to infer the underlying truth based on what we observe. In this study, real demand recovery is modeled as a maximum a posteriori (MAP) problem considering demand truncation, migration, and spatial correlation. The model details are provided in Section III. Moreover, due to the limitation of the system record, the ground-truth real demand of FFBS can't be obtained, which results in the difficulty of model validation. To address this issue, the simulation study is first conducted where the ground-truth data is obtainable to validate the model. Then the verified model is applied in real-world case to estimate the actual real-world demand of FFBS. The contributions of this study are as follows:

- A framework for Demand Recovery considering Truncation, Migration, and spatial Correlation (DRTMC) is proposed to estimate the real demand of FFBS. In particular, the real demand recovery under the framework of DRTMC is formulated as a maximum-a-posteriori (MAP) problem.
- The performance of the proposed DRTMC is validated and proved to be superior to the benchmarks through simulation experiments, then it is applied to the field dataset in real world for empirical study.

The remainder of this paper is organized as follows. A literature review of related work is presented in Section II. Section III presents the methodology, followed by the validation in Section IV. The case study in real world is then shown in Section V. Finally, the conclusion is provided in Section VI.

II. Literature Review

Research on demand recovery in the transportation field is widely reviewed in this section to provide a complete perspective. To begin with, a simple and straightforward method of recovering real demand or unmet demand is to conduct a survey. For instance, Ullman et al. conducted a state-wide survey of the inhabitants of Vermont to collect information about their long-distance travel, where nearly 22% of respondents reported they experienced unmet demands at least once yearly [29]. Further, they found that the number of unmet demands was larger than expected for females. Besides, Luiu and Tight focused on elderly people and used the National Travel Survey in England to investigate their travel difficulties and unmet demand. The result revealed that poor health, lack of access to transport resources, and gender were the main factors contributing to unmet travel demand [30]. Fields et al. designed the *MyAmble* app to record the unmet travel demand amongst low-income older adults. Their two-week study concluded that

16% of the travel demand failed to be met in this period [31]. Further, a survey was conducted by the Torbay Council in the UK to capture how much unmet taxi demand there is so that the taxi service levels could be improved [32]. Such survey methods can collect data easily but are time-consuming and limited to small sample sizes [33].

As an alternative, passively collected data can also be used for demand estimation. For example, using taxi GPS data in New York City, Zhang and Ghanem estimated the demand for street-hail taxis in each road segment using the non-stationary Poisson random field method [34]. Similarly, by utilizing street-hail taxi data in Xi'an City, Wang et al. developed a recursive decomposition probability model to estimate the real demand by incorporating the disaggregate passing time of each taxi [35]. Based on the car-sharing historical dataset in Boston and New York City, Fields et al. simulated the car-sharing demand truncation process with different levels of demand distribution, then compared the output of the simulation and the pickups in historical data to determine the most likely real demand distribution, the demand migration behavior was implicitly considered in the simulation while not explicitly formulated in the analytical model [23].

In addition, studies specifically related to the demand estimation of bike-sharing have also been conducted. To evaluate the service levels of FFBS and SBBS, Albiński et al. calculated the mean value of demand that is not truncated in the same period to replace truncated demands [18]. In addition, based on the CitiBike dataset, Jian et al. proposed a discrete event simulator to estimate the demand for each pair of stations in New York City [36]. Similarly, an iterative simulation-based inference method was established by Negahban to estimate the real demand for SBBS, proving that ignoring the demand truncation causes incorrect or suboptimal decisions [37]. To estimate the demand for FFBS, Wang et al. treated truncated demand as the missing value in the matrix and established a matrix completion solution to detruncate those values [21]. Moreover, the Tobit model is a common method applied for truncation research, and a representative study on bike-sharing was conducted by Gammelli et al., where the authors integrated the Tobit model with a Gaussian process and modeled the problem of real demand recovery as a time-series analysis to recover the real SBBS demand [22].

These approaches have been successfully applied in real demand recovery. However, the following aspects need to be explored further for FFBS. **(i)** The data structure used in these methods is commonly represented by a tuple (y, Tr), where y is the value of observed demand and Tr represents whether the real demand is truncated (i.e., $Tr = 1$) or not (i.e., $Tr = 0$). Thus, Tr is an "all-or-nothing" indicator. However, the demand could also be partially truncated in the focused time period. Therefore, in addition to indicator Tr, the duration of availability can be used to incorporate the information of truncation intensity. **(ii)** When a bike is unavailable in area *A*, it is possible for a user to migrate to nearby area *B* to pick up a bike. Such migration has rarely been investigated except for a few studies. To address these issues in the FFBS context, Wang et al. proposed the DTMP regression model to consider the demand truncation and migration process [24]. However, the demand correlation among nearby areas is ignored. Moreover, the regression model focuses on the trend of the whole dataset, although it reveals the causal influence of factors on real demand, it can't be used to estimate the real demand for each individual area. To overcome the limitation, this study tries to estimate the real demand of FFBS in each individual grid area by jointly considering the demand truncation, migration, and spatial correlation.

III. Methodology

This section introduces the DRTMC framework. First, we state the problem as a MAP estimation in Section *A*. Then, the truncation, migration, and correlation are modeled in detail in Sections *B* and *C*, and the demand estimation process is explained in Section *D*. The annotations used in this paper are summarized in Appendix *A* for ease of reference.

A. Problem Description

In this study, we aim to estimate the real FFBS demand based on the observable but truncated demand. To ensure that the problem is trackable and address the inconvenience of the free-floating feature, by following [21], [38], we first divide the study area into equal-sized grids of length 100 m (as in Fig.9). Then, the problem can be transferred to estimate the real demand in each grid.

Because the real demand in each day is a random value and its change is uncapturable, what we are interested in is the **expected value** of real demand, which is a statistical property for the daily real demand. The assumptions made before modeling are as follows:

Assumption 1: For each grid, the expected value of real demand is constant in the focused period (e.g., 7:00 am – 8:00 am). Since the expected value of real demand varies at different periods of the day, it is not possible to target the whole day in the study. Instead, we focus on a specific time period and assume that the expected value doesn't change within this period [36]. It is noted that the constant expected value doesn't mean the actual daily real demand is constant. Rather, the expected value of real demand is a statistical parameter for the daily real demand varying from day to day

Assumption 2: When the truncation happens, the unsatisfied demand will either leave the FFBS system or migrate to nearby to pick up the bike. The candidate grids for the migration include the nearby eight grids, which means the demand won't spill over to further grids. This assumption is relatively strong, but it is easy to be satisfied if we focus on the prosperous area (e.g., case study area in Fig. 9). This is because in those areas the supply of bikes is high, the situation that no bikes in all the eight nearby grids rarely happens.

Assumption 3: The observed demand and real demand both follow Poisson distributions but with different expected values. Particularly, the expected values of the real and observed demands can be related by considering the truncation and migration process. Then, we can assume that the observed demand and real demand follow two Poisson distributions with different but related parameters, respectively.

Assume that there are N grids and K days' observations. For grid $n \in N$, we denote the randomly observed demand in the specific period on the k^{th} day as $y_{n,k} \in \mathbf{Y}$. In particular, $y_{n,k}$ is a random variable that follows a Poisson distribution, $y_{n,k} \sim Psn(\lambda_n^o)$, where $Psn(\cdot)$ denotes the probability mass function of the Poisson distribution and $\lambda_n^o \in \boldsymbol{\lambda^o}, n = 1,2,\dots,N$ is the expected value of the observed demand in grid n. Similarly, let $\lambda_n^r \in \boldsymbol{\lambda^r},\ n = 1,2,\dots,N$ represent the expected value of the real demand in grid n. It is noted that $\boldsymbol{\lambda^r}$ is the quantity of interest that will be estimated in this study. In addition, denote $P(\boldsymbol{\lambda^r}|\boldsymbol{Y})$ as the probability that the expected value of real demand parameter is $\boldsymbol{\lambda^r}$ given that the observed demand is $\boldsymbol{Y}$. Here, $P(\boldsymbol{\lambda^r}|\boldsymbol{Y})$ is a posterior distribution. The target of this study is to maximize this posterior distribution, which will "guess" the most likely expected value of real demand conditional on the observed data. This is a MAP problem. According to the Bayesian theorem, $P(\boldsymbol{\lambda^r}|\boldsymbol{Y})$ can be written as,

$$P(\boldsymbol{\lambda^r}|\boldsymbol{Y}) = \frac{P(\boldsymbol{\lambda^r})P(\boldsymbol{Y}|\boldsymbol{\lambda^r})}{P(\boldsymbol{Y})}, \tag{1}$$

where $\boldsymbol{Y}$ is the given value and $P(\boldsymbol{Y})$ is a constant. In addition, the $P(\boldsymbol{\lambda^r})$ is the prior term, which models the common knowledge without needing the specific observed information. In particular, in the context of FFBS demand, it is common sense that the demand is spatially correlated and the real demand in a grid is similar to those of their neighbors. In other words, if the λ_n^r is similar to those of nearby grids, the prior term $P(\boldsymbol{\lambda^r})$ will be higher. Thus, the spatial correlation is considered as the prior information in $P(\boldsymbol{\lambda^r})$ when recovering the FFBS demand.

As for $P(\boldsymbol{Y}|\boldsymbol{\lambda^r})$, it is a conditional term that models the relationship between the expected real value λ_n^r and randomly observed information $y_{n,k}$. Although the real demand is different from the observed demand, they are not independent of each other. Specifically, $y_{n,k}$ is observed after the grids experience truncation and migration. Therefore, the observed demand $y_{n,k}$ is related to the truncated and migrated versions of the real demand λ_n^r. Accordingly, the conditional term $P(\boldsymbol{Y}|\boldsymbol{\lambda^r})$ is used to incorporate the truncation and migration process.

Overall, in real demand recovery, MAP estimation can be applied to estimate the expected value of the real demand $\boldsymbol{\lambda^r}$ conditional on the observed demand $\boldsymbol{Y}$. Moreover, because $P(\boldsymbol{Y})$ is a constant with regard to $\boldsymbol{\lambda^r}$, $\boldsymbol{\lambda^r}$ can be determined by maximizing the $P(\boldsymbol{\lambda^r})P(\boldsymbol{Y}|\boldsymbol{\lambda^r})$ directly, as follows,

$$\widehat{\boldsymbol{\lambda^r}} = \underset{\boldsymbol{\lambda^r}}{argmax}\frac{P(\boldsymbol{\lambda^r})P(\boldsymbol{Y}|\boldsymbol{\lambda^r})}{P(\boldsymbol{Y})} \propto \underset{\boldsymbol{\lambda^r}}{argmax}\big(P(\boldsymbol{\lambda^r})P(\boldsymbol{Y}|\boldsymbol{\lambda^r})\big). \tag{2}$$

Therefore, the core part of DRTMC is to model truncation and migration in $P(\boldsymbol{Y}|\boldsymbol{\lambda^r})$, and to model the spatial correlation in $P(\boldsymbol{\lambda^r})$, which will be elaborated in the next sections *B* and. For the convenience of modeling, a demand network, $G(V,E)\ v \in V\ \ e \in E$, is first constructed, as shown in Fig. 2. The constructed network can be viewed as a Markov network or Markov random field (MRF). This network contains two layers. The upper layer represents the expected value of the observed demand $\boldsymbol{\lambda^o}$, and the lower layer represents the expected value of the real demand $\boldsymbol{\lambda^r}$. Each layer consists of several points v, representing the expected value of real demand or observed demand in grids. It is noted that the points at the same location in the upper and lower layers represent the same grid but with different demand information. Without loss of generality, we focus on the point in the center in Fig. 2, whose values are denoted by λ_n^o and λ_n^r in the upper layer and lower layer, respectively. In addition, for the focused grids, there are eight migration grids, and their values are denoted by λ_m^o and λ_m^r, $m \in N(n)$. Here, $N(\cdot)$ is the neighborhood function containing the index of the migration grids. The links e in Fig. 2 reveal the relationships among different points v. For simplicity, we only show the links related to the center and ignore the other links. In particular, there are three types of links with different colors: demand truncation links e_t, demand migration links e_m, and demand correlation links e_c,

B. Truncation and Migration in Conditional Term $P(\boldsymbol{Y}|\boldsymbol{\lambda^r})$

The conditional term $P(\boldsymbol{Y}|\boldsymbol{\lambda^r})$ models the distribution between the observed demand $y_{n,k} \in \boldsymbol{Y}$ and the expected value of real demand $\lambda_n^r \in \boldsymbol{\lambda^r}$. Here, $y_{n,k}$ is a randomly distributed variable and λ_n^r is the expected value of real demand. To connect these two variables, the expected value of the observed demand $\lambda_n^o \in \boldsymbol{\lambda^o}$ plays a mediating role. In detail, $y_{n,k}$ follows the Poisson distribution with parameter λ_n^o, i.e., $y_{n,k} \sim Psn(\lambda_n^o)$. Meanwhile, based on the demand truncation and migration process, λ_n^o is a function of λ_n^r, i.e., $\lambda_n^o = f(\lambda_n^r, \boldsymbol{\lambda_m^r}), m \in N(n)$. Therefore, we can obtain $y_{n,k} \sim Psn(f(\lambda_n^r, \boldsymbol{\lambda_m^r}))$. The relationship between $y_{n,k}$ and λ_n^r can be expressed as,

$$\begin{aligned} P\big(Y = y_{n,k}|(\lambda_n^r, \boldsymbol{\lambda_m^r})\big) &= Psn\big(y_{n,k}|\lambda_n^o\big) \\ &= \frac{e^{-\lambda_n^o} * {\lambda_n^o}^{y_{n,k}}}{y_{n,k}!} = \frac{e^{-f(\lambda_n^r, \boldsymbol{\lambda_m^r})} * f(\lambda_n^r, \boldsymbol{\lambda_m^r})^{y_{n,k}}}{y_{n,k}!}. \end{aligned} \tag{3}$$

The problem then becomes how to model the function $\lambda_n^o = f(\lambda_n^r, \boldsymbol{\lambda_m^r})$ in (3). Once the $f(\lambda_n^r, \boldsymbol{\lambda_m^r})$ is formatted, the conditional term could be expressed analytically.

The expected value of the observed demand λ_n^o is consists of two parts. The first part, $\lambda_n^{o,tr}$, is the demand related to the truncation process, and it represents the demand that comes from grid n itself when the demand is not truncated (truncation link in Fig. 2(a)). The second part, $\lambda_n^{o,mig}$, considers the demand migration process and represents the demand that comes from nearby grid m after demand truncation occurs there (migration link in Fig. 2(b)):

$$\lambda_n^o = \lambda_n^{o,tr} + \lambda_n^{o,mig}. \tag{4}$$

Particularly, a higher value of $\lambda_n^{o,tr}$ is captured if the duration of availability in n is longer and vice versa. To represent $\lambda_n^{o,tr}$, the indicator a_n is defined to measure availability. It is expressed as $a_n = T_n^a/T$, where T_n^a is the duration for which the bike is available and not truncated in grid n, and T is the entire duration (in this study T is 1 h). Thus, the first part related to the truncation process can be expressed as,

$$\lambda_n^{o,tr} = a_n \lambda_n^r. \tag{5}$$

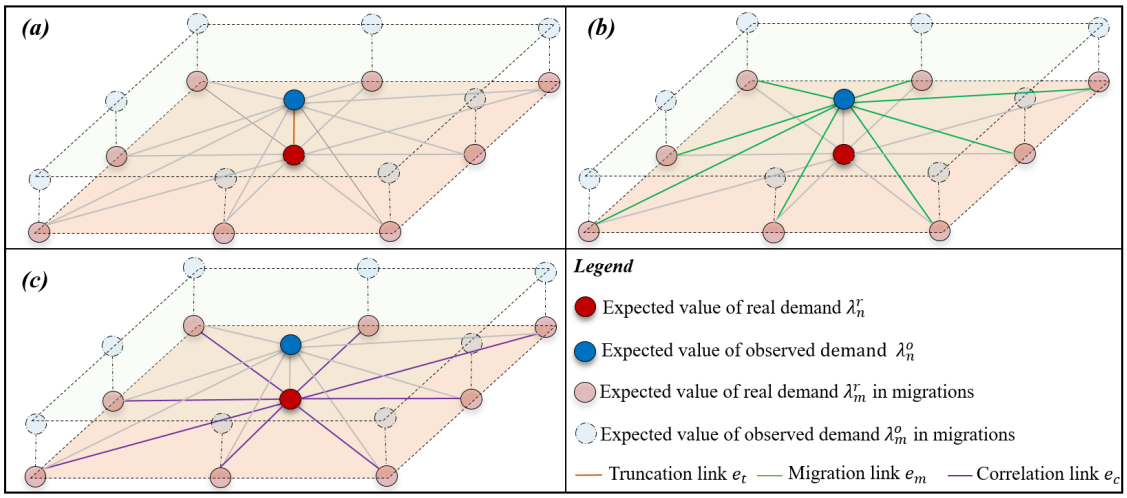


Fig. 2. Demand recovery considering (a) truncation, (b) migration, and (c) correlation.

The second part $\lambda_n^{o,mig}$ is related to the migration process. We choose one of the nearby grids $m \in N(n)$ for a later explanation. After demand in m is truncated, there are $\lambda_m^r(1-a_m)$ users cannot find a desired bike immediately, where a_m indicates the availability of m. Among those $\lambda_m^r(1-a_m)$ users, some of them may migrate to nearby grids to use the bike (i.e., case 1 in Fig. 3), whereas the remaining users leave the FFBS system and are counted as unmet demand (case 2 in Fig.3).

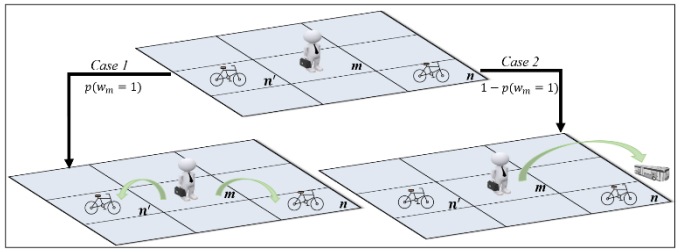


Fig. 3. User choices when the demand is truncated

In Fig. 3, $p(w_m = 1)$ represents the probability/willingness of migration. The w_m is a dummy variable, which takes 1 if the user in grid m decides to migrate to nearby grids when the first choice is truncated and 0 otherwise. In contrast, $1 - p(w_m = 1)$ in case 2 is the willingness of not migrate, which can be interpreted as the proportion of unmet demand with respect to the truncated demand. Consequently, the total expected demand spilled from m and captured by nearby grids is $\lambda_m^r(1-a_m)p(w_m = 1)$. Among all spillover demands, a proportion will be captured by grid $n \in N(m)$, and such proportion is denoted as $p(m \rightarrow n|w_m = 1)$. It is the conditional probability that a user migrates from m to n given that he or she decides still to use a bike. It is worth mentioning that $p(m \rightarrow n|w_m = 1)$ is zero if grid n has no bike across the entire study period. Thus, the expected demand migrates from m to n, $\lambda_n^{o,mig}(m)$, can be represented as,

$$\lambda_n^{o,mig}(m) = \lambda_m^r(1-a_m)p(m \rightarrow n|w_m = 1)p(w_m = 1). \quad (6)$$

To quantify $\lambda_n^{o,mig}(m)$, we need to model $p(w_m = 1)$ and $p(m \rightarrow n|w_m = 1)$. For $p(w_m = 1)$, it represents users' willingness of migration in the corresponding grid m, which is subjective information and can't be measured from the bike dataset. However, such kind of information can be surveyed from bike users and used as extra information when estimating the real demand (see *D* in this section).

As for $p(m \rightarrow n|w_m = 1)$, when users fail to find a bike at the desired grid, they will choose the bike at nearby grids. In this case, their goal is clear, i.e., to find the bike for trips. Compared with other factors (e.g., built environment, bike inventory level), they would only care about whether the bike is available or not. Thus, we assume that, as long as there is a bike inside a grid, this grid's attractiveness to users is the same as other nearby bike-available grids. Therefore, we can regard that the $p(m \rightarrow n|w_m = 1)$ is only related to whether the bike is available in nearby grids. To consider the influence of bike availability when calculating $p(m \rightarrow n|w_m = 1)$, we define the projected availability $pa_{m,n}$, which measures the duration in which the bike is available in n when no bikes are in grid m. The migrated demand is more likely to be captured by grid n if the corresponding $pa_{m,n}$ is higher. In other words, $p(m \rightarrow n|w_m = 1)$ is proportion to $pa_{m,n}$. Particularly, the $p(m \rightarrow n|w_m = 1)$ is calculated based on the dataset, which is macro information and not changeable for each user in $\lambda_m^r(1-a_m)p(w_m = 1)$. The calculation example of $pa_{m,n}$ is shown in Fig. 4.

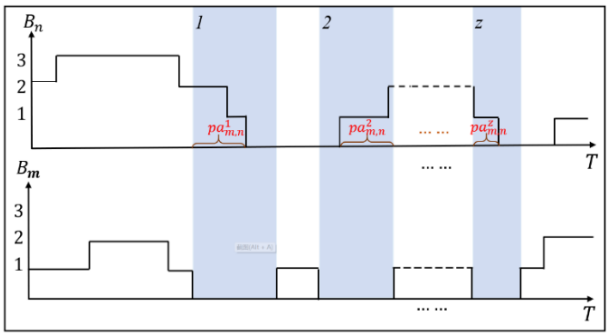


Fig. 4. Calculation of projected availability of nearby n

In Fig. 4, the B_n and B_m represent the numbers of available bikes in grid n and m, respectively. The duration of truncation of grid m is projected to demand-capturing grid n in the blue region. Then, $pa^z_{m,n}$ for grid n equals the z^{th} duration in which the bike is available in n but not available in m (i.e., $B_m = 0\ \&\ B_n \geq 1$), and $pa_{m,n}$ is the summation of $pa^z_{m,n}$, i.e., $pa_{m,n} = \sum_z pa^z_{m,n}$. The conditional migration probability $p(m \rightarrow n | w_m = 1)$ is proportional to $pa_{m,n}$, and can be represented as,

$$p(m \rightarrow n | w_m = 1) = \frac{pa_{m,n}}{\sum_{n' \in N(m)} pa_{m,n}} = \frac{\sum_z pa^z_{m,n}}{\sum_{n' \in N(m)} \sum_z pa^z_{m,n'}}. \quad (7)$$

By substituting (7) into (6), the demand migrating from m to n, $\lambda_n^{o,mig}(m)$, can be expressed as,

$$\lambda_n^{o,mig}(m) = \lambda_m^r (1 - a_m) \frac{\sum_z pa^z_{m,n}}{\sum_{n' \in N(m)} \sum_z pa^z_{m,n'}} p(w_m = 1). \quad (8)$$

Then, $\lambda_n^{o,mig}$ is the summation of all demands captured from nearby grids m, which can be expressed as,

$$\begin{aligned} \lambda_n^{o,mig} &= \sum_{m \in N(n)} \lambda_n^{o,mig}(m) \\ &= \sum_{m \in N(n)} \lambda_m^r (1 - a_m) \frac{\sum_z pa^z_{m,n}}{\sum_{n' \in N(m)} \sum_z pa^z_{m,n'}} p(w_m = 1). \end{aligned} \quad (9)$$

According to (4), (5), and (9), $\lambda_n^o = f(\lambda_n^r, \boldsymbol{\lambda_m^r})$ in (3) can be formulated as,

$$\begin{aligned} \lambda_n^o &= a_n \lambda_n^r + \\ &= \sum_{m \in N(n)} \lambda_m^r (1 - a_m) \frac{\sum_z pa^z_{m,n}}{\sum_{n' \in N(m)} \sum_z pa^z_{m,n'}} p(w_m = 1). \end{aligned} \quad (10)$$

Now, as $\lambda_n^o = f(\lambda_n^r, \boldsymbol{\lambda_m^r})$ is derived in (10), by substituting (10) into (3), the connection between the observed demand $y_{n,k}$ and the expected value of the real demand λ_n^r for one observation is established, which considers the demand truncation and migration process:

$$\begin{aligned} P\left(Y = y_{n,k} | \lambda_n^r, \boldsymbol{\lambda_m^r}\right) &= Psn(y_{n,k} | a_n \lambda_n^r + \\ \sum_{m \in N(n)} \lambda_m^r (1 - a_m) &\frac{\sum_z pa^z_{m,n}}{\sum_{n' \in N(m)} \sum_z pa^z_{m,n'}} p(w_m = 1)) \end{aligned} \quad (11)$$

The conditional term $P(\boldsymbol{Y} | \boldsymbol{\lambda^r})$ can then be formulated by multiplying the probability of each observation $y_{n,k}$,

$$\begin{aligned} P(\boldsymbol{Y} | \boldsymbol{\lambda^r}) &= \prod_k \prod_n P\left(Y = y_{n,k} | \lambda_n^r, \boldsymbol{\lambda_m^r}\right) \\ &= \prod \prod_k \prod \prod_n psn(y_{n,k} | a_n \lambda_n^r + \\ \sum_{m \in N(n)} \lambda_m^r (1 - a_m) &\frac{\sum_z pa^z_{m,n}}{\sum_{n' \in N(m)} \sum_z pa^z_{m,n'}} p(w_m = 1)). \end{aligned} \quad (12)$$

It should be pointed out that the demand in nearby truncation grids could migrate to the same grids. In other words, the migration grids would partially overlap if the truncation grids are located closely. Accordingly, $P\left(Y = y_{n,k} | \lambda_n^r, \boldsymbol{\lambda_m^r}\right)$ is not independent, and (12) somehow violates the independent condition. However, we keep this form for the following reasons. (i) Independent factorization is only satisfied by ignoring the migration process (i.e., no migration links in Fig. 2), but ignoring this process would leave out the important migration pattern of FFBS. (ii) We validated in Section IV that although the proposed model violates the independent condition to some extent, it is superior to that of the model satisfying the independent condition (i.e., the benchmark in (19), which contains no migration link).

C. Spatial Correlation in Prior Term $P(\boldsymbol{\lambda^r})$

We then model the prior term $P(\boldsymbol{\lambda^r})$, which considers the spatial correlation of the expected value of real demand. In this study, we assume that the demand λ_n^r is directly correlated with those in the nearest eight grids and has no direct correlation with the further grids. This assumption reflects the Markov property of the demand network and is widely used in the Markov random field. But it is noted that this assumption doesn't mean the grids have no correlation with the further grids. Instead, the further correlation is implicitly considered. For example, there is a grid *a* paired with the nearby grid *b*, and the demand is correlated. Meanwhile, grid *b* is also correlated with another grid *c,* which is a further grid for *a*. Through the mediate grid *b,* the correlation between grid *a* and further grid *c* is implicitly incorporated. In Fig. 2(c), for each pair of linked grids, we define the correlation function (or equivalently, the potential function in Markov random field) as follows,

$$P\left(e_c(\lambda_n^r, \lambda_m^r)\right) = \frac{1}{\sqrt{2\pi}} exp\left(-\frac{1}{2}\left(\frac{(ln(\lambda_n^r) - ln(\lambda_m^r))}{\sigma}\right)^2\right), \quad (13)$$

where σ is the "prior correlation parameter" in the prior term, which represents the extent of the spatial correlation. and $ln(.)$ is a logarithmic function, which is used to ensure that the expected demand λ_n^r is always positive. It is noted that (13) is not a normal distribution. We take its form since it is similar to the normal distribution and it satisfies the requirement of the potential function in MRF at the same time (i.e., to be always positive). In a MAP problem, the $P(e_c(\lambda_n^r, \lambda_m^r))$ will be maximized, which means that the values of $ln(\lambda_n^r)$ and $ln(\lambda_m^r)$ will be as similar as possible. Consequently, the increase in the real demand of one grid will make an increase in the real demand of another. Thus, an implicit positive correlation is incorporated. It is worth mentioning that the demand in nearby areas could also be substantially different as it highly depends on the built environment. But in this study, we divide the area into small grids with a length of 100 m. Thus, the built environment is similar and there would be few cases of substantially different demand levels. But then again, even if there is a substantial difference of demand resulting from different built environments, it will be implicitly captured by the observed demand and bike availability, which are then considered in the conditional term. As for the prior term, it should, by definition, only incorporate the prior knowledge. It is common sense that the demand in neighboring grids is positively correlated as those grids are close. In contrast, the

built environments in different locations are viewed as extra information, thus we think there is no need to consider it in the prior term.

In addition, as for σ, it reflects the operators/researchers' sense of the spatial correlation extent. As its name implies, the σ should be specified in advance as the prior information when estimating the demand. Although we qualitatively know from (13) that a high value of σ indicates a low spatial correlation and vice versa, the exact values for the different extents of spatial correlation are unknown. Thus, we need to first match the value of σ with the different extents of spatial correlation quantitatively. Then, the matched σ can be used as the reference list, and the operators can select appropriate σ based on their prior sense of spatial correlation. To this end, we conduct simulation studies, in which the specific extent of spatial correlation is first generated. Then, the exact value of σ can be matched to the given extent of spatial correlation. The example of matching the σ value is shown in Section IV.*C*.

Because each pair of points can be regarded as a clique in the demand network, according to the Hammersley–Clifford theorem in a Markov random field [39], the distribution of correlation links in the demand network can be independently factorized into the correlation function (or potential function) for each clique; thus, $P(\boldsymbol{\lambda}^r)$ can be formulated by multiplying all the correlation functions directly,

$$P(\boldsymbol{\lambda}^r) = \frac{1}{H_\sigma}\prod_{e_c} P(e_c) = \frac{1}{H_\sigma}\prod\nolimits_{e_c}\frac{1}{\sqrt{2\pi}} exp\left(-\frac{1}{2}\left(\frac{(l\,n(\lambda_n^r) - l\,n(\lambda_m^r))}{\sigma}\right)^2\right), \quad (14)$$

where H_σ is the normalizing constant and is usually called the partition function, it is added to ensure the range of (14) is between 0 and 1, which is consistent with the probabilistic model. The value of H_σ depends on the σ.

Finally, the log-likelihood function (objective function) for the real demand recovery under the DRTMC framework is formulated in (15). The truncation and migration are considered through the conditional term $P(\boldsymbol{Y}|\boldsymbol{\lambda}^r)$ and correlation is considered in $P(\boldsymbol{\lambda}^r)$,

$$LL = log\left(P(\boldsymbol{\lambda}^r)P(\boldsymbol{Y}|\boldsymbol{\lambda}^r)\right) = log\left\{\begin{array}{c}\frac{1}{H_\sigma}\prod_{e_c}\frac{1}{\sqrt{2\pi}} exp\left(-\frac{1}{2}\left(\frac{(l\,n(\lambda_n^r)-l\,n(\lambda_m^r))}{\sigma}\right)^2\right)\times \\ \prod_k\prod_n Psn\left(y_{n,k}\middle|\, a_n\lambda_n^r + \sum_{m\in N(n)}\lambda_m^r\right) \\ \times (1-a_m)\frac{\sum_z pa_{m,n}^z}{\sum_{n'\in N(m)}\sum_z pa_{m,n'}^z} p(w_m=1)\end{array}\right\}. \quad (15)$$

D. Demand Recovery

The real demand, observed demand, unmet demand, and migration demand are strongly related. Each of these four types of demands can be calculated if the other three are known. Commonly, the data obtainable for estimation are observed demand data. However, as proven by Vulcano et al., the real demand, unmet demand, and migration (substitution) demand cannot be estimated together based on the observed demand alone without extra information [25]. Therefore, in their study, the proportion of unmet demand was specified as an extra constraint to avoid multiple optimal results. By following the study [25], we assume the willingness of migration, i.e., $p(w_m = 1)$, can be accurately obtained through a questionnaire survey. Thus, the value of $p(w_m = 1)$ is specified as extra information when estimating the $\boldsymbol{\lambda}^r$. The $p(w_m = 1)$ could be different for each grid. However, the survey of $p(w_m = 1)$ for different grids is beyond the scope of the current study. For ease of calculation, we assume the $p(w_m = 1)$ in all grids happen to be the same from the survey and use one value to represent it. But in the real world, the operators could conduct a field survey. In this way, the heterogenous willingness of migration in different grids can be considered through the DRTMC model.

In general, when using the Markov random field technique, several classical solutions exist, such as Gibbs sampling, iterated conditional mode, min-cut/max-flow, expansion moves, and simulated annealing (SA) [40], [41]. Because the proposed model is not a standard Markov random field and the demand is coupled owing to the demand migration process, the common methods are not applicable here. Therefore, we adopt SA for demand recovery. The SA method has been widely applied in the transportation field to obtain approximate global solutions. It is a metaheuristic approach that mimics the annealing process in metallurgy [42], [43], [44], [45].

We illustrate how the SA is applied to the DRTMC problem in the context of FFBS. First, we classify the grids into two types: real demand grid and non-real demand grid. The former refers to the grids in which truncation does not occur, and it does not capture the migration demand as well (i.e., the demand is also not truncated in the nearby grids). Thus, for the first type, the expected value of its real demand can be directly represented as the average value of observed demand across K days and is not updated during the estimation process. For the second type of grid, which experiences the demand truncation or its nearby grid experiences truncation, the real demand is not equal to the observed demand and needs to be estimated by SA.

To implement SA for demand recovery, several parameters should be specified. **Temperature scheme:** The temperature scheme is set to follow a geometric pattern for consecutive iterations, i.e., $TP^q = \alpha TP^{q-1}$, where α is the cooling rate and TP^q is the temperature in iteration q. When the initial temperature TP^0 gradually decreases and reaches a predefined final temperature TP^f, the iteration stops. The values of TP^0, TP^f, and α are always empirically set based on the problem context, and in this study, they are empirically set as $TP^0 = 20$, $TP^f = 0.01$, and $\alpha = 0.95$ after pre-testing.

Initial solution and new solution generation: For initial solution $\{\lambda_1^{r,0}, \lambda_2^{r,0}, \dots, \lambda_n^{r,0}\} \in \boldsymbol{\lambda}^{r,\mathbf{0}}$, instead of using the average of observed demand $\{\bar{y}_1, \bar{y}_2, \dots, \bar{y}_N\} \in \bar{\boldsymbol{Y}}$ directly, we modify $\bar{\boldsymbol{Y}}$ according to the availability indicator a_n to reduce the truncation effect. The value of the average observed demand $\bar{y}_n$ is divided by its availability indicator a_n; that is, $\lambda_n^{r,0} = \bar{y}_n/a_n$. For new solutions at the z^{th} transition $\{\lambda_1^{r,z}, \lambda_2^{r,z}, \dots, \lambda_N^{r,z}\} \in \boldsymbol{\lambda}^{r,z}$, it evolves randomly from the solution $\boldsymbol{\lambda}^{r,z-1}$ at the last transition. In particular, for n, $\lambda_n^{r,z} = exp(\xi) * \lambda_n^{r,z-1}$, where ξ

is a term that follows a normal distribution with $\mu = 0$ and $\sigma_{SA} = 0.1$. The $exp\{...\}$ function ensures that the real demand is always positive. The process of new solution generation is only applied to the second type of grid in which the real demand needs to be updated.

Acceptance probability: When a new solution is generated, SA decides whether to accept the new solution. If the generated solution is better (i.e., the value of the objective function is higher), it is accepted; otherwise, it is accepted with a defined acceptance probability. The acceptance probability is specified using the Metropolis criterion $\Delta LL/\ TP^q$, where ΔLL is the difference between the objective values of the current and new solutions and TP^q is the current temperature. In addition, for the objective function, i.e., (15), the normalized term H_σ is constant with regard to different $\boldsymbol{\lambda}^r$ and can be ignored when calculating the objective function and the difference ΔLL.

With these implementation points, the SA repeatedly generates a new solution and decides whether to accept it. At each temperature, SA generates new solutions 50 times (i.e., $Z = 50$). When the temperature is high, the SA is more likely to accept the degrading solution, ensuring that the algorithm is trapped out of the local maxima. As the temperature gradually decreases, the system converges to a near-global or global optimum. The pseudocode of the cooling procedure for SA is shown in Algorithm 1 in Appendix *B*.

IV. The Validations In Simulation Study

Because the real demand for FFBS is not obtainable in the real world, we have no idea about the real demand information, and the model accuracy can't be directly evaluated. However, to ensure that the model is efficient and the estimated real demand is reliable, we first validate the model with a synthetic dataset obtained from the simulation, where the actual real demand can be obtained for comparison. Using the synthetic dataset for validation is a common approach to test the performance of the real demand estimation model for the shared travel modes [22], [23], [24]. In addition, another task of simulation is to match the value of σ with the specific extent of spatial correlation quantitatively, which can be used as the reference for the real-world case.

In the simulation, we first generate the expected value of real demand λ_n^r directly, which is regarded as the true value and stored as the reference for the model evaluation. Then, based on the λ_n^r generated before, the disaggregate pick-up and drop-off process for each day is simulated. During the simulation process, the interaction between the supply and demand is proceeded and the FFBS usage information is recorded. The output of the simulation in each grid contains the number of successful pick-ups and the available duration of bikes. It is noted that the output contains no real demand information since the real demand is not obtainable as well in the real world. Based on the outputs, the DRTMC model is applied to estimate real demand $\widehat{\lambda_n^r}$. Finally, the estimated result $\widehat{\lambda_n^r}$ is compared with the true value λ_n^r setting at the beginning. In this way, the model performance is evaluated.

The simulator for synthetic data generation is presented in detail in Section *A*, followed by the benchmarks in Section *B*. Section *C* illustrates the example of matching prior σ with a specific extent of spatial correlation. Finally, the validation results are shown in Section *C*.

A. Synthetic Dataset

A simulator is developed to synthesize the observed demand. The number of grids in the simulation is 196 (14 × 14). The simulation involves four steps, which are summarized in Fig.5.

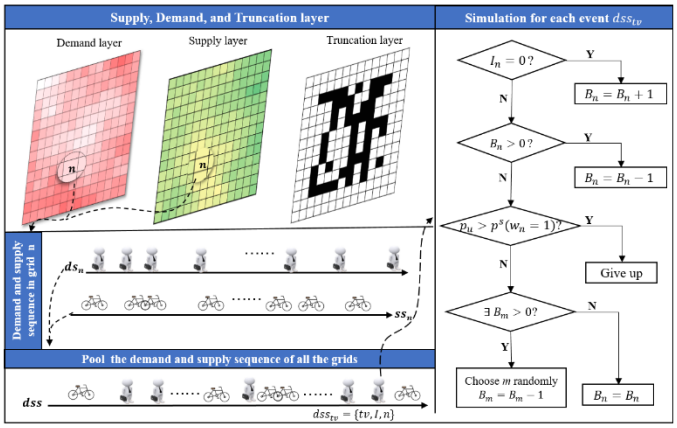


Fig. 5. Workflow of the FFBS simulator

In the first step, the expected value of real demand in each grid λ_n^r is generated. It is worth emphasizing that the λ_n^r is spatially correlated, and such patterns should be kept in the simulator. To this end, the "*gstat*" package in R is adopted to generate spatially correlated demand with the "*sph*" method (gstat: Spatial and Spatio-Temporal Geostatistical Modelling, Prediction and Simulation (r-project.org)), and we scale the demand λ_n^r in grids from 0 to 30. Similarly, the supply of FFBS in each grid, s_n, is also generated and scaled from 0 to 30. It is noted that, to reflect the imbalance of demand and supply patterns, λ_n^r and s_n are generated with different seeds, and their values are different. Directly generating the λ_n^r and s_n simplifies the simulation since it ignores the completed bike usage process (i.e., user picks up the bike at grid n and drops it off at another grid). However, the focus of this study is only the demand variation in each grid and there is no need to consider the demand flow in the whole FFBS system, Thus, this simplification is reasonable [36].

Then, we need to set the initial number of bikes in each grid. We could have generated a random initial number of bikes within a reasonable interval for each grid (e.g., 10-20). But in this way, the fraction of grids experiencing the demand truncation, which is denoted as $frac \in [0,1]$, will be different in each simulation. Consequently, the model can't be fairly evaluated as the comparisons are not conducted under the same situation. To address this issue and control the $frac$ the same, we first determine the value of $frac$ in advance. Based on $frac$, the grids experiencing demand truncation are randomly selected, which is shown as the truncation layer in Fig.5. For the selected grids, the initial number of available bikes is set to a small value (i.e., a random integer from 0 to 3). For the non-selected grids,

the number of available bikes was set to a large value (i.e., 100), so the demand would never be truncated when lots of bikes are prepared. Thus, only the selected grids would experience demand truncation. The distribution of the initial number of bikes set for the non-selected grids is extreme. However, it won't affect the model validations since only the observed demand and bike availability information are needed for demand estimation, the number of bikes remaining at non-selected grids after the simulation has no influence.

In the second step, by assuming that the actual demand and supply follow a Poisson distribution, the demand sequence $\boldsymbol{ds_n}$ and supply sequence $\boldsymbol{ss_n}$ in each grid were randomly generated with parameters λ_n^r and s_n, respectively. Here, the demand and supply sequences refer to the arrival of users and bikes, respectively.

In the third step, we pool the demand and supply sequence of all the grids and rearrange them in the time order to obtain the demand-supply sequence, which is represented as $\boldsymbol{dss} = \{dss_{t1}, dss_{t2}, \dots, dss_{tv}, \dots, dss_{tV}\}$ in the study area. Each dss_{tv} represents an "event," with information of {event time: tv; drop off or pick up indicator: I; grid index: n}. Indicator I is a dummy indicator with $I = 0$ meaning that the bike is dropped off, and $I = 1$ meaning that the bike is picked up.

In the final step, we process the demand-supply sequence $\boldsymbol{dss}$ in time order. For the v^{th} event dss_{tv}, we first judge its type (pick up or drop off); if dss_{tv} is a drop-off event, the number of bikes B_n in the corresponding grid n will increase by one. In contrast, when dss_{tv} is a pickup event, the user will pick up the bike directly if the bike is available ($B_n > 0$); otherwise, the user will decide whether to continue using the bike system based on the set willingness $p^s(w_n = 1)$ in the simulation. In detail, a uniformly distributed variable $p_u \in [0.1]$ is randomly generated for each user, which represents the disaggregate migration willingness. Then the generated p_u is compared with $p^s(w_n = 1)$. If a user decides still to use a bike, i.e., $p_u \leq p^s(w_n = 1)$, he or she will move to a nearby grid $m \in N(n)$ to pick up a bike if the bike is available; otherwise, he or she will decide to use another travel mode, i.e., $p_u > p^s(w_n = 1)$. For the former case, if there are multiple grids with available bikes, the user will choose the grid randomly because the attractiveness of grids is assumed to be the same. It is worth mentioning that in the simulation, the availability information (whether there is a bike in grids) dynamically changes as the behavior of each user is considered at the micro level, thus the candidate grids for migration may be different for different users. Particularly, the simulation proceeds with randomness and the output varies. When a bike is picked up in grid n or m, the number of available bikes B_n or B_m will decrease by one. If the user gives up using the bike system, the B_n will not change. Each time when B_n is updated, the information will be stored in the list $res_n = \{(B_{n,1}, t_1), (B_{n,2}, t_2), \dots, (B_{n,p}, t_P)\}$, where $B_{n,p}$ is the number of bikes in grid n at the time t_P and t_P is the time at which the pick-up or drop-off occurs.

There are two parameters determining the simulation scenarios: the percentage of grids experiencing demand truncation $frac$, and the willingness to migrate to nearby grids when bikes are not available, $p^s(w_n = 1)$. When the scenario is set after specifying $frac$ and $p^s(w_n = 1)$, the simulator can be run and the output of the simulator is res_n for each grid. From the output, we calculate the observed demand $y_{n,k}$ and the duration of availability a_n in each grid, which could be used for real demand estimation.

B. Benchmark Comparison

To test the model performance, we compare it with the following four benchmarks:

b1: Observed Pickups (OP), which ignores the demand truncation, migration, and spatial correlation process. As its name shows, it regards the observed demand as the real demand for each grid. This benchmark is selected to illustrate the consequence of directly treating the observed demand as real demand.

$$\lambda_n^r = \bar{y}_{n,k}. \tag{16}$$

b2: Simulation-based Optimization Approach (SOA), adopted from [36]. It is a data-driven optimization to infer the real demand of bikes. The SOA in [36] was originally designed for station-based bike-sharing, and we used it here for the context of FFBS. In detail, each grid is viewed as a station, of which the real demand will be estimated by SOA. The model is shown in (17). The $BAPT^O$ and $BAPT$ are the indicators extracted from the observed data and simulated data, respectively. $F(.)$ denotes the distribution pattern of the indicators. The $simu(.)$ is the simulation function generating the indicators $BAPT$ with different values of $\lambda_n^r \in \mathbb{S}$. The $simu(.)$ can be modified from Section *A* directly. The objective is to minimize the difference between $F(BAPT)$ and $F(BAPT^O)$, which is measured by the double-bootstrap test. The optimal λ_n^r with the minimal objective value is selected as the estimated real demand. The SOA in [36] incorporates the impact of the demand truncation (censoring) but ignores the demand migration and spatial correlation.

$$\begin{aligned} & MIN\ |F(BAPT) - F(BAPT^O)| \\ s.t. \quad & BAPT = simu(\lambda_n^r, \dots) \\ & \lambda_n^r \in \mathbb{S} \end{aligned} \tag{17}$$

b3: Demand Recovery Considering Truncation and Migration with No spatial Correlation (DRTM)*,* this model ignores the spatial correlation of the FFBS demand while considering the demand truncation and migration process. Thus, the objective function can be written as,

$$LL_{b3} = log\left\{ \begin{array}{l} \prod_k \prod_n Psn\left(y_{n,k} \middle| a_n \lambda_n^r + \sum_{m \in N(n)} \lambda_m^r \right. \\ \left. \times (1 - a_m) \frac{\sum_z pa_{m,n}^z}{\sum_{n' \in N(m)} \sum_z pa_{m,n'}^z} p(w_m = 1) \right) \end{array} \right\}. \tag{18}$$

b4: Demand Recovery Considering Truncation and spatial Correlation with No Migration (DRTC)*,* it considers the demand truncation process and incorporates the spatial correlation of demand but ignores the demand migration. The objective function is as follows:

$$LL_{b4} = log\left\{\frac{1}{H_\sigma}\prod_{e_c}\frac{1}{\sqrt{2\pi}}\exp\left(-\frac{1}{2}\left(\frac{ln\lambda_n^r - ln\lambda_m^r}{\sigma}\right)^2\right)\times \prod_k\prod_n Psn\left(y_{n,k}\middle|\, a_n\lambda_n^r\right)\right\}. \quad (19)$$

Moreover, to evaluate the performance of the proposed model compared with the benchmarks, the root-mean-squared error (RMSE) is adopted:

$$RMSE = \sqrt{\frac{1}{N}\sum_{i=1}^{N}\left(\widehat{\lambda_n^r} - \lambda_n^r\right)^2}, \quad (20)$$

where N is the number of grids and $\widehat{\lambda_n^r}$ and λ_n^r are the expected values of the estimated demand and real demand, respectively. In the validation, five indicators are used: $RMSE_{OP}$, $RMSE_{SOA}$, $RMSE_{DRTM}$, $RMSE_{DRTC}$, and $RMSE_{DRTMC}$ for the OP, SOA, DRTM, DRTC, and DRTMC methods, respectively. The benchmarks together with the DRTMC model follow an incremental process to consider the demand truncation, migration, and spatial correlation. In this way, the significance and the impact of incorporating truncation, migration, and spatial correlation can be revealed.

C. Matching the Value of σ with the Specific Extent of Spatial Correlation in Simulation

As mentioned before, σ is not determined arbitrarily as it has a physical meaning. Instead, it is learned from the given extent of correlation generated in the simulation. Fig. 6 shows the example of the real demand distribution generated in the simulation, it is obvious that the real demand is spatially correlated. We will search for the value of σ to match the extent of spatial correlation in Fig. 6.

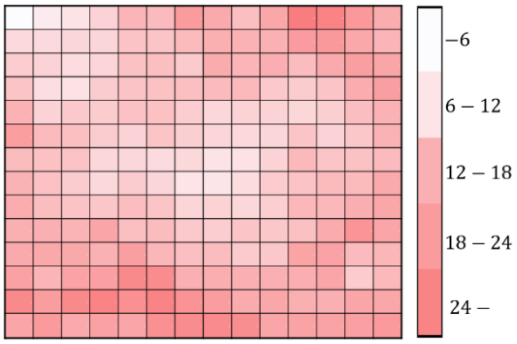


Fig. 6. Distributions of FFBS real demand in simulation

The idea is to search for the optimal value of σ, under which the $RMSE_{DRTMC}$ is minimized. In other words, other than the real demand parameters, the σ will be estimated as well in this part. However, we didn't estimate the demand parameter and σ jointly. This is because the term $\frac{ln(\lambda_n^r) - ln(\lambda_m^r)}{\sigma}$ in (15) makes the simultaneous estimation unstable and difficult. Instead, we used a line search to find the optimal σ value. In detail, we gradually increase the value of σ, the DRTMC model is then applied to estimate the real demand under each value of σ. After that, the $RMSE_{DRTMC}$ is calculated based on the estimated real demand. Finally, we select the σ with the minimal value of $RMSE_{DRTMC}$. To run the simulation, we use a specific case with $frac = 0.5$ and $p^s(w_n = 1) = 0.5$. It is worth mentioning that the σ is only related to the extent of correlation theoretically. Thus, it is not necessary to conduct more experiments with other values of $frac$ and $p^s(w_n = 1)$. For the set scenario, the simulation generates data for five days; that is, $K = 5$. The proposed model is then applied to the synthesized data to estimate the expected value of real demand. In the estimation, the value of $p(w_n = 1)$ needs to be specified in advance. Here, we set $p(w_n = 1) = p^s(w_n = 1)$ providing that we could accurately obtain the willingness of migration through the questionnaire survey. The search results are shown in Fig. 7. According to Fig. 7(a), the optimal value of σ is 1.5 with $RMSE_{DRTMC}$ = 0.79. It indicates that, if the extent of spatial correlation in the real world is similar to that shown in Fig. 6, the most suggested value of σ for operators is 1.5. In addition, Fig. 7(b) shows iterations of SA when $\sigma = 1.5$. The objective value increases from –12,723 to –11,394. After approximately 5,000 iterations, the objective value becomes stable and converges.

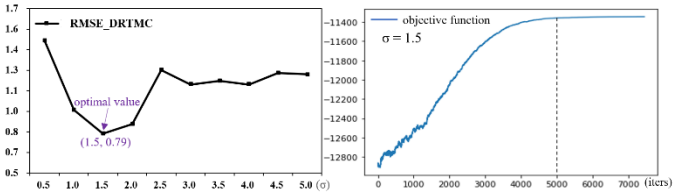


Fig. 7. (a) Matching the σ for the given extent of spatial correlation and (b) Estimating the expected value of real demand with SA

To illustrate this result further from the perspective of spatial correlation, we present the observed demand and estimated real demand for each grid in Fig. 8. The distribution of the observed demand is significantly different from that of the real demand. However, the distribution of the estimated real demand in Fig. 8(b) is much closer to that of the real demand, regardless of the spatial pattern or demand level. It indicates the searched σ value matches the extent of spatial correlation in Fig. 6 well. With similar procedures for other extents of spatial correlation, the reference list of σ can be made for the operators. Moreover, the superiority of the proposed model is confirmed as the $RMSE_{DRTMC}$ of the estimated demand is 0.79, which is 80.3% lower than the $RMSE_{OP}$ with 4.02.

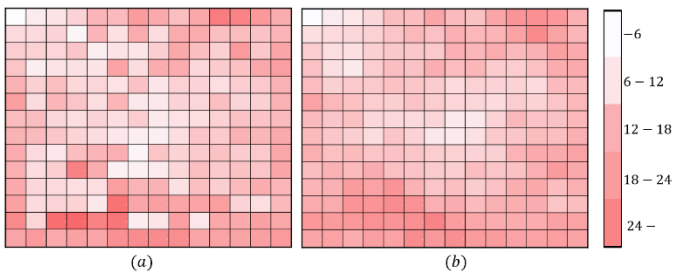


Fig. 8. Distributions of (a) observed demand ($RMSE_{OP}$ = 4.02), (b) estimated real demand ($RMSE_{DRTMC}$ = 0.79).

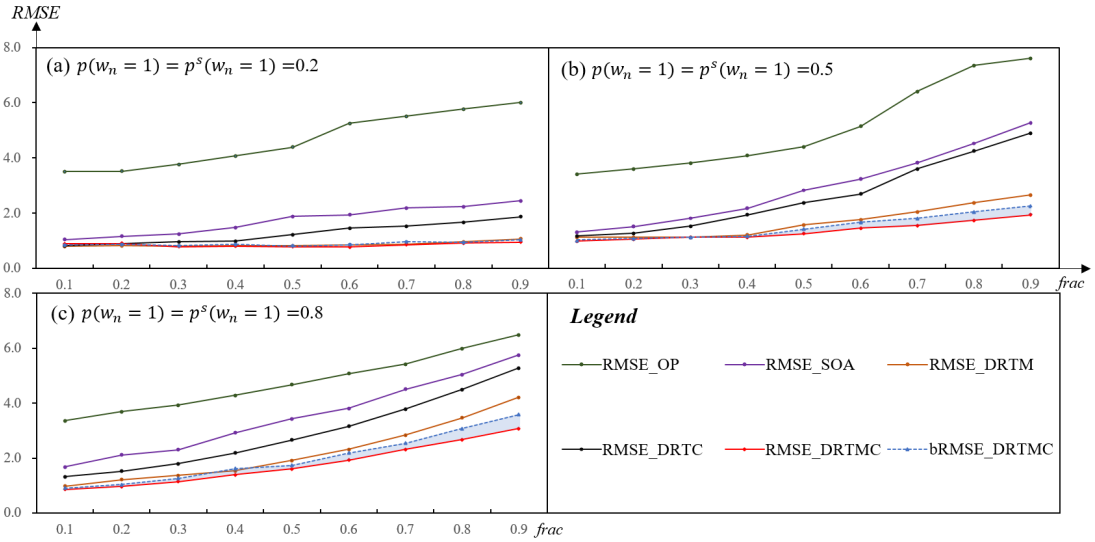


Fig. 9. Validation results under different scenarios

D. Validation Results

In this section, the model is validated in several scenarios by comparing the performance with those of the benchmarks using RMSE value. Particularly, we first choose $p^s(w_n = 1)$ to be 0.2, 0.5, and 0.8 for the low, middle, and high willingness of migration, respectively. Under each level of $p^s(w_n = 1)$, the percentage of grids that experience demand truncation increases from 10% to 90%.

In addition, different from above Section C, the σ in validation is specified in advance as prior knowledge here, which is the same as the real-world case. In detail, σ is first set as 1.5 providing that the operators are experienced in perceiving the extent of spatial correlation. Moreover, we also consider that σ is not 1.5 in case the FFBS operators have a biased sense of the spatial correlation. To this end, a 30% bias is considered. Accordingly, the value increases by 30%, i.e., $\sigma = 1.5 \times 1.3 = 1.95$, and decreases by 30%, i.e., $\sigma = 1.5 \times 0.7 = 1.05$, are adopted. The values far away from 1.5 are not considered since the matched value from the above Section *C* could provide the operators with a fair reference, so that the bias can be controlled within an acceptable rate. The result estimated with the biased σ would be inferior to the perfect prior σ. To see how the model performs when the spatial correlation is not perceived well, the worst RMSE value among the two results estimated with upward and downward biased priors σ is selected and denoted as $bRMSE_{DRTMC}$, which will be presented in the validation as well.

To avoid the effects of randomness, the estimation under each scenario is conducted 5 times, and the values are averaged. The validation results of the proposed DRTMC model, together with the benchmarks, are presented in Fig. 9. The conclusions are as follows. First, the proposed DRTMC model (red line) performs best in all scenarios, followed by benchmarks DRTM, DRTC, SOA, and OP. This finding demonstrates the superiority of the proposed model and that ignoring the demand truncation, migration, or spatial correlation impedes model performance. Second, the accuracy of the OP model (green line) is the lowest, indicating that directly treating the observed demand as real demand is not appropriate and will result in a high bias even when the fraction of demand truncation and willingness of migration are low. Third, both the parameters $frac$ and $p(w_n = 1)$ affect the model accuracy. As $frac$ and $p(w_n = 1)$ increase in different scenarios, the accuracy of all models decreases. However, unlike the other models, the increase in $RMSE_{DRTMC}$ only become significant under extreme conditions (i.e., $p(w_n = 1) \geq 0.8, frac \geq 0.7$), which rarely happens in the real world. Fourth, apart from the proposed DRTMC model, the performance of DRTM is also better than those of DRTC, SOA, and OP, which emphasizes the importance of considering demand migration when recovering the real bike-sharing system demand, but it has been almost completely ignored in previous studies. Finally, compared with the result with perfect prior σ, the model performance degrades when the biased σ is adopted, which is shown in the blue dash line ($bRMSE_{DRTMC}$). From the experiment, we notice that the worst RMSE values in the blue dash line come almost from the downward bias (i.e., $\sigma = 1.05$). It indicates the model is more sensitive in the high spatial correlation cases. The extent of degradation is illustrated in the blue area. This comparison indicates the necessity of choosing the appropriate prior spatial correlation parameter. The FFBS operators should be experienced in perceiving the extent of real demand spatial correlation. Otherwise, the performance of the DRTMC model will degrade or be even worse than other models if an unrealistic value of prior σ is specified.

V. Case Study

In this part, the model is applied to the central area of Shanghai city in real-world scenario. The Section *A* introduces data preparation, followed by the case study in Section *B*.

A. Preparing the Dataset for Case Study

We present a case study using the bike trip data from the Mobike company in Shanghai City. It is noted that there are multiple FFBS operators other than the Mobike company in this

city. Consequently, the users could pick up the bike from other companies when Mobike is unavailable. But it is expected that the government could have access to the whole dataset from all companies, so the completed dataset can be obtained by pooling those individual datasets. The data structure of the pooling dataset is the same as that of the Mobike dataset since the recording systems of all FFBS companies are the same. In this case, the proposed DRTMC model can be used for the whole FFBS system as well if it can be successfully applied in the Mobike system. In this study, we assume that there is only one FFBS provider and then apply the proposed model. Thus, the estimated real demand of Mobike can be regarded as the actual real-world demand of the FFBS system. As for the period used for estimation, the evening peak hours on weekdays were chosen. This is because the usage demand in the evening peak hour is higher and the demand truncation is more likely to happen, which results in higher unmet demand than other periods. Thus, the pickup and drop-off records of FFBS from 18:00 to 19:00 pm, ranging from September 2 to September 6, 2020, on five weekdays are selected.

As for the study area, we selected a prosperous location in the Huangpu district, which is near the center of Shanghai and covers around 2 km^2. The area is shown in Fig.10 (a), which is divided into 196 (14×14) grids, and the size of each grid is 100×100m. The grids that experienced demand truncation during the study period are indicated in purple. This study area includes two primary schools, a park (Gucheng Park), tourism spots (City God Temple and Yu Garden), and several shopping malls and residential locations. Since it is a prosperous area with a higher bike supply, the case that no bike is available in all eight nearby grids rarely happens. It shows that all the grids have the choice to migrate to a nearby area and don't need to go further, so Assumption 2 stated in the problem description part could be satisfied. In addition, we can get the fraction of truncation $frac$ in this case study is 15.3% (i.e., $frac = 30/196$). The average observed demand (successful pickups) is also illustrated in Fig. 10(b).

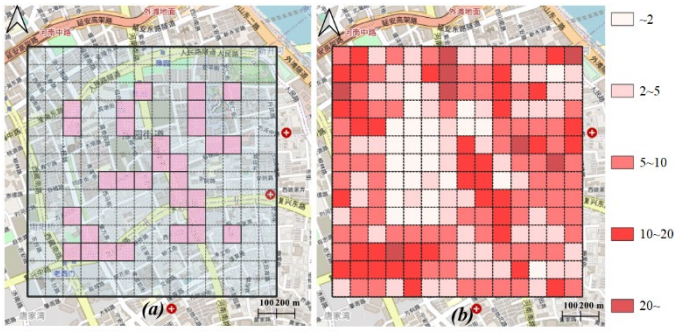


Fig. 10. (a) Chosen area and truncated grids and (b) Observed demand distribution in the real-world case study.

But it is noted that the truncation information in Fig.10 (a) (i.e., whether the demand is truncated) cannot be accessed from the raw dataset directly because of the limitation of the data structure of the FFBS system. Typically, each record in the FFBS dataset contains the unique ID of a bike, pick-up and drop-off times, pick-up or drop-off location (i.e., longitude and latitude), and lock status (i.e., being picked up or dropped off). Based on the above information, we obtain the number of pickups and drop-offs in different areas. However, the number of available bikes B_n remaining in grid n at any time tv is unknown. In other words, it is unknown whether the demand is truncated at time tv or not. To overcome the problem of no truncation information, Gammelli et al. regarded the historical observed demand as real demand, they manually set the truncation information [22]. However, this technique is inappropriate here for the FFBS system when considering the spatial correlation, and the reasons are as follows. The bike-sharing system studied in [22] is station-based, and they further gathered the stations into three hubs, which are far away from each other. Thus, the spatial correlation among hubs is ignored and they can be treated independently. However, here the demand is spatially correlated in the FFBS system. Accordingly, the truncation information should also be correlated. To retain such an intrinsic pattern, we could not generate truncation information manually or randomly. Thus, we proposed an algorithm to approximate the actual number of available bikes in grid n at time tv, which is shown in Appendix *C*.

B. Recovering the Real and Unmet Demand

After processing the data, we estimate the real demand using the proposed DRTMC model. The prior σ is set to 1.5 for illustration. Three values of $p(w_n = 1)$ representing low (0.2), middle(0.5), and high (0.8) willingness to migrate are adopted for the estimation. As a result, the estimated real demand in the whole area is 1614.3, 1570.7, and 1533.8 for the low, middle, and high willingness of migration, respectively. Compared with the average observed demand of 1443.3, the hourly unmet demand in the study area is 171.0, 127.4, and 90.5, respectively. It is noted that the unmet demand is lower when w_n is set to a higher value. This is consistent with common sense. For users who cannot pick up bikes immediately but have a high willingness to perform the migration, the demand is satisfied easily and consequently the unmet demand is low. In addition, for detailed illustration, the frequency distributions of the observed and estimated real demands are presented in Fig. 11. The difference between the distribution of the observed demand and that of the estimated real demand is significant. Regardless of the value of $p(w_n = 1)$, the frequency of observed demand in the 0–2 interval is much larger than that of estimated real demand. However, when the demand value is larger than 4, the frequency of estimated demand is almost larger than that of observed demand. Thus, estimating the real demand and distinguishing it from the observed demand is important.

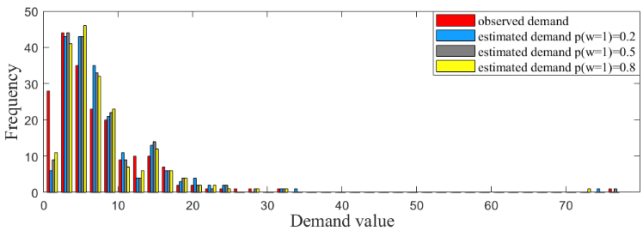


Fig. 11. Distribution of estimated real demand.

The purpose of estimating the real demand is to figure out the unmet demand in grids, which is vital information for FFBS management. To this end, we quantify the value of unmet demand u_n in each grid by calculating the difference between the estimated real demand and observed demand, i.e., $u_n = \lambda_n^r - \bar{y}_{n,k}$. We choose $p(w_n = 1) = 0.5$ for later presentation. Moreover, to illustrate the influence of spatial correlation, another value of σ representing relatively low spatial correlation is adopted as well ($\sigma = 5$). The distribution of unmet demand is shown in Fig. 12 (a) and (b). Grids with $u_n = 0$ are real demand grids, defined in Section *Demand Recovery*, their observed demands are equal to the real demand. It is noted that u_n could also be a negative value, which means that the estimated real demand is lower than the observed demand in this grid. This is because the observed demand in this grid captures not only the demand generated from itself but also the demand spilled over from other grids, resulting in a higher observed value and a negative value of u_n.

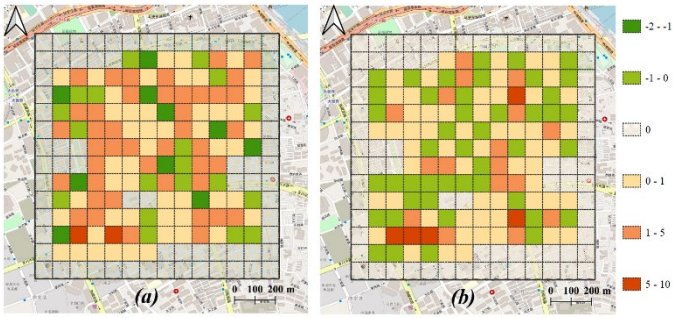


Fig. 12. Distributions of unmet demand under relative (a) high correlation ($\sigma = 1.5$) and (b) low correlation ($\sigma = 5$)

Compared with Fig. 12(b), the unmet demand in Fig. 12(a) shows an obvious spatial correlation. It indicates that the unmet demand tends to show a spatial correlation as well when the real demand is assumed to be highly correlated, which is consistent with common sense. However, as mentioned before, the spatial correlation is considered in the prior term and the parameter σ is determined by the operators' experience. In other words, if the operators regard the demand here is highly spatial correlated, a relatively small σ will be set, otherwise a higher σ will be used for demand estimation. Thus, we can't say which result is better. Instead, it requires the operators to be experienced in choosing suitable correlation parameters as prior information.

With the calculated unmet demand u_n, FFBS operators can determine whether they need to distribute bikes in a grid. No matter whether in high or low spatial correlation, the unmet demand distribution provides important guidance for the bike rebalance. For example, as shown both in Fig. 12 (a) and (b), the unmet demand in the bottom-left area is relatively high and the operators should distribute more bikes here. In addition, among the middle areas in Fig. 12 (b), the grids capturing the nearby demand ($u_n < 0$) are close and scattered among the grids with demand spillover ($u_n > 0$). Such an unbalanced demand is suitable for developing self-organized rebalance strategies [46]. Operators could incentivize users to go to nearby grids proactively to participate in the bike rebalance. Thus, they can avoid using the traditional trunk-based rebalance strategy, requiring hiring staff to redistribute bikes.

VI. Conclusion

In this study, we propose a framework for demand recovery considering truncation, migration, and spatial correlation (DRTMC) to estimate the real FFBS demand. First, the demand recovery is modeled as a maximum a posteriori (MAP) problem, and the problem is transferred to formulate the prior and conditional terms. Then, we establish the prior term by analyzing the spatial correlation of the demand and the conditional term by analyzing the demand truncation and migration process. Finally, a tailored simulated annealing approach is adopted to recover the real FFBS demand.

We then presented the results using synthesized data for validation and real data for empirical analysis. The validation results indicate that, compared with the observed demand, the real demand estimated by the DRTMC model is much closer to the set demand in all cases. Furthermore, compared with the other benchmarks, the proposed DRTMC model is more accurate and appropriate when recovering the real demand. For the empirical analysis, we use FFBS data from Shanghai City and choose a typical prosperous district covering approximately $2\ km^2$ for the case study. In particular, we adopt three values of the willingness of migration to estimate the real demand. As the results show, for the entire study area, the real demand is higher than the observed demand, and the distribution of the observed demand is significantly different from that of real demand, emphasizing the importance of recovering the real demand from the observed demand.

The proposed demand estimation model enables researchers to capture the real FFBS demand. Although we haven't conducted real rebalancing works, it is expected that the operation companies could develop more effective rebalance strategies to improve the FFBS service as long as accurate demand information is obtained. The limitations and future studies are as follows. First, this study considers only spatial correlation, but the temporal correlation is also important for demand estimation. Such features should be exploited in future studies. Second, we assumed that the users would only migrate to the nearby grids when they experienced the truncation. This assumption is only applicable to the prosperous area with higher bike supply. Thus, it is necessary to extend it and consider the possibility of migrating to further grids. So that the model can be applied to the more general cases. Moreover, this study focuses on recovering the real FFBS pickup demand, for the recovered demands, inferring their paired destination would also be valuable for managing the FFBS system [47]. In addition, the dataset used for the case study is provided by only one company, while there might be multiple companies operating the FFBS in the real world. If possible, we ought to ask for permission for the dataset from other FFBS companies and conduct a more complete analysis. Besides, since the ground-truth real demand is unavailable for model validation, the simulation is first conducted to generate the data for testing the model before conducting the real-world case study.

Although the validation results based on the synthetic dataset are satisfactory, the simulation involves some assumptions that simplify the real-world scenarios. In further studies, more factors (e.g. the multimodal interaction between FFBS and metro, the points of interest (POI) information, and the real-time bike rebalancing) need to be incorporated to increase the credibility of the validation in the simulation. Consequently, the result for the real-world scenarios will be more convincing and accurate and the FFBS rebalancing work will be more efficient.

APPENDIX

A. Summary for annotations

The annotations are summarized in the following Table I.

TABLE I
ANNOTATIONS

Notation	Definition	Notation	Definition
Sets and indices			
m, n	Indices of grids	N	Set of grids
k	Indices of days	K	Set of days
v	Indices of points in demand network	V	Sets of points in demand network
e	Indices of links in the demand network	E	Sets of links in demand network
e_t, e_m, e_c	Indices of truncation, migration, and spatial correlation links in demand network		
Variables and parameters			
$y_{n,k}$	Observed demand within a specific period on the k^{th} day in grid n	ΔLL	The difference between the objective functions of the current solution and the new solution in SA
$\bar{y}_n$	Average observed demand within a specific period in grid n	ξ	The random value following normal distribution for updating the solution in SA
$\mathbf{Y}$	Vector of $y_{n,k}$ in all grids and all days	α	Cooling rate parameter in SA
λ_n^r	The expected value of real demand in grid n	TP^q	The temperature parameter at iteration q in SA
$\boldsymbol{\lambda^r}$	Vector of λ_n^r in all grids	TP^0 , TP^f	The initial and final temperature parameters in SA
$\boldsymbol{\lambda_m^r}$	Vector of λ_m^r in the migration grids around grid n	$\lambda_n^{r,z}$	The generated new solution of λ_n^r at transition z under the given temperature in SA
λ_n^o	The expected value of observed demand in grid n	$\boldsymbol{\lambda^{r,z}}$	Vector of $\lambda_n^{r,z}$ in all grids in SA
$\boldsymbol{\lambda^o}$	Vector of λ_n^o in all grids	Z	Number of result transitions for each iteration in SA
$\lambda_n^{o,tr}$	Part of λ_n^o capturing the expected demand from grid n itself.	s_n	The expected value of dropping off in grid n set in the simulation
$\lambda_n^{o,mig}(m)$	Part of λ_n^o capturing the expected demand from nearby grid m	$frac$	Percentage of grids experiencing demand truncation set in the simulation
$\lambda_n^{o,mig}$	The sum of $\lambda_n^{o,mig}(m)$ for all nearby grids m	$p^s(w_n = 1)$	Willingness of migration when bikes are unavailable in grid n set in the simulation
T_n^a	The duration in which the bike is available in grid n	$\boldsymbol{ds_n}$	Pick up sequence in grid n in the simulation
a_n	Bike availability in grid n	$\boldsymbol{ss_n}$	Drop off sequence in grid n in the simulation
w_n	Dummy variable,1 if a user decides to migrate if the first choice is truncated in grid n, 0 otherwise	dss_{tv}	Pick up or drop off "event" at time tv in the simulation
T	The duration of the whole study period	I	Dummy indicator, 1 if the pick-up occurs and 0 drop-off occurs in the simulation
B_n	Number of bikes available in grid n	p_u	The willingness of migration for each user generated randomly in the simulation
$pa_{m,n}^z$	The z^{th} duration in which the bike is available in grid n but not available in grid m	res_n	The output information of grid n in the simulation
$pa_{m,n}$	The sum of $pa_{m,n}^z$ for all Z duration	$BAPT^o$	The indicators extracted from observed data for the benchmark *b2*
σ	Prior parameter for the extent of the spatial correlation	$BAPT$	The indicators extracted from simulated data for the benchmark *b2*
H_σ	Normalizing constant in the correlation function	$\widehat{\lambda_n^r}$	The estimated real demand in grid n
Functions			
$Psn(\cdot)$	Poisson distribution	$exp(\cdot)$	Exponential function
$N(\cdot)$	Neighborhood function containing the index of the nearby migration grids	$f(\cdot)$	Function mapping the expected value of real demand to the expected value of observed demand
$ln\ (\cdot)$	The ln function	$simu\ (\cdot)$	Simulation function in the benchmark *b2*
$F\ (\cdot)$	Distribution function in the benchmark *b2*		

B. Simulated annealing algorithm for real demand recovery under DRTMC

Algorithm 1 Simulated annealing for demand recovery

Initialization: Cooling rate α;
  Number of transitions for each iteration Z;
1 Initial temperature TP^0
  Final temperature TP^f;
  Mean μ and standard deviation σ_{SA} for vector $\boldsymbol{\xi}$
2 Calculate the initial solution $\boldsymbol{\lambda}^{r,0}$,
  Current solution: $\boldsymbol{\lambda}_c^r \leftarrow \boldsymbol{\lambda}^{r,0}$;
3 Current objective function: $ll_c \leftarrow ll(\boldsymbol{\lambda}^{r,0})$;
4 **while** $TP^p > TP^f$ **do**
5 **for** $z = 1: Z$ **do**
6 Generate new solution: $\boldsymbol{\lambda}_{\boldsymbol{new}}^r \leftarrow f(\boldsymbol{\lambda}_c^r, \boldsymbol{\xi}(\mu, \sigma_{SA}))$
7 Calculate new objective function: $ll_{new} \leftarrow ll(\boldsymbol{\lambda}_{\boldsymbol{new}}^r)$
8 **if** $ll_{new} > ll_c$ **then**
9 $\boldsymbol{\lambda}_c^r \leftarrow \boldsymbol{\lambda}_{\boldsymbol{new}}^r$
10 **else**
11 Generate random variable ϵ∼uniform[0,1]
12 **if** $\epsilon < exp\left[\frac{ll_{new} - ll_c}{TP^p}\right]$ **then**
13 $\boldsymbol{\lambda}_c^r \leftarrow \boldsymbol{\lambda}_{\boldsymbol{new}}^r$
14 **end if**
15 **end if**
16 **end for**
17 $TP^{p+1} = \alpha\, TP^p$
18 **end while**
19 **return** $\boldsymbol{\lambda}_c^r$

C. Approximating the number of available bikes in grid n at time tv

Algorithm 2 is developed to approximate the number of available bikes left in each grid n at time point tv. In detail, we first define $2T_\delta$ as the duration in which any bike will be used; in other words, the data out of the range of $[tv - T_\delta, tv + T_\delta]$ for tv can be ignored. T_δ is affected by the turnover rate of the bike in the focused area. Because a high turnover rate means that the bike is used more frequently, and in this case T_δ can be set shorter to reduce the computational efforts. But a higher T_δ could increase the accuracy. Here is the main idea, for each grid n, taking the drop-off record as an example, if the last record of a bike within $[tv - T_\delta, tv]$ is a drop-off, the bike is thought to be located at grid n at time tv. Similarly, we judge that a bike is at grid n at time tv if the first record of this bike within $[tv, tv + T_\delta]$ is a pickup.

It is possible that, at time tv, a bike is assigned to different grids n and m according to the drop-off record in $[tv - T_\delta, tv]$ and pick-up record in $[tv, tv + T_\delta]$, respectively. This indicates the rebalance is conducted. In this case, this available bike is assigned to grid n/ m if its drop-off/pick-up time is closer to tv. We then aggregate the number of available bikes for $[tv - T_\delta, tv]$ and $[tv, tv + T_\delta]$ and excluded duplicated bikes with the same ID. Consequently, we obtain the number of available bikes in grid n at time tv. The details of Algorithm 2 are as follows.

Algorithm 2 Calculating the number of available bikes

Initialization: Range of considered duration: $2T_\delta$;
  The list of focused time points: tv_list;
1 The available bikes in grid n: $B_n = [\,]$;
  The available bikes of all the grids $LB = [\,]$;
  The total dataset $\boldsymbol{D}$;
2 **for** grid $n = 1: N$ **do**
3 **for** time tv in tv_list **do**
4 Extract the records within $[tv - T_\delta, tv]$ and $[tv, tv + T_\delta]$ in grid n as $D_{n,-}$ and 1 $D_{n,+}$, respectively;
5 $\{D_{n,-}(1), D_{n,-}(2), \ldots D_{n,-}(P)\} \leftarrow groupbyID(D_{n,-})$
6 In each $D_{i,-}(p)$, extract the record $r_{i,-}(p)$ closet to the tv for bike p;
7 $r_{i,-} \leftarrow Pool\ (r_{i,-}(p))$
8 extract the drop-off record in $r_{i,-}$ as $r_{i,-,d}$
9 $\{D_{n,-}(1), D_{n,-}(2), \ldots D_{n,-}(Q)\} \leftarrow groupbyID(D_{i,+})$
10 In each $D_{i,+}(q)$, extract the record $r_{i,+}(q)$ closet to the tv for bike q;
11 $r_{i,+} \leftarrow Pool\ (r_{i,+}(q))$
12 extract the pickup record in $r_{i,+}$ as $r_{i,+,p}$
13 $pudp_i \leftarrow Pool\ (r_{i,-,d}, r_{i,+,p})$
14 Exclude the duplicated record with the same bike ID and get the $pudp_{i,e}$
15 $B_n(tv) \leftarrow |pudp_{i,e}|$
16 Add $B_n(tv)$ to B_n
17 **end for**
18 Add B_n to LB
19 **end for**
20 **return** LB

In this study, the evening peak chosen for analysis is 1 h, and we divide it with 10s to obtain $tv\_list = \{18{:}00{:}00, 18{:}00{:}10, \ldots, 18{:}59{:}50\}$. In other words, we check the number of available bikes in each grid every 10 s. In addition, $2T_\delta$ is set as 48 h, and thus, T_δ is 24 h. After executing Algorithm B, we obtain the matrix LB with a size 196×3600, where the values of row n and column tv in LB represent the number of available bikes at grid n at time tv. Through matrix LB, we could check whether the demand is truncated in this grid and, if so, how long the truncation is

REFERENCES

[1] Q. Chen, C. Fu, N. Zhu, S. Ma, and Q. C. He, "A target-based optimization model for bike-sharing systems: From the perspective of service efficiency and equity," *Transportation Research Part B: Methodological*, vol. 167, pp. 235–260, Jan. 2023, doi: 10.1016/j.trb.2022.12.002.

[2] J. Wang and Y. Wang, "A two-stage incentive mechanism for rebalancing free-floating bike sharing systems: Considering user preference," *Transp Res Part F Traffic Psychol Behav*, vol. 82, pp. 54–69, Oct. 2021, doi: 10.1016/j.trf.2021.08.003.

[3] L. Wu, W. Gu, W. Fan, and M. J. Cassidy, "Optimal design of transit networks fed by shared bikes," *Transportation Research Part B: Methodological*, vol. 131, pp. 63–83, Jan. 2020, doi: 10.1016/j.trb.2019.11.003.

[4] Y. Xing, K. Wang, and J. J. Lu, "Exploring travel patterns and trip purposes of dockless bike-sharing by analyzing massive bike-sharing data in Shanghai, China," *J Transp Geogr*, vol. 87, p. 102787, Jul. 2020, doi: 10.1016/j.jtrangeo.2020.102787.

[5] C. Xu, J. Ji, and P. Liu, "The station-free sharing bike demand forecasting with a deep learning approach and large-scale datasets," *Transp Res Part C Emerg Technol*, vol. 95, pp. 47–60, Oct. 2018, doi: 10.1016/j.trc.2018.07.013.

[6] Y. Liu, W. Y. Szeto, and S. C. Ho, "A static free-floating bike repositioning problem with multiple heterogeneous vehicles, multiple depots, and multiple visits," *Transp Res Part C Emerg Technol*, vol. 92, pp. 208–242, Jul. 2018, doi: 10.1016/j.trc.2018.02.008.

[7] X. Ma, Y. Ji, Y. Yuan, N. Van Oort, Y. Jin, and S. Hoogendoorn, "A comparison in travel patterns and determinants of user demand between docked and dockless bike-sharing systems using multi-sourced data," *Transp Res Part A Policy Pract*, vol. 139, pp. 148–173, Sep. 2020, doi: 10.1016/j.tra.2020.06.022.

[8] Y. Du, F. Deng, and F. Liao, "A model framework for discovering the spatio-temporal usage patterns of public free-floating bike-sharing system," *Transp Res Part C Emerg Technol*, vol. 103, pp. 39–55, Jun. 2019, doi: 10.1016/j.trc.2019.04.006.

[9] J. Song, L. Zhang, Z. Qin, and M. A. Ramli, "A spatiotemporal dynamic analyses approach for dockless bike-share system," *Comput Environ Urban Syst*, vol. 85, p. 101566, Jan. 2021, doi: 10.1016/j.compenvurbsys.2020.101566.

[10] Y. Shen, X. Zhang, and J. Zhao, "Understanding the usage of dockless bike sharing in Singapore," *Int J Sustain Transp*, vol. 12, no. 9, pp. 686–700, Oct. 2018, doi: 10.1080/15568318.2018.1429696.

[11] S.-H. Lee and H.-C. Ku, "A Dual Attention-Based Recurrent Neural Network for Short-Term Bike Sharing Usage Demand Prediction," *IEEE Transactions on Intelligent Transportation Systems*, vol. 24, no. 4, pp. 4621–4630, Apr. 2023, doi: 10.1109/TITS.2022.3208087.

[12] Y. Lv, D. Zhi, H. Sun, and G. Qi, "Mobility pattern recognition based prediction for the subway station related bike-sharing trips," *Transp Res Part C Emerg Technol*, vol. 133, p. 103404, Dec. 2021, doi: 10.1016/j.trc.2021.103404.

[13] F. Huang, S. Qiao, J. Peng, and B. Guo, "A bimodal Gaussian inhomogeneous poisson algorithm for bike number prediction in a bike-sharing system," *IEEE Transactions on Intelligent Transportation Systems*, vol. 20, no. 8, pp. 2848–2857, 2019, doi: 10.1109/TITS.2018.2868483.

[14] M. Dell'Amico, M. Iori, S. Novellani, and A. Subramanian, "The Bike sharing Rebalancing Problem with Stochastic Demands," *Transportation Research Part B: Methodological*, vol. 118, pp. 362–380, Dec. 2018, doi: 10.1016/j.trb.2018.10.015.

[15] C. Lv, C. Zhang, K. Lian, Y. Ren, and L. Meng, "A two-echelon fuzzy clustering based heuristic for large-scale bike sharing repositioning problem," *Transportation Research Part B: Methodological*, vol. 160, pp. 54–75, Jun. 2022, doi: 10.1016/j.trb.2022.04.003.

[16] J. Zhang, M. Meng, and D. Z. W. Wang, "A dynamic pricing scheme with negative prices in dockless bike sharing systems," *Transportation Research Part B: Methodological*, vol. 127, pp. 201–224, Sep. 2019, doi: 10.1016/j.trb.2019.07.007.

[17] J. Pfrommer, J. Warrington, G. Schildbach, and M. Morari, "Dynamic Vehicle Redistribution and Online Price Incentives in Shared Mobility Systems," *IEEE Transactions on Intelligent Transportation Systems*, vol. 15, no. 4, pp. 1567–1578, Aug. 2014, doi: 10.1109/TITS.2014.2303986.

[18] S. Albiński, P. Fontaine, and S. Minner, "Performance analysis of a hybrid bike sharing system: A service-level-based approach under censored demand observations," *Transp Res E Logist Transp Rev*, vol. 116, no. May, pp. 59–69, 2018, doi: 10.1016/j.tre.2018.05.011.

[19] M. Gong, Y. Hu, Z. Chen, and X. Li, "Transfer-based customized modular bus system design with passenger-route assignment optimization," *Transp Res E Logist Transp Rev*, vol. 153, Sep. 2021, doi: 10.1016/j.tre.2021.102422.

[20] J. Alonso-Mora, S. Samaranayake, A. Wallar, E. Frazzoli, and D. Rus, "On-demand high-capacity ride-sharing via dynamic trip-vehicle assignment," *Proc Natl Acad Sci U S A*, vol. 114, no. 3, pp. 462–467, Jan. 2017, doi: 10.1073/pnas.1611675114.

[21] S. Wang, H. Chen, J. Cao, J. Zhang, and P. S. Yu, "Locally Balanced Inductive Matrix Completion for Demand-Supply Inference in Stationless Bike-Sharing Systems," *IEEE Trans Knowl Data Eng*, vol. 32, no. 12, pp. 2374–2388, Dec. 2020, doi: 10.1109/TKDE.2019.2922636.

[22] D. Gammelli, I. Peled, F. Rodrigues, D. Pacino, H. A. Kurtaran, and F. C. Pereira, "Estimating latent demand of shared mobility through censored Gaussian Processes," *Transp Res Part C Emerg Technol*, vol. 120, p. 102775, Nov. 2020, doi: 10.1016/j.trc.2020.102775.

[23] E. Fields, C. Osorio, and T. Zhou, "A data-driven method for reconstructing a distribution from a truncated sample with an application to inferring car-sharing demand," *Transportation Science*, vol. 55, no. 3, pp. 616–636, May 2021, doi: 10.1287/trsc.2020.1028.

[24] J. Wang, T. Miwa, and T. Morikawa, "A Demand Truncation and Migration Poisson Model for Real Demand Inference in Free-Floating Bike-Sharing System," *IEEE Transactions on Intelligent Transportation Systems*, pp. 1–12, 2023, doi: 10.1109/TITS.2023.3275081.

[25] G. Vulcano, G. van Ryzin, and R. Ratliff, "Estimating Primary Demand for Substitutable Products from Sales Transaction Data," *Oper Res*, vol. 60, no. 2, pp. 313–334, Apr. 2012, doi: 10.1287/opre.1110.1012.

[26] X. Ding, C. Chen, C. Li, and A. Lim, "Product demand estimation for vending machines using video surveillance data: A group-lasso method," *Transp Res E Logist Transp Rev*, vol. 150, Jun. 2021, doi: 10.1016/j.tre.2021.102335.

[27] B. Chen and X. Chao, "Dynamic inventory control with stockout substitution and demand learning," *Manage Sci*, vol. 66, no. 11, pp. 5108–5127, Nov. 2020, doi: 10.1287/mnsc.2019.3474.

[28] A. G. Kök and M. L. Fisher, "Demand estimation and assortment optimization under substitution: Methodology and application," *Oper Res*, vol. 55, no. 6, pp. 1001–1021, Nov. 2007, doi: 10.1287/opre.1070.0409.

[29] H. Ullman and L. Aultman-Hall, "Exploring motivations and barriers for long-distance trips of adult women Vermonters," *Travel Behav Soc*, vol. 21, pp. 37–47, Oct. 2020, doi: 10.1016/j.tbs.2020.05.007.

[30] C. Luiu and M. Tight, "Travel difficulties and barriers during later life: Evidence from the National Travel Survey in England," *J Transp Geogr*, vol. 91, p. 102973, Feb. 2021, doi: 10.1016/j.jtrangeo.2021.102973.

[31] N. L. Fields, C. Cronley, S. P. Mattingly, E. M. Roark, S. R. Leat, and V. J. Miller, "Transportation mobility and health among older adults: Examining missed trips and latent demand," *J Transp Health*, vol. 21, p. 101069, Jun. 2021, doi: 10.1016/j.jth.2021.101069.

[32] Torbay Council, "Taxi unmet demand survey 2020," available at https://www.torbay.gov.uk/media/15909/taxi-unmet-survey.pdf

[33] X. Zhao *et al.*, "Estimating wildfire evacuation decision and departure timing using large-scale GPS data," *Transp Res D Transp Environ*, vol. 107, p. 103277, Jun. 2022, doi: 10.1016/j.trd.2022.103277.

[34] R. Zhang and R. Ghanem, "Demand, Supply, and Performance of Street-Hail Taxi," *IEEE Transactions on Intelligent Transportation Systems*, vol. 21, no. 10, pp. 4123–4132, 2020, doi: 10.1109/TITS.2019.2938762.

[35] J. Wang, T. Miwa, and T. Morikawa, "Recursive decomposition probability model for demand estimation of street-hailing taxis utilizing GPS trajectory data," *Transportation Research Part B: Methodological*, vol. 167, pp. 171–195, Jan. 2023, doi: 10.1016/j.trb.2022.11.014.

[36] A. Negahban, "Simulation-based estimation of the real demand in bike-sharing systems in the presence of censoring," *Eur J Oper Res*, vol. 277, no. 1, pp. 317–332, Aug. 2019, doi: 10.1016/j.ejor.2019.02.013.

[37] N. Jian, D. Freund, H. M. Wiberg, and S. G. Henderson, "Simulation optimization for a large-scale bike-sharing system," *Proceedings - Winter Simulation Conference*, vol. 0, no. 2013, pp. 602–613, 2016, doi: 10.1109/WSC.2016.7822125.

[38] Y. Ma, Z. Zhang, S. Chen, Y. Pan, S. Hu, and Z. Li, "Investigating the impact of spatial-temporal grid size on the microscopic forecasting of the inflow and outflow gap in a free-floating bike-sharing system," *J Transp Geogr*, vol. 96, no. January 2020, p. 103208, 2021, doi: 10.1016/j.jtrangeo.2021.103208.

[39] D. Koller and N. Friedman, *Probabilistic Graphical Models: Principles and Techniques*. MIT press, 2009.

[40] S. Prince, *Computer Vision: Models, Learning and Inference*. Cambridge University Press , 2012.

[41] T. Taniai, Y. Matsushita, Y. Sato, and T. Naemura, "Continuous 3D Label Stereo Matching Using Local Expansion Moves," *IEEE Trans Pattern Anal Mach Intell*, vol. 40, no. 11, pp. 2725–2739, Nov. 2018, doi: 10.1109/TPAMI.2017.2766072.

[42] J. Ma, D. Delahaye, M. Sbihi, P. Scala, and M. A. Mujica Mota, "Integrated optimization of terminal maneuvering area and airport at

the macroscopic level," *Transp Res Part C Emerg Technol*, vol. 98, pp. 338–357, Jan. 2019, doi: 10.1016/j.trc.2018.12.006.

[43] Y. Huo, D. Delahaye, and M. Sbihi, "A probabilistic model based optimization for aircraft scheduling in terminal area under uncertainty," *Transp Res Part C Emerg Technol*, vol. 132, Nov. 2021, doi: 10.1016/j.trc.2021.103374.

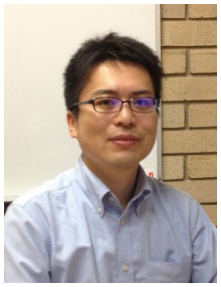
[44] X. Chang, J. Wu, H. Sun, G. H. de A. Correia, and J. Chen, "Relocating operational and damaged bikes in free-floating systems: A data-driven modeling framework for level of service enhancement," *Transp Res Part A Policy Pract*, vol. 153, pp. 235–260, Nov. 2021, doi: 10.1016/j.tra.2021.09.010.

[45] Y. Ancele, M. H. Hà, C. Lersteau, D. Ben Matellini, and T. T. Nguyen, "Toward a more flexible VRP with pickup and delivery allowing consolidations," *Transp Res Part C Emerg Technol*, vol. 128, Jul. 2021, doi: 10.1016/j.trc.2021.103077.

[46] J. Wang, F. Li, S. Yang, Y. Li, and Y. Wang, "A Real-Time Bike Trip Planning Policy With Self-Organizing Bike Redistribution," *IEEE Transactions on Intelligent Transportation Systems*, vol. 23, no. 8, pp. 10646–10661, Aug. 2022, doi: 10.1109/TITS.2021.3095177.

[47] M. Jiang, C. Li, K. Li, and H. Liu, "Destination Prediction Based on Virtual POI Docks in Dockless Bike-Sharing System," *IEEE Transactions on Intelligent Transportation Systems*, vol. 23, no. 3, pp. 2457–2470, Mar. 2022, doi: 10.1109/TITS.2021.3099571.

Jianbiao Wang received the B.E. degree in transportation engineering from Southeast University, China, in 2020. He is currently pursuing the Ph.D. degree with Nagoya University. His research interests include travel behavior analysis, travel demand estimation and public transit operation.



Tomio Miwa received the B.E., M.E., and D.E. degrees in civil engineering from Nagoya University, Japan, in 1998, 2000, and 2005, respectively. He is currently an Associate Professor at the Institute of Materials and Systems for Sustainability, Nagoya University. He specializes in fields, such as the analysis of driver's route choice behavior, traffic simulation, and transport management using ITS. He is the recipient of the Best Paper Award at the 11th EASTS International Conference.

Xinwei Ma received the B.E. degree in transportation planning and management from the Hebei University of Technology, in 2013, and the Ph.D. degree in transportation engineering from the School of Transportation, Southeast University, in 2020. He is currently a Lecturer at the Hebei University of Technology. His research interests include public transport, shared mobility, transport data analytics, and intelligent transport systems.

Dawei Li is currently an Associate Professor with the School of Transportation, Southeast University, Nanjing, China. His research interests are traffic behavior analysis, traffic demand prediction, traffic environment, big data and traffic information, traffic simulation, and network analysis. He is a member of Hong Kong Scholars Program, high level talent of "six peaks" in Jiangsu Province, chairman of the world Transport Conference (WTC), future oriented urban comprehensive transport technology committee, and a member of China highway society, automatic driving working committee.

Takayuki Morikawa received the B.E., and M.E. degrees in transportation engineering from Kyoto University, Japan, in 1981 and 1983, respectively, and the M.S. and Ph.D. degrees in civil engineering from Massachusetts Institute of Technology, U.S.A. He currently serves as a Professor at the Institute of Innovation for Future Society, Nagoya University. He specializes in transport planning, urban planning, transport policies, transport demand forecast, consumer behavior, and environmentally sustainable transport. He has written books on modeling travel behavior and is a well-known figure on Japanese government panels.