6-18-2022

Increasing the Business Value Of Free-Floating Carsharing Fleets By Applying Machine-Learning Based Relocations

Christoph Prinz
University of Goettingen, christoph.prinz@uni-goettingen.de

Malte Hellmeier
Fraunhofer Institute for Software and Systems Engineering ISST, malte.hellmeier@isst.fraunhofer.de

Mathias Willnat
University of Goettingen, mathias.willnat@uni-goettingen.de

Christine Harnischmacher
University of Goettingen, christine.harnischmacher@uni-goettingen.de

Lutz Kolbe
University of Göttingen, lkolbe@uni-goettingen.de

INCREASING THE BUSINESS VALUE OF FREE-FLOATING CARSHARING FLEETS BY APPLYING MACHINE LEARNING-BASED RELOCATIONS

Research Paper

Christoph Prinz, University of Göttingen, Göttingen, Germany, christoph.prinz@uni-goettingen.de

Malte Hellmeier, Fraunhofer Institute for Software and Systems Engineering ISST, Dortmund, Germany, malte.hellmeier@isst.frauenhofer.de

Mathias Willnat, University of Göttingen, Göttingen, Germany, mathias.willnat@uni-goettingen.de

Christine Harnischmacher, University of Göttingen, Göttingen, Germany, christine.harnischmacher@uni-goettingen.de

Lutz Kolbe, University of Göttingen, Göttingen, Germany, lkolbe@uni-goettingen.de

Abstract

Free-floating carsharing (CS) services provide customers with a fleet of vehicles distributed within an operation area. These services gained popularity because of their positive impact on societal and personal mobility. Understanding determinants of customer demand is a key challenge for developing and applying vehicle relocation strategies to prevent the formation of undersupply areas. In this study, we merge possible features from publicly available data sources with historical demand from CS services situated in three different-sized cities. We train and test a Random Forest (RF) regressor estimating demand based on the enhanced dataset. Building on this demand prediction, we developed a relocation strategy that optimizes vehicle availability at anticipated demand points. Our strategy improved the reservation acceptance ratio in all three reference systems between 7.1 % and 15.6 %. Furthermore, the number of relocations compared to a deterministic relocation strategy could be reduced by 82.3 % and 20.6 % in two cities.

Keywords: Free-Floating Carsharing, Relocation, Machine Learning, Big Data.

1 Introduction

CS services provide customers access to a fleet of shared vehicles, which are available for pick-up and drop-off within an operation area or at predefined stations and are mainly used for short-term rentals (Schmöller and Bogenberger 2020). Such services are a prime example of the sharing economy (Frey et al. 2019) and contribute to the necessary transformation of the mobility sector (Wittwer and Hubrich 2018). On a societal level, CS reduces the amount of public space needed for parking, car traffic congestion, carbon dioxide emissions, and noise pollution when considering electric vehicle fleets (Amatuni et al. 2020; Wappelhorst et al. 2014). On an individual level, CS systems offer access to heterogeneous vehicle types, discharge responsibilities for privately owned cars and cut mobility expenses (Jochem et al. 2020). Compared to schemes based on designated rental stations, free-floating systems concurrently increase access flexibility and egress convenience, constituting important success factors of such systems (Wittwer and Hubrich 2018) and acceptance constraints to the wider adoption of CS (Hahn et al. 2020). Vehicle availability suffers from spatial and temporal imbalances of supply and demand. When analyzing the historical booking data of a free-floating CS system, we determined that 85.9 % of the trips end at a different drop-off location compared to the pick-up location in a small-sized German city, which is likewise reported by Weikl and Bogenberger (2013). Various relocation strategies in which employees of the CS operator or incentivized users bring back the cars to high-demand locations are discussed in academia (Illgen and Höck 2019) to increase vehicle accessibility and address issues associated with drop-offs in low-demand areas. Thereby, the number and size of relocations must be economically viable to increase demand without losing sight of associated costs (Weikl and Bogenberger 2015).

Although start-ups and established providers are already active in the domain of carsharing, there is still a need to enhance understanding and advance research on vehicle balancing strategies (Nansubuga and Kowalkowski 2021). Especially the conjunction between accurate demand forecasts and the current vehicle supply is often neglected in vehicle relocation models (Jian et al. 2019).

The majority of existing research focuses on relocation optimization, mainly in station-based CS systems, while their business value is not validated or only validated within a single reference system (Illgen and Höck 2019). Most approaches primarily focus on classical baseline techniques instead of using the full potential of diverse data sources and new Machine Learning (ML) techniques. Existing research entailing neural networks or RF regressors to model CS mainly focuses on finding hyperparameters to minimize the quantitative prediction error (Cagliero et al. 2019; Cocca et al. 2020; Li et al. 2021). In contrast, such ML-based algorithms have not been turned into purposeful IT artifacts concerned with solving the business problem of vehicle relocations, including demonstrating economic value and managerial implications of the design method as suggested by Niederman and March (2012), and Rai (2017). Higher profitability through increased efficiency is essential to continuously expand CS networks and make them an integral part of sustainable future mobility (Illgen and Höck 2019). To address those research gaps and contribute to this field, our work aims at answering the following research question (RQ):

RQ: *How should data-driven vehicle relocations be designed to increase the business value of free-floating vehicle fleets?*

To answer the RQ, we design, develop, and evaluate a ML-based relocation approach based on three real-world carsharing booking datasets from a small city, a large city, and a metropolis. We structure this paper as follows: Section 2 provides an overview of the fundamental literature regarding vehicle relocation solutions and demand modeling approaches. In section 3, we describe our research methodology, which follows the design science paradigm introduced by Hevner et al. (2004). This also includes the derivation of requirements on the design of our relocation strategy. The structure of our datasets, the data preparation and enhancement process, the model training and hyperparameter optimization, and the instantiation of our vehicle relocation strategy are presented in section 4. We evaluate our final artifact in the form of a simulation in three real-world CS systems presented in section 5. Afterward, we discuss theoretical and practical implications, as well as limitations and future research possibilities in section 6. We finish our paper by concluding our work in section 7.

2 Literature Review

To evaluate the current status quo of vehicle relocation research and demand and supply prediction solutions in CS and related domains, we conducted a literature review based on the approach by Webster and Watson (2002). Our literature base combines the research considered in the reviews by Cepolina et al. (2014), Illgen and Höck (2019), and Jorge and Correia (2013), extended by a keyword search in the databases of AIS e-library, Science Direct, IEEE Xplore, Google Scholar and MDPI (keywords: (“carsharing“ ∪ “car-sharing“ ∪ “car sharing”) ∩ (“relocation” ∩ “machine learning”)). The results are filtered depending on their title and abstract referencing the implementation of vehicle relocation algorithms or supply, demand, and prediction-based topics. Then, we conducted a backward search by reviewing the citations of the identified publications. We only considered publications after 2010 to keep the focus on the latest relocation strategies. We analyzed the resulting set of 55 papers and clustered them into one of two categories *research on vehicle relocations* (section 2.1) and *demand prediction approaches* (section 2.2). The first category mainly focuses on solutions for the relocation problem (46 papers) while the second one focuses on the identification of supply and demand factors in CS systems (9 papers). Since both categories are not clearly distinguishable, some papers handle relocations as well as supply and demand factors, so they are assigned to the cluster of their main topic to avoid duplicates.

2.1 Vehicle Relocation Research

Since the service quality of CS systems suffers from supply and demand imbalances (Weikl and Bogenberger 2014), the main operational challenge of CS providers is to relocate cars from low to high-demand areas. The resulting increase in customer acceptance and reduction of vehicle idle times is a key success factor for the rise of CS as a sustainable mobility option (Schmöller et al. 2015). Besides the CS ecosystem, rebalancing of vehicles is also discussed as redistribution in other domains like taxi or on-demand mobility service systems (Lei et al. 2020; Marouf et al. 2014; Smith et al. 2013). However, the adaptability of such studies to the context of CS is limited (Illgen and Höck 2019; Schulte and Voß 2015): On the one hand, the physical relocation in such systems is less personal-intensive, because trucks can be used to relocate multiple vehicles (e.g. e-scooter or bikes). On the other hand, the fleet size can be increased more easily, because the investment requirements are less, fewer regulations, and fewer/no parking fees need to be paid.

According to the literature reviews of Cepolina et al. (2014), Illgen and Höck (2019), and Jorge and Correia (2013), relocation research commonly focuses on improving a CS system’s performance measures by finding decision variables that enhance the system’s capability to serve customers (e.g. leading to more accepted rental requests) and, subsequently, generate profits. The related customer behavior can be described with the transaction costs theory (Rindfleisch 2020). If customers face a decision for a means of transportation, they want to maximize their individual benefit (homo economicus). The walking distance to the next available vehicle can be considered a transaction cost. The gained utility of choosing the carsharing service must justify the initial effort to get to the car. Consequently, a carsharing system with effective relocations leads to small transaction costs and hence to sustainable business success.

CS relocation strategies are modeled using (mixed-)integer/linear programming (Carlier et al. 2015; Gambella et al. 2018; Santos et al. 2017; Weikl and Bogenberger 2015), simulation- (Alfian et al. 2017; Benarbia et al. 2020; Brendel et al. 2018; Kypriadis et al. 2020; Lopes et al. 2014), machine learning (Li et al. 2021) or combined approaches (Ait-Ouahmed et al. 2018; Wang et al. 2019; Zakaria et al. 2014).

The majority of presented solutions (34/46) target station-based CS systems, where customers can pick-up and drop-off vehicles only at predefined stations. However, free-floating CS systems are more attractive to customers since they offer more flexibility (Niels and Bogenberger 2017), by providing station-independent drop-offs and pick-ups within operation areas of cities. Nevertheless, only a minority of presented relocation strategies (12/46) are adapted to such distribution models.

Most presented research approaches are developed and evaluated based on historical booking data and system properties from a single real-world system (39/46). None of the presented strategies are applied to and evaluated in multiple systems. Furthermore, reference systems in cities with less than 200 000 citizens are underrepresented (3/46). This is especially problematic since customers in large cities can rely on multiple mobility and CS options, while customers in small cities and rural areas are often dependent on a single provider (Shaheen et al. 2020).

To conclude the status of carsharing relocation research, there is a gap in effective and adaptable relocation strategies for free-floating systems in small and medium sized cities. Presented research rarely uses real world data and is not challenged across multiple case studies. Thus, we focus on tackling the limitations of the presented research.

2.2 Data Driven Demand Prediction Approaches

To identify and reduce imbalances in CS systems, many solutions are based on an upstream supply and demand concept (Cocca et al. 2020; Nair and Miller-Hooks 2011; Repoux et al. 2015; Schulte and Voß 2015; Weikl and Bogenberger 2015). Especially the following ML-based approaches show promising results in predicting user demand.

Daraio et al. (2020) compare different ML techniques with classic deterministic approaches to predict car availability in a free-floating CS system in the following. They enhance historical booking data with additional weather, point-of-interest (POI), and neighborhood information. Cocca et al. (2020) aim to predict the demand in a free-floating CS system. In a first step, they compare ML techniques to predict future demand based on historical data. In a second step, the approach is extended by socio-demographic data to help CS operators to decide whether to expand the operation area into the neighborhood or not. Cagliero et al. (2019) address the problem of forecasting the number of cars in a free-floating CS system, tested with real-world data from the Car2Go operator in Portland, Oregon. With the help of geographical positions, POI, and weather data, the authors formulate a multivariate regression problem. By comparing ML and baseline techniques, they identify RF Regression as the best approach for the analyzed context. Schmöller et al. (2015) focus on an empirical analysis of a free-floating CS system in the German cities Berlin and Munich. The authors use and prepare historical data to classify supply and demand factors. They identify, for example, that weather influences demand in free-floating systems and high-demand areas are often related to a high density of companies. Further, free-floating CS systems are mainly used by young people in small households, and the demand is low on Mondays but increases during the week based on their analysis. Li et al. (2021) propose a non-parametric learning algorithm and two-stage stochastic programming modeling technique to reduce the vehicle relocation rate. They assume customer demand from New York taxi travels and evaluate their model based on a fictive station-based CS system.

The synthesis of features suggested by the five aforementioned studies builds the foundation of our data enhancement and informs our prediction design. Even though those studies suggest features and ML techniques to implement demand prediction models, none utilized and evaluated a suggested model built on real-world data into a relocation strategy. The majority of studies (4/5) evaluate their prediction performance only based on ML models and do not prove their effectiveness regarding potential business impact. Thus, we aim in presenting the first machine learning based relocation strategy, which has been evaluated against business value across three different scenarios.

3 Research Approach

We follow a design science research approach that constitutes an effective and efficient problem-solving paradigm that supports researchers in producing innovative ideas, practices, technical capabilities, and products for analyzing, designing, implementing, managing, and using information systems (Hevner et al. 2004). Our problem-solving process is executed in three iterations and structured with the interaction of relevance, design, and rigor cycles following Hevner (2007). We performed three iterations in total, starting with *exploratory booking data analysis & enhancement* (section 3.1), continuing with the *implementation of supply and demand-based relocations* (section 3.2), and completing with an *evaluation of our strategy in a carsharing simulation* (section 3.3).

3.1 Iteration 1: Perform Exploratory Booking Data Analysis & Enhancement

The first iteration aims to understand the problem domain, pre-process the CS datasets from three different cities, and analyze correlations between the included data features. Starting with a relevance cycle, we performed three expert interviews with practitioners related to the CS industry. Two of the selected experts are responsible for the strategic fleet operation at two different carsharing providers, where the first is a local provider and the second is a global one. The third expert is an experienced and practice-oriented researcher in the domain of carsharing relocations. They confirm that relocation algorithms are required to tackle the issue of low user acceptance and vehicle idle times. Furthermore, they state that the potential of available data sources goes beyond historical demand but is not unfolding completely yet. The requirements are summarized in Table 1 and build the foundation for the design of our data-driven demand prediction approach.

No.	Name	Explanation
R1	Publicity	External used third-party data sources are publicly available; It must not be required to use proprietary data sources
R2	Adaptability	Adaptable for different CS systems in different cities
R3	Specificity	Observance of the individual spatial and temporal factors

Table 1. Requirements for Data Enhancement.

Within the rigor cycle, existing solutions and possible supply and demand factors are identified. The results of the rigor cycle informed our artifact design and were already presented in section 2. Besides that, the input of the design cycle is historical booking data from three different free-floating CS systems situated in a small city, a large city, and a metropolis. Following Géron's (2019) ML procedure model, the collected data points were subjected to an exploratory data analysis in the first step. Secondly, a data procurement, clustering, and enhancement process is performed based on these data points. While applying an agile implementation model and exchanging knowledge continuously with our partners, the first demand prediction model is developed.

3.2 Iteration 2: Implement Supply and Demand based Relocations

We follow Hevner et al. (2004) to create an implementable artifact within the second iteration instead of a highly abstract concept. More specifically, we transform the outputs of our demand prediction model into relocation recommendations. As part of our second relevance cycle, we validated that the features building our prediction model match the experience of the CS provider. Brendel et al. (2018) determined requirements for CS relocation strategies which are presented in Table 2. The requirement set is confirmed in our interview and therefore, also guides the implementation of our relocation strategy.

No.	Name	Explanation
R4	Availability	Improving the vehicle acceptance ratio in the entire system
R5	Necessity	Only necessary relocations (e.g., a relocation should lead to a rental)
R6	Automation	Automation of decisions if and how a relocation should be performed
R7	Transparency	The system should supply humanly understandable outputs, providing transparency regarding their computation

Table 2. Requirements for Relocation Strategy.

Literature findings inform the artifact design about free-floating relocation strategies suitable to fulfill the requirement set. Like the first iteration, designing the relocation strategy is done by applying agile methods and validating its practicality with our partners.

3.3 Iteration 3: Evaluate Strategy in Carsharing Simulation

The goal of the third iteration is to evaluate whether our proposed relocation strategy fulfills the complete requirement set in the business contexts of the three CS providers. Therefore, we implement our artifact design into a Python based simulation environment. It determines the influence of relocation strategies on the CS system performance. To investigate the effectiveness of our suggested relocation strategy, we compare its performance by applying no relocations and a state-of-the-art deterministic approach. The goal of the optimization is to get the highest customer acceptance ratio by employing as less relocations as possible. To make the results generalizable, we do not consider any personnel, maintenance, or other system constraints for executing relocation recommendations. Our research project is completed with a rigor-cycle, where we present our findings to the audience of our paper.

4 Results

In the following sections, we present the key findings of our DSR iterations. This includes the *design of a demand prediction module* (section 4.1), the *design of our free-floating vehicle relocation approach* (section 4.2), and the *evaluation of findings in a simulation* (section 5).

4.1 Design of a Demand Prediction Module

For the training and test of our ML-based demand prediction module and basis for the simulation-based evaluation, we use a subset out of three historical booking datasets from different free-floating CS providers (R2). Dataset 1 and dataset 3 have been provided from the corresponding carsharing providers, while dataset 2 was crawled from a public booking platform. An analysis of dataset properties like median idle times, trip lengths, trip durations, and booking frequencies over time did not reveal significant differences to the empirical analysis of other carsharing data presented in literature (Costain et al. 2012; Schmöller et al. 2015). The properties of the considered CS systems are summarized in Table 3 and differ in timespan, number of rentals, city, and fleet size.

Property	**Dataset 1** **Small-City**	**Dataset 2** **Large-City**	**Dataset 3** **Metropolis**
Approx. City Population	120 000	600 000	3 700 000
Timespan	Jan 2017 – Jul 2019	Feb 2020 – Nov 2020	Jan 2020 – Oct 2020
Number of Bookings	64 812	180 143	1 361 138
Fleet Size (Number of Cars)	31	198	1500

Table 3. Overview CS Booking History Datasets

A booking is defined as a rental with vehicle movement. Short-time reservations without actual car usage and service maintenance tours were excluded for datasets 1 and 3. An exception is a minimal number of private bookings from employees. Every entry consists of a car ID, the pick-up and drop-off time, the pick-up and drop-off location as coordinates in the global positioning system (GPS), and the distance traveled by the rented vehicle in kilometers. To protect the privacy of their customers, the provider of dataset 3 applied a random noise of 200 meters on location related data columns.

In free-floating CS systems, users can pick-up and return cars from anywhere inside the provider's operation area (Wagner et al. 2015). We decide to set the spatial resolution of the prediction-model based on the user's average willingness to walk a distance of 500 meters to pick-up an available vehicle (Herrmann et al. 2014). Thus, the optimal grid size is $\frac{500}{\sqrt{2}} \approx 353.553m$ per edge length, resulting in a maximum walking distance of 500 meters within the same squared cell (R2).

The prediction model's performance improves, the more potential decision variables of CS customers are included in the training set. Therefore, we enhance all three historical booking datasets with a synthesis of the features suggested by Cagliero et al. (2019), Cocca et al. (2020), Daraio et al. (2020), and Li et al. (2021). Consequently, we enhance the dataset with hourly weather conditions like wind, clouds, rain, snow, and temperature from OpenWeather[1], as well as time-dependent data like the weekday and time of the day (Cagliero et al. 2019; Daraio et al. 2020; Schmöller et al. 2015) (R1, R2, R3). Furthermore, we determine the POI density of every grid element by counting the number of

[1] Available under https://openweathermap.org/history-bulk

POIs included in the SLIPO[2] dataset for every grid cluster (R1, R2, R3). Every booking is enhanced with the density of the drop-off and pick-up grid position (Cagliero et al. 2019; Daraio et al. 2020; Willing et al. 2017) (R2, R3).

We complete the data preparation phase by performing an exploratory data analysis on the enhanced dataset (Géron 2019). The number of bookings in the previously introduced squared cells is the dependent variable that has to be predicted in the latter course. Therefore, it is examined whether the extended information about weather, temperature, POIs, and time has an impact on the number of bookings.

Based on the findings of Cagliero et al. (2019), and Cocca et al. (2020), we cluster the enhanced dataset into a three-hour prediction horizon, which results in eight-time bins per day. Consequently, the sum of bookings per subset is added to every cluster (R2). The data preparation process is summarized in Figure 1.

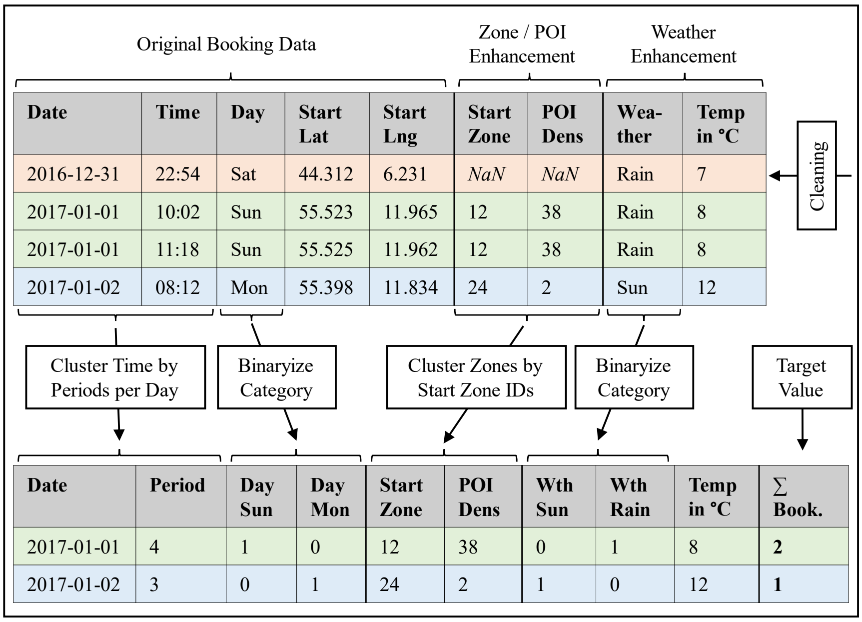

Figure 1. Summarized Data Preparation Process.

After enhancing, preparing, and clustering the data set, an ML model is trained to predict the future demand. Adapting the findings of Cagliero et al. (2019), Cocca et al. (2020), and Daraio et al. (2020), we use an RF as a regressor (R2). The future demand is represented by the number of bookings per period and cluster as a dependent variable. It is explained by the other features like temporal, zone, POI density, or weather data as independent variables (R2, R3).

To train the model, we apply the Gini impurity-based RF regressor included in the scikit learn Python package. The datasets are split randomly into two subsets for training and testing with a common 80 % / 20 % allocation (Géron 2019). The independent variables for grid-position and timeslot have to be provided to the RF regressor to predict the future demand. It has to be noted that the training data set encompasses past weather conditions and temperature data. For future predictions, an online weather forecast can be used to fill this time-variant feature.

To find the best performing hyperparameters for the RF regressor, a grid search in combination with cross-validation is used (Géron 2019). The winning hyperparameter configurations for the target-functions of minimizing the mean absolute error (MAE) or root mean square error (RMSE) mentioned in Vandeput (2018) are listed in Table 4.

[2] Scalable Linking & Integration of big POi, available under http://www.slipo.eu/

Optimizing Technique	N Estimators	Min Samples Split	Min Samples Lead	Max Depth	Max Features
Cagliero et al. (2019)	50	5	5	10	sqrt
MAE optimized	50	3	3	20	sqrt
RMSE optimized	700	16	5	26	sqrt

Table 4. Grid Search Results for Dataset 1 in Comparison

4.2 Design of a Free-Floating Vehicle Relocation Approach

With the help of future demand predictions throughout the RF model, a relocation strategy is built on top. It aims to improve vehicle availability (R4) and can be integrated into the backend of the generic vehicle relocation information system (GRIS) framework presented in Figure 2 (Brendel et al. 2018).

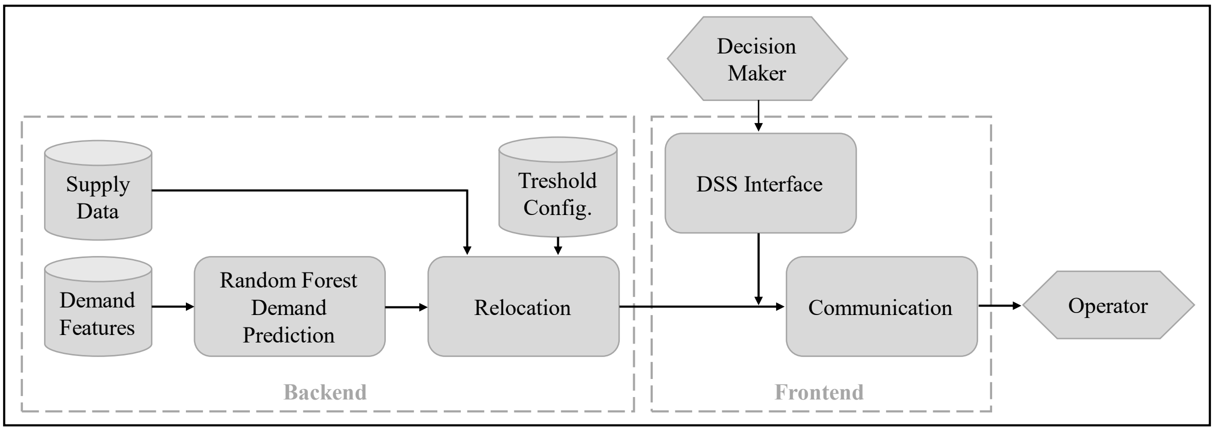

Figure 2: Generic Vehicle Relocation Information System (GRIS) adapted from Brendel et al. (2018)

The positions of all available vehicles are always given at present time and represent the current supply. The system suggests relocations based on the predicted demand situation at a future timeslot of three hours. Zones that contain fewer cars than bookings expected are marked as undersupply zones, while zones that contain more cars than bookings expected are marked as oversupply zones. Displaying the predicted level of supply with a heatmap over the operation area enables informed decisions for operation managers (R7). To implement the automated suggestion of relocations, providers can define a prediction threshold, which represents the minimum level of predicted demand to request a relocation (R6). Consequently, the strategy suggests relocations from oversupply zones to undersupply zones (R6). The prediction threshold causes the algorithm to only trigger relocations to regions with high demands (R4, R5) and allows the adaption of its sensitivity to the requirements of the system operation (R2). The zones are presented in ascending ordered by their potential impact that is defined by the difference between predicted bookings and system status quo (R7). Every performed relocation to a zone where the difference of predicted bookings and available cars is greater than one, will lead to a rental under the assumption that the prediction of the RF model is correct (R5). An example of the relocation algorithm is presented in Figure 3. To adapt the relocation strategy to a practice implementation, optimization and policy constraints like personnel availability and maintenance requirements have to be incorporated. Guidance for a holistic implementation of a decision support system for relocations is given by Clemente et al. (2018).

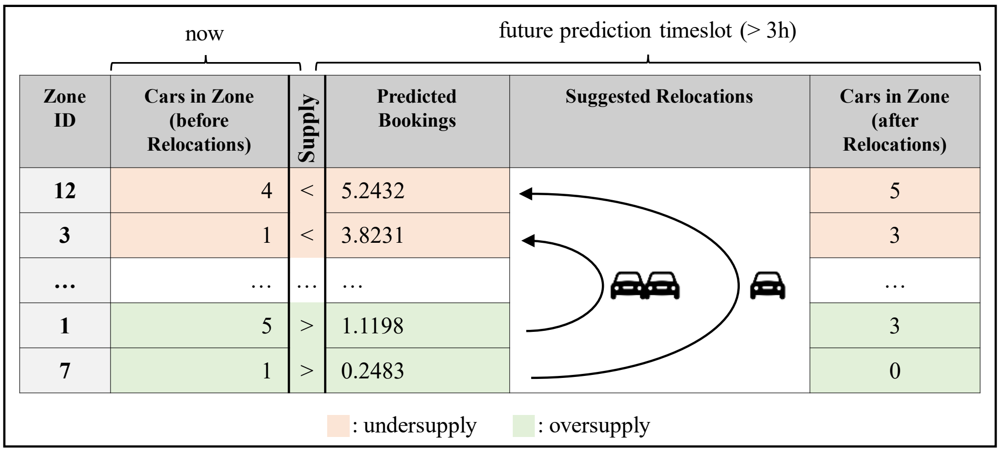

Figure 3. *Example of Relocation Algorithm.*

5 Evaluation

Calculating metrics like MAE and RMSE helps optimize the developed demand prediction model iteratively. However, it is questionable if the downstream relocation solution also shows better results in an economic sense. Following the suggestion of Prinz et al. (2021), we implement the algorithm in an event driven simulation, allowing us to optimize the hyperparameter of the RF regressor and the aforementioned prediction threshold against the economic performance indicator of acceptance ratio. The winning hyperparameter set is slightly different to the results from section 4.1. It is listed in Table 5 and taken as a parameter set for forthcoming simulations.

Simulation	**Prediction Threshold**	**N Estimators**	**Min Samples Split**	**Min Samples Lead**	**Max Depth**	**Max Features**
Small-City	1,2	50	16	5	26	sqrt
Large-City	1,4					
Metropolis	2,0					

Table 5. *Final Hyperparameters for RF Demand Prediction and Relocation Strategy.*

Furthermore, we compare the performance of our relocation algorithm RF with the baseline system behavior B (no relocations) and a deterministic relocation approach DT as suggested by Brendel et al. (2018). Therefore, all approaches are implemented as part of an event driven simulation using similar system setups. We run the simulation for all three datasets, ensuring that the data used for the simulation have not been part of the test- and training data used to build the RF model. To make the simulation robust against initial vehicle distribution, it was reiterated fifty times with random initial vehicle distributions. The average results of the simulated system behavior are presented in Table 6.

Simulation	**Small-City**			**Large-City**			**Metropolis**		
Fleet Size	30			198			400		
Timespan	01.01.2019 - 05.08.2019			01.10.2020 - 24.11.2020			01.09.2020 - 31.10.2020		
Strategy	***B***	***DT***	***RF***	***B***	***DT***	***RF***	***B***	***DT***	***RF***
Relocations	0	780	138	0	44	588	0	6 088	4 832
Acceptance	39.22 %	42.68 %	44.27 %	45.92 %	44.81 %	53.10 %	41.04 %	11.71 %	43.94 %

- **Strategies:** B = Base-Case, DT = Deterministic-Case, RF = Random Forest-Case
- **Fleet-Size:** The number of cars anticipated for the simulation.
- **Relocations**: Total number of relocations.
- **Acceptance:** Ratio of completed reservations and amount of booking requests (higher is better).

Table 6. *Results of the Simulation.*

The results of the simulation show that our strategy improved the acceptance ratio in all three reference systems between 7.1% and 15.6% (R2, R4). Furthermore, the number of relocations compared to the deterministic relocation strategy are reduced by 82.3% for the small city and by 20.6% for the metropolis (R2, R5). Moreover, it shows that the deterministic approach struggles to adapt to the large city and metropolis. The reason for that is, that it was originally designed for a small city system and that it needs further finetuning to better adapt.

The impact of the selected strategy on the business value of the fleet can be determined by offsetting the individual relocation costs against the increased revenue from additional future bookings. This is highly individual, depending on the price model and whether and how an operator implements user-based or operator-based relocations (Brendel et al. 2016; Di Febbraro et al. 2018). Especially for operator-based relocations other individual constraints like vehicle maintenance drives or certain staff availability have a high influence on the calculation of savings (Nourinejad et al. 2015). Effective relocation strategies also have indirect consequences for the future business value of a fleet, because vehicle availability is a key customer requirement on shared mobility systems (Herrmann et al. 2014). Customers who made bad experiences with vehicle availability will stop considering the carsharing service as a reliable means of transportation (Illgen and Höck 2019). In contrast, high vehicle availability also helps to attract new customer segments and consequently leads to the success of sustainable carsharing systems as part of future mobility.

6 Discussion

To answer our research question, we have addressed the challenge of vehicle relocations as part of the operation of free-floating CS systems. Within a structured literature review, we determined that datasets used in current solutions are mainly based on single system properties like historical bookings. Furthermore, traditional approaches are unreliable for relocation decisions, especially in free-floating CS systems, where demand is not cumulated on a few stations. To answer the research question, our DSR project with multiple evaluation steps demonstrated how big data could be used to improve the performance of free-floating CS systems from a customer and provider perspective. The following sections will present the study’s theoretical and practical implications and give its limitations and potential avenues for future research.

6.1 Theoretical Implications

We find developing data-driven relocation strategies for free-floating vehicle fleets is a rather novel and still ‘wicked’ problem (Gregor and Hevner 2013). Current publications in CS relocation research mainly focus on station-based CS systems, optimizing a single reference system, and are built on top of historical booking data. Thus, application domain maturity and solution maturity in vehicle relocation can both be regarded as low. The developed demand model and its application in a data-driven relocation strategy and instantiation as part of a simulation can be classified as improvements according to Gregor and Hevner (2013).

We synthesized the findings of Cagliero et al. (2019), Cocca et al. (2020), and Daraio et al. (2020) and demonstrated the development and evaluation of a data-driven relocation strategy for free-floating systems. Thus, we incorporated a machine-learning based demand prediction model with the GRIS framework presented by Brendel et al. (2018). In contrast to previously presented research, our presented relocation strategy is designed for free-floating systems and has proven effectiveness in three different sized case study settings based on real-world data. Implementing such a relocation strategy leads to a vehicle distribution where a customer’s transaction cost of picking up a car by foot does not exceed the gained utility for using the service. Relevance and rigor of our proposed design ensure a grounded contribution to the knowledge base that fits inside the green IS, operational research, and mobility domain.

6.2 Practical Implications

Important factors in IS DSR are the creation of practice-oriented solutions that are applicable for similar problems and other use cases in real-world scenarios (Hevner et al. 2004). Primarily, our study contributes to practice by offering a new data driven approach to suggest vehicle relocations for operators of free-floating CS fleets, helping to balance vehicle supply and demand. It addresses the whole range of different backgrounds and fields of applications like the technical side for implementation, the management role for generic resource planning of relocations, and the operative view for real vehicle movement operations. The developed RF strategy proved its adaptability to different CS systems, its positive impact on the acceptance-ratio, and its potential to decrease the number of relocations in comparison to other approaches. The practical application is guaranteed through humanly understandable outputs, which allows providers to build operator-based and user-based relocations on top of it.

In combustion-engine-based fleets, reducing the number of relocations compared to deterministic approaches leads to a reduced carbon-dioxide footprint. Increasing the vehicle availability in areas with high customer demands help CS providers to convince customers, and be established as a sustainable means of transportation. Consequently, the number of privately owned vehicles can be reduced, and the number of trips performed by vehicles (Amatuni et al. 2020; Jochem et al. 2020).

6.3 Limitations and Future Research

Despite the rigorous execution of the DSR project and its various evaluations, we are aware of some potential limitations. Also, our findings have implications for future research. Therefore, we will outline these limitations and opportunities for future research in the following paragraphs.

First, the applied evaluation was performed by a simulation based on three reference systems in Germany. The systems were seen isolated from other means of transport or competing for CS systems. Therefore, future research should engage practice partners to implement a field test to confirm the effectiveness of our relocation strategy. Future research could also consolidate our design knowledge into generalized design principles.

Second, the demand prediction model was trained against historical bookings, which does not reflect complete customer demand since booking requests that have been unsatisfied in the past are not part of the training set. Especially for regions with high demand, we expect a gap between historical booking data and real customer demand, because cars might have been unavailable nearby. Furthermore, our model has not considered factors like mass events or opening hours of shops and POI. Future research should work on strategies to increase the meaningfulness of classes and features of the prediction model. Entry points to identify unsatisfied demand could be the use of telemetry data from CS provider apps. Spatiotemporal features could be enhanced by using visitor frequencies or opening hours from data sources like Google Maps.

Third, our suggested relocation strategy ends by answering the question of which vehicle should be relocated to which position at a certain timeslot. The question of how the operation is performed is left unanswered, given that CS providers have limited relocation resources and must comply with other operational constraints like maintenance and recharging. Furthermore, including weather data in the model limits its long-term prediction to the availability and accuracy of weather forecasts.

7 Conclusion

Relocating vehicles in free-floating CS fleets is an operational challenge that requires data-driven managerial guidance. Overall, many approaches presented in related work focus on station-based CS systems or built data-driven optimization approaches that were not evaluated in a business domain. In this context, we tackled the problem of vehicle availability by developing a data-driven relocation strategy for free-floating CS systems by using real-world data. Based on requirements from practice and foundations from literature, we instantiated and evaluated a demand prediction model and a respective relocation strategy. As for the theoretical perspective, we addressed a relevant real-world problem with a set of design decisions. Regarding the practical perspective, our research shows the potential of increasing the acceptance ratio of shared vehicle fleets up to 15.6 %. Thus, CS gets more attractive to customers, accelerating its establishment as sustainable means of transportation. Future research could consolidate our design knowledge into generalized design principles. Furthermore, the performance of our relocation algorithm could be investigated in a field test.

References

Ait-Ouahmed, A., Josselin, D., and Zhou, F. 2018. "Relocation Optimization of Electric Cars in One-Way Car-Sharing Systems: Modeling, Exact Solving and Heuristics Algorithms," *International Journal of Geographical Information Science* (32:2), Taylor & Francis, pp. 367–398.

Alfian, G., Rhee, J., Ijaz, M., Syafrudin, M., and Fitriyani, N. 2017. "Performance Analysis of a Forecasting Relocation Model for One-Way Carsharing," *Applied Sciences* (7:6), p. 598.

Amatuni, L., Ottelin, J., Steubing, B., and Mogollón, J. M. 2020. "Does Car Sharing Reduce Greenhouse Gas Emissions? Assessing the Modal Shift and Lifetime Shift Rebound Effects from a Life Cycle Perspective," *Journal of Cleaner Production* (266), p. 121869.

Benarbia, T., Axhausen, K. W., and Farooq, B. 2020. "Modeling, Relocation, and Real-Time Inventory Control of One-Way Electric Cars Sharing Systems in a Stochastic Petri Nets Framework," *IEEE Transactions on Intelligent Transportation Systems* (Early Access), pp. 1–16.

Brendel, A. B., Brennecke, J. T., and Nastjuk, I. 2018. "Applying Econophysics in the Context of Carsharing - Development of a Vehicle Relocation Algorithm and Decision Support System," in *Thirty Nine International Conference on Information Systems*, San Francisco, pp. 1–17.

Brendel, A., Brauer, B., and Hildebrandt, B. 2016. "Toward User-Based Relocation Information Systems in Station-Based One-Way Car Sharing," in *Americas' Conference on Information Systems* (Vol. 13), San Diego, California: Association for Information Systems, August 10, pp. 1–10.

Cagliero, L., Chiusano, S., Daraio, E., and Garza, P. 2019. "CarPredictor: Forecasting the Number of Free Floating Car Sharing Vehicles within Restricted Urban Areas," in *2019 IEEE International Congress on Big Data (BigDataCongress)*, Milan, Italy: IEEE, July, pp. 72–76.

Carlier, A., Munier-Kordon, A., and Klaudel, W. 2015. "Mathematical Model for the Study of Relocation Strategies in One-Way Carsharing Systems," *Transportation Research Procedia* (10), pp. 374–383.

Cepolina, E. M., Farina, A., and Pratelli, A. 2014. "Car-Sharing Relocation Strategies: A State of the Art," *WIT Transactions on State of the Art in Science and Engineering* (78), pp. 109–120.

Clemente, M., Fanti, M. P., Iacobellis, G., Nolich, M., and Ukovich, W. 2018. "A Decision Support System for User-Based Vehicle Relocation in Car Sharing Systems," *IEEE Transactions on Systems, Man, and Cybernetics: Systems* (48:8), pp. 1283–1296.

Cocca, M., Teixeira, D., Vassio, L., Mellia, M., Almeida, J. M., and Couto da Silva, A. P. 2020. "On Car-Sharing Usage Prediction with Open Socio-Demographic Data," *Electronics* (9:1), p. 72.

Costain, C., Ardron, C., and Habib, K. N. 2012. "Synopsis of Users' Behaviour of a Carsharing Program: A Case Study in Toronto," *Transportation Research Part A: Policy and Practice* (46:3), pp. 421–434.

Daraio, E., Cagliero, L., Chiusano, S., Garza, P., and Giordano, D. 2020. "Predicting Car Availability in Free Floating Car Sharing Systems: Leveraging Machine Learning in Challenging Contexts," *Electronics* (9:8), p. 1322.

Di Febbraro, A., Sacco, N., and Saeednia, M. 2018. "One-Way Car-Sharing Profit Maximization by Means of User-Based Vehicle Relocation," *IEEE Transactions on Intelligent Transportation Systems*, pp. 1–14.

Frey, A., Trenz, M., and Veit, D. 2019. "Three Differentiation Strategies for Competing in the Sharing Economy," *MIS Quarterly Executive* (18:2), pp. 143–156.

Gambella, C., Malaguti, E., Masini, F., and Vigo, D. 2018. "Optimizing Relocation Operations in Electric Car-Sharing," *Omega* (81), pp. 234–245.

Géron, A. 2019. *Hands-on Machine Learning with Scikit-Learn, Keras, and TensorFlow: Concepts, Tools, and Techniques to Build Intelligent Systems*, (Second edition.), Sebastopol, CA: O'Reilly Media, Inc.

Gregor, S., and Hevner, A. R. 2013. "Positioning and Presenting Design Science - Types of Knowledge in Design Science Research," *MISQ* (37:2), pp. 337–355.

Hahn, R., Ostertag, F., Lehr, A., Büttgen, M., and Benoit, S. 2020. "'I like It, but I Don't Use It': Impact of Carsharing Business Models on Usage Intentions in the Sharing Economy," *Business Strategy and the Environment* (29:3), pp. 1404–1418.

Herrmann, S., Schulte, F., and Voß, S. 2014. "Increasing Acceptance of Free-Floating Car Sharing Systems Using Smart Relocation Strategies: A Survey Based Study of Car2go Hamburg," *Computational Logistics* (8760), Lecture Notes in Computer Science, (R. G. González-Ramírez, F. Schulte, S. Voß, and J. A. Ceroni Díaz, eds.), pp. 151–162.

Hevner, A., March, S., Park, J., and Ram, S. 2004. "Design Science in Information Systems Research," *MIS Quarterly* (28:1), pp. 75–105.

Hevner, A. R. 2007. "A Three Cycle View of Design Science Research," *Scandinavian Jorunal of Information Systems* (19:2), pp. 1–6.

Illgen, S., and Höck, M. 2019. "Literature Review of the Vehicle Relocation Problem in One-Way Car Sharing Networks," *Transportation Research Part B: Methodological* (120), Elsevier Ltd, pp. 193–204.

Jian, S., Rey, D., and Dixit, V. 2019. "An Integrated Supply-Demand Approach to Solving Optimal Relocations in Station-Based Carsharing Systems," *Networks and Spatial Economics* (19:2), pp. 611–632.

Jochem, P., Frankenhauser, D., Ewald, L., Ensslen, A., and Fromm, H. 2020. "Does Free-Floating Carsharing Reduce Private Vehicle Ownership? The Case of SHARE NOW in European Cities," *Transportation Research Part A: Policy and Practice* (141), pp. 373–395.

Jorge, D., and Correia, G. 2013. "Carsharing Systems Demand Estimation and Defined Operations: A Literature Review," *European Journal of Transport and Infrastructure Research* (13:3), pp. 201–220.

Kypriadis, D., Pantziou, G., Konstantopoulos, C., and Gavalas, D. 2020. "Optimizing Relocation Cost in Free-Floating Car-Sharing Systems," *IEEE Transactions on Intelligent Transportation Systems* (21:9), pp. 4017–4030.

Lei, Z., Qian, X., and Ukkusuri, S. V. 2020. "Efficient Proactive Vehicle Relocation for On-Demand Mobility Service with Recurrent Neural Networks," *Transportation Research Part C: Emerging Technologies* (117), pp. 1–17.

Li, X., Gao, J., Wang, C., and Huang, X. 2021. "A Data-Driven Approach for Vehicle Relocation in Car-Sharing Services with Balanced Supply-Demand Ratios," *International Journal of Intelligent Transportation Systems Research*.

Lopes, M. M., Martinez, L. M., and Correia, G. H. de A. 2014. "Simulating Carsharing Operations through Agent-Based Modelling: An Application to the City of Lisbon, Portugal," *Transportation Research Procedia* (3), pp. 828–837.

Marouf, M., Pollard, E., and Nashashibi, F. 2014. "Automatic Parallel Parking and Platooning to Redistribute Electric Vehicles in a Car-Sharing Application," in *2014 IEEE Intelligent Vehicles Symposium Proceedings*, pp. 486–491.

Nair, R., and Miller-Hooks, E. 2011. "Fleet Management for Vehicle Sharing Operations," *Transportation Science* (45:4), pp. 524–540.

Nansubuga, B., and Kowalkowski, C. 2021. "Carsharing: A Systematic Literature Review and Research Agenda," *Journal of Service Management* (32:6), Emerald Publishing Limited, pp. 55–91.

Niederman, F., and March, S. T. 2012. "Design Science and the Accumulation of Knowledge in the Information Systems Discipline," *ACM Transactions on Management Information Systems* (3:1), pp. 1–15.

Niels, T., and Bogenberger, K. 2017. "Booking Behavior of Free-Floating Carsharing Users: Empirical Analysis of Mobile Phone App and Booking Data Focusing on Battery Electric Vehicles," *Transportation Research Record: Journal of the Transportation Research Board* (2650:1), pp. 123–132.

Nourinejad, M., Zhu, S., Bahrami, S., and Roorda, M. J. 2015. "Vehicle Relocation and Staff Rebalancing in One-Way Carsharing Systems," *Transportation Research Part E: Logistics and Transportation Review* (81), pp. 98–113.

Prinz, C., Lichtenberg, S., Willnat, M., Brendel, A. B., and Kolbe, L. M. 2021. "CASSI: Design of a Simulation Environment for Vehicle Relocation in Carsharing," *AIS THCI Journal* (13:1), pp. 1–26.

Rai, A. 2017. "Diversity of Design Science Research," *MIS Quarterly* (41:1), pp. iii–xviii.

Repoux, M., Boyacı, B., and Geroliminis, N. 2015. "Simulation and Optimization of One-Way Car-Sharing Systems with Variant Relocation Policies," in *Transportation Research Board 94th Annual Meeting*, Washington DC, United States, pp. 1–18.

Rindfleisch, A. 2020. "Transaction Cost Theory: Past, Present and Future," *AMS Review* (10:1), pp. 85–97.

Santos, A. G., Cândido, P. G. L., Balardino, A. F., and Herbawi, W. 2017. "Vehicle Relocation Problem in Free Floating Carsharing Using Multiple Shuttles," in *2017 IEEE Congress on Evolutionary Computation (CEC)*, pp. 2544–2551.

Schmöller, S., and Bogenberger, K. 2020. "Carsharing: An Overview on What We Know," in *Demand for Emerging Transportation Systems*, Elsevier, pp. 211–226.

Schmöller, S., Weikl, S., Müller, J., and Bogenberger, K. 2015. "Empirical Analysis of Free-Floating Carsharing Usage: The Munich and Berlin Case," *Transportation Research Part C: Emerging Technologies* (56), pp. 34–51.

Schulte, F., and Voß, S. 2015. "Decision Support for Environmental-Friendly Vehicle Relocations in Freefloating Car Sharing Systems: The Case of Car2go," *Procedia CIRP* (30), pp. 275–280.

Shaheen, S., Cohen, A., Chan, N., and Bansal, A. 2020. "Sharing Strategies: Carsharing, Shared Micromobility (Bikesharing and Scooter Sharing), Transportation Network Companies, Microtransit, and Other Innovative Mobility Modes," *Transportation, Land Use, and Environmental Planning*, pp. 237–262.

Smith, S. L., Pavone, M., Schwager, M., Frazzoli, E., and Rus, D. 2013. "Rebalancing the Rebalancers: Optimally Routing Vehicles and Drivers in Mobility-on-Demand Systems," in *2013 American Control Conference*, pp. 2362–2367.

Vandeput, N. 2018. *Data Science for Supply Chain Forecast*, Nicolas Vandeput.

Wagner, S., Willing, C., Brandt, T., and Neumann, D. 2015. "Data Analytics for Location-Based Services: Enabling User-Based Relocation of Carsharing Vehicles," in *International Conference on Information Systems 2015*, Fort Worth, TX, December 14, pp. 1–16.

Wang, L., Liu, Q., and Ma, W. 2019. "Optimization of Dynamic Relocation Operations for One-Way Electric Carsharing Systems," *Transportation Research Part C: Emerging Technologies* (101), pp. 55–69.

Wappelhorst, S., Sauer, M., Hinkeldein, D., Bocherding, A., and Glaß, T. 2014. "Potential of Electric Carsharing in Urban and Rural Areas," *Transportation Research Procedia* (4), pp. 374–386.

Webster, J., and Watson, R. T. 2002. "Analyzing the Past to Prepare for the Future: Writing a Literature Review.," *MIS Quarterly* (26:2), pp. xiii–xxiii.

Weikl, S., and Bogenberger, K. 2013. "Relocation Strategies and Algorithms for Free-Floating Car Sharing Systems," *IEEE Intelligent Transportation Systems Magazine* (5:4), pp. 100–111.

Weikl, S., and Bogenberger, K. 2014. "Nachfragephaenomene von Free Floating Car Sharing-Systemen: Raeumlich-Zeitliche Angebots-Nachfrage-Asymmetrie," *Forschungsgesellschaft Für Straßen- Und Verkehrswesen* (002:106), pp. 591–606.

Weikl, S., and Bogenberger, K. 2015. "A Practice-Ready Relocation Model for Free-Floating Carsharing Systems with Electric Vehicles – Mesoscopic Approach and Field Trial Results," *Transportation Research Part C: Emerging Technologies* (57), pp. 206–223.

Willing, C., Klemmer, K., Brandt, T., and Neumann, D. 2017. “Moving in Time and Space – Location Intelligence for Carsharing Decision Support,” *Decision Support Systems* (99), pp. 75–85.

Wittwer, R., and Hubrich, S. 2018. “Free-Floating Carsharing Experiences in German Metropolitan Areas,” *Transportation Research Procedia* (33), XIII Conference on Transport Engineering, CIT2018, pp. 323–330.

Zakaria, R., Dib, M., Moalic, L., and Caminada, A. 2014. “Car Relocation for Carsharing Service: Comparison of CPLEX and Greedy Search,” in *2014 IEEE Symposium on Computational Intelligence in Vehicles and Transportation Systems (CIVTS)*, pp. 51–58.