Strategic Management of Limited Transportation Recourses to Support Mobility of Disadvantaged and Disabled Travelers during the COVID-19 Pandemic or Similar Situations

Final Report

by

Yu Qian
Associate Professor
Department of Civil and Environmental Engineering
The University of South Carolina
300 Main Street-C228
Columbia, SC 29208
Office: (803)777-8184
Email: yuqian@sc.edu

Ye Liu, The University of South Carolina
Gurcan Comert, Benedict College
Negash Begashaw, Benedict College

October 2022

Center for Connected Multimodal Mobility (C^2M^2)

200 Lowry Hall, Clemson University
Clemson, SC 29634

DISCLAIMER

ACKNOWLEDGMENT

The research team greatly thanks C^2M^2 for partially supporting this project.

Technical Report Documentation Page

1. Report No.	2. Government Accession No.	3. Recipient's Catalog No.
4. Title and Subtitle Strategic Management of Limited Transportation Recourses to Support Mobility of Disadvantaged and Disabled Travelers during the COVID-19 Pandemic or Similar Situations		5. Report Date October, 2022 6. Performing Organization Code
7. Author(s) Yu Qian, Ph.D., ORCID: 0000-0001-8543-2774 Ye Liu, ORCID: 0000-0002-2268-9545 Gurcan Comert, Ph.D., ORCID: 0000-0002-2373-5013 Negash Begashaw, Ph.D., ORCID: 0000-0002-4192-3069		8. Performing Organization Report No.
9. Performing Organization Name and Address University of South Carolina 300 Main St. Columbia, SC 29208		10. Work Unit No. 11. Contract or Grant No.
12. Sponsoring Agency Name and Address Center for Connected Multimodal Mobility (C^2M^2) Clemson University 200 Lowry Hall, Clemson, SC 29634		13. Type of Report and Period Covered 14. Sponsoring Agency Code
15. Supplementary Notes		

16. Abstract

The COVID-19 lockdown has reduced public transportation service to the disadvantaged and disabled people who urgently need adequate mobility to obtain essential supplies. This report aims to improve the life quality of people with disabilities and elderly people by addressing social exclusion, accessibility, and mobility issues. Demand-responsive transport services are frequently offered in the context of door-to-door transportation of the elderly and persons with disabilities. We study and compare two frameworks. We apply Sample Average Approximation (SAA) and Rolling Horizon (RH) to optimize a car-sharing system for the total cost, including initiation cost and operation cost, after the fleet size is determined. The model is generated with given geographic conditions and other local information to be tailored for specific applications for local communities. Given that no historical data is available, random sample data is generated to simulate expected demands. We consider three types of probability distributions for daily demand data, and the results generated using three different distributions are being examined and compared. The research shows that the demand data following a normal distribution, results in the minimum total cost. Additionally, we study the impact of several factors on total cost, including demand fulfillment rates and operation hours. Our results suggest that the effect of fulfillment rate on fleet size is exponential after a threshold under all three types of daily demand data, and extended operation hours can significantly reduce the total cost. Finally, the paper provides applicable frameworks for city planners, Non-profit Organizations (NPOs), and policymakers to better allocate limited resources to implement the car-sharing system when little to no historical travel information is available for low-density population areas. It is anticipated that the outcome of this research will benefit disadvantaged and disabled travelers during COVID-19 or similar difficult situations in the future.

17. Keywords	18. Distribution Statement
Car-sharing, Stochastic Programming, Mixed integer programming	No restrictions.

19. Security Classif. (of this report)	20. Security Classif. (of this page)	21. No. of Pages	22. Price
Unclassified	Unclassified		NA

LIST OF TABLES

LIST OF FIGURES

CHAPTER 1

Introduction

1.1 Research Motivation and Literature Review

In many areas, basic amenities are in the suburbs, while low-income populations reside in rural areas and central cities. Criden (2008) pointed out that access to transportation has become limited to low-income individuals and families. Limited access to transportation can pose challenges to those underserved individuals and households in many aspects. Álvaro Aguilera-García (2022) pointed out that the individuals with a higher usage of current car-sharing facilities are usually described as male, young, wealthy, well-educated, and those who reside in denser urban areas. Specific populations, such as elderly people and people with disabilities, are more vulnerable to transportation barriers (Dabelko-Schoeny, 2021).

The car-sharing and ridesharing services have generated a wide variety of impacts in different aspects, including the environment, human behavior, and the economy (Shaheen and Cohen, 2020). However, since the pandemic, ride-hailing services provided by Uber and Lyft have decreased with fewer available drivers. The pandemic has led to city lockdowns, resulting in restrictions on billions of people worldwide. Among them, a group of disadvantaged and disabled people are facing more significant challenges, especially those who cannot own a vehicle or cannot drive a vehicle. Oluyede (2022) suggests that transportation barriers contribute to poor health outcomes as lack of transportation prevents people from reaching healthcare providers in a timely manner, resulting in canceled appointments, and delayed or missed medical attention. Rozenfeld (2020) found that transportation insecurity and living in senior communities are associated with higher odds of initial infection of COVID-19.

Meanwhile, low-density areas have some favorable characteristics for establishing car-sharing services, including high car dependence due to few or absence of public transportation and large parking availability. However, it faces significant challenges in generating profit given low demand. After conducting a case study of a rural area in Italy, Rotaris and Danielis (2018) concluded that the “station-based” business model, which allows users to pick-up and drop-off a car only in stations, operated on the social grounds and provides services to low-income families, was the only viable business model in less-densely populated areas.

To the best of our knowledge, the kinds of literature regarding providing car-sharing services to low-income or under representative populations stops on a strategic level to propose the most efficient policies to ensure social equity (Pan, Martin, & Shaheen, 2022), (Golub, Satterfield, Serritella, Singh, & Phillips, 2019), (Wong, Broader, & Shaheen, 2020), (Vermesch, Boisjoly, & Lachapelle, 2021). Only a few have gone beyond the strategic level to provide the implementing works aimed at such social-oriented car sharing system: Yu & Shen (2019) proposed a shared ride system that can optimize the pick-up routes under a random travel time by solving a stochastic mixed-integer program using decomposition algorithm. However, it did not offer insights on how to determine the initial fleet size for such car sharing system, nor did it provide a detailed analysis on the impacts of different variables that are critical to the success of the system.

In response to the issues mentioned above, we propose a social-oriented car-sharing system that can help local government and Non-profit Organizations (NPOs) provide essential mobility services to populations most vulnerable to COVID-19 and other similar diseases or disasters. The paper aims to provide easy-to-follow, step-by-step instruction on the planning (fleet size, unit trip fare, operational hours) of a car-sharing system to address the equality and fairness concerns of disadvantaged groups regarding mobility needs. Our project focuses on designing a car-sharing system for underserved populations in the City of Columbia in South Carolina or areas having similar conditions.

The proposed car-sharing system is station-based, with service stations that are close to senior housing, hospitals, shopping malls, and local traffic management authorities such as train stations and airports. Due to the COVID-19 or similar pandemic situations, the fleet is supposed to be supported by the local government with the emergency relief fund, volunteers, or a mix of both. Considering that we focused on car-sharing systems in less densely populated areas, parking is assumed to be abundant. Lastly, since real-world data about car-sharing is unavailable in many underserved communities, we generate simulation datasets to reflect uncertain travel demands.

In this study, we formulate and solve a temporal-spatial network problem to maximize ride demand fulfillment with a given budget. The duration of the car-sharing system is assumed to be three months. We feed the simulation datasets to a multi-stage stochastic mixed integer linear programming (MILP) model to SAA model and a Rolling Horizon (RH) model to 1) determine the initial fleet size in each service station and 2) suggest the ideal daily operating hours. We further compare these two frameworks to determine the better practice in building the such car-sharing system.

The rest of the report is organized as follows: Chapter 2 overviews two optimization approaches and reviews the relevant studies in both frameworks. Chapter 3 presents the problem formulation, which introduces the spatial-temporal network, describes our approach to generate the random ride request when no real-world data is available, and our implementation of the two models. Chapter 4 conducts a case study and provides insights by applying both frameworks to solve the car sharing problem in the city of Columbia, SC, and demonstrates the differences in the computational efficiency for different framework models. Chapter 5 concludes the paper with our findings, limitations, and future research directions.

CHAPTER 2

Overview of Optimization Approaches

2.1 Stochastic Programming Problem

Two-stage stochastic mixed-integer programming has significantly been implemented to solve the optimization problems for the vehicle/bicycle sharing system. A fundamental two-stage stochastic programming problem divides the decision variables into first and second-stage variables. The problem can be stated as follows:

$$\min_{x \in X} \boldsymbol{c}^{\boldsymbol{T}} \boldsymbol{x} + \mathbb{E}[Q(\boldsymbol{x}, \widetilde{\boldsymbol{d}})]$$
$$s.t. \boldsymbol{A}\boldsymbol{x} = \boldsymbol{b} \qquad \text{2-1}$$
$$x_i \geq 0,\ i=1,\ 2,\ \ldots,\ I$$

where

$$Q(\boldsymbol{x}, \boldsymbol{d}) = \min_{y \in Y} \xi(\boldsymbol{d})^T \boldsymbol{y}$$
$$s.t.\ \ \boldsymbol{T}\,\boldsymbol{x} + \boldsymbol{W}\boldsymbol{y} = \boldsymbol{h} \qquad \text{2-2}$$
$$d_k \geq 0, k = 1,2, \ldots, K$$

In model 2-1, $\boldsymbol{c}^{\boldsymbol{T}}\boldsymbol{x}$ denotes the first-stage cost, $\mathbb{E}[Q(\boldsymbol{x}, \widetilde{\boldsymbol{d}})]$ denotes the expected cost of the second stage, and $\boldsymbol{d}$ denotes the discrete random variables in different scenarios for the second-stage problem. $\boldsymbol{x}$ and $\boldsymbol{y}$ are the solutions we are looking for that optimize first-stage and second-stage problems, respectively. $\boldsymbol{A}\boldsymbol{x} = \boldsymbol{b}$ in model 2-1 and $\mathbf{T}\boldsymbol{x} + \boldsymbol{W}\boldsymbol{y} = \boldsymbol{h}$ in model 2-2 are constraints to consider when solving the optimization models, where $\boldsymbol{A}, \boldsymbol{b}, \boldsymbol{T}, \boldsymbol{W}$ are all deterministic matrices. The optimization model in 2-1 finds the optimal value of $\boldsymbol{x}$, which minimizes the first-stage costs and the expected cost of the resource function before the demand becomes known. Once the uncertainties are realized, the second stage decisions, also called the *recourse decisions*, are carried out. Given the optimal determined values of $\boldsymbol{x}$, the model in 2-2 returns the optimal cost in the second stage with the realization of the known demands.

For example, in this report, the first stage decision variable $\boldsymbol{x}$ can be the initial number of vehicles and/or parking slots allocated to each fixed station, **c** can be the Cost of purchasing a vehicle or/and a parking slot. In the second stage, the variables $\boldsymbol{y}_{\boldsymbol{ij}}$ in **y** represents the vehicle movement from region i to region j, and $\boldsymbol{d}$ ($\widetilde{\boldsymbol{d}}$) represents the realized (expected) car rental demands.

In a case where the demand $\boldsymbol{d}$ is known, the objective functions in a 2-stage stochastic problem turn into a linear process. However, car/bike-sharing companies often need to purchase vehicles and parking slots or permits way before the demands are known. The *expected* demand $\widetilde{\boldsymbol{d}}$ is then a random variable, and the probability distribution of $\widetilde{\boldsymbol{d}}$ is unknown. In the case of the car-sharing system, while vehicle and parking slot/permit purchases happen only once, the relocation actions' distribution may happen repetitively on a monthly, weekly, or even daily basis. Therefore, the average optimal Cost of the second stage resource optimization model 2-2 converges to the expected resource function $\mathbb{E}[Q(\boldsymbol{x}, \widetilde{\boldsymbol{d}})]$.

In a *scenario-based approach*, the random variables $\widetilde{\boldsymbol{d}}$ has a finite number of distributions, $\widetilde{\boldsymbol{d_1}}, \widetilde{\boldsymbol{d_2}}, \dots, \widetilde{\boldsymbol{d_k}}$ with probabilities $\boldsymbol{p_1}, \boldsymbol{p_2}, \dots, \boldsymbol{p_k}$. Model 1-1 and Model 1-2 can then be written as follows:

$$\min_{x \in \boldsymbol{X}, \boldsymbol{y_1}, \boldsymbol{y_2}, \dots, \boldsymbol{y_k}} \boldsymbol{c^T x_k} + \sum_{k=1}^{K} p_k \, \xi\left(\widetilde{\boldsymbol{d_k}}\right)^{\boldsymbol{T}} \boldsymbol{y_k} \qquad \text{2-3}$$

Further, in a car-sharing system, the demands happen over a finite horizon of *T* periods. The problem turns into a *multi-stage stochastic* problem with each $\widetilde{\boldsymbol{d_t}} = [d_{t1}, d_{t2}, \dots, d_{tk}]$, $t = 1, .., T, k = 1, 2, \dots, K$ a random vector of car rental demands. At stage $t = 1$, the problem is a *two-stage stochastic* problem and can be written as:

$$\min_{x_0 \geq 0} \boldsymbol{c^T x_0} + \mathbb{E}[\boldsymbol{Q_1}(\boldsymbol{x_0}, \widetilde{\boldsymbol{d}}_{\boldsymbol{t=1}})] \qquad \text{2-4}$$

At stage $2 \leq t \leq T$, the *expected* resource function turns into a *dynamic programming* function:

$$\phi(x_{t-1}) = \mathbb{E}\{\, \xi_t\left(\widetilde{\boldsymbol{d_t}}\right)^T \boldsymbol{y_t}\left(\widetilde{\boldsymbol{d_t}}\right) \mid \widetilde{\boldsymbol{d}}_{\boldsymbol{t-1}} = \boldsymbol{d_{t-1}}\} \qquad \text{2-5}$$

To simplify the dynamic programming function, we can assume that ξ_t is stagewise independent of other ξ. When the number of purchased vehicles are the same across different scenarios, we can add the constraints $\boldsymbol{x_0} = \boldsymbol{x_1} = \cdots = \boldsymbol{x_s}$

There are two ways to obtain the random demand sample. First, the random demand sample can be viewed as historical data of k observations of $\widetilde{\boldsymbol{d}}$. Second, it can be generated by *Monte Carlo simulations* (Lampa & Samolejová, Fleet Optimization Based on the Monte Carlo Algorithm, 2020), (Takes & Kosters, 2010). To estimate the *expected* resource function, the Sample Average Approximation (SAA) (Kim, Pasupathy, & Henderson, 2014) is implemented.
The objective function in model 2-2 can be written as:

$$\widehat{\boldsymbol{Q}}(\boldsymbol{x}, \widetilde{\boldsymbol{d}}) \coloneqq \frac{1}{S} \min_{x \in \boldsymbol{X}} \sum_{s=1}^{S} \xi\left(\widetilde{\boldsymbol{d_s}}\right)^{\boldsymbol{T}} \boldsymbol{y_s} \qquad \text{2-6}$$

The challenge to solving the SAA problem is the computational complexity. The number of scenarios would grow exponentially as the number of random parameters $\widetilde{\boldsymbol{d}}_s, s \in \boldsymbol{S} = \{1, .., S\}$ grows. For multi-stage problems, the computational complexity grows exponentially with the increase in the number of stages.

Lu et al. (2017) proposed a two-stage stochastic integer programming problem to approximate the results of a multi-stage dynamic model to determine the fleet allocations for both reservation-based and free-floating car-share systems. The uncertain travel demand is generated as independent and identically distributed (*i.i.d.)* samples using Monte Carlo sampling, and the SAA technique is employed to construct the problem. The model is further modified to optimize vehicle relocation in a reservation-based system in a rolling-horizon approach which shows a significant advantage over a benchmark policy in terms of profitability and quality of service (QoS). Like Lu et al. (2017), Shu, et al., (2013) proposed a stochastic network model with proportionality constraints to address the problems such as the deployment and redistribution in the bike-sharing network. Shu, et al., (2013) first test the model without considering the rebalancing of the bike-sharing system if the number of demands $\boldsymbol{r_{ij}(t)}$ from region, i to region j follows a Poisson process. It defines a linear programming (LP) to determine the optimal initial number of bikes assigned to each station and the bicycle utilization rate $\boldsymbol{\alpha(t)}$, with the objective function to maximize the fulfilled demands at a fixed $\boldsymbol{\beta}\ (\boldsymbol{t}) = \sum_{\boldsymbol{t}} \boldsymbol{\alpha(t)}$ **.** The model shows the importance of choosing an appropriate utilization rate, below which the system will need a drastic increase in the number of bicycle deployments to increase the fulfilled requests. It further brings in new decision variables to represent rebalancing arcs in a rolling-horizon format.

Carrese et al. (2020) employed binary linear programming (BLP) and genetic-algorithms to help the local government decide the optimal parking slots provided to the rental companies to maximize the total profit. He, Hu, and Zhang (2019) proposed a stochastic dynamic program for the fleet repositioning problem to minimize the repositioning cost and the penalty cost given the observed distribution of the fleet size. They used real-world multi-region data and developed a “myopic” 2-stage model which only considered the cost that emerged in the current period. They used a multi-stage enhanced linear decision rule (ELDR) which utilizes the auxiliary information in a 2-region system. This is illustrated in Figure 2.1. The demand is assumed to be joint probability distributed and independent over periods.

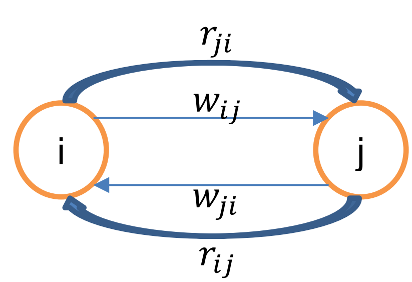


Figure 2.1 A 2-region Subgraph of A Car-Share System Network (He et al. 2019)

In Figure 2.1, w_{ij} represents the number of fulfilled one-way trips from region *i* to region *j*, w_{ii} and w_{jj} represent the number of fulfilled two-way trips, and r represents the number of repositions, which is a decision the companies need to make. Recently, many studies applied a two-stage stochastic program approach on various topics including, relocation problems for electric vehicle (EV) (Lin & Kuo, 2021), EV charging station location problems, (Li et al. 2022), EV power distribution through microgrids (Tan et al. 2022; Hou et al. 2022).

As we mentioned earlier, most literature focus on for-profit car-sharing or bike-sharing systems in the city or urban areas where the population is dense. Our study applies a multi-stage stochastic model to set up a car-sharing system for underserved populations with uncertain demands, limited budgets, unlimited parking space, and unavailable real-world data. The paper later applies Rolling Horizon Heuristic, also mentioned in section 2.2, and compares two models for their results and computational efficiency.

Moving bikes and vehicles around in the city can be expensive. The operating system needs to make efficient rebalancing decisions to maintain service quality while minimizing the repositioning cost. Henderson et al. (2016) explored the rebalancing problems of bike sharing system for both during and after rush-hour planning in the city of New York. The challenge for rebalancing the bikes during rush hour is that the dynamic approach can change drastically, so the system status may no longer hold when the truck arrives. Based on historical trip data, Henderson et al. (2016) solved the problem by pre-balancing the bike during the overnight shift before rush hour begins. For an overnight rebalancing problem, they tested both the greedy algorithm and the mixed-integer programming (MIP). They found that the greedy algorithm tended to generate better solutions as the number of trucks increased.

However, the problems stated in the above literature mainly focus on static relocation approaches after the initial fleet size is determined, whereas our literature considers integrating planning and operations for car sharing to determine initial fleet size to purchase.

2.2 Rolling Horizon Heuristic

Rolling horizon (RH) is a decision-making model adopted mainly in making immediate decisions in a dynamic stochastic situation. It was first introduced by Baker (1977) that proves the effectiveness of rolling schedules for production planning. The Rolling Horizon approach aims to speed up the solution process of a problem with a significant time structure by breaking the overall time structure into reasonable time frames. The rolling horizon approach has been used to solve many production planning and scheduling problems in different applications.

Cordeau et al. (2015) introduced the rolling horizon framework to plan the delivery of vehicles to dealers by auto-carriers to optimize the travel distances, costs of operations, and penalties for deliveries with known demands and an initial number of carriers. Local search procedures with a CPU time are invoked in this study to optimize routing plans to a dynamic and multiple-day setting. Ma & Koutsopoulos (2022) introduced the concept of the Near-on-Demand service to obtain future requests in advance. The study applied the Rolling Horizon approach to make rebalancing decisions for ride-pooling to minimize the impact on congestion and vehicle miles traveled. Other studies (Wang and Szeto, 2021; Yang et al 2021; Wu et al 2022, Zaneti et al. 2022) apply the rolling horizon approach to solve various problems, including repositioning of bike/vehicle and charging schedules of EVs.

Most of the literature proposed RH as a solution at a detailed decision-making level to relocate vehicle when the project is being executed after certain resources (e.g., initial fleet size) is determined. Hartleb & Schmidt (2020) applied the RH to determine the minimum fleet size needed to fulfill given demanded trips. However, the optimal number of vehicles to be purchased before the car-sharing system is to be executed is different from the minimum fleet size when other factors, such as fulfillment rate (that is smaller than 100%), idle cost, penalty cost, repositioning cost etc., are considered. We explore the impact of different settings on the system here to help potential organizations, such as the local governments and NPOs, to quickly set up an efficient temporary car-sharing system with little or no historical demand data for reference. Our study mainly focuses on solving a multi-stage stochastic programming problem using SAA and RH frameworks to make one-time purchase decisions before the operation starts with the redistribution impact considered simultaneously. When RH framework is used to solve a multi-stage problem, our paper demonstrates the effect of a delicate issue of what horizon length to include in the subproblems.

Our contribution and main results are summarized as follows:

1. We develop demands based on an average of three probability distributions to simulate uncertain demands when real-world data is unavailable.

2. Our study compares a two-stage stochastic model and a rolling horizon model to determine initial car fleet size to satisfy uncertain car-share demands. We also quantify runtime for both models and recommend the rolling horizon model for its flexibility if runtime is a concern.

3. Via numerical experiments using different self-generated data, we conduct sensitivity analysis and show the impact of other variables on the total costs, including initial vehicle purchasing and operating costs. The optimal fleet size has a linear relationship with the total cost. The fulfillment rate is shown to have a two-phase impact on total cost - the relationship between the actual fulfillment rate and total cost is linear before a certain threshold and turns to be exponential after. Our study also suggests that extended daily service hours can significantly bring down the required initial fleet size and the total cost.

CHAPTER 3

Problem Formulation

3.1 A Spatial-Temporal Network

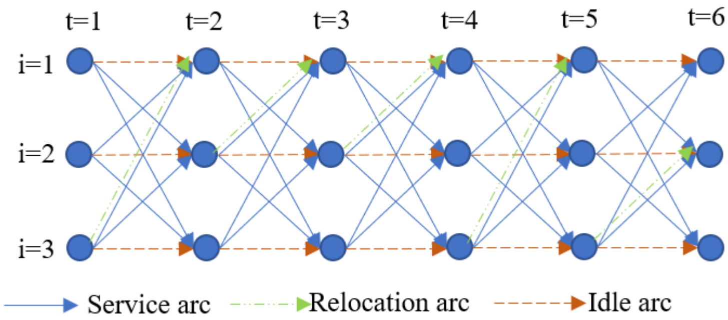
In the case study, we construct a spatial-temporal network G (N, E) to model the movement of vehicles among the service stations as flows. The network G, also illustrated in Figure 3.1, contains two important parts. Node $n_{i,t} \in N$ denotes a region $i \in I, I = \{1,2,\dots,I\}$ at time $t \in T, T = \{1,2,\dots,T\}$, whereas arc $e_{i,j} \in E$ represents the flow from region $i \in I$ to region $j \in I, I = \{1,2,\dots,I\}$. There are three different types of arcs in the network G:

1) *Service arc (ε_R)*: flows on the service arcs represent fulfilled travel requests. The flow on a service arc $(n_{i,t}, n_{i',t'}) \in \varepsilon_R$ represents fleet that is dispatched from region i in period t and returned to region i' in period t'. For a single arc $e \in \varepsilon_R$, let $\underline{w_e}$ be the arc's capacity, i.e., the maximum number of dispatched vehicles. Here we define the arc's capacity to be the travel requests. The capacity of each arc equals the smaller of the travel requests and fleet size.
   a) If the fleet size at $n_{i,t}$ is larger than the number of travel requests, the arc's capacity $\underline{w_e}$ on the service arc equals to the travel request.
   b) If the fleet size is smaller than the number of service requests, the arc's capacity $\underline{w_e}$ on the service arc equals to the travel fleet size.

On each service arc, the "reward" (in other words, the "positive social impact") r is generated for each fulfilled service request. Besides the "positive social impact" generated from fulfilled service requests, a "penalty" (in other words, the "negative social impact") δ occurs for each unfulfilled service request if the service region does not have enough fleet inventory when a customer makes a request.

2) *Reposition Arc (ε_L)*: To better allocate the fleet size at each region at a specific time period to maximize the fulfillment rate, an optimization model repositions the vehicles across regions at the lowest operation cost. The operation costs, including labor and fuel cost, occur when the repositioning happens. We denote the operational cost per flow unit on the reposition arcs to be c_L .

3) *Idle Arc (ε_I):* If there are more available service vehicles than the consumer service requests in a service region, the decision-maker wastes the limited resources, in other words, pays the holding cost, such as the cost of looking for appropriate parking spaces and the labor costs for the surplus fleets. We use the flow units on an idle arc $e = (n_{i,t}, n_{i',t'}) \in \varepsilon_I$ to represent idle surplus fleets in region i from period t to $t+1$, and c_I to represent the Cost per flow unit on idle arcs.



Arc Type	Unit Flow Cost
Service Arc $(n_{i,t}, n_{i',t'})\hat{I}e_R$	Positive social impact: $-r$ Negative social impact: θ
Reposition Arc $(n_{i,t}, n_{i',t'})\hat{I}e_L$	Reposition cost: c_L
Idle Arc $(n_{i,t}, n_{i',t'})\hat{I}e_I$	Idle cost: c_I

Figure 3.1 Spatial-Temporal Network

3.2 Monte Carlo Sampling Process

Algorithm 1

Input: Number of scenarios *S*, number of average trips N_{trip}, Standard deviation std, Total time period T, Travel time matrix and T_{mat}
Output: Demand hash table $d = \{ t_o, t_d, i_o, i_d : n\}$

#Step 1:
#Initialize daily travel size as an empty hash table $demand\ dict$, and daily travel request data as another hash table d $demand\ dict$= {}
$d = \{\}$

#Step 2: Add each valid trip as a hash table key and assign the value as 0
While $1 \leq t_o \leq$T:
    While $1 \leq i_o \leq$I:
        While $1 \leq i_d \leq I$:
            $t_d \leftarrow calculated\ based\ on\ t_o, i_o\ and\ i_d$ and T_{mat}
            $d[\ t_o, t_d, i_o,\ i_d] = 0$

#Step 3: Generate normal distribution samples among the scenarios
$demand\ = \{\} \leftarrow Normal\ distribution\ (S, N_{trip}, std)$

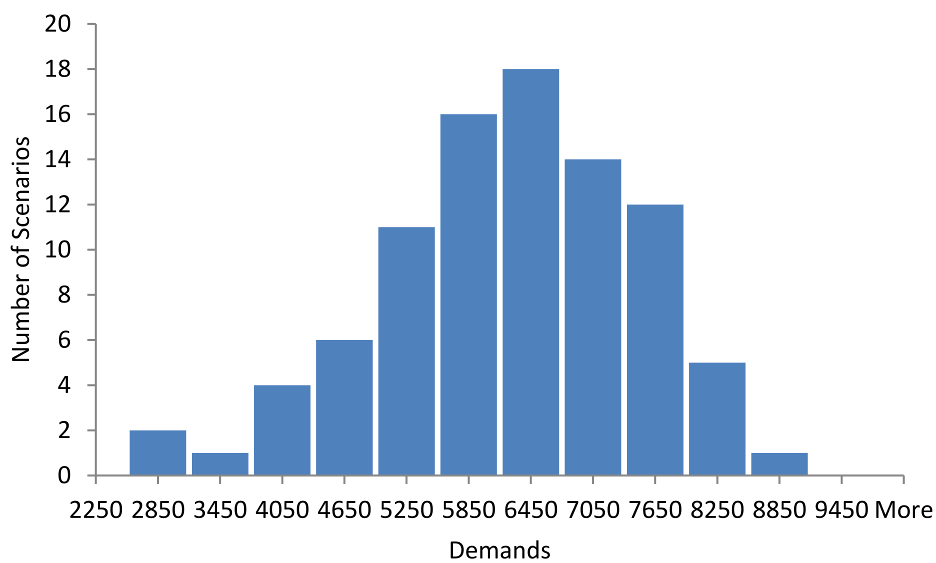
#Step 4: Update demand hash table d
for $trip\ size$ in $demand\ dict$:
    while $trip\ size \geq 1$:
        $t_o \leftarrow randomly\ pick\ a\ start\ time\ from\ \boldsymbol{T} = [1,2, \dots, T]$
        $i_o\ \leftarrow randomly\ pick\ a\ start\ region$ (origin)
        $i_d \leftarrow randomly\ pick\ an\ end\ region$ (destination)
        $t_d \leftarrow calculated\ based\ on\ t_o, i_o, i_d$ and T_{mat}
        $if\ \ t_o, t_d, i_o,\ i_d\ in\ d$:   #To make sure the demand is valid
        $d[\ t_o, t_d, i_o,\ i_d]$+=1

return d

Algorithm 1 illustrates the detailed steps to generate the daily travel request data for every scenario $s \in \boldsymbol{S}$. The data for daily trip size is stored in a dictionary, or a hash table, $demand\ dict = \{s_1 = \boldsymbol{d_1}, s_2 = \boldsymbol{d_2}, \dots, s_{90} = \boldsymbol{d_{90}}\}$, with its *key* representing each scenario, and its *value* indicates daily request data, which is stored in hash table d_s.

For each scenario s, there is a corresponding travel request dataset that $\boldsymbol{d_s}$. Its *key* contains four pieces of information: start time (pick-up time) t_o, end time (drop-off time) t_d, start region (pick-up location) i_o, and end region (destination) i_d, whereas its *value* shows the size of the specific request trip. For example, d$[1,2,3,4]$=10 suggests that there are ten requests for the trip with pickup time $t_o = 1$, drop-off time $t_d = 2$, pick-up location $i_o = 3$, and the destination $i_d = 4$.

The number of scenarios can also be viewed asthe car-sharing system's period of operation. We consider 90 scenarios (i.e., S = 90) to represent a 3-month operation period and keep the computational time at a manageable scale. Average daily trip size N_{trip} should be estimated based on the local targeting populations and their mobility needs. In our case, we approximate the *expected* average daily trip size $N_{trip} = 6000$. Typically, the observed data fit a probability distribution among the 3-month operation period, as shown in Figure 3. The requested trip size for each scenario was generated using Monte Carlo Sampling process from a normal distribution.



Trip sizes following normal distribution with mean = 6000, variance = 1200, and number of scenarios S = 900

Figure 3.2 Trip Sizes Following Normal Distribution

After the *size* of the expected daily demand is generated for each scenario, the hash table for each scenario $\boldsymbol{d_s}$ is generated in such a way that start time t_o (origin), start region i_o, and end region i_d (destination) are all randomly selected whereas end time t_d is calculated based on the corresponding t_o, i_i, i_d and travel time matrix T_{mat} in Table 3.1. In Chapter 4.2.3, we further explored how different types of request travel data impact the overall car-sharing system.

Table 3.1 Travel Time Matrix

Region (min)	1	2	3	5	6	7	8	9
1	0	15	30	30	30	30	15	30
2	15	0	15	30	30	30	15	30
3	30	30	0	30	15	30	30	30
4	30	30	30	0	30	15	30	30
5	30	30	15	30	0	15	15	30
6	30	30	30	15	15	0	30	30
7	15	15	30	30	15	30	0	30
8	30	30	30	30	30	30	30	0

Algorithm 2

Input: Instance $P = (I, T, S, r, c, h_r, h_i, \delta, B)$, expected demands $\boldsymbol{d}$ (see Algorithm 1)
Output: Total vehicle fleet size at the initial time $\sum_i^I x_{i,t}$ $where\ t = 1$

#Set up the Objective function: model 2-1

$$prob = \underset{x \in X}{minimize}\ v \sum_{i=1}^{I} x_{i,1} + \frac{1}{S} \sum_{s=1}^{S} [c_I y_{\varepsilon_I} + c_L y_{\varepsilon_L} - rd + \delta(\underline{d} - d)]$$

#Set up the variables with initial boundaries
$x_{i1} \geq 0, y_{\varepsilon_I} \geq 0, y_{\varepsilon_L} \geq 0$

#Iterate through all scenarios
For s in range (1, S+1):
    #In each scenario, iterate through the time
    For t in range (1, T+1):
        #In each time, iterate through each location i
        For i in range (1, I+1):
            #Subject to the following constraints:
            # Flow constraint:
            #Inbound: model 2-4
            If t< T:
            $x_{i,t}^s - \sum_{e \in n_{i,t(out)}} (y_{\varepsilon_I} + y_{\varepsilon_L} + d) = 0$
            #Outbound: model 2-5
            if t> 1:
                $\sum_{e \in n_{i,t+1(in)}} (y_{\varepsilon_I} + y_{\varepsilon_L} + d) - x_{i,t+1}^s = 0$

    #Budget constraint: model 2-2
    $c \sum_{i=1}^I x_{i,1} \leq B$
    #Fleet size constraint: $\sum_i^I x_{i,1}$ share the same value for all scenarios
    if s>1:
        $\sum_i^I x_{i,1}^s - \sum_i^I x_{i,1}^{s+1} = 0$
    #Fulfillment rate constraint: model 2-6
    $\alpha \sum \underline{d}_{i,t} \leq \sum d_{i,t} \leq \sum \underline{d}_{i,t}$
#Solve the optimization problem and update $x_{i,1}$ using CBC (Coin-or branch and cut)
$x_{i,1} \leftarrow Solve\ prob$

Return $\sum_i^I x_{i,1}$

The pseudocode for the mixed-integer linear programming (MILP) model is presented in Algorithm 2. The algorithm reflects the multi-stage stochastic model, which is constructed in Chapter 3.3. PuLP is an open-source package that integrates with many linear and mixed-integer models, including commercial solvers such as CPLEX and Gurobi (Mitchell et al 2011), and free open-source solver such as CBC. PuLP can default to its own choice of available solvers, depending on the problem structures. In our case study, we select CBC as our default solver given the target users are mainly government or NPG. Algorithm 2 was coded in Python as a MILP and was solved by CBC. The algorithm is run on Cloud Google Colab.

3.3 A multi-stage Stochastic Programming Model

We employ a multi-stage stochastic model where the demand for the trips is represented by $\boldsymbol{\widetilde{d}_t} = \{\boldsymbol{\widetilde{d}_1}, \dots, \boldsymbol{\widetilde{d}_n}\}$. Assuming that the *expected* $\boldsymbol{\widetilde{d}}$ follows a certain probability distribution (e.g., Normal distribution) before the *actual* demand $\boldsymbol{d}$ is known, the company will purchase the vehicles at the fixed Cost of v per vehicle. The optimization model below describes the first stage problem.

$$\min_{x \in \boldsymbol{X}} v \sum_{i=1}^{I} x_{i,1} \; + \mathbb{E}[\boldsymbol{Q}(\boldsymbol{x}, \boldsymbol{\widetilde{d}})] \tag{3-1}$$

$$s.t. \; x_i \in \boldsymbol{X} = \{\boldsymbol{x} \in \mathbb{Z}_+ : v \sum_{i=1}^{I} x_i \le B, \; i = (1,2, \dots, I)\} \tag{3-2}$$

where $x_{i,t}$ denotes the supply level of vehicles at node $n_{i,t}$. Specifically, $x_{i,1}$ is the first stage decision which denotes the number of vehicles to purchase and deploy in region *i* at $t = 1$. $\boldsymbol{\widetilde{d}}$ is the vector of *expected* demand $\tilde{d}_{\mathrm{ii}'\mathrm{tt}'} \in \boldsymbol{\widetilde{d}}, i, i' \in I = \{1, 2, \dots I\}, t, t' \in T = \{1, 2.., T\}$ before the demand is known. $\mathbb{E}[\boldsymbol{Q}(\boldsymbol{x}, \boldsymbol{\widetilde{d}})$ is the Optimal value of the expected resource function, and B is the budget. The decision process follows the form:

$$expectation \; (\boldsymbol{\widetilde{d_1}}) \rightarrow decision \; (\boldsymbol{x_1}) \rightarrow expectation \; (\boldsymbol{\widetilde{d_2}}) \rightarrow decision \; (\boldsymbol{x_2}) \\ \rightarrow expectation \; (\boldsymbol{\widetilde{d_3}}) \rightarrow \cdots expectation(\boldsymbol{\widetilde{d_{T-1}}}) \rightarrow decision \; (\boldsymbol{x_{T-1}})$$

The goal is to minimize the objective function in 3-1, which is the total Cost of purchasing vehicles at the Service Capacity Organization Stage and the total expected Cost at the operating stage, subject to the constraints in Equation 3-2. Equation 3-2 specifies that the total Cost of purchasing does not exceed the given budget B. After the demand is observed, true demand is denoted by $\boldsymbol{d} = \{\boldsymbol{d^1}, \boldsymbol{d^2}, \dots, \boldsymbol{d^S}\}$. The company decides how to make the optimal reposition decisions. The resource decision is denoted by $y_e^s \in \boldsymbol{Y^s}$ and represents flows on arc $e \in E = \varepsilon_R \cap \varepsilon_L \cap \varepsilon_I$, in scenario $s \in S = \{1,2, \dots, S\}$. The optimization model below describes the problem for stage t = 2, 3, …, T.

$$Q(\boldsymbol{x}, \boldsymbol{d}) = \min_{y \in \boldsymbol{Y}} \left\{ \sum_{s \in S} \left[\sum_{e \in E} \boldsymbol{c_e} \, \boldsymbol{y_e^s} \; + g\,(\boldsymbol{Y}) \right] \right\} \tag{3-3}$$

$$s.t.\ \ \boldsymbol{Y}\,(\boldsymbol{x},\boldsymbol{d}) \coloneqq \{\, y_e^s \geq 0\ \ \forall e \in E = \ \varepsilon_R \cap \varepsilon_L \cap \varepsilon_I, s \in \boldsymbol{S} = \{1,2,\dots,S\}: \quad x_{it} = \ \textstyle\sum_{e\in\, n_{it(out)}} y^s\,, \forall\, s \in \boldsymbol{S}\,\} \tag{3-4}$$

$$x_{it} = \textstyle\sum_{e\in\, n_{it(in)}} \boldsymbol{y^s}\,,\ \ \forall s \in \boldsymbol{S} \tag{3-5}$$

$$\alpha \textstyle\sum \boldsymbol{d_{i,t}} \leq \sum\ \underline{\boldsymbol{d}}_{\boldsymbol{i,t}} \leq\ \ \sum \boldsymbol{d_{i,t}}\,,\ \forall s \in \boldsymbol{S} \tag{3-6}$$

Where $\boldsymbol{d}$ is the observed demand vectors $\boldsymbol{d_{ij}^s} \in \boldsymbol{d} = \{\boldsymbol{d^1}, \boldsymbol{d^2}, \dots, \boldsymbol{d^S}\}, s \in \boldsymbol{S} = \{1,2,\dots,S\}$, $i,j \in \boldsymbol{I} = \{1, 2, \dots I\}$, here we sometimes also use ($y_{\varepsilon_R}^s = d^s$), and $\boldsymbol{c}$ is the Cost per unit flow; $g(\boldsymbol{Y})$ is the penalty from unfulfilled demands and is given by

$$\boldsymbol{g}(\boldsymbol{Y})\ = \delta(\textstyle\sum \underline{\boldsymbol{d}} - \sum \boldsymbol{d});\tag{3-7}$$

δ is the unit cost for each unfulfilled demand; and $\underline{d}$ is “True” fulfilled demand.

The objective in 3-3 is to minimize the total Cost at the *Operating Stage* given the optimal $x_{i,1}$ decided in the first stage. By implementing SAA, the *expected* resource function can be written as:

$$\widehat{\boldsymbol{Q}}\left(\boldsymbol{x}, \widetilde{\boldsymbol{d}}\right) \coloneqq \min_{x\in X} \{\ \frac{\boldsymbol{1}}{\boldsymbol{S}} \textstyle\sum_{s=1}^{S}\ \ [\sum_{e\in E} \boldsymbol{c_e}\, \boldsymbol{y_e^s} +\ \boldsymbol{g}(\boldsymbol{Y})\,]\} \tag{3-8}$$

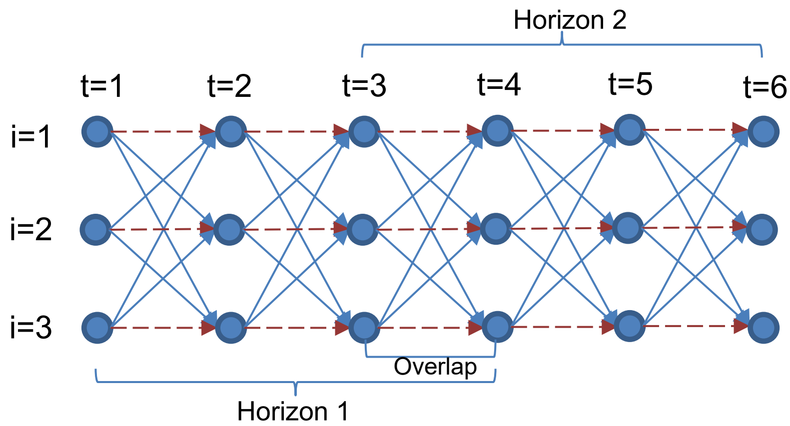
For every scenario s, constraints 3-4 and 3-5 define the vehicle movement at each node n_{it}. Constraint 3-4 defines total outbound flow, whereas constraint 3-5 describes the total inbound flow, which should equal to the supply level x_{it} at node n_{it}. Additionally, constraint 3-6 introduces the fulfillment rate α to ensure that the service flow satisfies the observed demand $\boldsymbol{d}$ at or above the fulfillment rate α.

3.4 Rolling Horizon Model

The Rolling Horizon model is implemented in both *Service Capacity Organization Stage (Stage I)* and *Operating Stage (Stage II).* In *Service Capacity Organization Stage,* the RH determines the optimal fleet size with the *expected* request data before the car-sharing system starts to operate. During *Stage II,* the RH is applied to make repositioning decisions that minimize the total operational costs according to the *actual* request data with the given fleet size determined from *Stage I*.

Instead of solving the problem by a multi-stage stochastic model, RH divides the total operation time into manageable batches, known as horizons, so that the spatial-temporal model is solved repeatedly for a horizon length of *H* by a linear programming model similar to *Algorithm 2*. Each $\boldsymbol{horizon}$ is being overlapped by a certain period, defined by $\boldsymbol{o}.$ Figure 4 shows a rolling horizon framework example with the horizon length equal three time periods and overlapping time equals one time period. The overlapping period o shall be no less than two times the maximum single trip length minus one time period, as shown by the following inequality.

$$o \geq 2(max(t_d - t_o)) - 1 \qquad \text{3-9}$$



Rolling horizon framework with $i = 3$ and $T = 6$. Since the maximum single trip takes only one period, overlapping time is set to be 1

Figure 3.3 Rolling Horizon Framework Example

To determine the optimal fleet size, the *expected* request daily trip is generated using the same sampling technique illustrated in section 3.2, after which the RH framework is implemented with the objective function defined in model 3-1 for every horizon $H_i \in \boldsymbol{H} = \{H_1, H_2, \dots, H_n\}$. The model is subject to the same constraints defined in model 2-4, 2-5, and 2-6. Additionally, for a horizon $h_i, i \geq 2$, the total initial fleet size should be greater than that in H_{i-1}. A simple RH framework is defined in *Algorithm 3*, whereas the linear programming function, which is very similar to *Algorithm 2*, is defined in *Algorithm 4*.

Algorithm 3

Input: Instance $P = (I, T, S, r, c, h_r, h_i, \delta, B, H)$, expected demands $\boldsymbol{d}$ (see Algorithm 1)
Output: Total vehicle fleet size at the initial time $\sum_i^I x_{i,t}$ $where\ t = T - H$

#The Rolling Horizon Framework
#Set the initial total fleet size to 0
$fleet\_size = 0$
for t_o in range $(1, T \times S - H + 1, H - o)$:
    $fleet\_size = LP(t_o, P, fleet\_size)$ #See Algorithm 4
return $fleet\_size$

Algorithm 4

#Define function $LP(t_o, P, fleet\_size)$
$LP(t_o, P, fleet\_size)$:
#Set up the Objective function: model 2-1

$$prob = \underset{x \in X}{minimize}\ v \sum_{i=1}^{I} x_{i,1} + \frac{1}{S} \sum_{s=1}^{S} [c_I y_{\varepsilon_I} + c_L y_{\varepsilon_L} - rd + \delta(\underline{d} - d)]$$

#Set up the variables with initial boundaries
$x_{i1} \geq 0, y_{\varepsilon_I} \geq 0, y_{\varepsilon_L} \geq 0$

#Iterate through all scenarios
for t in range $(t_o, t_o + H)$:
#In each time, iterate through each location i
For i in range (1, *I*+1):
#Subject to the following constraints:
#Flow constraints:
#Inbound: model 3-4
If t< $t_o + H$:
$x_{i,t}^s - \sum_{e \in n_{i,t(out)}} (y_{\varepsilon_I} + y_{\varepsilon_L} + d) = 0$
#Outbound: model 3-5
if t> t_o:
$\sum_{e \in n_{i,t+1(in)}} (y_{\varepsilon_I} + y_{\varepsilon_L} + d) - x_{i,t+1}^s = 0$
#Budget constraint: model 3-2
$c \sum_{i=1}^{I} x_{i,1} \leq B$
#Fleet size constraint:
$\sum_i^I x_{i,t=t_o} \geq$ fleet_size
#Fulfillment rate constraint: model 3-6
$\alpha \sum \underline{d}_{i,t} \leq \sum d_{i,t} \leq \sum \underline{d}_{i,t}$
#Solve the optimization problem and update $x_{i,1}$ using CBC (Coin-or branch and cut)
$x_{i,1} \leftarrow Solve\ prob$

Return $\sum_i^I x_{i,1}$

After the optimal fleet size is determined, the rolling horizon framework can be later used to make real-time reposition decisions at *Stage II* to minimize the total operational costs when the car-sharing system starts to operate.

CHAPTER 4

Application of The Developed Optimization Model: Case Study in Columbia SC

In this chapter, we present an application of the optimization model developed in this study to a case study utilizing the geographic information in Columbia, South Carolina, and the typical fleet operation cost information from the literature.

4.1 Numerical Settings

4.1.1 Fleet Stations

In this case study, we propose a station-based model with eight fleet locations ($I = 8$) selected from Columbia, the capital city of South Carolina, as stations. The eight fleet locations were chosen to be close to the service areas, which include senior housing, clinics, medical center, shopping center, train stations, and airports. The service routes include all the exact addresses for each service area, as shown in Table 4.1. Figure 4.1 shows the approximate locations of the eight service areas.

Table 4.1 Selected Service Areas in the City of Columbia

NO.	NAME	ADDRESS	TYPE
1	Merrill Gardens at Columbia	2205 Gregg St, Columbia, SC 29207	Retirement Community
2	Still Hopes Episcopal	1 Still Hopes Dr, West Columbia, Sc 29169	Retirement Community
3	Lexington Medical Center	2720 Sunset Blvd, West Columbia, Sc 29169	Hospital
4	Columbiana Centre	100 Columbiana Cir, Columbia, Sc 29212	Shopping Mall
5	Total Dental Care Of South Carolina	1061 St Andrews Rd, Columbia, Sc 29210	Dental Care
6	Seven Oaks Senior Citizens Center	200 Leisure Ln Columbia, SC 29210	Senior Service Center
7	Amtrak Station	850 Pulaski St, Columbia, Sc 29201	Transportation Hub
8	Columbia Metropolitan Airport	3250 Airport Blvd, West Columbia, Sc, 29170	Transportation Hub

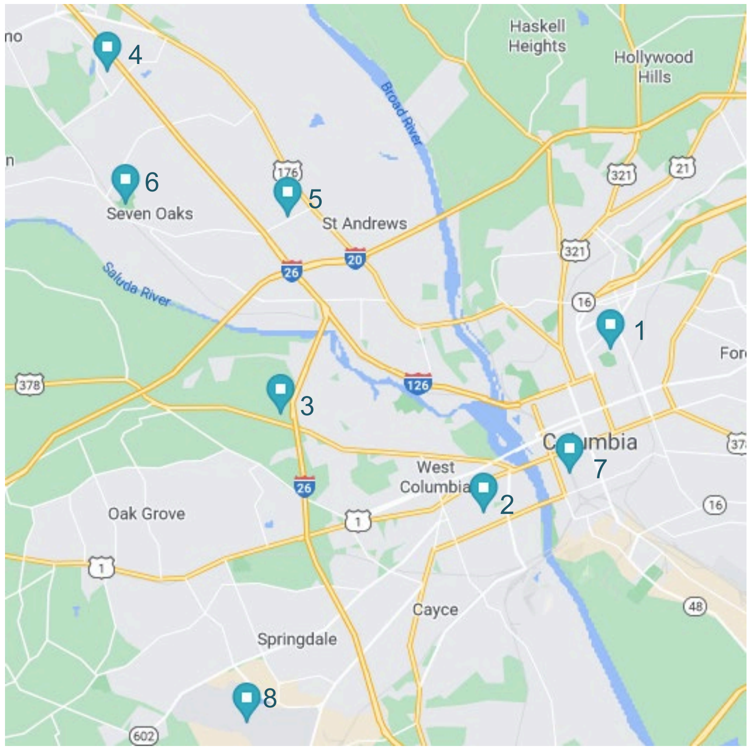


Figure 4.1 Selected Service Areas marked on Google Map

4.1.2 Operation Stage Hour

The demand-responsive transport service is planned to offer door-to-door transportation of the elderly and persons with disabilities in a three-month period, during which the operation hour is chosen to be from 10:00 AM to 3:00 PM daily with one time period equals to 15 minutes, resulting in a total of 20 time periods daily (time node $T = 21$). The static travel time (minutes) for each service route without considering traffic is estimated using Google map.

4.1.3 Cost Categories

A summary of other hypothetical cost categories is described below:

- Vehicle acquisition cost: $v = \$16,000$ per vehicle. However, we use $v = \$54$ as the daily depreciation of the vehicle when considering the resale value.
- Cost per unit flow $\boldsymbol{c}$:
    - Vehicle idle cost: $c_I = \$0.5$ per vehicle per time period (15 minutes)
    - Vehicle reposition cost: $c_L = \$3$ per vehicle per time period
- To simplify the calculation, the positive impact (revenue) generated by fulfilled ride requests is set as $r = -\$5$ per trip per time period
- Penalty due to unfulfilled ride request: $\delta = \$2$ per request
- Budget B=$500,000 is the constraint for the *total Cost* of the three-month (90 days) service *after* the resale of the vehicle. The daily budget is therefore calculated as $5,500.

4.2 Numerical Results

Table 4.2 shows different initial fleet sizes with their associated cost categories. In the table, the negative sign indicates that the car-sharing system model is making a profit, the second and third columns with the names “Min Fulfill Rata” and “Actual Fulfill Rate” will be further explained in section 4.2.1. We also generate a random travel request dataset with its daily trip size following gamma distribution using the same mean (6000 trips) and variance (1200 trips). The results for initial fleet sizes under different fulfillment rate constraints α are very similar to the model with the trip size following a normal distribution. We used the normal distribution trip size data for the rest of the study.

Table 4.2 Different Initial Fleet Sizes with Their Associated Cost Categories

Fleet size	Min Fulfill Rate	Actual Fulfill Rate	Idle Cost ($)	Reposition Cost ($)	Revenue ($)	Penalty ($)	Daily Total Cost ($)
485	66.5%	90.0%	473.29	36.57	43,165.83	1,125.84	-15,157.30
520	70.0%	92.8%	659.05	41.93	44,707.83	805.60	-14,911.58
571	75.1%	95.9%	995.51	48.14	46,319.33	461.93	-13,739.02
621	80.0%	97.8%	1,395.55	46.92	47,308.83	242.13	-11,855.62
672	85.1%	98.9%	1,859.87	39.38	47,881.17	121.38	-9,375.65
728	90.1%	99.6%	2,395.84	34.47	48,206.33	44.76	-6,246.90
790	95.0%	99.9%	3,017.80	25.55	48,333.67	10.91	-2,491.66
890	100.0%	100.0%	4,033.31	14.76	48,379.28	0.00	3,802.60

4.2.1 Impact of Fulfillment Rate

Constraint 3-6 brings in a fulfillment rate constraint to make sure that for every scenario $s_i, i = 1,2,\ldots,90$, a minimum fulfillment rate α is met. Figure 4.2 shows the *actual* minimum fulfillment rate α_{min}, given by:

$$\alpha_{min} = \min\left\{\underline{\alpha_1}, \underline{\alpha_2}, \ldots, \underline{\alpha_s}\right\}, S = \{1,2,\ldots,90\}$$

as well as the *overall* fulfillment rate $\alpha_{overall}$, given by:

$$\alpha_{overall} = \frac{\sum_{s\in S}\sum_{e\in E}\mathrm{d}_e^s}{\sum_{s\in S}\sum_{e\in E}\underline{d_e^s}}, \quad S = \{1,2,\ldots,90\}$$

under different minimum fulfillment rate constraints α. It is worth noticing that the actual minimum fulfillment rate after solving the LP is 66.5%, and the actual overall fulfillment rate is already 90.0% without considering the fulfillment rate constraint (α = 0). As the minimum fulfillment rate constraint α keeps increasing, the growth of the actual overall fulfillment rate $\alpha_{overall}$ becomes slower.

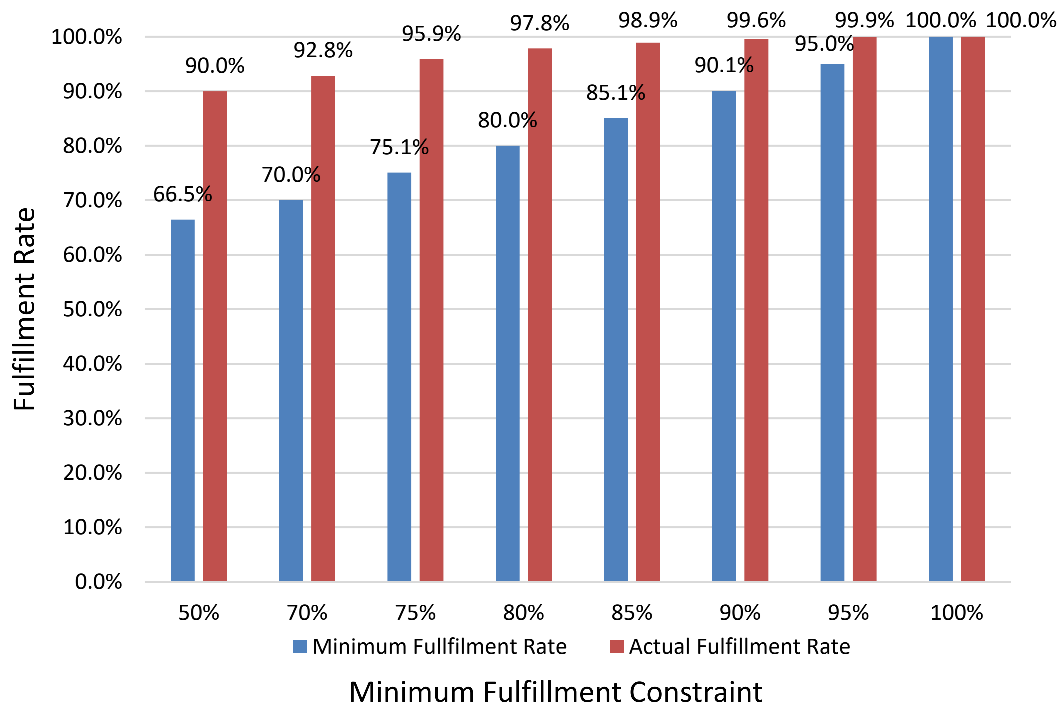
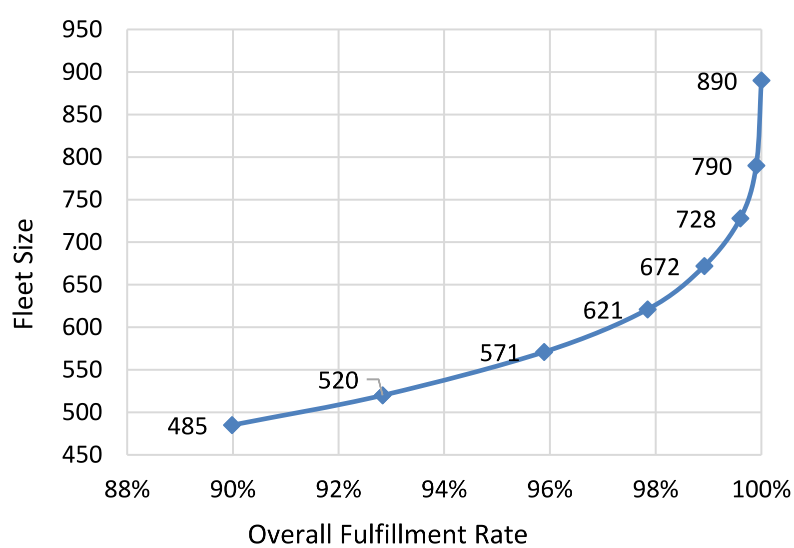
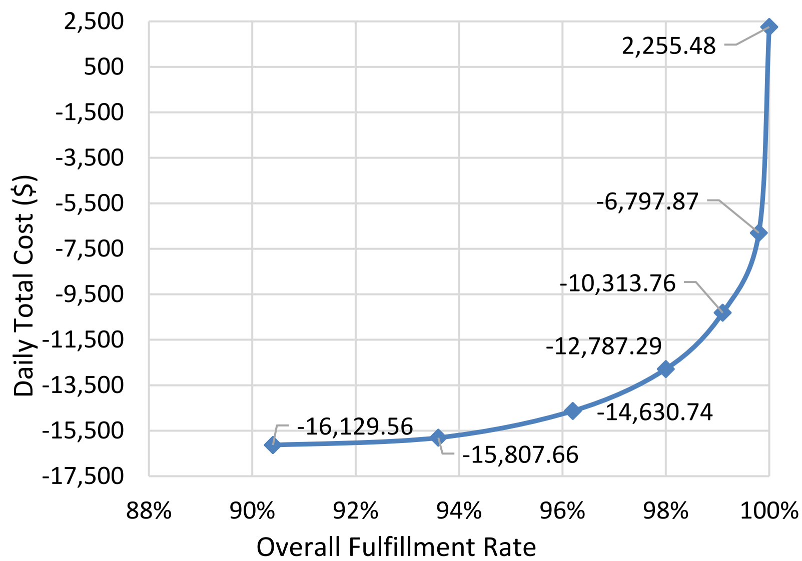


Figure 4.2 Different Measurements of Fulfillment Rates

Figure 4.3 a) suggests that the total fleet size $\sum_{i=1}^{I} x_{i,1}$ and the actual fulfillment rate $\alpha_{overall}$ has a linear relationship before $\alpha_{overall}$ reaches 96%, after which the fleet size increases exponentially as $\alpha_{overall}$ increases, indicating a significant increase in fleet size can only bring up the fulfillment rate $\alpha_{overall}$ by a small amount. Figure 7 b) shows that the optimal total cost, defined by model 1-1, increases exponentially as actual overall fulfilment $\alpha_{overall}$ increases. Finally, Figure 4.4 shows the relationship between fleet size and daily cost. The relationship is linear by large. It also shows that the car-sharing system stops generating profit as the fleet size reaches 900.



(a) Impact of Different Actual Overall Fulfillment Rate $\alpha_{overall}$ on Optimal Fleet Size



(b) Impact of Different Actual Overall Fulfillment Rate $\alpha_{overall}$ on the Daily Total Cost

Figure 4.3 Impact of Different Actual Overall Fulfillment Rate $\alpha_{overall}$

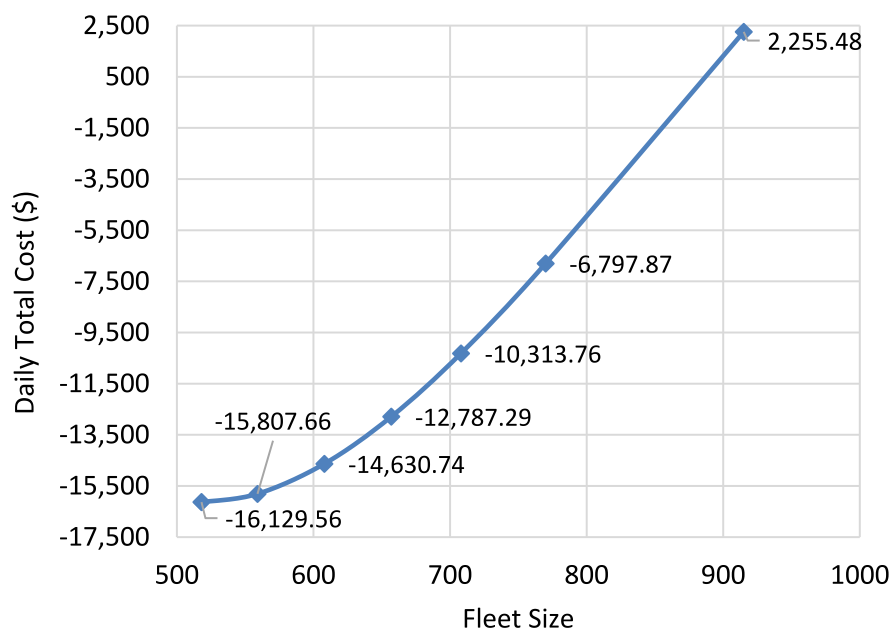


Figure 4.4 Total Daily Cost at Different Daily Fleet Sizes

4.2.2 Impact of "Positive Impact"

We further studied the profitability of the car-sharing system with different unit positive impacts r, measured in dollars per period. The study of the impact of positive impact r was performed with minimum fulfillment rate constraint α set to 0.8 ($\alpha_{overall} = 0.971$), right before the relationship between the fleet size and fulfillment rate changes from linear to exponential. The impact of the unit positive impact r is shown in Figure 4.5.

The car-sharing system starts to make a profit as the unit positive impact r reaches approximately \$2.6 per time period (15 min). Figure 4.5 shows that the lowest positive impact value can be \$2.2 per hour to meet all the constraints without exceeding the daily budget.

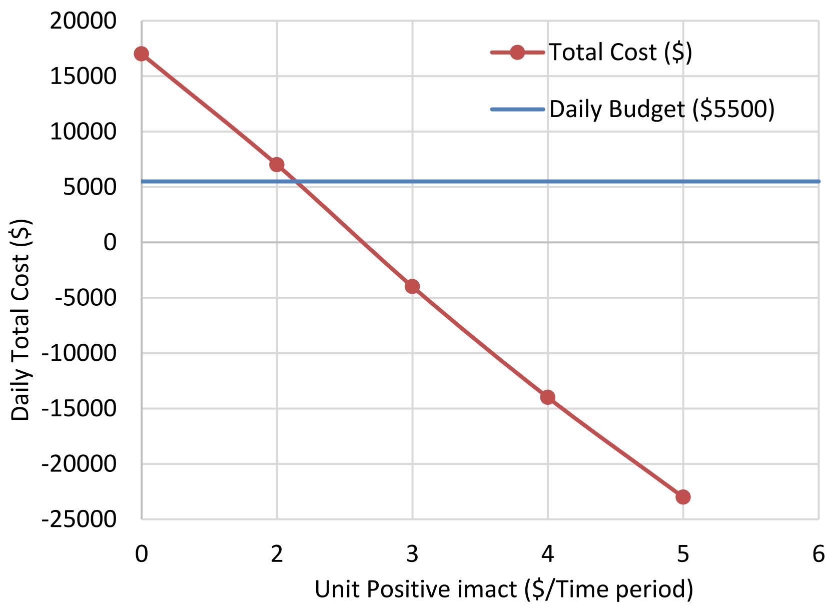
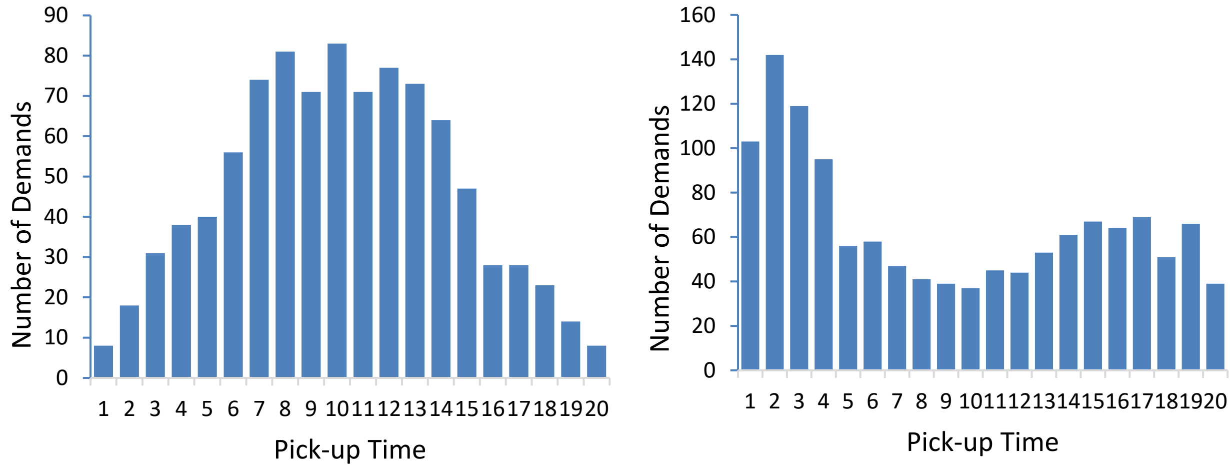


Figure 4.5 Total Stage-One Cost Plus Expected Operation Cost at The Different Unit Positive Impact

4.2.3 Impact of Pick-up Times Following Several Types of the Probability Distribution

The case study shown above studies the car-sharing when the pick-up (start) time follows a uniform distribution for every scenario. However, the actual pick-up (start) time hardly follows a uniform distribution. Therefore, two additional sets of request data, which follow two types of probability distributions, are generated to reflect the expected demands better. Figure 4.6 a) shows that the daily request data is Normally distributed, with the peak pick-up time at around noon. Figure 4.6 b), on the other hand, shows a Bimodal Gamma distribution with two peak pick-up times at around 10:00 AM and 2:00 PM ($T = 16$).



(a) Pick-up time follows Normal distribution

(b) Pick-up time follows Bimodal Gamma distribution

Figure 4.6 Pick-up Time Data Follows Two Different Probability Distributions

Table 4.3 and Table 4.4 show the summarized information regarding the car-sharing system when demands within a single day follow a normal distribution and a bimodal gamma distribution, respectively. Figure 4.7 shows the impact of the demand distribution on the relationship between the service quality, quantified by the actual overall fulfillment rate $\alpha_{overall}$ and the total daily cost. When the total trip size remains the same, the daily demand data following a uniform distribution requires the least total cost to support the car-sharing system, whereas a model following daily bimodal gamma distribution and normal distribution require a similar cost. Since the actual daily demand distribution is unknown before the car-sharing system begins to operate, it is reasonable to consider all three types of distribution and assign equal weight to all three types of request datasets. To make sure the system operates under the budget while maintaining the desired service level ($\alpha_{overall} \geq 0.85$), it is recommended to start with **556** vehicles. Later, as more actual data is available, more vehicles can be added according to the same car-sharing model to maintain a designated level of fulfillment.

Table 4.3 Different Initial Fleet Sizes with Their Associated Costs When Demands within A Single Day Follows A Normal Distribution

Fleet size	Min Fulfill Rate	Actual Fulfil Rate	Idle Cost ($)	Reposition Cost ($)	Revenue ($)	Penalty ($)	Total Cost ($)
473	60.3%	77.5%	1,068.29	16.36	36,305.78	2,530.22	-7,067.09
608	70.1%	87.4%	1,845.85	21.98	41,928.11	1,412.29	-5,706.07
694	75.0%	92.1%	2,441.24	27.77	44,466.72	890.89	-3,491.99
785	80.0%	95.3%	3,167.80	28.83	46,272.28	523.64	-17.87
887	85.0%	97.6%	4,068.08	27.26	47,499.94	267.82	4,897.49
1006	90.0%	99.1%	5,196.16	20.59	48,236.67	100.58	11,507.63
1158	95.0%	99.8%	6,701.46	10.88	48,572.50	26.80	20,753.06
1421	100.0%	100.0%	9,338.11	2.71	48,668.00	0.00	37,420.34

Table 4.4 Different Initial Fleet Sizes with Their Associated Costs When Demands within A Single Day Follow A Bimodal Gamma Distribution

Fleet size	Min Fulfill Rate	Actual Fulfil Rate	Idle Cost ($)	Reposition Cost ($)	Revenue ($)	Penalty ($)	Total Cost ($)
432	57.9%	77.7%	580.86	24.64	36,947.61	2,510.96	-10,379.96
575	70.1%	86.9%	1,487.20	26.24	42,135.56	1,472.18	-7,968.71
648	75.0%	89.8%	2,072.89	19.41	43,708.33	1,149.47	-5,377.54
725	80.0%	92.1%	2,725.84	16.07	44,952.44	883.31	-2,096.89
835	85.0%	94.6%	3,711.34	11.25	46,207.89	611.58	3,272.53
1018	90.0%	97.3%	5,414.42	8.84	47,546.72	302.69	13,195.46
1324	95.0%	99.5%	8,387.69	4.57	48,476.94	58.93	31,493.12
1778	100.0%	100.0%	12,913.70	0.20	48,661.00	0.00	60,265.90

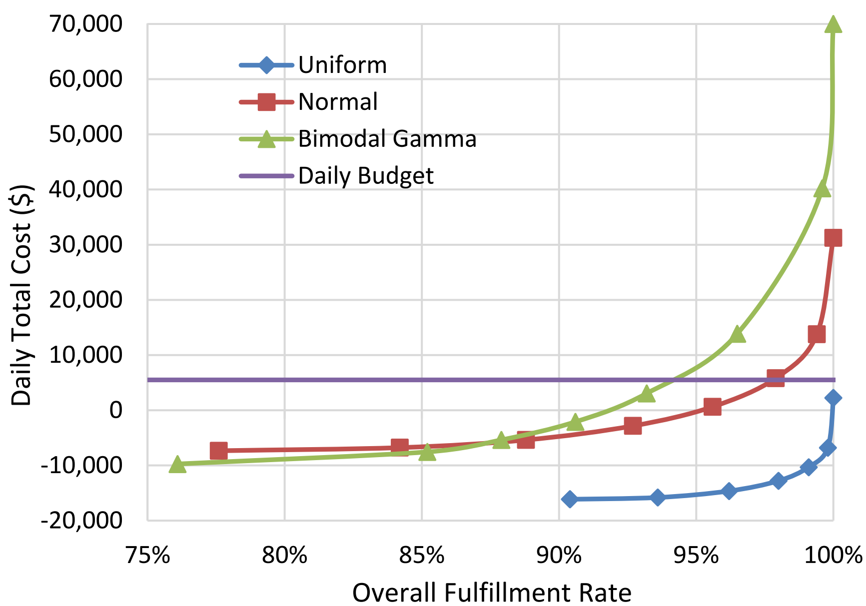


Figure 4.7 Comparison Among Three Different Distributions

4.3 Further Exploration on Rolling Horizon Heuristic

4.3.1 Impact of Horizon Length

During the *Service Capacity Organization Stage*, it is anticipated that as the length of the horizon increases, the total daily cost would decrease. The logic behind this is that as the model "sees" further into the future, the LP model can make better decisions on how to optimize repositions. Under the same numerical settings, Table 4.5 gives the optimal total daily cost at different horizon lengths. It shows the results when the overlapping time period is fixed at $o = 3$ and the fulfillment rate at $\alpha = 1$. The model confirms that total cost decreases as the length of the horizon increases. Meanwhile, the total runtime decreases at first and increases almost linearly as the horizon length increases, shown in Figure 4.8. From Table 4.5, we can identify that when the overlapping time period is set to 3, the optimal runtime happens at H=9.

Note that the total runtime can be broken down as the average runtime per epoch times the number of epochs. Figure 12 indicates that the average runtime per epoch (unit runtime) in our experiment fits into a 2-degree polynomial function. The number of epochs can be calculated using $\frac{S(T-1)}{H-o}$. The degree of the polynomial function can be different, however, as the input size or type of MIP solvers change. Therefore, the general total runtime (*R*) has a close-to estimation to model 4-1.

$$R=\frac{\boldsymbol{S(T-1)}}{\boldsymbol{H-o}}\times(a_1H^n+a_2H^{n-1}+\cdots+a_{n+1}) \qquad \textit{4-1}$$

where a's are real numbers and n's are integers.

Figure 4.9 shows the numerical comparison between two different approaches. For the RH approach, the horizon length and overlap are set to 9 and 3, respectively. All other numerical settings are identical. We can see that the two results are very similar. The multi-stage approach is generally comparable to an RH approach with a defined horizon length and overlapping time (*H=T-1,* and *o=0).* Thus, the runtime for multi-stage approach can be written as $S\times(a_1H^n+a_2H^{n-1}+\cdots+a_{n+1})$. It increases polynomially as *T* increases whereas RH is flexible in choosing horizon length, significantly affecting affects total running time.

At this point, it is safe to assume that the RH approach is comparable to multi-stage in producing acceptable results. We will continue to explore the impact of daily service hours using the RH framework.

Table 4.5 Optimal Total Daily Cost at Different Horizon Lengths

Fleet size	Horizon Length	Idle Cost $	Reposition Cost $	Revenue generated $	Total cost $
1421	5	16,071.65	0	48,668.00	44,137.65
1421	6	14,010.69	0.17	48,602.72	42,142.14
1421	7	13,031.62	0.67	48,602.72	41,163.57
1421	8	12,439.73	0	48,629.17	40,544.56
1421	9	12,016.69	0.17	48,602.72	40,148.14
1421	10	11,720.65	0.67	48,577.39	39,877.93
1421	20	10,654.53	0.83	48,104.67	39,284.69
1421	30	10,338.35	0.83	48,032.78	39,040.41
1421	40	10,162.36	0.83	47,599.17	38,893.88
1421	50	10,253.93	0.97	48,468.83	38,520.06
1421	60	10,210.99	3.03	48,512.33	38,435.69

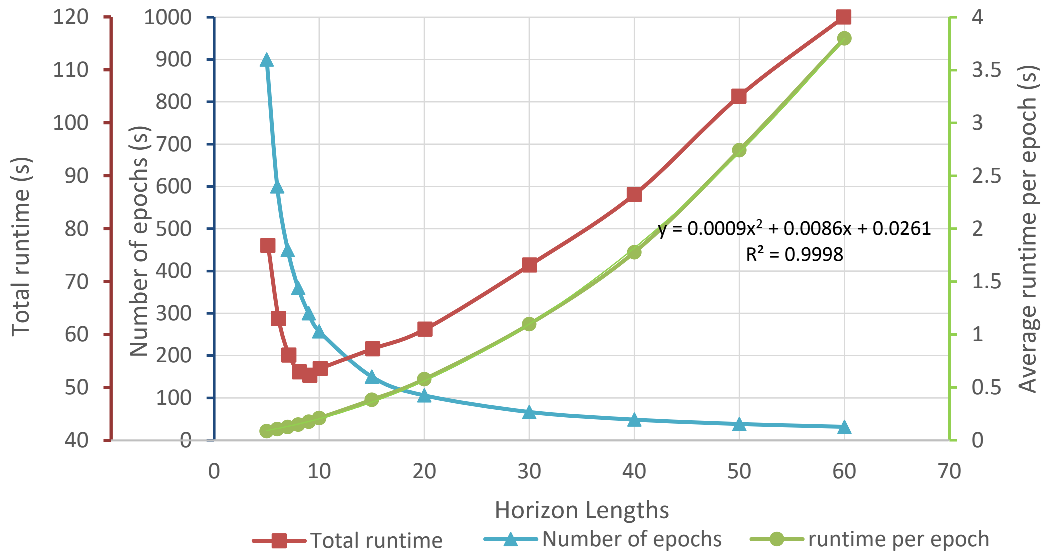


Figure 4.8 Run Time Measured at Different Horizon Lengths for T=21

As daily operating time T increases, the computational complexity only increases linearly when the rolling horizon framework is implemented, and the results are comparable to that when a multi-stage model is applied, which is shown in Figure 4.9. For example, when daily demand data follow a normal distribution, the fulfillment rate's impact still holds on the optimal fleet size, as shown in Figure 4.10. According to the rolling horizon approach, the optimal fleet size that can fulfill the 85% fulfillment rate should be around 600.

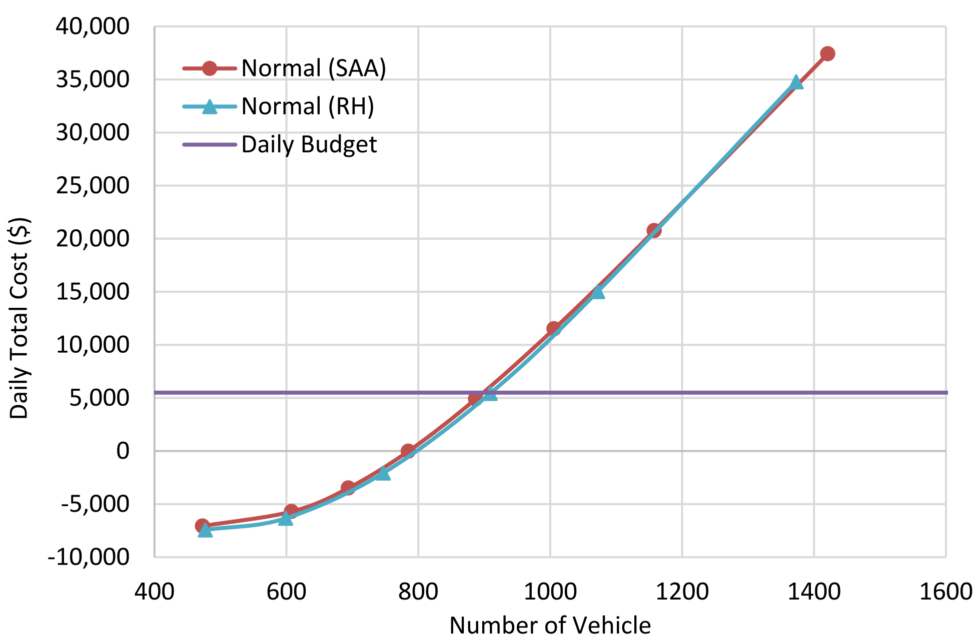


Figure 4.9 Comparison Of Multi-Stage and RH Approaches for Daily Trips under A Normal Distribution

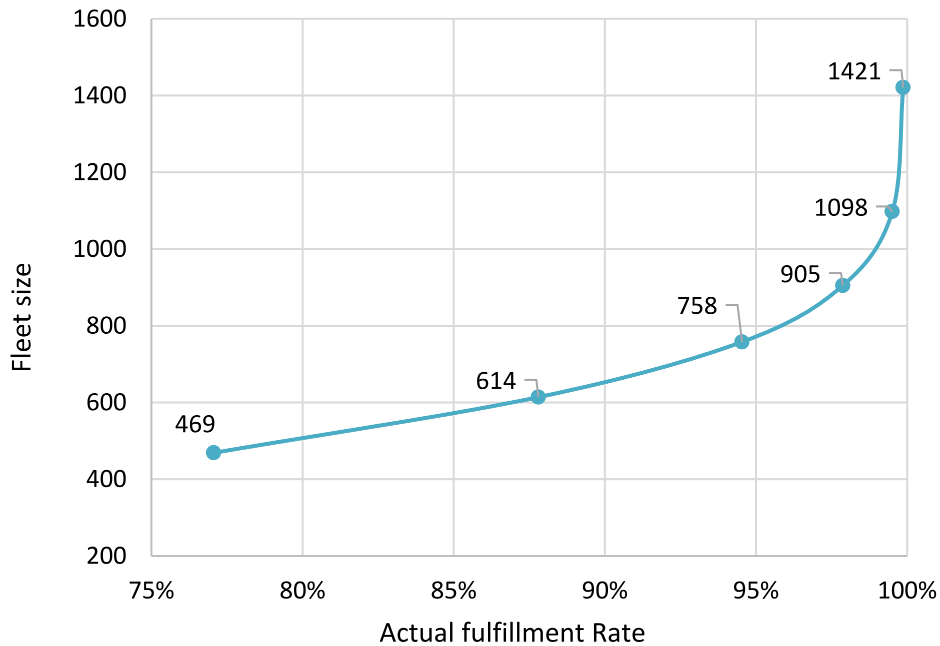


Figure 4.10 Optimal Fleet Size at Different Actual Overall Fulfillment Rate $\alpha_{overall}$

4.3.2 Impact of Daily Service Hours

To cover typical business hours ("9am to 5pm") for most businesses, we propose to extend the operating hours of the car-sharing system from 8:00 AM to 8:00 PM ($T = 48$). When daily trip size remains the same, extending operation hours can significantly reduce the fleet size while maintaining the same overall fulfillment rate. Hence the cost at *Service Capacity Organization Stage* is lowered. Figure 4.11 shows the comparison among three car-sharing models with daily data following a normal distribution. It suggests extending operating hours can significantly bring down the daily total cost.

For the extended service hour model with *T=48*, it is recommended to reduce the fleet size to 400, at which the model is generating profits, and the fulfillment rate reaches around 95%.

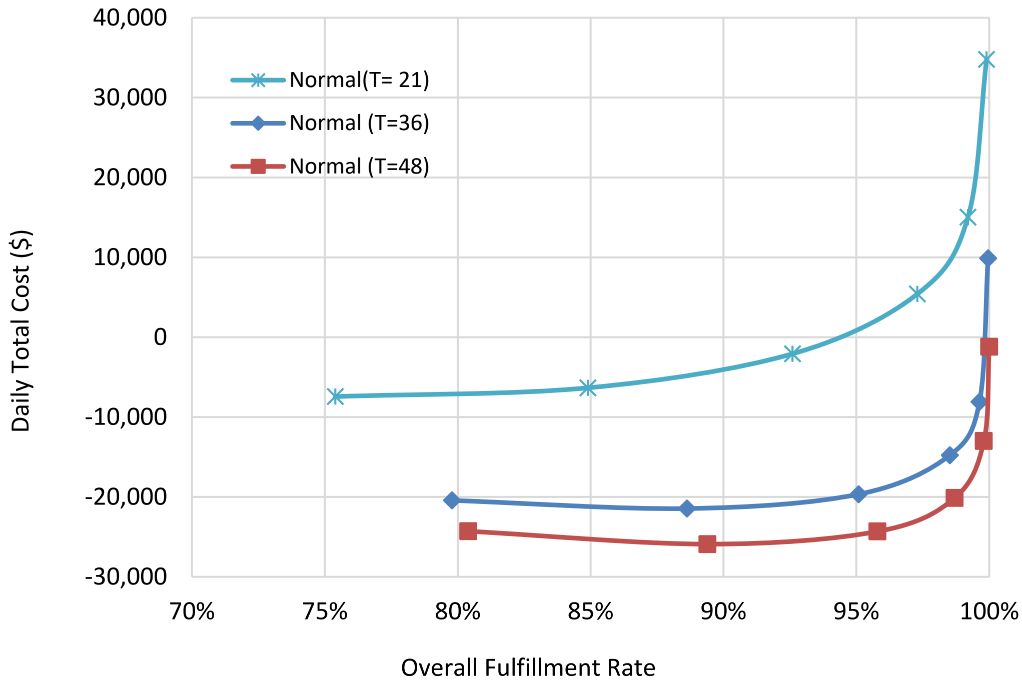


Figure 4.11 Comparison of The Impact of Two Different Service Hours

CHAPTER 5

Conclusions

In this study, a multi-stage stochastic integer programming model is developed to determine the initial fleet size for a temporary car-sharing system when the total cost includes *Service Capacity Organization Stage (Stage I)* and *Operating Stage (Stage II),* is optimal. The model is being tested on randomly generated data whose daily demands follow three different probability distributions and are later being solved by two different approaches. Though the demand data following a normal distribution, results in a relatively lower total cost, our results indicate that the demand distribution has little impact on the daily cost of the system as long as the average daily demands remain the same. Our results also suggest that the impact of fulfillment rate on fleet size is exponential under all three types of daily demand data, and the number of vehicles has a linear impact on the total cost. Furthermore, the Rolling Horizon (RH) is shown to be more flexible, with a guaranteed linear runtime with a fixed horizon length as daily operation times T increases than a multi-stage stochastic model and the results generated by the two approaches are very similar. We test the impact of the daily operating hour using the Rolling Horizon framework, and the results indicate that the daily operating hour has a significant impact on the fleet size when the total trip remains the same; that is, the longer the daily operating hours, the less the vehicle size is needed. Thus, understanding the optimal operating hours can significantly bring down total cost while fulfilling the needs of our target users.

The case study has two major limitations. First, we choose CBC (Coin-or branch and cut) as our default solver here as it is open source so that the local governments and NPOs who usually have limited funding can start building out their car-sharing system at their earliest convenience. The choice implies the managerial and practical nature of our study. However, CBC as a solver has many limitations itself. CBC's relatively slower speed compared to that of other commercial solvers which use the most advanced implementations of the latest algorithms will limit itself to solving small-sized problems due to computational complexity. The second limitation lies in the problem formulation portion of the study. First, the traveling time among the stations is based on static traffic information. It is important to consider real-time traffic as the travel time may have a significant impact on the results. Second, the model is being tested on demand data following three types of probability distribution to simulate real-life situations, considering real-time demand data is expensive and sometimes impossible to obtain, given there might be no active users to begin with. However, it is necessary our future research must test our model on creating real-time demand data so that it can reflect the actual need of our target audience and draw more insights. Furthermore, as electric cars are becoming less expensive and having increased driving range, having an electric fleet for our car-sharing system will positively impact the environment and society. Our future research plans to study electric or hybrid car-sharing system and compare it to the car-sharing service with traditional gas cars.

REFERENCES

ÁlvaroAguilera-García, Gomez, J., Antoniou, C., & Vassallo, J. M. (2022). Behavioral factors impacting adoption and frequency of use of car-sharing: A tale of two European cities. *Transport Policy*, 55-72.

Baker, K. R. (1977). An experimental study of the effectiveness of rolling schedules in production planning. *Decision Sciences, 8*(1), 19-27.

Carrese, S., D'Andreaiovanni, F., Tommaso, G., Nardin, A., & Zamberian, L. (2020). An optimization model and genetic-based matheuristic for parking slot rent optimization to car-sharing. *Research in Transportation Economics*.

Cordeau, J.-F., Dell'Amico, M., Falavigna, S., & Iori, M. (2015). A rolling horizon algorithm for auto-carrier transportation. *Transportation Research Part B: Methodological, 76*, 68-80.

Criden, M. (2008). *The Stranded Poor: Recognizing the Importance of Public Transportation for Low-Income Households.* National Association for State Community Services Programs.

Dabelko-Schoeny, H., Maleku, A., Cao, Q., White, K., & Ozbilen, B. (2021). “We want to go, but there are no options”: Exploring barriers and facilitators of transportation among diverse older adults. *Journal of Transport & Health*.

Golub, Aaron, Satterfeld, Vivian, Serritella, Michael, Singh, Jai, & Phillips, Senna (2019). Assessing the barriers to equity in smart mobility systems: A case study of Portland, Oregon. *Case Studies on Transport Policy, 7*(4).

Hartleb, J., & Schmidt, M. (2020). A Rolling Horizon Heuristic with Optimality Guarantee for an On-Demand Vehicle Scheduling Problem. *ATMOS*.

He, L., Hu, Z., & Meilin, Z. (2019). Robust Repositioning for Vehicle Sharing. *22*(2), 223-428.

Henderson, S. G., O'Mahony, E., & Shmoys, D. B. (2016). (Citi)Bike Sharing. *Operations Research*.

Hou, H., Wang, Q., Xiao, Z., et al. (2022). Data-driven economic dispatch for islanded micro-grid considering uncertainty and demand response. *International Journal of Electrical Power & Energy Systems*.

Kim, S., Pasupathy, R., & Henderson, S. G. (2014). A Guide to Sample-Average Approximation. *Handbook of Simulation Optimization*, 207-243.

Lampa, M., & Samolejová, A. (2020). Fleet Optimization Based on the Monte Carlo Algorithm. *7*(1), 17-21.

Li, M., Zhang, C., Ding, M., & Lv, R. (2022). A two-stage stochastic variational inequality model for storage and dynamic distribution of medical supplies in epidemic management. *Applied Mathematical Modelling*, 35-61.

Lin, D., & Kuo, J. (2021). The vehicle deployment and relocation problem for electric vehicle sharing systems considering demand and parking space stochasticity. *Transportation Research Part E: Logistics and Transportation Review, 156*.

Lu, M., Chen, Z., & Shen, S. (2017). Optimizing the Profitability and Quality of Service in Car-share Systems Under Demand Uncertainty. *Manufacturing & Service Operations Management*(https://doi.org/10.1287/msom.2017.0644).

Ma, Z., & Koutsopoulos, H. N. (2022). Near-on-demand mobility. The benefits of user flexibility for ride-pooling services. *Transportation Research Part C: Emerging Technologies*.

Mitchell, S., O'Sullivan, M., & Dunning, I. (2011). PuLP: A Linear Programming Toolkit for Python.

Oluyede, L., Cochran, A. L., Prunkl, L., Wang, J., Wolfe, M., & Mcdonald, N. C. (2022). Unpacking transportation barriers and facilitators to accessing health care: Interviews with care coordinators. *Transportation Research Interdisciplinary Perspectives*.

Pan, A., Martin, E., & Shasheen, S. (2022). Is access enough? A spatial and demographic analysis of one-way car-sharing policies and practice. *Transport Policy, 127*

Rotaris, L., & Danielis, R. (2018). The role for car-sharing in medium to small-sized towns and in less-densely populated rural areas. *Transportation Research Part A: Policy and Practice, 115*(https://doi.org/10.1016/j.tra.2017.07.006), 49-62.

Rozenfeld, Y., Beam, J., Maier, H., Haggerson, W., Boudreau, K., Carison, J., & Medows, R. (2020). A model of disparities: risk factors associated with COVID-19 infection. *Internaltional Journal for Equity in Health*.

Shaheen, S., & Cohen, A. (2020). Innovative Mobility: Car-sharing Outlook. *Transportation Sustainability Research Center - University of California, Berkeley*.

Shapiro, A., Dentcheva, D., & Ruszczynski, A. (2014). Lectures on Stochastic Programming: Modeling and Theory, Second Edition.

Shu, J., Chou, M. C., Liu, Q., Teo, C.-P., & Wang, I.-L. (2013). Models for Effective Deployment and Redistribution of Bicycles Within Public Bicycle-Sharing Systems. *Operations Reseach* (http://dx.doi.org/10.1287/opre.2013.1215).

Takes, F. W., & Kosters, W. A. (2010). Applying Monte Carlo Techniques to the Capacitated Vehicle Routing Problem.

Tan, B., Chen, H., Zheng, X., & Huang, J. (2022). Two-stage robust optimization dispatch for multiple microgrids with electric vehicle loads based on a novel data-driven uncertainty set. *International Journal of Electrical Power & Energy Systems*.

Vermesch, P., Boisjoly, G., & Lachapelle, U. (2021). Commuting mode share and workplace-based public transport services: An equity perspective. Case Studies on Transport Policy, 9(2).

Wang, Y., & Szeto, W. (2021). The dynamic bike repositioning problem with battery electric vehicles and multiple charging technologies. Transportation Research Part C:Emerging Technologies.

Wu, M., Yu, C., Ma, W., et al. (2022). Joint optimization of timetabling, vehicle scheduling, and ride-matching in a flexible multi-type shuttle bus system. Transportation Research Part C: Emerging Technologies, 139.

Yang, S., et al. (2021). Double-balanced relocation optimization of one-way car-sharing system with real-time requests. Transportation Research Part C: Emerging Technologies.

Yu, M., & Shen, S. (2018). Designing and Optimizing an Integrated Car-and-ride Sharing System for Mobilizing Underserved Populations.

Zaneti, L., Arias, N., Almeida, M., & Rider, M. (2022). Sustainable charging schedule of electric buses in a University Campus: A rolling horizon approach. Renewable and Sustainable Energy Reviews, 161.