Pamukkale Üniversitesi Mühendislik Bilimleri Dergisi

Pamukkale University Journal of Engineering Sciences

User-based relocation strategy for free floating car-sharing system: an Istanbul case

Serbest gezen araba paylaşım sistemi için kullanıcı temelli yer değiştirme stratejisi: bir İstanbul örneği

Mısra ŞİMŞİR[1], U. Mahir YILDIRIM[2], Doruk SEN[2]*

[1]Department of Industrial Engineering, Yildiz Technical University, Istanbul, Turkey.
msimsir@yildiz.edu.tr
[2]Department of Industrial Engineering, Istanbul Bilgi University, Istanbul, Turkey
mahir.yildirim@bilgi.edu.tr, doruk.sen@bilgi.edu.tr

Received/Geliş Tarihi: 15.08.2023 Revision/Düzeltme Tarihi: 18.11.2023 doi: 10.5505/pajes.2024.76299
Accepted/Kabul Tarihi: 05.01.2024 Research Article/Araştırma Makalesi

Abstract

In a world with limited resources, it is crucial for individuals to utilise shared systems and develop strategies to optimise their usage. To cope with this, 'servicizing' has emerged as a rapidly growing promising solution, especially in car-sharing systems. These systems can be split into two: station-based and free-floating. The latter introduces more flexibility to the customers as free-floating systems allow users to pick up and drop off vehicles anywhere within predetermined operational zones. This flexibility may come with an additional cost by bringing a potential imbalance between demand and supply. This imbalance can harm the company's profitability and customer satisfaction. In this study, the imbalance problem of the system of free-floating car sharing is considered. A mixed integer linear programming model is developed and tested with real data for free floating car sharing systems to solve this problem. The proposed system consists of four modules: clustering, forecasting, optimization model, and relocation strategy. According to the results, it is observed that the system is more balanced with satisfying 9% more demand and more profitable with earning 6% more. The study was conducted on a car-sharing company that is based in Istanbul, but the results can be applied to any free-floating car-sharing system. This ensures customer satisfaction by meeting demand and balancing the system.

Keywords: Car-Sharing, Free-floating, User-based Relocation, Machine Learning

Öz

Sınırlı kaynaklara sahip bir dünyada, bireylerin paylaşımlı sistemleri kullanma ve kullanımlarını optimize etmek için stratejiler geliştirmesi hayati öneme sahiptir. Bu duruma cevap olarak, 'hizmetleştirme' özellikle araç paylaşım sistemlerinde hızla büyüyen umut verici bir çözüm olarak ortaya çıkmıştır. Bu sistemler ikiye ayrılmaktadır; istasyon tabanlı ve serbest dolaşan. Serbest dolaşan sistemlerin belirlenmiş operasyonel bölgeler içinde araçları herhangi bir yerden alıp bırakmaya izin verdiği bilindiği için müşterilere daha fazla esneklik sunar. Bu esneklik, talep ile arz arasındaki potansiyel bir dengesizlik getirerek ek bir maliyete neden olabilmektedir. Bu çalışmada, serbest dolaşan araç paylaşım sisteminin dengesizlik problemi ele alınmıştır. Bu problemi çözmek için serbest dolaşan araç paylaşım sistemleri için bir karma tamsayılı doğrusal programlama modeli geliştirilmiş ve gerçek verilerle test edilmiştir. Önerilen sistem dört modülden oluşmaktadır: kümeleme, tahmin, optimizasyon modeli ve yeniden konumlandırma stratejisi. Sonuçlara göre, sistemin %9 daha fazla talebi karşılayarak daha dengeli ve %6 daha fazla kazanç sağlayarak daha karlı olduğu gözlemlenmiştir. Çalışma İstanbul merkezli bir araç paylaşım şirketi üzerinde gerçekleştirilmiştir, ancak sonuçlar herhangi bir serbest dolaşan araç paylaşım sistemine uygulanabilir. Bu, talebi karşılayarak ve sistemi dengeleyerek müşteri memnuniyetini sağlamaktadır.

Anahtar kelimeler: Araç Paylaşımı, Serbest-Gezen, Kullanıcı Tabanlı Yeniden Konumlandırma, Makine Öğrenmesi

1 Introduction

The finite resources and the challenges brought with them necessitate the inclusion of the sharing economy while the pressure on limited resources strengthens. This pressure has led to proposing models to optimize resource allocation and scheduling in different industries. Such a model has been proposed for device-to-device communication [1]. An increasing number of individuals seek goods and services. So, the sharing economy is brought into discussion for obtaining a valid solution with the maximization of the use of resources. This way, the overall demand can be diminished for the scarce resources, and a positive attitude can be developed to avoid overconsumption. The sharing economy encourages individuals to cooperate and share, promoting mutual responsibility. Moreover, the sharing economy serves as another means to fulfill consumer needs and assures the long-term use of resources. Home-sharing, ridesharing, second-hand item sharing, entertainment sharing, car-sharing, and bike-sharing exemplify the sharing economy. For example, the crowd-shipping model in the transportation sector, which is one of the most used areas of the sharing economy, recommends users to transport packages between service points on their way during their travels [2]. Another model of the sharing economy of the transportation sector is Car-sharing, which is a model where individuals can access and use vehicles that are provided by a company on a short-term basis, often by the hour or minute, without any ownership. Users can effortlessly hire these vehicles via a mobile app. These users can even rent a car together for the same trip, which is called ridesharing [3], [4]. Within the car-sharing system, a company offers a group of vehicles that members share by picking up and dropping off. The literature identifies three types of car-sharing services: one-way, two-way, and free-floating. These three car-sharing models differ in demand patterns and alternating purposes of customers [5]. One-way and two-way car-sharing

*Corresponding author/Yazışılan Yazar

systems are types of station-based models of car-sharing services. In station-based models, companies have predetermined stations where users can pick up and drop off vehicles within the service area. In one-way car-sharing systems, users can rent and drop cars at stations in the operational area. Therefore, origin and destination stations are not necessarily the same for such models. On the other hand, in a two-way car-sharing system, users are obliged to drop the car at the same station that they rented the car from. Therefore, these systems are not considered flexible. In the free-floating model, users can pick up or drop off vehicles anywhere in the operational area without restrictions. However, this flexibility can cause a problem, namely, imbalanced supply and demand. To address this problem, relocation strategies are introduced in the literature. Relocation problems can be defined as the field of determining the number of vehicles to be transferred from one region to another in determined time intervals.

The relocation of vehicles is studied as operator-based, user-based, and hybrid models. In an operator-based system, the company allocates some of its workers to move the cars to different locations, incurring labor costs for this task. However, in a user-based system, users are encouraged to move vehicles to desired locations through incentives like earning extra minutes for their next trip or receiving price discounts. Most of the literature focuses on operator-based relocation strategies for a station-based model to solve this problem [6]–[12]. Using operators is a deterministic approach compared to a user-based relocation strategy. The main reason behind this is the certain number of employees in the company that are assigned solely to this job. On the other hand, user-based relocation is a probabilistic approach that depends on the user's will. However, this model does not include additional staffing costs. There are some works in the literature that studied user-based relocation strategies for station-based car-sharing models [13]–[16]. Even though free-floating car-sharing systems are a relatively new concept, there have been a few studies on the subject in the literature. In this context, the work of Weikl & Bogenberger divided the operational area into 15 macroscopic and 478 microscopic zones to try to relocate the vehicles in an operator-based free-floating car-sharing system [17]. The authors developed an algorithm to find the optimal movement of vehicles between zones for both conventional and electric vehicles (EVs). For the first time in literature, a model was created and evaluated in a real-world setting. The study includes historical rental data but not the estimated demand model. One-way car sharing is also studied for repositioning EVs with the operator-based approach in which a multi-objective mathematical model for EVs with reservations is developed [6]. Trip reservations are made in advance on their model as it is favorable to know the demand before renting. Working with reservations or knowing the demand in advance makes problems relatively easier to solve because when demand is unknown, a demand estimation problem should be solved, which is entirely another subject to investigate. When directing vehicles based on incorrect estimates, a company can face larger losses instead of running a profitable business. Furthermore, free-floating car-sharing systems are also studied along with genetic algorithm [18]. Relocations in the study are carried out by operators who share a shuttle. Therefore, the authors focused on optimizing the shuttle route.

The main contribution of this study is that a mixed integer linear programming model is developed and tested with real data for free floating car sharing systems. Also, the necessary parameters for the model, such as demand and virtual stations, were obtained using machine learning algorithms. At the same time, segmentation strategies have been suggested to enable user-based relocation. This means that each necessary step of the relocation problem for rental companies starting from forecasting and ending with relocation strategies, was addressed simultaneously in this study, on the contrary to the current studies in the literature.

The rest of the paper is organized as follows. Section 2 gives information about the description of the problem, the main assumptions, and our methodology for solving this problem. Then, Section 3 discussed the results and foundations of the model. Finally, we conclude our findings and make recommendations for future research in Section 4.

2 Problem description & methodology

Free-floating is considered the most flexible type of car-sharing model. In a free-floating car-sharing system, a vehicle rented by a user can be dropped off anywhere within the predefined operational area. This flexibility of the system causes some problems, such as vehicle accumulation in specific regions and hours, called as the imbalance in the literature. An imbalance of supply and demand makes low profit and less customer satisfaction for the company. The imbalance of supply and demand is related to the limited number of vehicles in the regions with high demand and vice versa. So, the vehicles that are not rented and have been idle for a long time should be directed to areas with high rental performance, as not doing so prevents potential earnings. The sustainability in the availability of cars is one of the most significant issues for users. Hence, the main goal should be to have at least one vehicle around the user when searched. Lack of vehicles causes the user to search for other transportation methods, and in parallel, decreases customer satisfaction. For such cases, there are several issues to be addressed, such as pricing and repositioning. In line with this, some studies contribute to the literature with the development of dynamic pricing strategies to allow different pricing schemes based on regions and hours [19]–[21].

In this study, a solution approach for free floating car-sharing system that utilizes relocation of vehicles is proposed. The proposed approach consists of four modules: (i) clustering, (ii) forecasting, (iii) optimization model, and (iv) user-based relocation strategy. Figure 1 shows the methodology of the study in detail and the relationship between each module. First, machine learning algorithms are used for clustering and forecasting purposes. The results of these methods are used as inputs for the mathematical model. In parallel with these, a relocation strategy is recommended to attract users.

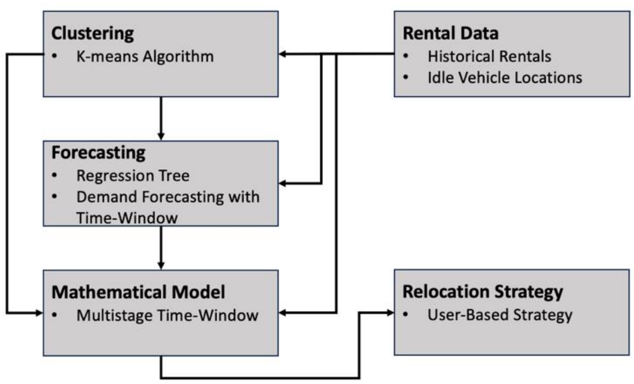


Figure 1. Methodology of study

2.1 Mathematical model

As mentioned, this study aims to ensure customer satisfaction by effectively meeting demand and ensuring an adequate presence of vehicles in the required areas. To achieve this goal, a mathematical model is built that helps to optimize vehicle relocations under demand uncertainties. First, the demand and supply characteristics of the system are described to help us understand the problem easily and make formulations for initializing mathematical models.

Vehicles: The total number of active vehicles in the system varies because sometimes vehicles are required to be repaired or maintained, or new vehicles can be provided to the system. The input was provided to the model by considering these instant changes.

Clusters/Stations: The service area defined by the company in Istanbul is divided into 100 virtual clusters. For this, the k-mean clustering algorithm is applied. Each cluster has a center and is distinctive based on the area, user demography, etc. For example, some regions are preferred for shopping, restaurants, or nightclubs, while some are industrial zones.

Operations:

Rental: Rentals can start and finish at any time during the day within the defined service area of the company. If a nearby car is available, users can begin their rental immediately.

Relocation: This process involves taking a vehicle from its current location and delivering it to a specific destination. Users of the system handle the task of relocating the car.

Costs: Generally defined as operating this service to users. During the relocation process, two types of costs arise. These are the incentive costs to be given to customers for relocation and the fuel cost between the customer's destination and the directed location. Operating costs are considered as maintenance or service costs, which is incurred when a vehicle experiences a breakdown due to accidents or requires routine monthly and during the rental process, fuel consumption cost and insurance cost are considered.

Penalty Cost: Lost rental revenue when demand cannot be satisfied. This cost is calculated as the lost profit earned from an average rental. The penalty cost is assumed to be the same for each lost demand.

Rental Revenue: The revenue that is earned from the rental fee of users is calculated on average based on the money that comes from the distance between two clusters.

Model parameters and variables used in the model are shown in Table 1. The objective function is calculated as revenue earned from each rental minus the total cost of the company while operating this process. Constraints (2) state that the sum of the total number of satisfied demands and unsatisfied demands must be equal to the forecasted demand for each cluster and time interval. Constraints (3) set the number of available vehicles at the beginning of the first-time interval to the number of initial idle vehicles for each station j. Constraints (4) indicate the amount of supply for each station for the initial time interval. Rentals and relocations that started before time interval 1 and will arrive at station j at the beginning of time interval 1 can be used to satisfy the demands from station j to another station in time interval 1 along with the vehicles that are already at station 1. Constraints (5) represent that when the time interval t is greater than 1, for each time interval and each station, the total number of demands from station j to other clusters and the relocation from cluster j to other clusters during the time interval t can be met by the number of cars available at the beginning of the time interval t. Any vehicles that are left will be used in the next interval. Finally, constraints (6) represent non-negativity.

$$max\ z \sum_{j} \sum_{l} \sum_{t} [(R_{jl} - TC_{jl}) * d_{jlt} - (TC_{jl} * r_{jlt})] - \sum_{j} \sum_{l} \sum_{t} (m_{jlt} * PC) - VFC * v \quad (1)$$

$$d_{jlt} + m_{jlt} = OD_{jlt} \quad \forall j,l,t \quad (2)$$

$$k_{j1} = w_j^0 \quad \forall j \quad (3)$$

$$\sum_{l} (d_{jl1} + r_{jl1}) + k_{j2} = w_j^0 + \sum_{l} (s_{lj}^0 + q_{lj}^0) \ \forall j \quad (4)$$

$$\sum_{l} (d_{jlt} + r_{jlt}) + k_{jt+1} = k_{jt} + \sum_{j} (d_{jlt-1} + r_{jlt-1}) \ \forall j, 1 < t < 8 \quad (5)$$

$$k_{j,t}, d_{jlt}, m_{jlt}, r_{jlt} \geq 0 \quad \forall j,l,t \quad (6)$$

able 1. Model parameters and variables

Sets	Station/Cluster	j, l (1, 2, … , 100)
	Time Period	t (1, 2, … , 8)
Decision Variables	d_{jlt}	# of accepted rentals from station j to station l for time interval t
	r_{jlt}	# of relocation from station j to station l for time interval t
	m_{jlt}	# of unserved demand from station j to station l for time interval t
	k_{jt}	# of available vehicles at station j at the beginning of time interval t
Parameters	OD_{jlt}	# of forecasted demand from station j to station l for time interval t
	R_{jl}	Trip revenue from station j to station l
	TC_{jl}	Trip cost between station j and station l
	PC	Penalty cost for loss demand
	VFC	Vehicle fixed cost (insurance, maintenance etc.)
	v	Number of vehicles available in the system
	w_j^0	# of initial idle vehicles at station j at the beginning of the planning horizon
	s_{lj}^0	# of vehicles that started to relocate from cluster l to j before time interval 1 and will arrive at cluster j during time interval 1
	q_{lj}^0	# of vehicles are rented before time interval 1 started from cluster l to j

2.2 Empirical studies

2.2.1 Data analyses

There are two types of data in car-sharing platforms. The first is the search data which corresponds to searching for a car in the application and logging out if none can be found. The second one is the rental data, which keeps a record of successful rents. In this study, both search and rental data are used for analysis. Demand loss is considered along with the real demand as one of the main targets of this research is to reduce the loss demand by implementing rebalancing techniques. The main reason behind not meeting demand is considered to be the imbalance of the system in which vehicles are not available at the right place at the right time. This is one of the main contributions of this study, as the literature focuses heavily on historical rental data, to the best of our knowledge. Search data makes the model more realistic but brings several problems. The main problem is that the same user can be connected to the application more than once in the same location within a short time interval. This duplication must be eliminated and singularised for searches by the same users within 30 minutes in the same location. This 30-minute limit is determined by the company based on experience to take customer satisfaction into account. The information on the data source used in this research is as follows:

- The company has 415,000 registered customers as of May 2020.
- There are 1,650 vehicles in its fleet.
- There are 1,397,000 rental records.
- The number of search requests per month is around 2M.

Data contain information on customers, searches, and rentals such as membership date, phone number, membership ID, rental time, vehicle ID, usage time, km traveled, cleaning score, customer reviews, and rental location as latitude and longitude. Table 2 shows the number of rentals on a day/month basis.

According to the analyzed data, the number of rentals on weekends is higher than on weekdays. 33.40% of all rentals occur on Saturday and Sunday. The number of Saturday rentals is the highest of the week. On weekdays, Fridays and Tuesdays have significantly higher numbers of rentals compared to the others. As can be inferred from the data, a significant change in demand occurs depending on the day of the week. Table 3 indicates the number of rentals according to time. The duration of each time interval is three hours, and an operational day is divided into eight-time intervals (12 p.m.–3 a.m., 3 a.m.–6 a.m., . . ., 9 p.m.–12 p.m.).

As represented in Table 3, most rentals are realized during rush hours, such as 3 p.m. - 6 p.m. and 6 p.m.-9 p.m. intervals.

Table 2. Monthly rental data (hourly based)

Day/Month	Average(%)
Monday	12.71
Tuesday	14.17
Wednesday	12.82
Thursday	13.01
Friday	13.89
Saturday	16.96
Sunday	16.44

The main reason behind this is the high density of workplaces and offices in the specified regions. This can be further explained by the user's tendency to start their commute to work. In addition, public transportation can be very busy during these hours, and people do not want to use public transportation due to crowdedness. It can also be seen that the lowest number of rentals occurs between 3 a.m. and 6 a.m.

Table 3. Monthly rental data

Time/Month	Average(%)
12 p.m.- 3 a.m.	7.95
3 a.m.- 6 a.m.	2.67
6 a.m.- 9 a.m.	12.23
9 a.m.- 12 a.m.	12.01
12 a.m.- 3 p.m.	15.25
3 p.m. - 6 p.m.	18.73
6 p.m. - 9 p.m.	19.89

2.2.2 Clustering

First, the regions where vehicles will be redirected must be determined. To achieve that, Istanbul regions should be defined as demanders or suppliers. To be able to define such a region, the operational area is divided into clusters. In this way, clusters can be considered as the stations of one-way or two-way systems. As in the literature, this process is one of the most common approaches for free-floating car-sharing systems [6], [17]. The clustering technique divides Istanbul into zones based on historical rental data. Since free floating model is not station-based, it is necessary to create virtual stations in order to address the source and destination of car rentals in the model. Figure 2 shows the clusters of Istanbul. To achieve this, a k-means clustering algorithm was introduced, i.e. creating virtual stations in Istanbul. As a first step, the algorithm randomly determines *k* centroids based on the number of clusters given as input and assigns each data point to the nearest cluster to minimize the total distance. Euclidean Distance is used in this research as a distance measure. At each iteration, until it converges, every centroid is updated to a new location, which is the average of the assigned data points. To do this, the number of clusters must be given as input to the algorithm. The defined clusters reflect the patterns and must cover the entire operational area. The k-means clustering algorithm used in this study needs only one parameter, which is the number of clusters/stations. This parameter is determined by taking a 15-minute walking distance which is accepted as convenient in the free-floating car sharing systems literature. Also, the company addressed in this study uses 15-minute reservation time as well. Based on that, the number of virtual stations in Istanbul is determined as 100. Therefore, the k-means algorithm is run with this parameter.

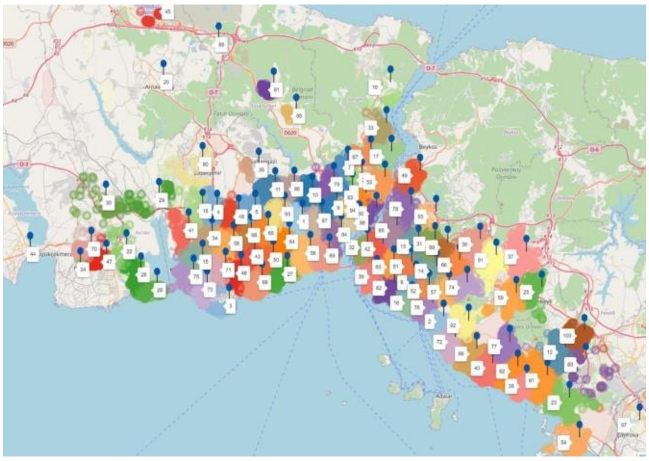


Figure 2. Cluster map

2.2.3 Forecasting

To make a relocation that can provide balance to the system, it should be known exactly how much demand is in which region and at what time. Otherwise, we may cause an accumulation of extra vehicles in regions where there is less demand or a shortage of vehicles in the regions where there is high demand. Therefore, the accuracy of the forecast is a very important part of the recommended method in terms of the correctness of the relocation strategy, which depends on the quality of the demand forecasts. This paper is different in the way that it uses a real demand forecast module instead of working directly with historical rental data or working with known reservations in advance. Working with historical data may be unrealistic because demand has been changing over time, and many factors are affecting that. In addition, search data is also taken into account, which is also not seen in the literature. Using machine learning algorithms, we understand both the real demand patterns and the determinants or features that affect them. Also, a bike-sharing system is studied along with different factors that affect the demand, such as season, weather, rush hours, etc [22]. Based on these factors, the validation distribution of vehicles is performed for zone categorization. A regression tree is also used as a forecasting technique in the literature for the free-floating car-sharing system to monitor the number of cars [23]. In this study, supervised machine learning methods are used to estimate demand. More than one method was tested, and based on the test error of the models, the one with the minimum test error was chosen to predict demand. Since the company was recently established during this study, the model includes short-term data. Data include the number of daily rentals from August 1, 2019, to January 30, 2020. The real-life data which is used for the computational experiments spans a 6-month period. The first five months' data is used for the training and last month data is used for the test. The demand for a week is estimated after the model is finalized, and the mathematical model is run with the forecast results for a week and the data for the real-time idle vehicles.

The features of the model are determined based on the data analysis. As explained in Section 2.2.1., there is an obvious change in demand depending on the days of the week that are added to the model as categorical variables. Other variables are determined by intuition. The model has many features, and feature selection methods are used to eliminate irrelevant ones. The final list of the variables that have an impact on dependent variables is Days of the Week, Temperature and Precipitation, Number of Vehicles in the Fleet, and 15-minute-based Price. Weather conditions, such as rainy or sunny, affect the number of rentals. The unit of the temperature variable is Celsius, and the rainfall variable is grams of rain per square meter area (gr/m^2). The number of vehicles in the fleet gives the maximum number of vehicles available for operation on a single day. The total number of vehicles a company provides to the users can vary. This means that the number of vehicles in the fleet changes every day due to their maintenance or repair requirements. It is an effective feature on the number of rentals because more vehicles mean increased availability to users, and therefore, increased service level of the company. The number of vehicles in the fleet affects the availability of the vehicles to customers. Studies in the literature show that there is a positive correlation between the availability of vehicles and the frequency of use [24]. Also, the study states that availability is required to generate rentals, and customers likely value high-availability systems for their reliability [25]. In addition, [26] reports that 80% of the customers complete the trip, while the remaining 20% willing to use carsharing were forced to shift to other modes due to car unavailability. This shows the effect of the number of vehicles in the fleet on the demand. The fee of a specific rental is calculated with two parameters in the system, which are a dynamic fee per 15 minutes depending on vehicle models and a fee per extra mileage that is made during the rental. For example, there are 5 kilometers to use for free in a 15-minute rental, and if the user exceeds 5 kilometers during the 15-minute rental, extra money for extra-kilometer usage must be paid by the user.

Multiple linear and nonlinear forecast methods were implemented. The root mean square error (RMSE) value was calculated to see the test errors in all models. At the end of the analysis, it was decided to use the regression tree method which has the lowest RMSE, to estimate the number of rentals. Details of the model with the lowest error will be reported in the findings and discussion section.

2.2.4 Computational Details

All clustering and forecasting studies are implemented in R-Studio software. Computation time is significant for the developed model because the input includes the number of idle and active vehicles in the system and to get these inputs immediately due to dynamic demand and fluctuations in vehicle status. The mathematical model is run on GAMS Distribution 30.3.0 software, and the computation takes approximately 10 seconds, including the data import and export results from/to an Excel spreadsheet. The computation time to obtain this result satisfies the requirements, which means that the model can run at any time that is needed with the GAMS Solver license. The computer that is used for computation is a 15-inch MacBook Pro (Late 2016) manufactured by Apple Inc. has a 2.6GHZ quad-core Intel Core i7-6700HQ CPU, Radeon Pro 450 VRAM 2GB GPU, and 16GB LPDDR3 (2133 MHz) RAM.

3 Findings and discussion

In this section, all findings of the proposed method are explained. First, the relationship between features and demand is shown with correlation matrices. Figure 3 shows that there is a strong and positive correlation between weekends (Saturday and Sunday) and the number of rentals. On the other days, there is a negative correlation between predictors and responses. As mentioned in Section 2, the number of rentals on Saturdays and Sundays is higher than on other days. Therefore, it seems that Saturday and Sunday positively affect the number of rentals.

Figure 4 is related to other predictors that are used in the model. Temperature and precipitation data are provided by the weather services. According to the correlation matrix, the number of rentals is decreasing when the temperature is high. This is not an expected result. There is a negative correlation between the number of rentals and temperature with precipitation. This result may be related to the unwillingness of most people to go out on rainy days. Additionally, when the number of vehicles in the fleet increases, the number of rentals increases. The reason behind that is that people want to find an available vehicle nearby; otherwise, they may prefer to use other transportation options. Therefore, we can say that the availability of vehicles near users increases the number of rentals. In addition, it seems that the number of rentals increases as the price variables increase. This result is related to the regular growth of the system. Normally, the number of rentals is expected to decrease when prices increase. However, it can be explained that it has been two years since the company owned this brand, which has grown very fast, so the number of rentals increases regularly.

The last finding of this analysis is related to the sudden increase and decrease in demand due to system interruptions and advertisements. The reason behind this sudden drop may be system interruption. People cannot terminate their rentals when there is an interruption in the system. Interruptions other than regular system maintenance cause rental numbers to drop suddenly. An interruption in the system during the day can cause a sudden drop, as in the graph. One of the reasons that increases the error of the model is advertisements. Sudden increases in the number of rentals are observed on the advertise-supported days. Large-budget ads and sponsorships can suddenly increase the number of rentals. For example, airport product placement ads increased the number of airport rentals. As a second example, a huge audience in television ads sees the brand and registers on the same day to try it out. Therefore, the number of rentals suddenly increases. The fluctuating nature of the system can affect the accuracy of the forecast model.

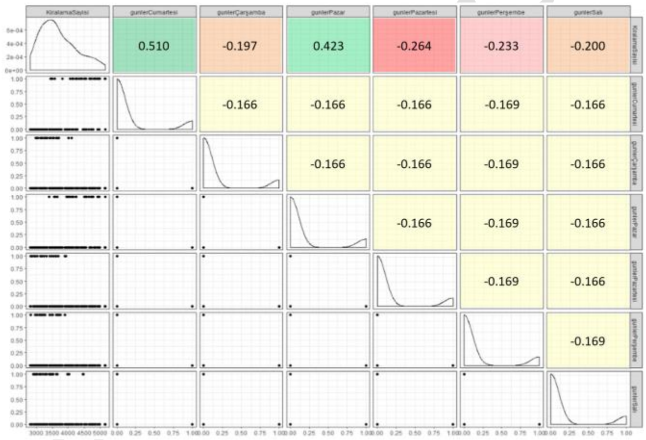


Figure 3. Correlation Matrix

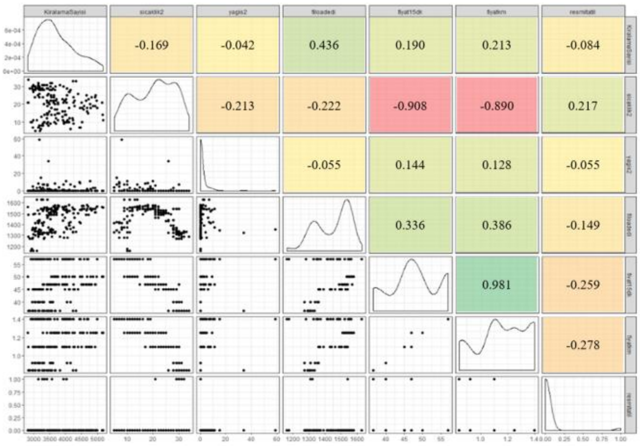


Figure 4. Correlation Matrix 2

3.1 Regression tree

The Regression tree method is one of the nonlinear methods. It branches the variables with the tree method, which creates a rule set to estimate the number of rentals. Model as in the formula below:

$$\begin{aligned} Number\ of\ Rentals \\ &= \alpha + \beta_1 Day\ + \beta_2 Temperature \\ &+ \beta_3 Fleet + \beta_4 15minPrice \end{aligned}$$

Figure 5 shows a regression tree. In the first node, the model separates weekdays from weekends using a day variable. When doing this, it performs a classification operation with the minimum error value. As expected, the model was divided into two on weekdays and weekends. The reason is that the number of rentals realized on weekends is higher than on weekdays. Since the day variable is categorical, it is divided into two branches: yes and no. Since other variables are numerical, other variables are divided into two according to a numerical value. The order of the variables used to create the model determines the importance of that variable. We can say that the most important variable for us is the day variable.

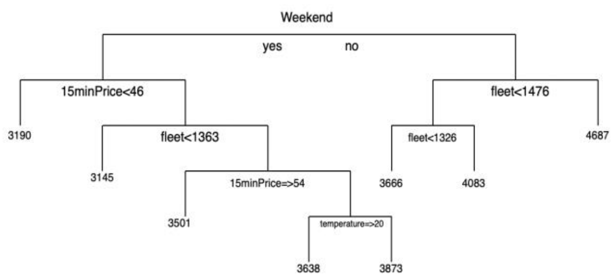


Figure 5. Regression tree

The regression tree yields RMSE, MAE, and MAPE values of 327, 257.8, and 7.1% respectively. Figure 6 indicates the regression tree output for which the blue line represents the actual number of rentals, while the red line shows the predicted.

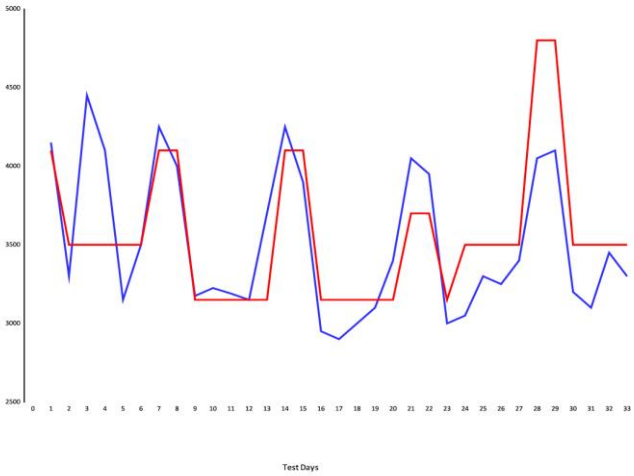


Figure 6. Difference between observed and test values

3.2 Model Outputs

The model is run for each day of the week, and each day begins with a time interval of 12 a.m. - 3 a.m. and ends with 9 p.m.- 12 a.m. There are eight time intervals in total. Moreover, to show the effect of relocation on profit, the model is run with an additional constraint (7) without altering the rest of the model.

$$r_{jlt} = 0 \quad \forall\, j, l, t \tag{7}$$

Constraint (7) states that the relocation from one cluster to another must be zero for every cluster and every time interval. Then, two model outputs can be compared. Table 4 shows that the system is more balanced and profitable with relocation. The number of satisfied demands and the profit of each day with relocation is always higher than those of the model without relocation. In addition, lost demand is extremely increasing without relocation, leading to a decrease in profit. From the beginning, the main purpose of the study was to show how profit can be increased by satisfying more demand using relocation.

When calculating the objective function, an incentive cost was not included, but a user-based relocation strategy can be is successful only by incentivizing customers. This was not considered because the incentive could be given specifically to user groups, and the model could not be built based on customers' specifications. The model output was built to reflect the number of relocations that are made. Also, the user-based relocation system is an emerging system, so it is not clear whether the users will accept or not to be relocated. Based on customer segmentation, which is discussed in the part of the user-based relocation strategy, users will be prioritized for the sustainability of this strategy. A part of the profit from relocation must be reserved as an incentive budget for the strategy by marketing team decisions.

The model tries to balance the distribution of the current positions of the vehicles by relocation. After the evaluation, the vehicles the model decides to relocate may need to be processed based on the sequence of waiting times in descending order. This is because vehicle waiting times may exceed more than three days with respect to current operations. Therefore, vehicle waiting times should be considered for this situation because one of the goals of this study is to decrease vehicle waiting times. Due to prioritization, the waiting time for vehicles can be minimized, and management can implement effective results for the system.

According to the results, relocations are mostly made between the adjacent clusters, considering the overall cost of the relocations. Depending on the day and time intervals, this direction may vary oppositely. The model considers the revenue and costs of that trip while determining whether the demand should be satisfied. If there is a demand that will make long trips, such as from the European side to the Asian side, the model makes relocation from neighbouring clusters to demand clusters to satisfy that demand.

Table 4. Model outputs

	With Relocation			Without Relocation		
Days	PSD (%)	PUD (%)	Profit (K TL)	PSD (%)	PUD (%)	Profit (K TL)
Monday	91	9	167.2	81	19	152.4
Tuesday	89	11	95.9	85	15	93.3
Wednesday	93	7	172.7	87	13	166.6
Thursday	90	10	133.7	82	18	122.9
Friday	91	9	180.5	85	15	179.8
Saturday	89	11	251.5	84	16	250.3
Sunday	92	8	269.8	76	24	232.5
Total	91	9	1,271.5	82	18	1,197.8

PSD: Percentage of Satisfied Demand

PUD: Percentage of Unsatisfied Demand

In Istanbul, public transportation lines are generally horizontal, such as the Metrobus line along the D100 highway, and people may have difficulties reaching their destinations vertically. Figure 7 shows the map view of the D100 highway with bold yellow line and vertical relocations with green arrowed lines. Commuters generally use at least two transportation options to reach their destination points. Therefore, with limited access to public transportation, car-sharing applications are used. In Figure 7, it is clearly stated that the general way of relocation is vertical lines, not horizontal.

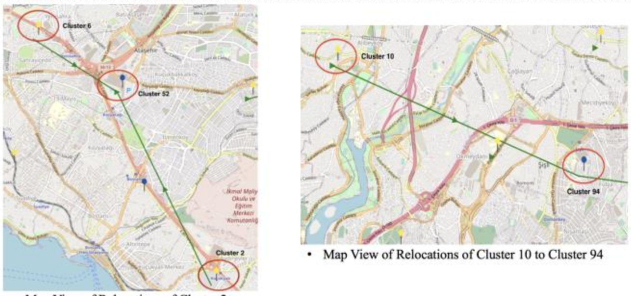
Figure 7. Map view of vertical relocations

3.3 Model Outputs Analyses

Some observations are classified related to the output of the model. Clusters are categorized according to customer trip characteristics.

3.3.1 Effect of shopping centers

To explain how shopping centers affect users' behaviors, clusters 52 and 10 can be analyzed. This region can be classified by hosting a large shopping center that has a large parking area, and people visit the center early with public transportation or other commuting options. It can be mentioned that shoppers usually prefer to visit there around 11 a.m. - 12 p.m. on weekends. Then they prefer to leave after having dinner, so the demand for vehicles increases in the evenings. In this cluster, the vehicles are not relocated and continue to cover the demand for upcoming intervals. Moreover, some relocations are made from its neighbors to meet demands. Figure 8 shows that relocations from its adjacent clusters, which are 76, 2, and 6, are needed to satisfy customers who need a comfortable trip home after shopping because they may need to carry heavy bags, which can make them feel tired.



- Map View of Relocations of Cluster 2, Cluster 6 to Cluster 52
- Map View of Relocations of Cluster 10 to Cluster 94

Figure 8. Map view of relocations

Investigating cluster 10, it is already known that many students living in the region use the car-sharing application frequently depending on the survey that is conducted, as many universities are in this region and public transportation is known to be insufficient. There is no need to relocate within weekdays because cluster 10 is a starting and ending cluster, making it balanced. For this reason, on weekdays, the vehicles in this cluster are not relocated and kept considering the demand of the next period. On the other hand, a decrease in demand on weekend mornings leads to a surplus in cluster 10, so relocation is made from cluster 10 to adjacent clusters 9, 94, and 86. It can be said that cluster 10 is surrounded by these clusters with many shopping centers. Therefore, people want to spend their time on weekends going to the gym, restaurants, and stores that are in shopping centers, such as Trump Towers and Cevahir Mall in Cluster 94, and Vialand in cluster 86. Also, cluster 94 has the highest number of relocations that are made into this cluster on weekends due to the popularity of the entertainment areas.

3.3.2 Effect of airports

According to model outputs, vehicles are kept in cluster 45, which covers Istanbul Airport. The reason for keeping these vehicles where they are is that the outgoing vehicles of the cluster and the incoming vehicles to the cluster are balanced. Also, the parking problem is not an issue for this cluster. Due to the availability of parking areas, cars are kept for the supply of next time intervals, and relocation is not needed with the model outputs. Therefore, the number of vehicles that should be kept to satisfy the demand for the next time interval is known. On the other hand, this is not the same for cluster 83, which covers Sabiha Gokcen Airport, because there is an imbalance of demand and supply. Figure 9 shows a map view of adjacent clusters and Sabiha Gokcen Airport. If necessary, relocation can be made at the right time, parking problems will not be an issue for this cluster, and demand can be satisfied.

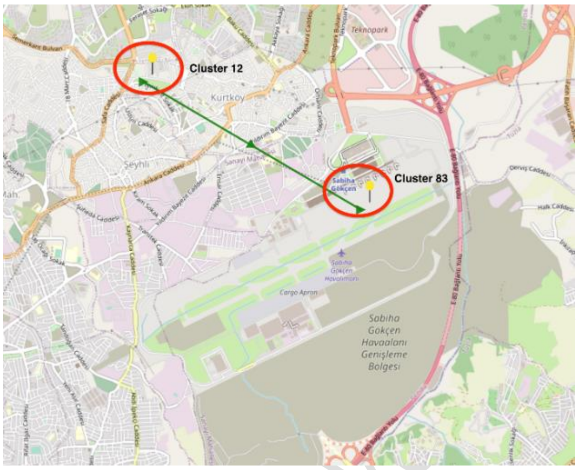


Figure 9. Example of relocation

3.4 User-based relocation strategy

In the literature, most of the relocation strategies studied were based on operators. Operator-based strategy means that the total number of operators of the company is known, and it is certainly known that the operator will relocate the car at any time needed because it is a controllable variable for the company. On the other hand, when a user-based strategy is considered, it is much more difficult to implement. The demand is an uncertain variable, making it difficult to solve the problem. In addition to this, future demand is necessary to relocate at the same time. It is a probabilistic approach and adds another uncertainty to the problem while making it more realistic. Using this strategy is better from a cost perspective because operators are very costly to the company. As a solution approach, a user-based relocation strategy is recommended. This strategy, which is based on the participation of customers in the system by giving incentives, does not involve all users of the system. It gives incentives to the customers who benefit from the system, like having no records on the debt list, good driving records, clean usage, etc., because the total cost of a customer exceeding the company benefit is an undesired case.

Customer segmentation is the first step for a user-based relocation strategy. Customers are segmented according to usage frequency and recency, which means the last usage of the customer and the total amount they paid. It can be said that behavioral segmentation is applied. For the strategy's applicability, it is important to have both a usage rate and customer loyalty to the application. With this method, the most accurate campaign will be published in the right customer group. A technique developed in the work of Kotler and Armstrong classifies the customers into four groups to build the right relationship with the right customers, as the segments are shown in Figure 10a [27]. The result of this segmentation enables accurate customer-type definitions. The primary customer group consists of individuals who rent cars frequently and spend substantial amounts on these rentals. These are called True Friends; the company does not want to lose them to maintain the system. Barnacles segment consists of users with short-term rentals with frequent car usage. This class of customers is important to help with vehicle relocation. "Butterflies" are customers who spend generously but have infrequent app usage. This group should be encouraged to make more leases with campaigns. Strangers group provides small returns with renting less often. If they frequently create inconvenience (accidents, traffic tickets, etc.) to the system, they will be removed from the system. For a user-based relocation strategy, customers who are always with the company, meaning that they are often connected to the system by using the application, are important to achieve the sustainability of the strategy. The company can focus on True Friend and Barnacle segments for that purpose. Therefore, loyalty is one of the important criteria for an effective and sustainable relocation strategy. It is important to stay in touch with customers who bring high returns by phone calls to increase their usage frequency, as they are close to becoming True Friends.

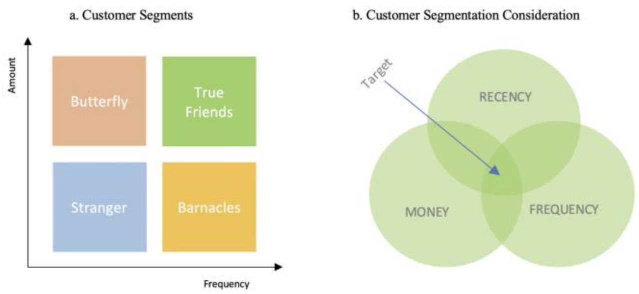


Figure 10. Customer segmentation

When campaigns are offered to users, the frequency and amount of money to be paid are the most important indicators, but the recency status of the customers is also an indicator. Therefore, we should not offer campaigns to users who have not used the application recently. The intersection of these clusters will be to target customers for our strategy. Figure 10b shows the target customers.

Furthermore, a gamification strategy could be developed for the system's sustainability, as the idea allows users to earn badges and points based on their rental status within the application. The point-based system can be converted into rewards for upcoming rentals. Therefore, a suitable relocation system can be generated by providing rewards on a gamified system for enhanced entertainment. Furthermore, potential users may want to be part of this system based on user feedback.

4 Conclusion

A relocation strategy is introduced for a free-floating car-sharing system. The model handles clustering and forecasting parameters as inputs. The outputs of the model are the number of idle vehicles at the beginning of time interval t, the number of satisfied demands, the number of unsatisfied demands, and the number of relocations that are made based on clusters. The relocation model allows the company to allocate resources (vehicles) to high-demand points from low-demand points effectively concerning the balance of cost and revenue. The user-based strategy has less cost than the operator-based in practice. Still, it may create some difficulties in terms of reliability and the time it takes for customers and companies to get used to the system. Although there are numerous difficulties, this system may potentially bring new customers to the system, as well as increase the satisfaction of existing users.

Free-floating is the recent model of the types of car-sharing system. In this setup, a company's fleet of vehicles is shared among system members, allowing for pick-up and drop-off. Since 2019, FFCS has been developed in Istanbul. From the literature point of view, FFCS is still improving in terms of the accessibility of cars for people who are using this system. The convenience for customers to leave vehicles wherever they want adds uncertainty to the system and makes it more challenging to manage. However, this uncertainty creates convenience and ease of use for users. In addition, considering the costs of alternatives such as taxi or private car ownership, car-sharing platforms offer affordable prices to their customers. Therefore, customers decide to use these platforms due to the low-cost system. Another issue is environmental effects. Traffic and crowds are major problems in Istanbul. Today, there has been an increase in the number of private cars. This increases traffic and CO_2 emissions, causing air and noise pollution. Therefore, such systems are beneficial for avoiding traffic jams while decreasing CO_2 emissions. Considering these alternatives, a customer-based relocation strategy is expected to be beneficial to the case company.

For future studies, some enhancements and improvements may be considered. For example, due to the bird flight distance of clusters, the cost function does not effectively work because the impact of the traffic jam and the real distances are not considered. Hence, the total cost may increase. In some cases, this is also caused by different rentals with the same start-end points and may cover different distances in terms of the preferences of different routes. The cost and revenue functions should be developed to achieve more accurate results and make the model more realistic. In this way, clusters can be prioritized for the demand that must be covered. Also, one of the future extensions can be the stochastic setting for demand. Since our study covers deterministic demand based on forecast methods using historical data. Stochastic demand setting may be more realistic, as the better performance of the model depends on how accurately the demand is predicted. As a result, the introduced relocation model can be applied to any free-floating sharing system. With the adoption of this system, companies can increase their revenues, as well as customer loyalty and satisfaction.

5 Author contribution statements

All authors who have worked together on this study and contributed equally.

6 Ethics committee approval and conflict of interest statement

Ethics committee approval is not needed for this article. Also, there is no conflict of interest with any person/ institution in the article prepared.

7 References

[1] B. Kartal Çetin and N. K. Pratas, "Resource sharing and scheduling in device-to-device communication underlying cellular network," *Pamukkale University Journal of Engineering Sciences*, vol. 27, no. 5, pp. 604–609, 2021, doi: 10.5505/pajes.2021.71597.

[2] B. Yıldız, "Crowd-shipping service network design problem," *Pamukkale University Journal of Engineering Sciences*, vol. 28, no. 1, pp. 104–116, 2022, doi: 10.5505/pajes.2021.83809.

[3] N. Agatz, A. Erera, M. Savelsbergh, and X. Wang, "Optimization for dynamic ride-sharing: A review," *Eur J Oper Res*, vol. 223, no. 2, pp. 295–303, 2012.

[4] Z. T. Qin, H. Zhu, and J. Ye, "Reinforcement learning for ridesharing: A survey," in *2021 IEEE International Intelligent Transportation Systems Conference (ITSC)*, 2021, pp. 2447–2454.

[5] V. A. Alencar, F. Rooke, M. Cocca, L. Vassio, J. Almeida, and A. B. Vieira, "Characterizing client usage patterns and service

demand for car-sharing systems," *Inf Syst*, vol. 98, p. 101448, 2021. [22]

[6] B. Boyacı, K. G. Zografos, and N. Geroliminis, "An optimization framework for the development of efficient one-way car-sharing systems," *Eur J Oper Res*, vol. 240, no. 3, pp. 718–733, 2015. [23]

[7] B. Boyacı, K. G. Zografos, and N. Geroliminis, "An integrated optimization-simulation framework for vehicle and personnel relocations of electric carsharing systems with reservations," *Transportation Research Part B: Methodological*, vol. 95, pp. 214–237, 2017.

[8] M. Bruglieri, A. Colorni, and A. Luè, "The relocation problem for the one-way electric vehicle sharing," *Networks*, vol. 64, no. 4, pp. 292–305, 2014.

[9] M. Bruglieri, A. Colorni, and A. Lue, "The vehicle relocation problem for the one-way electric vehicle sharing: an application to the Milan case," *Procedia-Social and Behavioral Sciences*, vol. 111, pp. 18–27, 2014.

[10] M. Bruglieri, F. Pezzella, and O. Pisacane, "Heuristic algorithms for the operator-based relocation problem in one-way electric carsharing systems," *Discrete Optimization*, vol. 23, pp. 56–80, 2017.

[11] M. Bruglieri, F. Pezzella, and O. Pisacane, "An adaptive large neighborhood search for relocating vehicles in electric carsharing services," *Discrete Appl Math (1979)*, vol. 253, pp. 185–200, 2019. [24]

[12] D. Jorge, G. H. A. Correia, and C. Barnhart, "Comparing optimal relocation operations with simulated relocation policies in one-way carsharing systems," *IEEE Transactions on Intelligent Transportation Systems*, vol. 15, no. 4, pp. 1667–1675, 2014.

[13] M. Barth, M. Todd, and L. Xue, "User-based vehicle relocation techniques for multiple-station shared-use vehicle systems," 2004.

[14] M. Clemente, M. P. Fanti, A. M. Mangini, and W. Ukovich, "The vehicle relocation problem in car sharing systems: Modeling and simulation in a petri net framework," in *International Conference on Applications and Theory of Petri Nets and Concurrency*, 2013, pp. 250–269.

[15] [15] A. Di Febbraro, N. Sacco, and M. Saeednia, "One-way carsharing: solving the relocation problem," *Transp Res Rec*, vol. 2319, no. 1, pp. 113–120, 2012.

[16] [16] A. Di Febbraro, N. Sacco, and M. Saeednia, "One-way car-sharing profit maximization by means of user-based vehicle relocation," *IEEE Transactions on Intelligent Transportation Systems*, vol. 20, no. 2, pp. 628–641, 2018.

[17] [17] S. Weikl and K. Bogenberger, "Integrated relocation model for free-floating carsharing systems: Field trial results," *Transp Res Rec*, vol. 2563, no. 1, pp. 19–27, 2015.

[18] [18] W. Herbawi, M. Knoll, M. Kaiser, and W. Gruel, "An evolutionary algorithm for the vehicle relocation problem in free floating carsharing," in *2016 IEEE Congress on Evolutionary Computation (CEC)*, 2016, pp. 2873–2879.

[19] [19] D. Jorge, G. Molnar, and G. H. de Almeida Correia, "Trip pricing of one-way station-based carsharing networks with zone and time of day price variations," *Transportation Research Part B: Methodological*, vol. 81, pp. 461–482, 2015.

[20] [20] S. Ren, F. Luo, L. Lin, S.-C. Hsu, and X. I. Li, "A novel dynamic pricing scheme for a large-scale electric vehicle sharing network considering vehicle relocation and vehicle-grid-integration," *Int J Prod Econ*, vol. 218, pp. 339–351, 2019.

[21] [21] M. Xu, Q. Meng, and Z. Liu, "Electric vehicle fleet size and trip pricing for one-way carsharing services considering vehicle relocation and personnel assignment," *Transportation Research Part B: Methodological*, vol. 111, pp. 60–82, 2018.

[22] S. Reiss and K. Bogenberger, "Validation of a relocation strategy for Munich's bike sharing system," *Transportation Research Procedia*, vol. 19, pp. 341–349, 2016.

[23] L. Cagliero, S. Chiusano, E. Daraio, and P. Garza, "CarPredictor: forecasting the number of free floating car sharing vehicles within restricted urban areas," in *2019 IEEE International Congress on Big Data (BigDataCongress)*, 2019, pp. 72–76.

[24] F. Ferrero, G. Perboli, M. Rosano, and A. Vesco, "Car-sharing services: An annotated review," *Sustainable Cities and Society*, vol. 37. Elsevier Ltd, pp. 501–518, Feb. 01, 2018. doi: 10.1016/j.scs.2017.09.020.

[25] C. Hardt and K. Bogenberger, "Empirical Analysis of Demand Patterns and Availability in Free-Floating Carsharing Systems," in *IEEE Conference on Intelligent Transportation Systems, Proceedings, ITSC*, Institute of Electrical and Electronics Engineers Inc., Dec. 2018, pp. 1186–1193. doi: 10.1109/ITSC.2018.8569700.

[26] L. M. Martínez, G. H. de A. Correia, F. Moura, and M. Mendes Lopes, "Insights into carsharing demand dynamics: Outputs of an agent-based model application to Lisbon, Portugal," *Int J Sustain Transp*, vol. 11, no. 2, pp. 148–159, Feb. 2017, doi: 10.1080/15568318.2016.1226997.

[27] P. Kotler and G. Armstrong, *Principles of marketing*. Pearson education, 2010.