NATIONAL INSTITUTE FOR CONGESTION REDUCTION

FINAL REPORT
DECEMBER 2021

Enhancing Equity and Access in Dockless Micromobility Services:

Case Study, Performance Metrics and Rebalancing Model

Daniel Rodríguez-Román, PhD
Alberto M. Figueroa-Medina, PhD, PE
Benjamín Colucci-Ríos, PhD, PE, JD
Carlos del Valle González, PhD

For:
National Institute for Congestion Reduction
University of South Florida
Center for Urban Transportation Research | University of South Florida

4202 E. Fowler Avenue, ENG030, Tampa, FL 33620-5375
nicr@usf.edu

NICR
NATIONAL INSTITUTE FOR CONGESTION REDUCTION

NICR NATIONAL INSTITUTE FOR CONGESTION REDUCTION
4202 E. Fowler Avenue, ENG 030
Tampa, FL 33620-5375
www.nicr.usf.edu

DISCLAIMER

The contents of this report reflect the views of the authors, who are responsible for the facts and the accuracy of the information presented herein. This document is disseminated in the interest of information exchange. The report is funded, partially or entirely, by a grant from the U.S. Department of Transportation's University Transportation Centers Program. However, the U.S. Government assumes no liability for the contents or use thereof.

Technical Report Documentation Page

1. Report No.	2. Government Accession No.	3. Recipient's Catalog No.

4. Title and Subtitle	5. Report Date
Enhancing Equity and Access in Dockless Micromobility Services: Case Study, Performance Metrics and Rebalancing Model	September 30, 2021
	6. Performing Organization Code
7. Author(s)	8. Performing Organization Report No.
Daniel Rodríguez-Román, Alberto M. Figueroa-Medina, Benjamín Colucci-Ríos Carlos del Valle-González, Lina Villa Zapata, Gabriela Yáñez González, Andrés G. Camacho Bonet, Charmelis Reyes Cruz, Juan E. Flórez-Coronel, Juan M. González-López, Joshua R. Santiago-Ibarra	
9. Performing Organization Name and Address	10. Work Unit No. (TRAIS)
University of Puerto Rico at Mayagüez Department of Civil Engineering and Surveying Call Box 9000 Mayagüez, PR 00681-9000	11. Contract or Grant No. 69A3551947136
12. Sponsoring Organization Name and Address	13. Type of Report and Period Covered
U.S. Department of Transportation University Transportation Centers 1200 New Jersey Avenue, SE Washington, DC 20590 United States	Final Research Report (January 2020 – September 2021)
	14. Sponsoring Agency Code

15. Supplementary Notes

16. Abstract

A case study of the Mayagüez dockless e-scooter service (MDES) is presented. The objectives of the case study were to i) explore the characteristics of MDES users and nonusers, ii) examine the relationship between the sociodemographic characteristic of the study area and observed trip levels, iii) propose spatial access indicators and apply them in the MDES service area, and iv) discuss the equity in spatial access and the congestion reduction potential of MDES. This analysis provides insights on a micromobility system that operates in a congested, auto-centric city with limited public transportation alternatives. Among other things, the study found that the main users of the system were university students and more likely to be young and male, and that MDES potentially reduced auto trips. Also, a spatial equity analysis suggests differences in access to MDES vehicles.

In addition, this project proposes an optimization-based approach to generate dockless micromobility vehicle rebalancing plans that reflect equity concerns. In the proposed method, target vehicle distributions are identified such that the predicted efficiency and equity performance of the system is maximized. Then a pickup and delivery problem is applied that balance efficiency and equity objectives when searching for the optimal spatial distribution of the vehicle fleet. The application of this two-step methodology is illustrated through numerical experiments.

17. Key Words	18. Distribution Statement
Micromobility, Scooter, Equity, Rebalancing	

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

ACKNOWLEDGEMENTS

The research team would like to thank Skootel for providing data from their micromobility services and answering questions about their system. The researchers also thank Mr. Fernando A. Acosta Pérez for his collaboration in the cluster analysis performed in this project.

Table of Contents

NICR NATIONAL INSTITUTE FOR CONGESTION REDUCTION
4202 E. Fowler Avenue, ENG 030
Tampa, FL 33620-5375
www.nicr.usf.edu

Tables

Figures

Acronyms

ABM	Activity-Based Model
AUC	Area Under the (cumulative relative frequency) Curve
DBS	Dockless Bikesharing Service
DES	Dockless E-scooter Service
GIS	Geographic Information System
GPS	Global Positioning System
MDES	Mayagüez Dockless E-scooter Service
ML	Machine Learning
PDP	Pickup and Delivery Problem
SBS	Station-based Bikesharing Service
TDP	Target Distribution Problem
UPRM	University of Puerto Rico at Mayagüez

NICR
NATIONAL INSTITUTE FOR CONGESTION REDUCTION
4202 E. Fowler Avenue, ENG 030
Tampa, FL 33620-5375
www.nicr.usf.edu

Executive Summary

Shared micromobility services provide clean mode choice alternatives (e.g., shared bicycles, scooters) that could potentially help cities shift towards auto independence. There are, however, concerns that the benefits of these services are inequitably distributed among different population groups. To address these concerns, studies have attempted to understand the travel demand patterns associated with micromobility services, and researchers have proposed models for the planning and operations of micromobility systems based on equity considerations. This research report contributes to this literature by presenting a case study of a dockless e-scooter system in Mayagüez, Puerto Rico, and by proposing an optimization-based framework to guide the vehicle rebalancing operations of dockless micromobility services according to efficiency and equity objectives.

The case study of the Mayagüez dockless e-scooter service (MDES) uses data collected through an online survey, data provided by the micromobility operator, and data retrieved from US Census databases. The objectives of the case study were to i) explore the characteristics of MDES users and nonusers, ii) examine the relationship between the sociodemographic characteristic of the study area and observed trip generation levels, iii) propose spatial access indicators and apply them in the MDES service area, and iv) discuss the equity in spatial access and the congestion reduction potential of MDES. This analysis provides insights on a micromobility system that operates in a congested, auto-centric city with limited public transportation alternatives and that serves a population facing downward economic and demographic trends. The main findings of the case study are:

- Service users tended to be young and male. In an online survey, 38% of women and 48% of males indicated that they were service users. An analysis of the survey responses suggests that people in the 18-to-26 age group are at least two times more likely to be e-scooter users than people in other age groups.
- Trips were concentrated in and around a university's campus. Approximately 78% of e-scooter trips started or ended at UPRM. Outside UPRM, trips primarily concentrated in neighborhoods with large student populations.
- Costs, inadequate pavement conditions, lack of dedicated space for on-street riding, and safety concerns were identified by survey respondents as the main reasons for not using the service.
- Traffic congestion and lack of parking spaces were identified among the main reasons for using MDES, which suggest that the service could have reduced auto trips in the service area. The magnitude of the auto trip reduction needs to be estimated for the case of Mayagüez, but the results point to the potential capacity of micromobility services to replace short auto trips, particularly in cities that lack effective public transportation services.
- There were differences in spatial access among neighborhoods. This observation could be explained by the spatial distribution of the main system users (university students), the sociodemographic characteristics of the population in Mayagüez (generally low-income and older), and the rebalancing operations being aimed at satisfying user demand.
- Regression analyses suggest that income levels, mixed land use, and university locations are positively associated with e-scooter trip demand.

Inequality in spatial access to dockless micromobility services can be mitigated by implementing equity-conscious vehicle relocation strategies. To this end, the second main contribution of this project is a methodology that integrates two optimization models to generate vehicle rebalancing plans that reflect equity concerns. In the first model, target vehicle distributions are identified such that the predicted efficiency and equity performance of the system is maximized. Two objective function formulations were

proposed to account for spatial and social equity considerations. The second model is a pickup and delivery problem that balances efficiency and equity objectives when searching for the optimal spatial distribution of the vehicle fleet. The application of this two-step methodology is illustrated through numerical experiments. Simulation results suggest that, relative to efficiency-focused rebalancing, there are scenarios in which equity-focused rebalancing operations could result in minor reductions in total trips and significant improvements in spatial access to micromobility services. As the simulations were based on the MDES characteristics and on assumed behavioral parameters, additional studies are needed to reach generalizable conclusions on the likely effects of equity-focused rebalancing operations on the overall performance of dockless micromobility services.

NICR
NATIONAL INSTITUTE FOR CONGESTION REDUCTION
4202 E. Fowler Avenue, ENG 030
Tampa, FL 33620-5375
www.nicr.usf.edu

1. Introduction

Shared micromobility services could help communities reduce their dependence to private vehicles, encourage physical activity, drive technological innovations, and foster economic growth, among other potential benefits. Naturally, there are also potential costs and concerns associated with shared micromobility. Chief among these concerns is the possibility that disadvantaged communities will not have equitable access to emerging travel alternatives. In response to this concern, governments have enacted policies intended to ensure an equitable allocation of micromobility resources. For example, in Chicago's e-scooter pilot program, companies were required to distribute half of their e-scooter fleets within underserved community areas, both in their morning fleet distributions and subsequent rebalancing operations (City of Chicago, 2020). The mixed results from the Chicago program suggest that much remains to be studied regarding the policies and methods that have been developed to ensure equitable access to micromobility. The research community has also engaged with these equity issues, in part by studying the response of different communities to micromobility services, characterizing micromobility-related travel patterns, and proposing methods to plan and operate these systems efficiently and equitably. This research report contributes to these efforts.

The objective of this report is twofold. *The first main objective is to present a case study of the dockless e-scooter rental service in Mayagüez, Puerto Rico (PR)*. Besides documenting the micromobility experience of a city within the understudied Latin American context, this is also a case study of a system that operates in an auto-oriented urban environment with limited public transportation options, and that serves communities facing a depressed economic outlook, downward demographic trends, and the aftermath of recent hurricanes and earthquakes. The primary research questions related to this first objective are:

i. What were the user characteristics in the Mayagüez dockless e-scooter service (MDES)?
ii. What sociodemographic and land-use factors explain the demand levels for MDES?
iii. What were the spatiotemporal patterns of e-scooters trips in MDES?
iv. Did MDES have an impact on traffic congestion?
v. Was there equitable spatial access to MDES?

To help answer the last question, methods for quantifying spatial access in the context of dockless micromobility systems were developed and applied using MDES data. The case study contributes to research on travel behavior and micromobility usage patterns.

The second main objective of this report is to present an optimization-based decision framework that considers efficiency and equity goals in the vehicle rebalancing operations of dockless micromobility services. The proposed approach uses models to predict future states of the micromobility system. Based on these states rebalancing decisions are made so that the vehicle distribution reflects a target (ideally optimal) distribution, as defined by equity and efficiency performance objectives. Approaches to quantifying accessibility and equity from a spatial and social perspective are considered in the model formulation. The spatial equity measure proposed considers distance-based access to the vehicle fleet at the level of building units, while the social equity measure proposed accounts for both the distribution of vehicles and the characteristics of individuals. Efficiency is considered from the perspectives of a firm interested in maximizing the number of trips and of a planning agency interested in achieving objectives related to the transportation system performance, such as reducing traffic congestion.

NICR NATIONAL INSTITUTE FOR CONGESTION REDUCTION
4202 E. Fowler Avenue, ENG 030
Tampa, FL 33620-5375
www.nicr.usf.edu

Besides this introduction, this report is composed of three additional chapters.

- *Chapter 2: User Characteristics, Travel Patterns, and Spatial Access in MDES.* The MDES case study is presented in this chapter. Also, network-based measures used to quantify spatial access to dockless micromobility services are proposed and applied to the MDES case.
- *Chapter 3: Vehicle Rebalancing to Improve Equity in Access to Dockless Micromobility Systems.* In this chapter, the rebalancing models are discussed. The results from numerical tests performed with the proposed methods are presented.
- *Chapter 4. Conclusions and Future Research Directions*

Methods used to adjust trip record data, train machine learning models, and solutions to the proposed optimization models are discussed in the report's appendices. In this report, the term "shared micromobility" refers to the shared use of bicycles, scooters, or other low-speed vehicles that can be reserved or rented for relatively short periods (Shaheen and Cohen, 2019). Although studies that examine station-based micromobility services will be reviewed, a dockless e-scooter service is the focus of the case study and the rebalancing method proposed applies primarily to dockless micromobility services, as previously mentioned.

2. User Characteristics, Travel Patterns, and Spatial Access in MDES

The study of micromobility user characteristics and spatiotemporal travel patterns can have multiple objectives, including advancing basic knowledge of human travel behavior; supporting the development of travel demand models that can be used by communities and their agencies to plan, design, or evaluate transportation projects or policies; and guiding the investment and operational decisions of transportation service firms. This project adds to the growing body of literature that examines the experiences of communities with emerging transportation mode alternatives, and it presents methods to assess spatial access to these new systems. The main objectives of the research presented in this chapter are to:

i. *Explore the characteristics of MDES users and nonusers.* An online survey was conducted in which respondents were asked about their opinions and experiences with MDES. The survey responses were analyzed using exploratory data analysis methods and logistic regression. Additionally, the travel behavior of MDES users was examined using trip data provided by the system operator. This data was analyzed using clustering methods.
ii. *Examine the relationship between the sociodemographic characteristic of the study area and observed trip generation levels in MDES.* A database was created that combined US Census sociodemographic data and land use information of the study region with the trip arrival and departure counts at the level of cells in a grid-based zonal system. Regression analysis methods were used to find relationships between sociodemographic and land-use variables and the trip arrival and departure counts.
iii. *Evaluate the spatial access to MDES of the neighborhoods in the study area.* Network-based methods for computing spatial access indicators are proposed and applied using the MDES trip data.
iv. *Discuss the equity in spatial access and the congestion reduction potential of MDES.*

The sections that follow discuss in detail the data and methods used to accomplish the research objectives. The rest of the chapter is divided into seven sections. In the next section, literature related to the equity and spatiotemporal analyses in the context of docked and dockless micromobility is reviewed, as well as previous studies that examine the definitions of three key concepts: equity, justice, and accessibility. The second section gives background information on the study area and MDES. This is followed by a description of the data sources used in the analysis (third section) and the methods used in the study (fourth section). Results are presented in the fifth section. In the last two sections, a general discussion of the results and closing remarks are offered.

2.1. Literature Review on Travel Patterns in Micromobility Services and Equity

2.1.1. Definitions of Equity and Accessibility in Transportation

Research on micromobility services has explored their operational characteristics, the characteristics of micromobility users, and the related travel activity patterns, as well as the equity and policy questions that arise with the arrival of new transportation alternatives in the public space. This equity research builds on previous work that examines the fairness in the distribution of benefits and costs of transportation infrastructure and services. Karner et al. (2020) note that transportation equity research and practice is generally concerned with identifying who benefits and who is burdened by a transportation project, which in turn requires the selection of measures relative to which equity can be assessed (i.e., equity "of what"). They also highlight the distinction between transportation justice and transportation equity, the former being primarily concerned with the transformation of social structures that give rise to inequality, while the latter is primarily concerned with "processes and distribution of social goods and opportunities". Pereira et al. (2017) explain that concepts of equity draw from theories of justice and that the academic literature often uses the terms equity and justice interchangeably, as they did in their work. Relevant theories of justice include utilitarianism, libertarianism, intuitionism, Rawlsian egalistarism, and the capabilities approach. Pereira et al. argue that, regarding distributive justice concerns, the focus on transportation-related analyses should be on accessibility as a human capability.

Previous research focused on the distributive outcomes associated with transportation systems that have explored how benefits and costs are distributed among individuals of different groups (vertical equity) and individuals of the same group (horizontal equity) (Litman, 2005). The concept of spatial equity is also relevant to micromobility studies, which accounts for the spatial distribution of impacts of a project, service, or policy (Levinson, 2010). In the context of public transportation options, fairness is often evaluated in terms of accessibility, which can be understood as the potential that individuals have for reaching destinations or opportunities via a transportation alternative (Guo et al., 2020). As Páez et al. (2012) discuss, there are multiple ways of measuring accessibility, which could be distinguished, for example, based on their normative (e.g., how far it is reasonable to travel) and positive aspects (e.g., how far people actually travel).

In this project, equity is discussed in terms of accessibility. Specifically, in the context of the dockless micromobility system considered, accessibility is defined relative to the ease with which a person can access a micromobility vehicle. It is posited that a person has access to the system if they have opportunities to use the system, and an opportunity arises if there are vehicles close to the person and if the person has the capacity to use a vehicle. This perspective will be referred to as equity in access. In this chapter, measures are applied and proposed to address the spatial access component of equity. In Chapter 3, models are proposed to measure equity in terms of the capacity of individuals to use the system.

2.1.2. Equity and Spatiotemporal Analysis in the Context of Docked Micromobility

Considerable work has been completed to understand station-based (or docked) bike-sharing systems (SBS), one of the earliest forms of micromobility (Fishman, 2016). In addition to exploring operator characteristics (Shaheen et al., 2013), researchers have examined the sociodemographic of registered bike-share users and of the cities in which these systems operate. They have found differences in the usage patterns and access of people of different genders, age groups, races, educational attainment, and income levels. For example, in a London SBS, Beecham and Wood (2014) found that males performed more commuting trips, whereas females performed more leisure-oriented trips. Wang and Akar (2019) observed in a New York City BSS that, on average, males completed close to 1,500 trips per hour versus approximately 500 trip completions per hour by females; similar gender gaps have been observed in other systems, and factors that explain these differences have been explored (Shaheen et al., 2013). Using data from an online survey, Bachand-Marleau et al. (2012) determined that the primary reason people used a BSS in Montreal, Canada was that it was useful for one-way trips. The researchers also estimated a logistic regression model that suggested that being in proximity to a station and being a bus user were positively related to BSS usage, while females and recreation-only cyclists were less likely to use the system.

Besides observing a gender gap, Wang and Lindsey (2019) found that age is inversely correlated with the use of an SBS in the Minneapolis-St. Paul area, and that 30-day and annual SBS members living in neighborhoods with higher concentrations of minorities and classified as having lower socioeconomic status perform more trips than those in wealthier, predominately white neighborhoods. However, they note that the data in the study does not link demographic information to trips. In an analysis of 29 BSSs in the US, Barajas (2018) found that higher proportions of White residents are observed in block groups within a 5-minute walk of a bike-sharing station. Barajas also concluded that the systems served higher-paying skilled jobs. Using data from Chicago, Qian and Jaller (2021) analyzed the destination choice behavior of population groups classified as disadvantaged, and they concluded that these groups make longer SBS trips to access opportunities, such as grocery stores and places of work and that they are more sensitive to the costs structures. Generally speaking, earlier studies also suggest an overrepresentation of white, higher income, and college-educated individuals among SBS users (Shaheen, 2012; Smith et al., 2015).

However, studies also suggest that, relative to the general cyclist population, young females from lower household incomes are more likely to be users of SBS (Buck et al., 2013), as well as that an extension of SBS to economically disadvantage neighborhoods increases the participation of residents in those communities (Goodman and Cheshire, 2014). Ursaki and Aultman-hall (2015) suggested several strategies to improve equity in micromobility programs, including allocating sufficient vehicles to low-income neighborhoods, providing discounted rates for people who need them, and installing bicycle infrastructure in diverse communities.

The study by Chen et al. (2019) is noteworthy because it proposes a methodology to assess SBS accessibility and equity at the personal level, as opposed to the more common zone-based analyses found in the literature. In Chapter 3, the subject of person-level measures of accessibility is revisited in the context of dockless micromobility systems.

2.1.3. Equity and Spatiotemporal Analysis in the Context of Dockless Micromobility

Dockless bike-sharing services (DBS), which have rapidly expanded worldwide in recent years, appear to improve user experience relative to the docked alternative (Chen et al., 2020). Shen et al. (2018) concluded that, in Singapore, DBS trips are positively associated with high-density commercial areas, diverse economic activity, and supportive cycling infrastructure, among other factors. According to Gu et al. (2019), in the context of China, DBS users tend to be young and college-educated, and the gender gap is nearly nonexistent. In Seattle's DBS program, no neighborhood was consistently excluded from access to the system, although the spatial equity analysis of Mooney et al. (2019) revealed that neighborhoods with a higher proportion of college-educated, high-income residents, and more community resources had more bikes. The researchers quantified neighborhood access to the DBS by computing the average number of bikes available per resident per day; bike idle times were also measured. Hirsch et al. (2019) report that users of Seattle's DBS program were disproportionally young, White, male, and bicycle owners, and they lived closed to the city center. In a study of Boston's DBS, Gehrke et al. (2021) found evidence of insufficient spatial access to dockless bikes in neighborhoods with a higher share of renter-occupied housing and Black residents. They also found that neighborhoods with a high share of carless households were positively associated with dockless bike-share access and trip generation, thus suggesting that micromobility services can help foster residential and transportation choices that do not depend on private autos. In Rethymno, Greece, Bakogiannis et al. (2019) conducted a survey that revealed that the principal factors that limit the use of the city's DBS were the lack of adequate cycling infrastructure and traffic safety concerns.

Studies that have compared Washington DC's SBS and dockless e-scooters systems (DES) have found differences in the temporal and spatial distribution trips (McKenzie, 2019; Younes et al., 2020). Using hierarchical clustering analysis and data from a German DES, Degele et al. (2018) identified four types of users, which were characterized by their level of participation in the system, age, and average travel distance, among other factors. Using data from Singapore, Zhu et al. (2020) compared a DBS with a station-based e-scooters service; the researchers found the e-scooters had a better performance in terms of utilization, as well as a more compact spatial distribution. Also, in a comparison of docked and dockless services, Qian et al. (2020) concluded that the dockless system provided greater access and attracted more trips in communities of concern in San Francisco.

In a comparison of e-scooters usage patterns in Austin, TX, and Minneapolis, MN, Bai and Jiao (2020) determined that in both cities e-scooters were predominantly used in the downtown areas and university campuses, but there were marked differences in the temporal distribution of trips. In Indianapolis, IN, e-scooter trips were also observed to be concentrated in downtown and campus areas, and the temporal distribution of e-scooter trips did not follow the standard morning/afternoon peak-hour patterns observed for other modes (Mathew et al., 2019). The temporal difference was also observed by Zou et al. (2020), who used e-scooter trajectory data from Washington DC's DES to explore the travel paths of e-scooter trips. Among other things, they found that arterials and local streets with heavy traffic are associated with a higher share of e-scooter use. Caspi et al. (2020) used spatial regression analysis to examine the influence of the built environment, land use, and demographics on e-scooter trip generation; they found that high rates of e-scooter usage were associated with zones with a high density of students.

2.2. Description of Study Area and E-Scooters System

This project provides a case study and analysis of a DES that operates in the Municipality of Mayagüez, located in the western region of Puerto Rico, a territory of the United States of America. According to the US Census, the municipality has lost more than 10,000 residents in the last decade; its current population estimate is approximately 77,000. The median age of residents is 40, 52% of the population is female, 25% has a bachelor's degree or more, 53% live below the federal poverty line, the median household income is $14,120, and the unemployment rate is 27%. Around 93% of the population reports driving alone or as part of a carpool as their mode of transportation, with only 1.3% using public transportation (Census, 2019.). These statistics reflect the economic and demographic reality of most municipalities in PR outside of the San Juan Metropolitan Area. Mayagüez, however, is a college town, which results in unique economic dynamics not observed in neighboring municipalities. Its largest academic institution is the University of Puerto Rico at Mayagüez (UPRM), located on the northern outskirts of the city center (the Pueblo borough). UPRM has a college population of approximately 13,000 students and 2,500 employees. Among students, 40% are estimated to live in the Pueblo borough, with the UPRM-adjacent communities of Mayagüez Terrace, Bosques, and Trastalleres being the most densely student-populated areas. Non-motorized modes of transportation are used by 48% of students that reside in the Pueblo borough (Arroyo, 2020). Lastly, students are generally observed to return to their hometowns during the weekends and holidays.

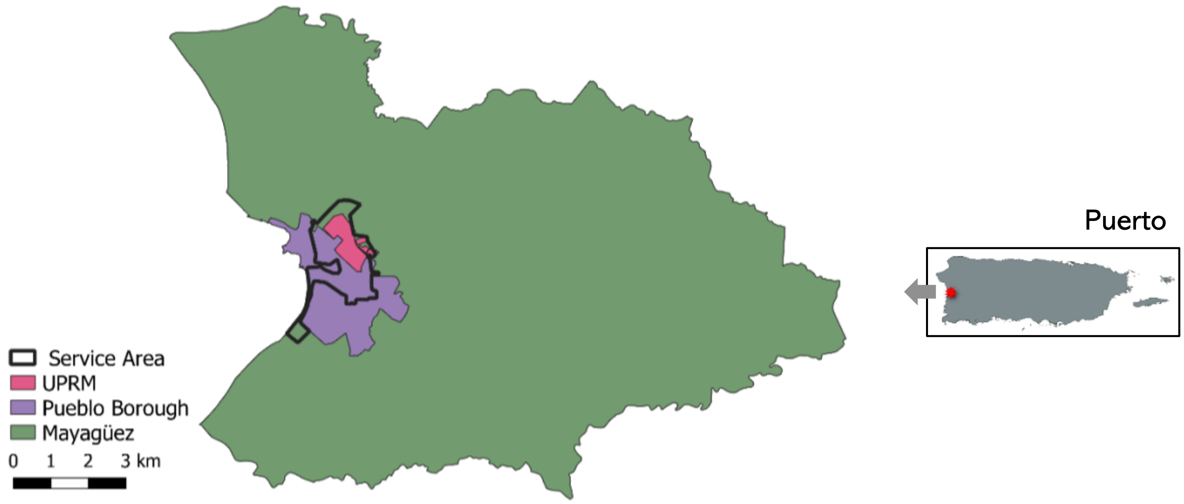

Figure 1. MDES service area with respect to UPRM and the Pueblo borough, and Mayagüez's location within PR

On August 3, 2019, a dockless e-scooter rental service (MDES) began operating in Mayagüez within a service area that included UPRM, its adjacent neighborhoods, and the center of the Pueblo borough (Figure 1). MDES is the first micromobility service to operate in PR. The service area has been periodically expanded since its inception. The boundary depicted in Figure 1 represents the 3.5 km^2 area in which there was service during the 2019-2020 academic year. Skootel, a local micromobility company, owns and operates the system. The price of a trip is $1 to activate the e-scooter plus 20 cents per travel minute. The hours of operation of the e-scooter service are from 6:00 AM to 8:00 PM; in the morning, the operator located e-scooters within the service area, and they were collected after 8:00 PM. Rebalancing operations were conducted during the day. The number of units in daily operation was approximately 90 e-scooters during weekdays and 30 e-scooters during weekends. The service paused its operations due to the COVID-19 pandemic.

2.3. Description of Data Sources

This section describes the three data sources used in the study: the e-scooters trip data, the region's demographic and network data, and data obtained from an online survey. Additionally, a method used to disaggregate the available sociodemographic data is discussed.

2.3.1. E-scooters Data and its Processing

The operator provided records on approximately 66,000 e-scooter trips completed during the 2019-2020 academic year. A trip record consisted of the trip date, the trip starting and ending times, the trip starting and ending coordinates, the price paid, and unique identifiers for the user and the e-scooter. The main source of error in the data was the trip end location, as it appears that the system was registering intermediate GPS coordinates as the ending coordinate of trips, meaning that subsequent locations trip ends and starts of the same e-scooter would not be reasonably close even considering normal GPS error. Approximately 51% of the ending and starting coordinates of successive trips were 150 meters or more apart. To address this problem and standard GPS errors, an algorithm was created to match trip starts and ends when possible, or it would preserve the coordinate difference, and label it as the result of a rebalancing (see Appendix A for details on the algorithm). Less than 1% of trip records were not considered in the study because the ending coordinate was missing. The trip records do not contain the users' sociodemographic characteristics.

2.3.2. Zonal Systems and Sociodemographic Data

Figure **2** presents the neighborhood zones identified for the spatiotemporal analysis of trip patterns (hereafter, the neighborhood zonal system). These zones aggregate sub-borough zones within the Pueblo borough. Zone 8's southwest "tail" is a scenic path along a coastline park leading to various sports facilities, recreational areas, and public housing facilities; a negligible number of records are associated with this section. In Figure 3, the grid-based zonal system used for the regression analysis is presented (grid-based zonal systems are commonly used in regression analyses, especially if it is reasonable to expect spatial dependencies). The cells in the grid have dimensions of 100 meters by 100 meters. The relatively small size of the cells responds to the relatively small size of the MDES service area.

Zone ID	Name
1	UPRM
2	Bosque
3	Terrace
4	Trastalleres
5	Paris
6	Norte
7	Centro
8	Playa
9	Plaza

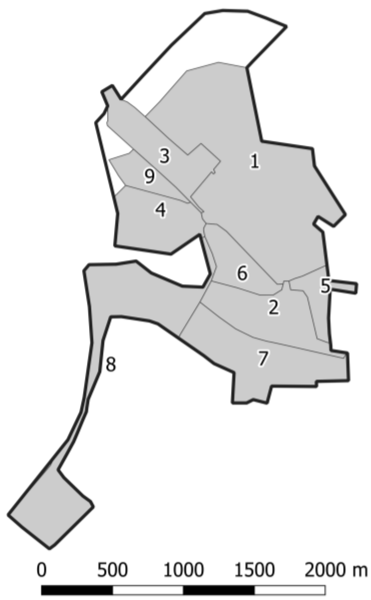

Figure 2. Neighborhood zonal system within the MDES service boundary

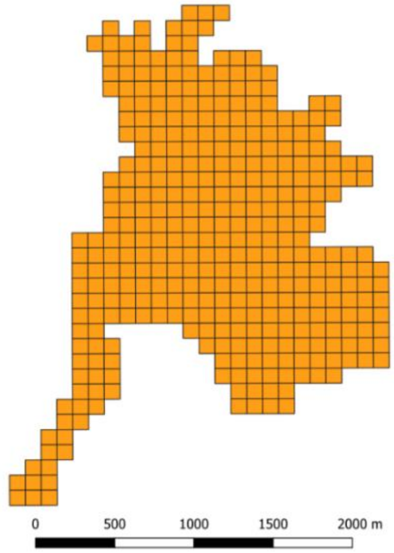

Figure 3. Grid zonal system for regression analysis

Sociodemographic information was collected from the American Community Survey (ACS) 2014-2018 five-year estimates for the Census block groups in Mayagüez (Census, 2019). This is the most disaggregated data available, except for the 2010 Census block information of the region. The ACS information was used as its more likely to reflect the dramatic demographic shifts observed in PR due to the island's economic crisis, the effects of the 2017 hurricanes, the effects of the 2020 earthquakes, and the related migration to mainland US. ACS estimates of population, median age of the population, employment and unemployment levels, college population, and income per capita were collected. Besides the standard uncertainty associated with ACS estimates, there are at least two basic sources of error that likely affect the sociodemographic data used in this study. First, there is a substantial floating student population whose size is not insignificant relative to the Pueblo borough population; there are only preliminary high-level estimates of the spatial distribution of the UPRM student population (Arroyo, 2020). Second, the MDES service area does not match standard zonal classifications and there are only 22 Census block groups in the region of interest, which hinders the estimation of the characteristics of people that reside within or close to the area.

The data aggregation problem was addressed through a data disaggregation procedure. GIS tools were used to represent the footprint of every building in the service area and within a 187.5-meter buffer beyond the service boundary (187.5 meters is approximately the distance covered by a pedestrian walking at 1.25 m/s for 2.5 minutes (Schimpl et al., 2011)). Using Google Earth's Street View functionality and satellite images, each building was assigned a land-use classification (e.g., residential, commercial) and a floor count. This task was completed for 5,058 structures. Finally, the population of each Census block group was disaggregated proportionally to each residential area polygon within its boundary. Based on this procedure, it was determined that 10,280 residents live within the service area. The population of each zone in the neighborhood zonal system and each cell in the grid system was computed by aggregating the residential polygon population within their respective boundaries.

Additional GIS analyses were performed to compute cell-level values for median age, employment levels, unemployment levels, college population, and income per capita based on the US Census data. The sociodemographic attributes assigned to each cell were the same as those of the Census block in which it was contained unless the cell intersected more than one Census block. In such cases, each sociodemographic attribute (e.g., median population age) was computed by taking the weighted average of the Census block attribute value, with the weight of each block value being the block's population within the cell as indicated by the residential polygons. In Table 1, descriptive statistics are presented for the selected variables. For comparison, the median age, unemployment rate and average income in Puerto Rico are 43 years, 14.1%, $20,539, respectively (Census, 2019).

Table 1. Descriptive statistics for the cell-level socioeconomic and land use variables

Variable	Average	St.Dev.	Min	Max
Population (count)	52.7	45.8	0.1	231
Age (years)	29.1	9.3	21.4	58
Employment (job/km^2)	0.5	0.3	0.1	1.4
Unemployment (%)	9.0	6.1	1.8	24.8
Income (US$)	6,366	3,617	357	13,343
College Students (count)	21.2	32.2	0.0	150
Commercial Land Use (1000×km^2)	0.9	2.0	0.0	14.2
Mixed Land Use (1000×km^2)	0.6	1.6	0.0	13.7
UPRM Land Use (1000×km^2)	0.9	2.7	0.0	17.4
Industrial Land Use (1000×km^2)	0.1	1.3	0.0	22.0
Public Land Use (1000×km^2)	0.3	1.2	0.0	9.3

Notes: The statistics for the socioeconomic variables were computed without taking into consideration cells in which there was no population (204 cells had population). For the statistics of the land use variables, the variable values for 350 cells were used. St.Dev. stands for "Standard Deviation".

2.3.3. Transportation Networks

Two networks were created, namely, a pedestrian network and an e-scooter network. Both networks were developed using the US Census Topologically Integrated Geographic Encoding and Referencing (TIGER)/Line road shapefiles for the region. The e-scooter network consisted of the road links plus pedestrian paths used by e-scooter users inside the UPRM campus (Figure 4). The pedestrian network incorporated the link additions made to the e-scooter network, as well as pedestrian bridges and paths not accessible to e-scooters. In addition to the nodes associated with the travel links, centroid nodes representing each building in the service area were included in the network and linked via centroid connectors (connectors linked to the closest network node). To ensure proximity between building centroids and their connecting node, links were subdivided (i.e., intermediate nodes added) so that all links had a length of 30 meters or less.

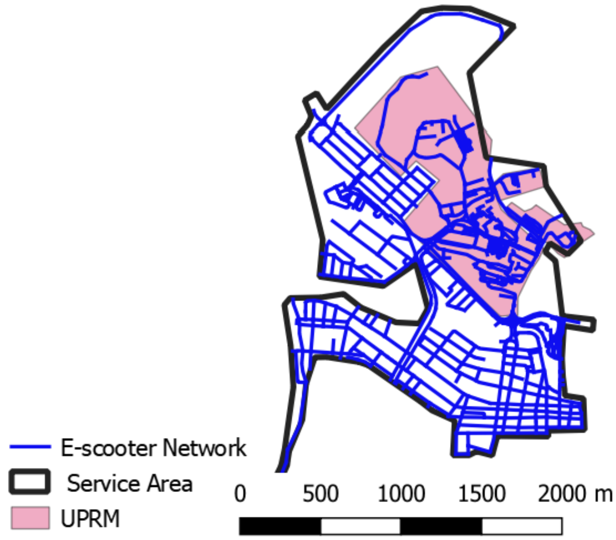

Figure 4. Main section of e-scooter network

2.3.4. Online Survey

A survey about the MDES was distributed via email and social media from March to May 2020. The survey targeted the UPRM community and Mayagüez residents. In addition to a section with sociodemographic questions, the survey consisted of four sections that explored the extent to which the responder used the system and the reasons for their usage levels, among other factors. A total of 417 responses were received, of which 11% were removed from the analysis given inconsistent responses. Among the accepted responses, 56% of responders indicated that they had never used or had used once the service, 48% identified as female, 50% identified as male, 79% indicated that they were UPRM students, and 77% indicated that they were 26 years old or younger. The online survey helped to explore the sociodemographic characteristics of MDES users and nonusers.

2.4. Methods

This section discusses the methods used to:

i. explore the characteristics of MDES users,
ii. explore the spatiotemporal patterns of trips in MDES,
iii. quantify the relationship between observed e-scooter trip levels and sociodemographic characteristics of the region, and
iv. quantify the spatial access of locations in the MDES service area.

2.4.1. Methods to Explore the MDES User Characteristics and Usage Patterns at the Person-Level

A binary logistic regression model was estimated using the survey data to explore the relationship between a person's sociodemographic characteristics and the likelihood of using MDES. The model's dependent variable had a value of 1 if the person indicated that they had never used the system or had used it only once, and 0 otherwise. The considered independent variables include binary (dummy) variables for age group (18-20, 21-26, 27-49, 50 and over), gender (1 if female, 0 otherwise), UPRM relationship (student, employee, neither), location of residence (e.g., Terrace, outside Mayagüez), and the availability of travel modes to the person (bicycle, auto, transit).

Hierarchical clustering was used to group users according to their trip behavior, as recorded in the MDES data. For this cluster analysis, each user was characterized according to the following statistics: the number of trips performed, the average time between trips (in hours), the standard deviation of the time between trips, presence in the system (in days, the time difference between the first and last recorded trip), and the probability of performing at least one more trip in a day given that the user had already performed a trip in that particular day. Each variable was standardized to ensure consistent features. The clustering was performed by applying Ward's method (Ward, 1963). Data from August 3, 2019, to December 15, 2019, was used in this analysis; data after this period corresponds to the holiday season and to a period in which user behavior was affected by the earthquake events that occurred in PR in January 2020.

2.4.2. Methods to Relate Sociodemographic and Land Use Factors with the Generation of MDES trips

Regression analysis methods were applied to quantitatively relate observed trip generation levels with the selected sociodemographic data. Two dependent variables were considered in the analysis: total number of cell-level trip departures and the total number of cell-level trip arrivals. Sociodemographic variables and land-use variables were used in the analysis. The variables in the sociodemographic category are population, median age, employment levels, unemployment levels, and two binary variables related to income per capita and college student population (these variables were created after tests with the direct, untransformed variables were not successful). The income per capita dummy variable has a value of one if a cell's income per capita is greater than $3,900, and zero otherwise. The $3,900 value is the median income for full-time dependent students in the academic year 2015-2016 (Radwin et al., 2018). The college student dummy variable equaled 1 if a cell's student population represented more than 26.5% of the total population of a grid cell, and 0 otherwise. This percentage was determined by analyzing the number of students within the population of 50 college towns in the US. Five land-use variables, for five land-use types, were used in the regression analysis. The variables consisted of the building space (measured as floor area) dedicated to commercial, mixed-use, industrial, public, and UPRM land use categories. The public land use category includes schools, churches, and government buildings. Naturally, the UPRM land use category refers to buildings that are located within the UPRM campus.

Models using linear, Poisson, zero-inflated Poisson, negative binomial, spatial lag, and spatial error regression methods were considered in the analysis. Spatial regression was used as it is an approach commonly used in the literature (e.g., Caspi et al., 2020). Poisson, zero-inflated Poisson, negative binomial models were estimated as these are standard models used in analyses with count data (Washington, Karlafti, & Mannering, 2011). The count models, including the negative binomial model, were estimated using the statsmodel library in Python (Seabold & Perktold, 2010) and the spatial regression models were estimated using the GeoDa software, a spatial data analysis tool (Anselin, Syabri, & Kho, 2006). For the spatial regression models, the first-order queen contiguity weight matrix was used. Ultimately, the results for the negative binomial and spatial error models are reported in the results section (Section 2.5.4). The negative binomial model was selected given that count data is being analyzed and overdispersion was detected using the test proposed by Cameron and Trivedi (2001). The spatial error model was selected over the spatial lag model given their relative measures of goodness-of-fit.

2.4.3. Methods for Quantifying Spatial Access to Dockless E-scooters

Spatial access to the MDES e-scooters was quantified from the perspective of the building centroids, with the centroids serving as a proxy for the location of people. Two spatial access indicators are proposed: average distance to the K-closest scooters and the area under the scaled cumulative relative frequency curve (AUC). In addition, the number of scooters per person is computed in this study based on node-level measures. For this analysis, e-scooter locations were mapped to network nodes for each day and each period of the day (5-minute periods were used to discretize time in a day). The indicators were computed for each day, each period, and each node of interest in order to reflect that, in the context of dockless micromobility services, a neighborhood's spatial access to the service varies as vehicles are used during the day.

For each building centroid**,** and each day d and period (e.g., 5-minute intervals), the average distance to the K-closest e-scooter was determined, where K is a parameter set by the analyst. Computing this distance for a contiguous set of periods offers an indication of the travel cost incurred in reaching an e-scooter. Distances between nodes were determined using Dijkstra's shortest path algorithm.

The scaled AUC indicator is proposed as a more general spatial access indicator that considers all the distances from a node to each e-scooter. The AUC indicator is the area under the cumulative frequency curve of the shortest path distances from a node (e.g., a building unit) to all e-scooters in the system. To illustrate the AUC concept, consider the example in Figure 5. Figure 5(a) presents the cumulative relative frequency curve of the distances from node e to all e-scooters on day d at period p. The AUC for node e's curve (685.5) is greater than the AUC of node u's convex-like curve (328.5; Figure 5(b)). In general, as a node's cumulative frequency curve is more concave, more area will be under the curve, which in turn implies that the e-scooters are closer to the node; lower area values indicate that the fleet is generally farther from a node. Being overall closer to the fleet could imply greater chances of finding a parked e-scooter in the proximity of a location (i.e., more spatial access). The indicator is termed "scaled" as one could select an upper limit on the distances considered, scale all distances relative to this threshold, and discard all observations, if any, beyond the threshold (in this case, it is possible that the curve would not reach the 100% accumulation value). This procedure results in a unit area, and the value of the AUC would be bounded between 1 (100% of e-scooters located at 0 meters from the node) and 0 (no e-scooter located within the distance threshold).

A common measure of spatial access to a system is the number of vehicles in a zone. Based on this idea, the last spatial access indicator is scooter density, defined here as the node-level number of e-scooters per person. To compute a node-level density, a graph tree is constructed with all nodes within a threshold distance l for the (root) node being analyzed (Figure 6). Then, the node's scooter density is computed by dividing the total number of scooters in the nodes in its tree by the total population of the building centroids incident to the tree. This can be repeated for all nodes, days, and periods. This approach circumvents the boundary issues created by density measures computed at the zone level. To obtain zone-level e-scooter densities, one could average the density values of the nodes contained by each zone.

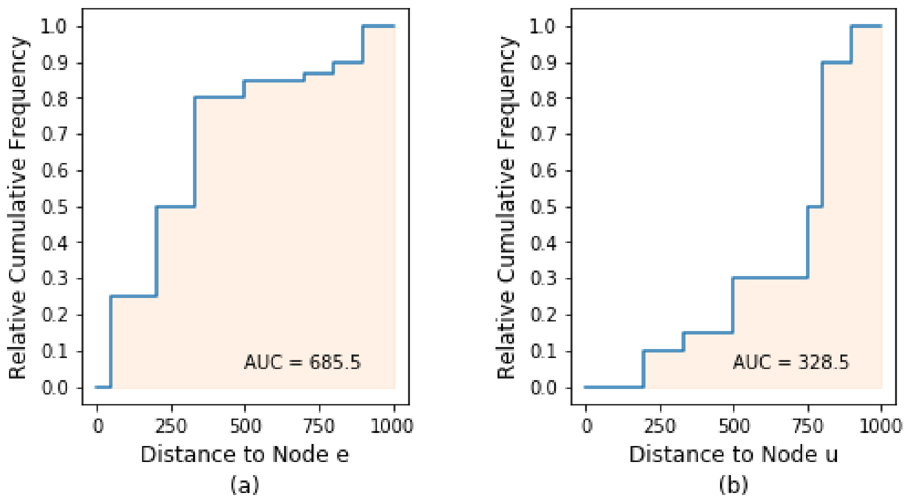

Figure 5. Area under the relative cumulative frequency curve (AUC) for the e-scooters distances of node e (a) and node u (b)

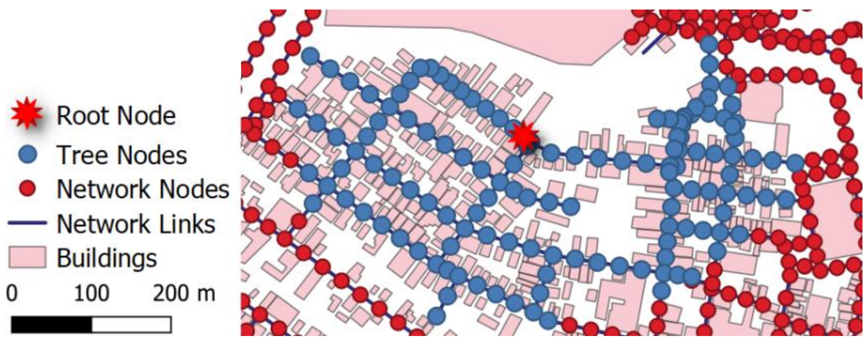

Figure 6. Tree constructed from a root node (l =375 meters)

2.5. Results

This section discusses the results of the analyses performed to understand the usage patterns of MDES, the regression analyses, and the spatial access analyses using the proposed indicators.

2.5.1. Factors Associated with the Use of MDES and User Clusters

The main characteristics of MDES users and nonusers are illustrated in Figure 7. Sixty-two percent of the female responders were nonusers, and 52% of the UPRM students indicated that they were users. As the age of the responder increases, the probability of being an MDES user decreases. On a question that allowed survey participants to select more than one option, users indicated that the three main reasons for using MDES were: travel time savings, avoidance of traffic congestion, and the lack of auto parking spaces (options selected by 58%, 37%, and 34% of responders, respectively). Note that the last two reasons suggest that MDES reduced auto trips for a section of the population. From the opposite perspective, the three main reasons given for not using or for using MDES less than desired were: the lack of space on the roads for e-scooters, trip costs, and unavailability of e-scooters where needed (options selected by 40%, 35%, and 23% of responders, respectively. In addition, 21% of responders also indicated that e-scooters were not safe. In terms of trip purpose, the survey revealed that the three main MDES trip purposes were: to attend classes, work, or visit a restaurant (options selected by 78%, 44%, and 43% of responders, respectively).

Table 2 presents the estimated coefficients for the logistic regression model. The model suggests that a person is less likely to use MDES if they have available a car or a bicycle, although the coefficients for these parameters are not significant at the 95% confidence level. The model also suggests that as a person's age increases, they are less likely to use MDES. For example, relative to the 27-50 age group, persons in the 18-20 and 21-26 age groups are, respectively, five and two times more likely to be MDES users. With respect to gender, females are 1.7 times less likely to use MDES.

Table 2 Coefficients and Odds-Ratios for Logistic Regression Model

Variable	β	p-value	Odds-ratio $(\exp(\beta))$	1/Odds-ratio $(1/\exp(\beta))$
Intercept	-0.32	0.52	-	-
Car Availability	-0.55	0.08	0.57	1.74
Bike Availability	-0.37	0.20	0.69	1.45
Age: 18-20	1.70	0.00	5.47	0.18
Age: 21-26	0.93	0.02	2.53	0.39
Age: $\geq$50	-1.35	0.06	0.26	3.86
Female	-0.55	0.02	0.58	1.73

Number of observations: 371
Pseudo R-squared: 0.12
Log-likelihood: -223.10
Log-likelihood ratio test p-value: <0.001

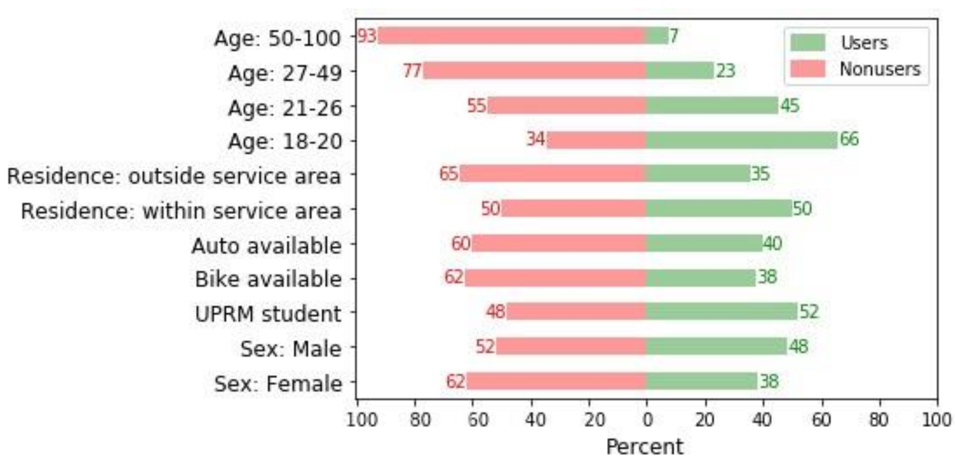

Figure 7. User and nonuser representation within population groups based on survey responses

In the clustering analysis, three clusters were identified using the dendrogram and selecting the number of clusters that maximize the distance metric across the different clusters. Table 3 reports the average value of the computed attributes of the members within each cluster. Users from cluster 1 are primarily users that made one or two trips during the period of analysis. They represent 46% of registered MDES users, 60% of whom used the system within the first three weeks of operations. The users in cluster 2 performed, on average two trips per month, had a 30% probability of performing a second trip, and had an average time between trips of 208 hours. Cluster 3 corresponds to the high-frequency users, which represent 14% of users. The members of this cluster made around 65% of e-scooter trips. On average, they made 49 trips during the period of analysis and, if they made a trip, their probability of performing additional trips was 0.42, on average.

Table 3. Average User Travel Characteristics by Cluster Group

Cluster ID	Percent of registered users	Number of trips	Average time between trips (hours)	STD of time between trips (hours)	Presence (days)	Subsequent Trips Probability
1	46	1.8 (1.1)	1128.0 (894)	6.5 (25.3)	5.3 (13.4)	0.00 (0.0)
2	40	10.3 (6.4)	207.6 (194)	239.4 (202.1)	58.6 (33.3)	0.30 (0.2)
3	14	49.2 (27.6)	62.1 (28.0)	88.1 (42.4)	102 (16.5)	0.42 (0.2)

Note: Standard deviation in parenthesis; users that used the system once were assigned a value of 1800 hours for "average time between trips", which represents the maximum feasible time in the period considered.

2.5.2. Spatiotemporal Patterns of Trips in MDES

As expected, MDES trips are mostly associated with UPRM and the neighborhoods in which students live. The regions with a high concentration of trip origins are identified in Figure 8. The illustrated density values were determined by kernel density estimation using the original start latitudes and longitudes of all trips. As can be seen, most trips start from the UPRM, Terrace, and Bosque zones; all other hotspots in the map are associated with student residences, except for a Centro zone hotspot that corresponds to the city's main plaza, where the City Hall and major religious facilities are located. The origin-destination heatmap presented in Figure 9 confirms this observation: most trips start, end, or start and end in the university. Specifically, 35% of trips start and end at UPRM, 25% have the university only as its destination, and 18% have it only as its origin.

The temporal patterns also suggest MDES dependence on the UPRM community. The boxplots in Figure 10 indicate that trip rates plummet during weekends when most students return to their hometowns (there are 87% fewer trips on weekends). On average, Wednesday is the day of the week with most trips. The MDES trip distribution by the time of day does not exhibit the standard peak-hour pattern observed in other modes. Based on the median distribution, the peak period during weekdays was observed between 8:00 AM and 1:00 PM (Figure 11 and Figure 12 plot the data of each type of day for the analysis period, in addition to the quartiles). On average, 50% of weekday trips are completed by noon. For weekends, a peak period is not observed (Figure 12) and 50% of trips are completed by 3:00 PM. This analysis was performed using data from the August 3, 2019, to December 15, 2019, period. The demand for MDES reduces to close to zero when the university is not in session (in Figure 13, the period between the end of December and the end of January). The maximum number of trips was observed during the week of August 19-23, which coincides with the first week that e-scooters were given permission to operate inside the UPRM campus.

Zone ID	Name
1	UPRM
2	Bosque
3	Terrace
4	Trastalleres
5	Paris
6	Norte
7	Centro
8	Playa
9	Plaza

Figure 8. Density map of trip starts

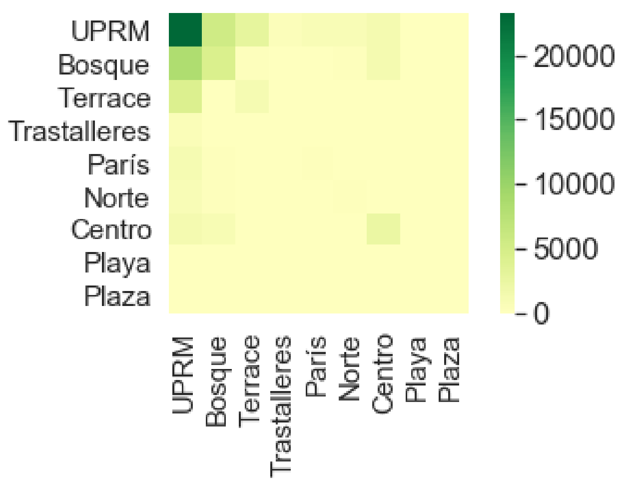

Figure 9. Origin-destination matrix of MDES trips

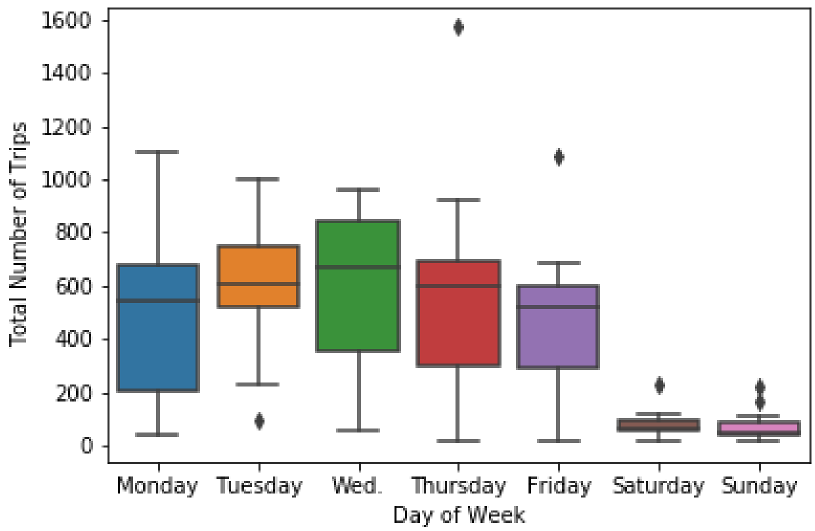

Figure 10. Total e-scooter trips by day of week

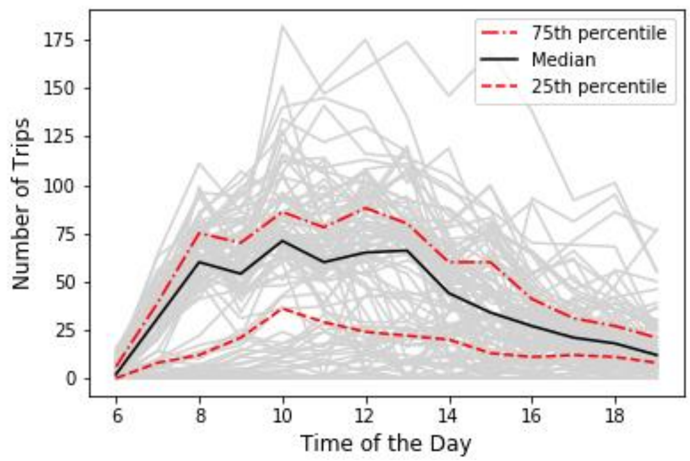

Figure 11. Hourly trip frequency for weekdays

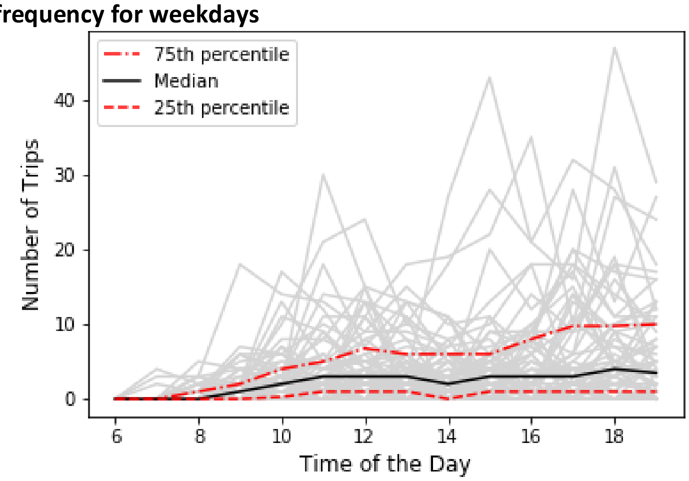

Figure 12. Hourly trip frequency for weekends

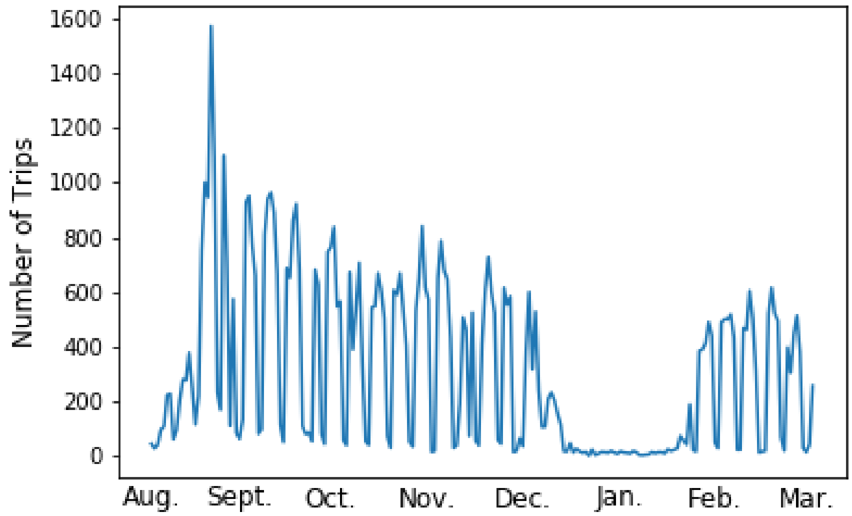

Figure 13. Daily e-scooter trip count (August 2019 – March 2020)

2.5.3. Indicators of Spatial Access

In
Figure **14** and
Figure **15**, the average minimum walking distance to an e-scooter (K=1) and to the ten closest e-scooter (K=10) is depicted for each node (only nodes connected to a building centroid considered) and each day for the period of analysis. Noticeable in both figures is the increase in the distance between a node and e-scooters during weekends and the December-January break (and earthquake period) due to the reduction of the available fleet in those periods, which in turn responds to the lower demand for the service. The exceptions to this observation are the Bosque, Centro, and París zones, as these zones have businesses next to which the operator placed e-scooters even when UPRM was not in session. But note that within the nodes of each zone there is variability, particularly for relatively large zones that contain a large number of nodes (e.g., there are 800 nodes in UPRM versus the seven, closely located nodes in Plaza. On average, nodes in the Bosque neighborhood have the lowest minimum walking distance to the ten closest e-scooters (378 ± 171 meters), while nodes in the Playa zone have the largest walking distance (1388 ± 635 meters). Note that the color contrast between
Figure **14** and
Figure **15** communicates information that could be of interest to planners. For instance, observing the Terrace heatmap, a marked contrast can be observed in the average minimum and 10-minimum distances on the weekends (see bright bands that appear in
Figure **15**), which indicates a general scarcity of scooters in that zone; again, this scarcity reflects the relatively few students in Terrace during weekends. There are zones (e.g., UPRM and Terrace) for which four regimes can be observed: the periods prior to MDES access to UPRM, the fall semester, the winter break and earthquake event, and finally the late start of the spring semester.

Relative to the previous figures, the heatmap depiction of the AUC values in Figure 16 provides a similar, yet more general picture of a node's proximity to the scooter fleet. In this analysis, the scaling factor was set to 2,000 meters. The Bosque and París nodes have similar AUC patterns, as well as the highest values (0.58 ± 0.14 and 0.55 ± 0.10, respectively). Also noteworthy is the Norte zone, which has a relatively high and constant AUC value (0.55 ± 0.08), probably due to the central location of this zone. Once again, the Playa zone has the worse indicator (AUC value of 0.21 ± 0.16). The AUC indicator can be used to examine the patterns at the node level by time of day, as illustrated in Figure 17. In this example, the AUC values for the most significant UPRM destination node, the most significant Bosque origin nodes, and a Playa node are depicted in Figure 17 (a), (b), and (c), respectively. This figure was generated by plotting the three nodes' weekday data for the period of August 3, 2019, to December 15, 2019. A two-peak curve is observed for Figure 17(a), which is likely related to student's class schedules in that UPRM region, while the curve in Figure 17(b) shows how the fleet is closest to that node in the morning (the operator places e-scooters there at the start of the day), and it relatively disperses from it as the day progresses. Figure 17(c) suggests that the e-scooter fleet is generally far from the Playa node, and it gets farther from the node as the day progresses.

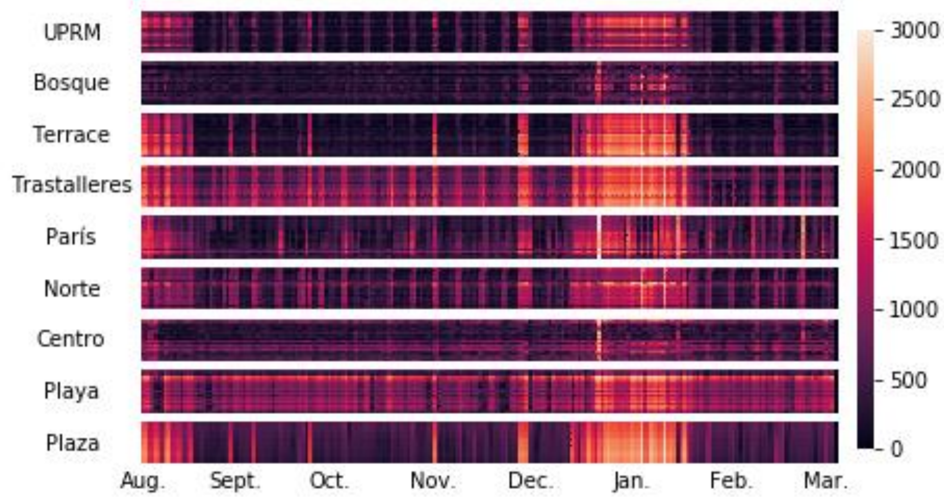

Figure 14. Average minimum walking distance (meters) to an e-scooter for each node and each day

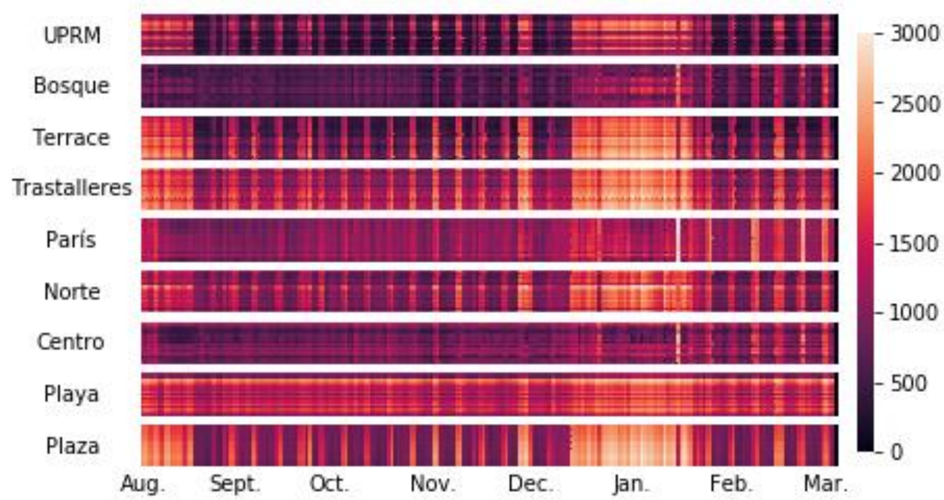

Figure 15. Average walking distance (meters) to the 10 closest e-scooters for each node and each day

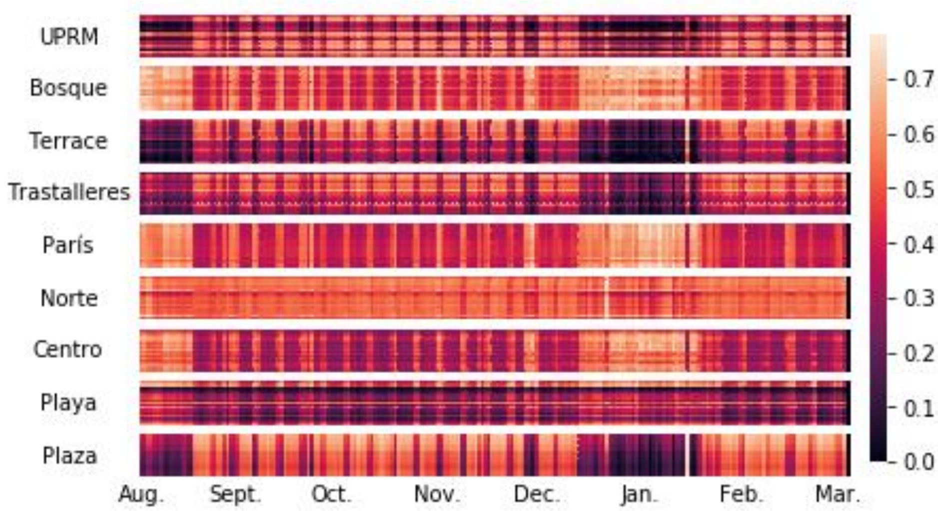

Figure 16. Average AUC value for each node and each day

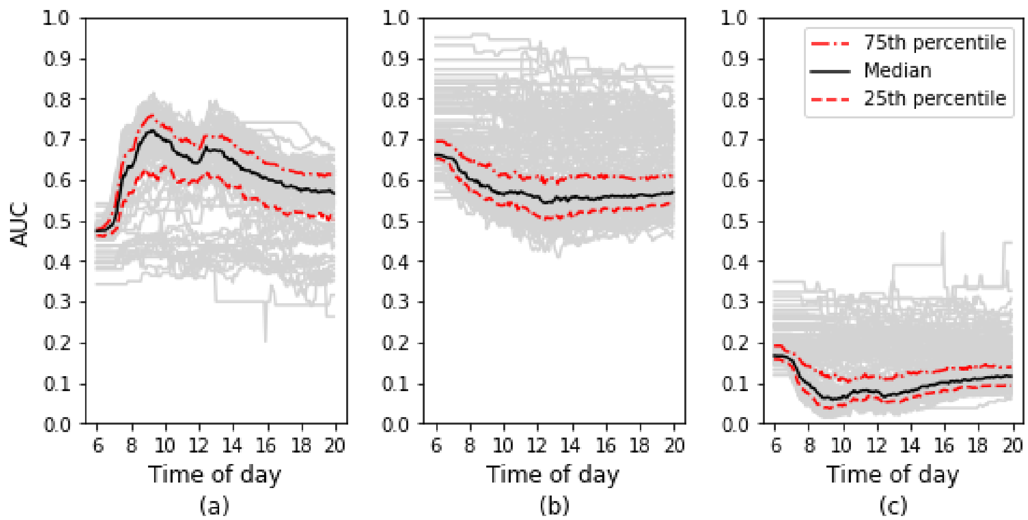

Figure 17. AUC values by time of day for selected nodes

Figure 18 shows the average daily e-scooter per 100 people (density) for each zone during the period of interest. The plots in this figure represent the median density value for network nodes connected linked to centroids of residential buildings (the UPRM and Playa zones were removed from this analysis given their lack of residential buildings). The l parameter was set to 187.5 meters (corresponding to a 2.5-minute walk). The Norte and Bosque zones had the highest average e-scooter densities (1.19 ± 1.55 and 1.04 ± 0.64). The Norte zone has a relatively low population (70 people), yet a relatively high number of e-scooters were located in it as it has sidewalks with significant pedestrian activity, it is close to a UPRM pedestrian entry point, and it contained several business establishments. The zones with the lowest e-scooters densities are Trastalleres and Playa, with densities approaching zero.

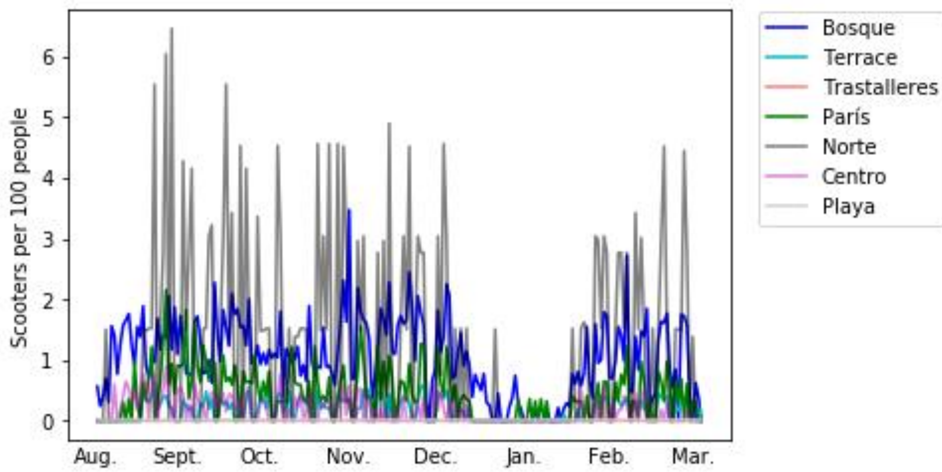

Figure 18 Average daily e-scooters per person for each zone

2.5.4. Regression Analysis of Trip Generation, Sociodemographic, and Land Use Variables

The results from the negative binomial regression models are reported in Table 4. For both dependent variables, all independent variables are significant except for employment and public land use. The results for the sociodemographic variables suggest that the population, income dummy, and college student dummy variables are positively associated with e-scooter trips for both dependent variables. The positive correlation between the income variable and trip activity highlights the possible difficulties that low-income residents of Mayagüez may have in accessing the e-scooter system. Median age and unemployment have negative coefficients, suggesting older and unemployed people could have less access. The positive coefficient for the college student dummy suggests that, as in previous studies, university students are major users, in relative terms, of micromobility services. These results also align with the results obtained in the online survey and spatiotemporal analyses presented in the previous sections.

For the land-use variables, the models suggest that commercial, mixed, and UPRM land uses have a positive association with both the number of departures and arrivals in a cell. In contrast, industrial land use is negatively associated with e-scooter trip activity. Among the land-use variables, the coefficients with the highest magnitudes are the ones for mixed and UPRM land uses. These results agree with previous studies that show a positive correlation between micromobility trip generation and mixed and university land uses. As most mixed-use buildings are in densely student-populated neighborhoods such as Terrace, Pueblo, and París, the results also support the observations made in the previous sections.

In Table 5 the results for the spatial error models are reported. In contrast to the negative binomial model, in this model, only the population, mixed land use, and UPRM land-use variables are statistically significant at the 95% confidence level for the trip departures models. For the trip arrival model, only the mixed land use and UPRM land-use variables are statistically significant (at the 95% and 90% confidence level, respectively). The lambda is the spatial autoregressive coefficient, which is positive and significant for both models.

The negative binomial and spatial error model results differ on which variables are significantly related to trip generation in MDES. This is unsurprising given the differences in model structure and their underlying assumptions. For example, the spatial error model accounts for spatial dependence (through the error terms) among observations, whereas spatial dependence structures are not part of the negative binomial model. However, the negative binomial regression model is used for count data, whereas the spatial error model was not designed as a count data model. In any case, both types of models confirm that the principal factor driving demand for MDES is UPRM and its students.

Table 4. Negative Binomial Regression Models for Trip Generation in MDES

Variables	Trip Departures Model		Trip Arrival Model	
	Coefficient	*p Value*	*Coefficient*	*p Value*
Constant	4.685	0.000	4.876	0.000
Population (people)	0.010	0.000	0.007	0.000
Median Age (years)	-0.123	0.000	-0.091	0.000
Employment (job/km²)	0.493	0.205†	0.127	0.728†
Unemployment (%)	-0.045	0.006	-0.072	0.000
Income Dummy	1.894	0.000	1.295	0.000
Student Dummy	2.176	0.000	1.747	0.000
Commercial LU (km²)	95.497	0.020	81.630	0.036
Mixed LU (km²)	292.363	0.000	250.958	0.000
UPRM LU (km²)	234.381	0.000	223.507	0.000
Industrial LU (km²)	-163.478	0.004	-99.072	0.055†
Public LU (km²)	-81.435	0.173†	-49.732	0.381†
Number of observations: 350	Log-likelihood: -1743.9 χ^2 statistic: 1090 (p-value: <0.001)		Log-likelihood: -1810 χ^2 statistic: 775 (p-value: <0.001)	

Note: LU stands for land use. † $p > 0.05$.

Table 5. Spatial Error Models Regression Models for Trip Generation in MDES

Variables	Trip Departures Model		Trip Arrival Model	
	Coefficient	*p Value*	*Coefficient*	*p Value*
Constant	120.1	0.053	147.7	0.016
Population (people)	1.344	0.048*	0.856	0.155
Median Age (years)	-1.580	0.733	-1.412	0.737
Employment (job/km²)	0.783	0.996	1.152	0.993
Unemployment (%)	0.558	0.934	-0.843	0.892
Income Dummy	-56.810	0.616	-28.9	0.779
Student Dummy	72.306	0.368	33.4	0.644
Commercial LU (km²)	-4236.1	0.725	-1308.8	0.901
Mixed LU (km²)	45512.7	0.003**	24254.0	0.071*
UPRM LU (km²)	25824.1	0.004**	21132.8	0.007**
Industrial LU (km²)	-4031.6	0.805	-5726.4	0.690
Public LU (km²)	5827.2	0.765	451.3	0.979
Lambda	0.592	0.000	0.661	0.000
Number of observations: 350	Pseudo R-squared: 0.361 Log-likelihood: -2578 Log-likelihood ratio test: 41.57		Pseudo R-squared: 0.429 Log-likelihood: -2535 Log-likelihood ratio test: 57.20	

Note: LU stands for land use. * $p < 0.1$. ** $p < 0.01$.

2.6. Discussion

The MDES user characteristics and usage patterns are similar to those discussed in the reviewed literature: users tend to be young, male, and relatively young; most trips occur on weekdays, and within the day there are no evident peak hours, and the main driver behind the generated trips are activities in a university. Perhaps not as common is the drastic drop of e-scooter trip demand when there are no classes at UPRM, particularly during the weekends, which suggest that the floating population of students, rather than permanent residents, are the users of the system. The analysis of trip arrivals and departures and the computed spatial access indicators shows an inequitable spatial distribution of e-scooters among neighborhoods in the MDES service area. There are probably several obvious reasons why e-scooter trips were mainly made by students and were concentrated in UPRM and its adjacent neighborhoods. Reasonable candidate explanations include that:

i. the operator primarily located e-scooters in and around UPRM,
ii. a significant proportion of Mayagüez residents face economic and age-related health barriers that impede their participation in the system,
iii. permanent residents in the study region perform activities (e.g., work, groceries) outside the study region and, therefore, have little use for the system, and
iv. there are fewer students and visitors in the MDES service area during the weekends, and therefore there is a low demand for the system, given that the Pueblo borough has relatively few attractive activities during the weekends (as previously mentioned, outside a few restaurants, most establishments, and offices in Pueblo are closed during the weekend).

The spatial equity analysis performed in this study would need to be complemented by a vertical equity analysis to provide a definitive explanation of the underlying causes driving the differences in e-scooter spatial distribution. It is reasonable to expect that, beyond the age and gender differences observed in this study, there are important social group factors that affect the observed e-scooter travel patterns. Chief among these are the markedly different income levels among neighborhoods in the service region (which is not clearly reflected in the available Census data). Income differences are not only associated with the amount of money that people allocate for travel, but they are also associated with access to the cashless means of payment that the MDES accepts.

Another noteworthy result of this study is the observation that MDES might be reducing the auto trips of a proportion of its users. As indicated by survey respondents, among the main reasons for the use of MDES are the lack of parking spaces in UPRM and the traffic congestion in the region. This suggests that MDES could have a positive impact on the congestion in the study area, but the magnitude of this impact is unclear. Further study is necessary to explore the potential impact of MDES on traffic congestion in the study region. The level of demand for the system and the impact that MDES could have on traffic congestion is probably negatively impacted by Mayagüez's built environment. The sidewalks and local streets in Mayagüez are relatively narrow, the street space is mostly occupied by moving or parked motor vehicles, pavements are not in good condition, and there is close to zero infrastructure dedicated to bicycle and e-scooter users. This in turn creates the safety concerns identified in the online survey. As the reviewed literature suggests, and it would be reasonable to expect, creating space in the built environment for non-auto modes could improve MDES demand and the demand for bicycling, reduce congestion, and potentially enhance equity and access.

A possible limitation of the analysis of the online survey, including the logistic regression model, is the possibility it suffers from sampling biases. The survey activities in this project were affected by the COVID-19 crisis. The crisis prevented on-the-field observations and interviews as MDES paused operations and in-person classes and research were suspended at UPRM. This situation hindered plans for a more controlled approach to the selection of survey participants (essentially, opportunity sampling was used given the changes in research activities). Given that the survey was distributed using webtools (e.g., UPRM email service), self-selection bias could arise if, for example, people interested in MDES were more likely to participate, while those not interested in the service ignored the survey. Also, sampling bias could be introduced, for example, by the lack of participation from people who do not check their emails frequently and never saw the survey announcements. The results of the online survey analysis, however, align with the results from a 2019 pilot study in which field observations and interviews were made, so the survey provides information that is consistent with other observations.

2.7. Closing Remarks

The first shared micromobility service in Puerto Rico started operations in the Municipality of Mayagüez in August 2019. It is operated by a local micromobility startup, which has expanded operations to four cities in PR. Given the public transportation options in PR, it is conceivable that micromobility services could help reduce short auto trips in the markets in which they operate, and consequently help reduce traffic congestion.

The MDES experience would suggest that the benefits of micromobility services in PR are likely to be distributed inequitably, particularly in areas, like Mayagüez, in which the local population has, on average, low-income levels. However, more research is required to determine how transferable the MDES experience is in PR and elsewhere. A key area of research is the impact of the built environment on micromobility demand in an auto-centric society that lacks access to reliable public transportation alternatives.

In this chapter methods, for generating network-based spatial access indicators were proposed and applied in the MDES case. The proposed method can be used to analyze, at a granular node level, equity in spatial access of micromobility systems. As will be shown in the next chapter, the proposed indicators can also be used as part of rebalancing procedures that account for equity objectives.

3. Vehicle Rebalancing to Improve Equity in Access to Dockless Micromobility Systems

One of the main challenges in micromobility services is the concentration of vehicles in suboptimal locations, which is why rebalancing operations are necessary. From the perspective of operators, a vehicle distribution of vehicles can be suboptimal if it results in, for example, fewer trips than possible. From a societal perspective, however, a vehicle distribution can be considered suboptimal if it largely excludes disadvantaged communities. These equity concerns have been reflected in city-wide rebalancing requirements, as mentioned in Chapter 1.

In this chapter, a two-step model is proposed that can be used to account for equity objectives in the rebalancing operations of dockless micromobility services. The proposed approach incorporates optimization models to define target vehicle distributions according to efficiency (e.g., number of trips) and equity (e.g., accessibility) performance objectives. Equity in access indicators, both from a spatial and social equity perspective, are proposed as objective functions in the decision-making process. These target distributions are then used as inputs in a multi-objective pick-up and delivery problem (PDP) whose solution defines the vehicle relocation plan. In addition to model formulations, a heuristic based on the differential evolution algorithm is presented to solve the proposed problems. Numerical examples are presented to illustrate the application of the proposed methods.

This chapter is organized as follows. Previous rebalancing models for station-based and dockless systems are reviewed next, followed by a discussion of the proposed optimization models and solution methods. In Section 3.3, the application of the proposed model is illustrated using travel simulation models. In the last section, future research directions are discussed.

3.1. Literature Review on Rebalancing Models

Two types of vehicle rebalancing (relocation) strategies are considered in the literature: the customer incentives approach and the operator-based approach. In the first approach, users are incentivized to relocate vehicles to selected locations, while in the operator-based approach the relocation process is performed by service staff. PDPs are often used to developed models for operator-based rebalancing methods. The operator-based rebalancing operations can be further categorized as either dynamic or static (Chemla et al., 2013). Static rebalancing refers to rebalancing that occurs before or after the service is in operations or when there is low demand, while dynamic rebalancing is performed during the operation hours when the status of the system is rapidly changing. Shui and Szeto (2020) present a review of bicycle-sharing service planning problems that defines key terms and modeling approaches directly relevant to rebalancing problems.

3.1.1. Rebalancing Models for Station-based Micromobility Services

Benchimol et al. (2011) and Chemla et al. (2013) presented single-vehicle PDPs without time constraints that sought to minimize the total vehicle travel distance required to complete a rebalancing plan. Chemla et al. (2013) presented a branch-and-cut algorithm for solving such a problem, from which a feasible solution was obtained via a tabu search. Both papers assumed a known target inventory level for each station in the system. Raviv and Kolka (2013) and Regue and Recker (2014) developed a Markov chain formulation to define the minimum and maximum inventory level as an input for a bike-sharing repositioning plan. Raviv and Kolka (2013) proposed a proactive rebalancing approach, instead of a reactive one. Intended to prevent the occurrence of inefficient system states. They developed a framework to solve the dynamic bike-sharing repositioning problem by integrating four models: a demand forecasting model, a station inventory model, a redistribution needs model, and a vehicle routing model. Similar modeling approaches are observed in models that extend these works (e.g., Zhang et al. 2017, Schuijbroek et al. 2017).

3.1.2. Rebalancing Models for Dockless Micromobility Services

Naturally, rebalancing models for dockless (also known as free-floating or stationless) services differ from those proposed for station-based services, in part because station capacity constraints are not a problem. Liu and Xu (2018) proposed an optimization-based model that considered clustering analysis to solve the rebalancing problem for dockless bikes-haring services. Similarly, Caggiani et al. (2018) proposed a framework that uses spatiotemporal clustering as part of a demand forecasting methodology. The proposed framework had the objective of maximizing user satisfaction and minimizing the number of lost users. Additionally, Barabonkov et al. (2020) presented a mixed integer programming problem for rebalancing operations in dockless bike-sharing systems. The objective of their model was to minimize lost profit, which was computed as a function of lost demand; the proposed methodology also required a model to forecast demand levels. From the customer incentives perspective, Pan et al. (2019) proposed a deep reinforcement learning framework for incentivizing users to rebalance dockless systems.

3.1.3. Equity and Optimization-based Models in Transportation

An extensive number of models have been proposed to plan and operate transportation systems considering equity objectives. These include equity-related models for road pricing (Levinson, 2010), bus service design (e.g., Ferguson et al., 2012), road network design (Caggiani, Camporeale, & Ottomanelli, 2017), and traffic signal control (Han, Liu, Gayah, Friesz, & Yao, 2015), to name a few. In the context of micromobility, Caggiani et al. (2020) proposed an optimization-based approach to determine the number and layout of bike stations and their respective racks to minimize the implementation and operation cost, and they included spatial equity constraints to control for the relative number of bicycles and walking times between zones. Caggiani et al. (2020b) extended this work, in part, by considering the presence of multiple modes. The model's objective was the minimization of accessibility inequalities as computed by the Theil inequality index.

3.2. Rebalancing Model

A model is proposed to rebalance vehicles in dockless micromobility systems according to efficiency and equity objectives. The following assumptions are made in the proposed model:

i. The rebalancing model is invoked to generate vehicle pickup and drop-off recommendations (i.e., the rebalancing plan) only during periods $t \in \boldsymbol{T}$.
ii. Models are available to predict the micromobility system performance for a given time horizon (see Appendix B).
iii. Models are available to estimate the probability that a person rents a micromobility vehicle.

The proposed methodology consists of two steps that are executed in each period $t \in \boldsymbol{T}$:

Step 1. Define target vehicle distributions: In this step, optimization modes are used to determine the vehicle distributions that maximize system performance metrics for a given time horizon. The optimization model used in this step is called the target distribution problem.

Step 2. Define vehicle pickup and delivery locations. With the target distributions as inputs, a PDP is used to define the vehicle pickup and delivery locations and quantities.

As explained in the next subsections, both the target distribution problem and the PDP can be used to reflect system efficiency and equity considerations in the development of rebalancing plans.

3.2.1. Target Distribution Problem

The target distribution problem (TDP) is proposed to identify, for a given period, vehicle distributions that result in optimal system performance in a future time horizon (e.g., next two hours). For each period $t \in \boldsymbol{T}$, the target distribution problem (TDP) assumes that prediction models can be used to forecast the performance of the system as a function of the vehicle distribution $\boldsymbol{s}_t$ and of complementary information contained in $\boldsymbol{\Lambda}_{z,t}$ (e.g., system activity patterns in previous periods, weather, day of the week, time-of-day). Let F_z be a model that predicts the system's performance according to objective z, $\boldsymbol{J}$ be the set of all locations (e.g., zones, neighborhoods, regions) in the service area, s_j be the number of scooters in zone j ($j \in \boldsymbol{J}, s_j \in \boldsymbol{s}$), and h be the size of the vehicle fleet. Then, the target distribution is determined by finding a solution to:

$$\max F_z\left(\boldsymbol{s}, \boldsymbol{\Lambda}_{z,t}\right) \tag{1}$$

$$subject\ to$$

$$\sum_{j \in \boldsymbol{J}} s_j = h \tag{1.1}$$

$$s_j \in \mathbb{Z}^+ \cup \{0\} \qquad \forall j \in \boldsymbol{J} \tag{1.2}$$

Constraint (1.1) ensures that the selected target distribution adds to the available vehicle fleet size, while constraints (1.2) are the integrality constraints. Note that the dimensions of $\boldsymbol{s}$ are $1 \times |\boldsymbol{J}|$. In the context of this project, the TDP would have to be solved at least twice to separately identify the target distributions for the efficiency and equity objectives. Consequently, for each objective, a separate $F_z\left(\boldsymbol{s}, \boldsymbol{\Lambda}_{z,t}\right)$ model would have to be used. Previous studies have developed modeling approaches that can be applied to predict efficiency-related objectives, such as the total number of trips in the system (see Appendix B for a review of machine learning approaches to forecasting demand). In the following subsection, two approaches are presented to define equity performance metrics and objectives for the rebalancing operations of dockless micromobility services. In appendix C, a heuristic to find solutions to the TDP is proposed.

3.2.1.1. Specification of Equity Performance Metrics and Objective Functions

As discussed in Section 2.1.1, the term equity can be understood from multiple perspectives. In the proposed rebalancing model, equity is quantified using an indicator that reflects the inequalities in the distribution of opportunities to access the system. That is, equity is examined in terms of opportunities to access the service of interest. Let A_t be the inequality indicator for the system at period t and a_{tu} be a proxy that captures the opportunity to access the system for entity $u \in \boldsymbol{U}$ at period t. A_t reflects the inequalities in the distribution of the a_{tu} values in the entities of interest.

Two measures are proposed to define a_{tu}. The first measure ($a_{ut;l}$) is defined as the AUC indicator proposed in Section 2.4.3. As previously discussed, this indicator defines the proximity of a vehicle fleet to a node as the area under the cumulative relative frequency curve of the shortest path distances from the node to all vehicles in the network. Therefore, $a_{ut;l}$ depends on a network representation of transportation infrastructure (e.g., sidewalks, roadways) and land uses. The entities of interest for $a_{ut;l}$ are building units (e.g., commercial establishments, apartment units) that are represented as nodes in a network.

The second measure ($a_{ut;K}$) is developed from the social equity perspective. $a_{ut;K}$ is computed for each person in a sample of the population as a function of the person's attributes. The attributes are captured in an estimate of each individual's probability of selecting a micromobility service. Let w_{jt} be the number of vehicles in zone j at time t per the expected demand for those vehicles and p_{kijt} be the probability that a person u in zone i ($i \in \boldsymbol{J}$) uses a vehicle in zone j and period t. Then the person-level measure of access is:

$$a_{ut;K} = \sum_{j \in J} p_{uijt} w_{jt} \tag{2}$$

This measure is analogous to the cumulative opportunity measures discussed by Kwan (1998). The w_{jt} term reflects the relative opportunities that person u has for using the service and the p_{uijt} term (which ranges from zero to one) reflects how accessible those opportunities are to the person. The w_{jt} term could be directly tied to the probabilities of use as follows:

$$w_{jt} = \frac{s_{jt}}{\sum_{g \in \boldsymbol{U}} \sum_{i \in \boldsymbol{J}} p_{gijt}} \tag{3}$$

where the denominator is the expected demand for the scooters at zone j at period t. As stated in assumption (iii), here it is assumed that models exist (e.g., logit models) to compute mode choice probabilities.

In practice, computing the spatial access measure $a_{ut;l}$ is inexpensive, both in terms of data requirements and computational resources. The social access measure $a_{ut;K}$, however, might require new data-sharing paradigms in micromobility services. For example, an operator could create an equity-oriented program in which individuals voluntarily participate by sharing data (e.g., real-time location for a period, but also limited sociodemographic characteristics), and this information is used to guide rebalancing so that vehicles are distributed considering each person's $a_{ut;K}$.

Given the selected access measure, the inequality indicator that aggregates this information can be computed. Several inequality indicators can be found in the literature (e.g., see Ramjerdi, 2006). Here, the Atkinson index of inequality is selected for the simulation tests discussed in Section 3.3. In the current context, the index is defined as:

$$A_t = 1 - \left[\frac{1}{|\boldsymbol{U}|} \sum_{u \in \boldsymbol{U}} \left(\frac{a_{tu}}{\bar{a}_t} \right)^{1-\varepsilon} \right]^{\frac{1}{1-\varepsilon}} \tag{4}$$

where ε is an inequality aversion parameter that can range from zero to infinity, but typically it is set to a value of 0.25 to 2 (excluding 1 for the given expression), with larger values indicating that more importance is given to the transfer of the access measure at the lower end of the distribution (Levy, Chemerynski, & Tuchmann, 2006). The Atkinson index value ranges from zero, representing perfect equality, to one, representing the opposite. Therefore, the performance indicator (alternatively, the equity objective) based on this indicator would have to be defined as $F_{A_t} = -A_t$ in model (1) (as the analyst is interested in minimizing A_t, and model (1) is a maximization problem). The Atkinson index was selected as it has several attractive features, including that it satisfies the Pigou–Dalton transfer principle and it is subgroup decomposable (Levy et al., 2006). However, note that the proposed methodology is flexible in that it does not depend on which index is selected.

3.2.1.2. Considering Congestion Reduction Objectives in the Target Distribution Problem

Given the models that compute the mode choice probabilities of individuals, a natural approach to considering a traffic congestion objective in the TDP is to define a function that captures the number of auto trips given the micromobility vehicle distribution. Let $q_{ut}(\boldsymbol{s}_t, \boldsymbol{\Lambda}_{ut})$ be the probability of performing an auto trip for person u at period t characterized by the information in $\boldsymbol{\Lambda}_{ut}$, and let $\boldsymbol{U}_q \in \boldsymbol{U}$ be the set of auto users in the system. Then, the congestion reduction performance metric F_q can be defined as:

$$F_q = -\sum_{u \in \boldsymbol{U}_q} q_{ut}(\boldsymbol{s}_t, \boldsymbol{\Lambda}_{ut}) \tag{5}$$

3.2.2. Minimum Level-of-Service Constraints

The TPD assumes that a system-wide minimization of inequity can be pursued by achieving a system state (i.e., vehicle distribution) that results in a reduction of an inequality index during a given time horizon. In practice, government actors have established simpler equity requirements for micromobility companies, including spatial distribution standards that must be achieved during vehicle relocation operations (Hirsch, Stratton-Rayner, et al., 2019). Generally, these types of requirements can be included in optimization models as level-of-service constraints. For example, in the PDP discussed in the next section, a constraint is included to ensure that a given set of zones (e.g., neighborhoods) has a minimum number of vehicles at the end of the rebalancing operations. More complex constraints could be included to achieve more specific or local level-of-service standards. From a network perspective, the model constraints can be formulated so that the generated rebalancing plan satisfies requirements related to, for example, the:

- minimum pedestrian travel times from a set of nodes to their respective K closest vehicles,
- service area coverage implied by the vehicle distribution and pedestrian travel time, or
- average values of an access measure in previous periods of a day.

More complex constraints could mean a harder-to-solve optimization problem, which in turn would require the development of heuristics that enable the practical application of the proposed model.

3.2.3. Pickup and Delivery Model

The target vehicle distributions generated via the TDP serve as parameters of the PDP. Let $\boldsymbol{\lambda}$ and $\boldsymbol{\mu}$ be the target vehicle distributions according to the efficiency and equity objectives, respectively. Given that the $\boldsymbol{\lambda}$ and $\boldsymbol{\mu}$ distributions are likely to be different and that there is a fixed number of vehicles in the system, the goal of the PDP is to produce a rebalancing plan that satisfies both the efficiency and equity targets to some degree. If a target is not met in a zone i (i.e., if $\lambda_i - s_i > 0$ for the efficiency objective or $\mu_i - s_i > 0$ for the equity objective), then there is a vehicle deficit in the zone, and the PDP attempts to minimize this deficit.

The objective of the proposed PDP is to minimize the weighted sum of the zone-level vehicle deficits, relative to the $\boldsymbol{\lambda}$ and $\boldsymbol{\mu}$ distributions, and the transportation costs that result from the rebalancing operations. The proposed PDP also includes a constraint that ensures that a minimum number of vehicles is present in selected zones, which guarantees a minimum level of service. Next, the model notation and formulation are presented and discussed. For simplicity, a single rebalancing vehicle (hereafter, truck) is assumed to be available to execute the PDP's rebalancing plan (as in the MDES operations). Note that the main ideas behind the PDP do not depend on the number of trucks available for the rebalancing operation. Some of the notation in Schuijbroek et al. (2017) is adopted here.

NICR NATIONAL INSTITUTE FOR CONGESTION REDUCTION
4202 E. Fowler Avenue, ENG 030
Tampa, FL 33620-5375
www.nicr.usf.edu
CUTR Center for Urban Transportation Research
Berkeley University of California
Texas A&M Transportation Institute
UPR Recinto Universitario de Mayagüez

Sets

$\boldsymbol{J}$: set of zones

Parameters:

s_i : number of vehicles in zone $i \in \boldsymbol{J}$
k : truck capacity
λ_i : target number of vehicles for zone $i \in \boldsymbol{J}$ according to the efficiency objective
μ_i : target number of vehicles for zone $i \in \boldsymbol{J}$ according to the equity objective
c_{ij} : cost of travel from zones i to j
m_i : minimum number of vehicles that must be present in zone i according to equity or efficiency considerations
ω_z : weight given to objective z ($z = (1,2,3)$, with 1 referring to the transportation cost component, 2 to the efficiency objective, and 3 to the equity objective)

Variables:

x_{ij} : 1 if the truck goes from zones i to j; 0 otherwise
y_{ij} : number of micromobility vehicles moved from zones i to j
n_{iz} : vehicle deficits in zone i based on objective z

Objective Function:

$$Minimize\ Z = \omega_1 \sum_{i \in \boldsymbol{J}} \sum_{j \in \boldsymbol{J}} c_{ij} x_{ij} + \omega_2 \sum_{i \in \boldsymbol{J}} n_{i2} + \omega_3 \sum_{i \in \boldsymbol{J}} n_{i3} \tag{6}$$

Constraints:

$$\sum_{j \in \boldsymbol{J}} x_{0j} = 1 \tag{6.1}$$

$$\sum_{i \in \boldsymbol{J}} x_{i0} = 1 \tag{6.2}$$

$$\sum_{i \in \boldsymbol{J}} x_{ir} - \sum_{j \in \boldsymbol{J}} x_{rj} = 0 \qquad \forall r \in \boldsymbol{J} \tag{6.3}$$

$$\sum_{i \in \boldsymbol{J}} x_{ii} = 0 \qquad \forall i \in \boldsymbol{J} \tag{6.4}$$

$$k x_{ij} \geq y_{ij} \qquad \forall i, j \in \boldsymbol{J} \tag{6.5}$$

$$s_i + \sum_{j \in \boldsymbol{J}} y_{ji} \geq \sum_{j \in \boldsymbol{J}} y_{ij} \qquad \forall i \in \boldsymbol{J} \tag{6.6}$$

$$s_i + \sum_{j \in \boldsymbol{J}} y_{ji} - \sum_{j \in \boldsymbol{J}} y_{ij} \geq m_i \qquad \forall i \in \boldsymbol{J} \tag{6.7}$$

$$n_{i2} \geq \lambda_i - \left(s_i + \sum_{j \in \boldsymbol{J}} \left(y_{ji} - y_{ij} \right) \right) \qquad \forall i \in \boldsymbol{J} \tag{6.8}$$

$$n_{i3} \geq \mu_i - \left(s_i + \sum_{j \in \boldsymbol{J}} (y_{ji} - y_{ij}) \right) \qquad \forall i \in \boldsymbol{J} \qquad (6.9)$$

$$y_{0j} = 0 \qquad \forall i \in \boldsymbol{J} \qquad (6.10)$$

$$x_{ij} \in \{0,1\} \qquad \forall i,j \in \boldsymbol{J} \qquad (6.11)$$

$$n_{i2}, n_{i3}, y_{ij} \in \mathbb{Z}^+ \cup \{0\} \qquad \forall i,j \in \boldsymbol{J} \qquad (6.12)$$

Objective function (6) minimizes the costs associated with the truck routing and deficits resulting from not meeting the target distributions. Constraint (6.1) and (6.2) ensure that the truck starts and ends in zone 0 (the depot or parking zone). Constraints (6.3) ensure flow conservation. Constraints (6.4) prevent dwelling or loops in the same zone. Constraints (6.5) ensure that pickups and drop-offs only occur when leaving or arriving at a zone and that the number of vehicles moved does not exceed the truck capacity. Constraints (6.6) ensure that the vehicle does not pick up more than the number of vehicles in each zone and they prevent negative inventory. Constraints (6.7) ensure that a given vehicle supply m_i is available in each zone i ($m_i \geq 0$), which, as previously mentioned, is an alternative approach to capture equity considerations. Constraints (6.8) and (6.9) give the deficit value at the end of the route based on the respective targets. These constraints consider the zones' initial vehicle supply, the vehicle movements in and out to the zones, and the target levels. Constraint (6.10) blocks the flow of vehicles from the depot. Finally, constraints (6.11) and (6.12) define the characteristics of the decision variables.

The ω_z weights can be defined using different analysis perspectives. For example, the transportation costs and the efficiency deficit terms could be combined into a measure of loss of profit. Or the problem could be solved with different sets of weights and the rebalancing decision would be made based on the different rebalancing plans generated (the feasibility of this idea would depend on the computational cost of solving a single PDP instance).

3.3. Numerical Experiments

Two sets of experiments were conducted to illustrate the application of the proposed models. In the first set of tests, the outputs of the PDP were examined, and in the second set of tests, the two-step methodology was incorporated in a simulation of e-scooter demand in MDES. In both sets of tests, OR-Tools, and specifically the CBC solver, were used to solve the PDP (Perron & Furnon, 2019).

3.3.1. PDP Tests

In the PDP tests, it was assumed that the service of interest operates in a 20-zone service area, the vehicle fleet has a size of 220, the service truck could carry up to 10 vehicles, and the minimum number of vehicles m_i for each zone was set to 60% of their λ_i values. Details on the scenario considered in the PDP experiments, including the number of scooters in the zones and the λ_i and $\mu_i$values used, can be found in Appendix D. The PDP was solved ten times, each time using a different set of weights. In all trials, ω_1 was set to zero; ω_2 sequentially varied from 0 to 1 with increments of 0.1, and ω_3 was defined as $1 - \omega_2$. This trial configuration was used to focus on the tradeoffs between the efficiency and equity objectives. Figure 19 reports the results of the PDP tests. In the experiments, as more weight is given to the equity objective, the predicted deficit from the efficiency perspective increases, and vice versa. Extreme deficits are observed when either ω_2 or ω_3 equal 1, but interestingly the deficit values cluster for all other weight sets, with the efficiency deficit being lower in 90% of cases given the definition of the m_i terms; had m_i been defined based on the μ_i values the opposite would have been true. Naturally, this is not a generalizable observation. The results do not generally imply that meeting equity objectives results in a substantial loss in efficiency, or vice versa, (as illustrated in the next set of numerical experiments) or that zero deficit values are always achievable.

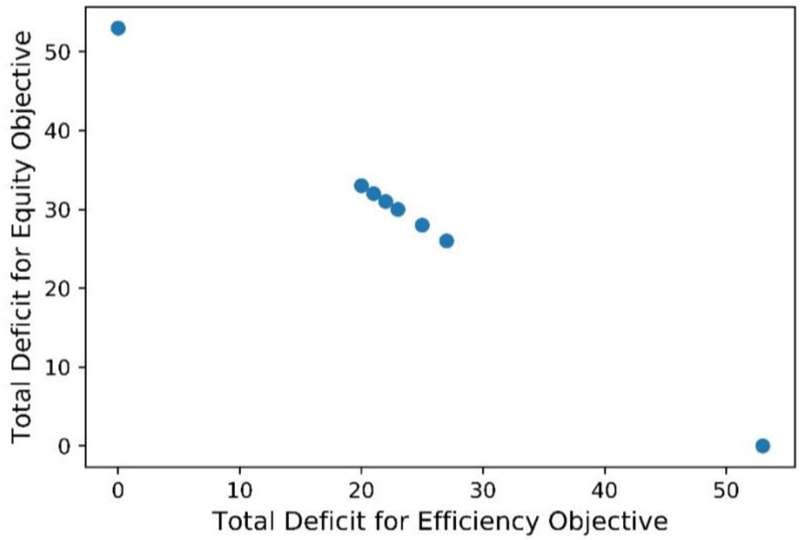

Figure 19. Efficiency and equity deficits in PDP tests

3.3.2. Simulation Tests using Agent-Based Model

The rebalancing methodology was embedded within a computer program that simulated e-scooter travel demand using MDES as the background setting. The program's inputs were generated using data associated with MDES, including the region's transportation network and land use information, and the e-scooter travel patterns. As there are no detailed travel behavior data for the region, model structures and parameters were assumed to set up the travel behavior models.

The e-scooter fleet size was set to 100, similar to the fleet size in MDES. The zonal system used in the simulation is presented in Figure 20. There are 219 zones in which simulated agents act (identified as the "zones" in the figure) and 15 regions that serve as the pickup zones from the optimization models' perspective (more details on these regions next). The e-scooter pickup and drop-off decisions were made at the level of the 15 regions (and therefore $\boldsymbol{\lambda}$ and $\boldsymbol{\mu}$ have dimensions 1×15). Note that there is a drop-off zone associated with each region.

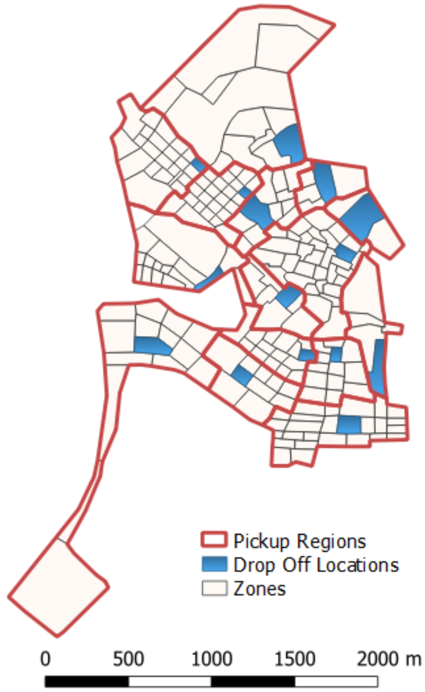

Figure 20. ABM zonal system

Figure 21 presents the structure of the simulation (in the figure, t_{max} refers to the last simulation period, and go is a Boolean variable that indicates if a rebalancing plan is active or not). Time in the computer model was discretized into 5-minute blocks, with the day starting at 6:00 AM and ending at 8:00 PM. In each period t (i.e., program iteration), the travel behavior of synthetic agents was simulated, the location of e-scooters in the system was updated, and then the program executed a rebalancing plan if one was active, or it generated a rebalancing plan if the iteration corresponded to a period within the set $\boldsymbol{T_r}$ of predetermined rebalancing periods. The rebalancing period set was defined as $\boldsymbol{T_r}$ ={8:00 AM, 10:00 AM, 12:20 PM, 2:20 PM, 4:00 PM}. If $t \in \boldsymbol{T_r}$, the target distributions were generated using the TDP. The efficiency target distribution $\boldsymbol{\lambda}$ was defined as the vehicle distribution expected to maximize the number of trips in a two-hour time horizon, as predicted by a trip departure model generated by the XGBoost algorithm (see Appendix B). The equity target distribution $\boldsymbol{\mu}$ was defined as the distribution that minimized the AUC-based Atkinson inequality index, with ε set to 0.75. The AUC was computed as discussed in Chapter 1. Instead of searching for the optimal $\boldsymbol{\mu}$ each time the rebalancing model was invoked, a fixed $\boldsymbol{\mu}$ was determined. It was assumed that the $\boldsymbol{\mu}$ that minimized the inequality in spatial access could be treated as the static minimum number of e-scooters required by each region. The PDP's m_i values were set to zero in the simulation. The TDPs were solved using the heuristic discussed in Appendix C.

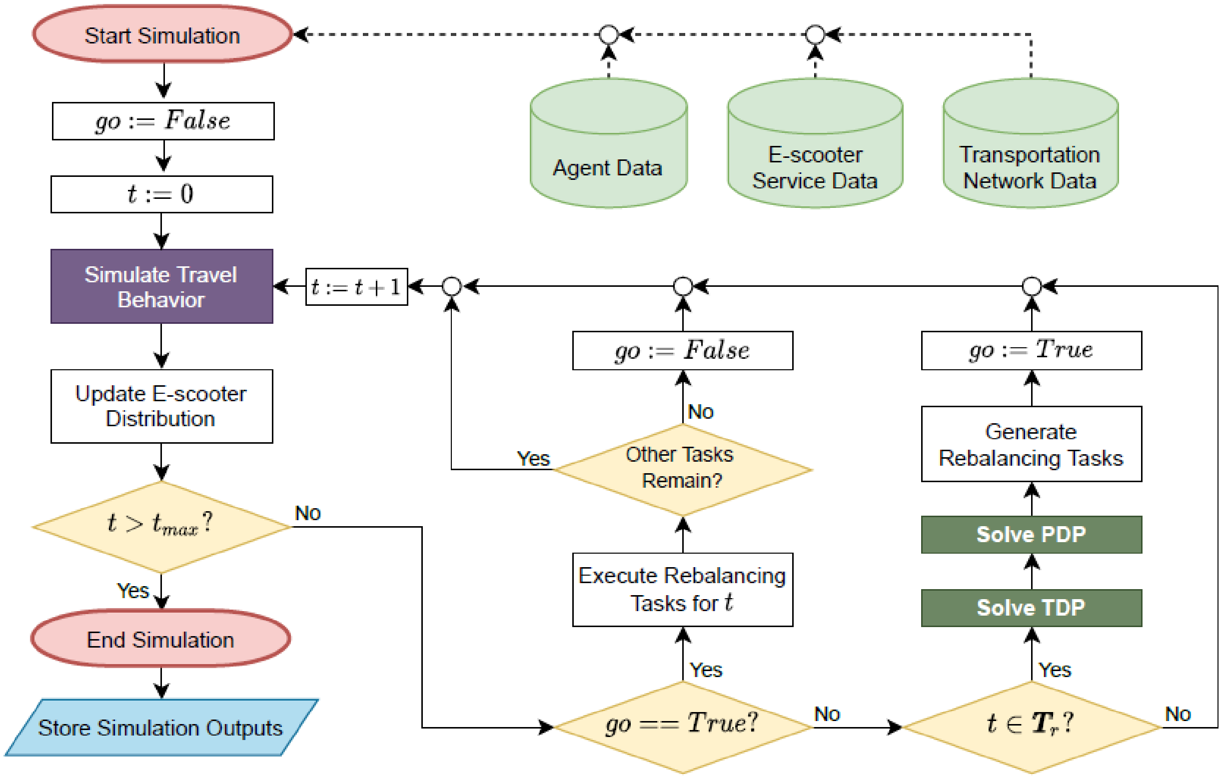

Figure 21. Framework of the computer simulation program

Given the $\boldsymbol{\lambda}$ and $\boldsymbol{\mu}$ e-scooter distributions, the PDP was solved, and the implementation of the vehicle pickup and drop-off plan generated by the model was simulated. The travel time between pickup and delivery zones was accounted for in the simulation, although, for simplicity, instantaneous vehicle pickup and drop-offs were assumed. Vehicle pickups were performed at the level of the 15 pickup regions, meaning that if, for example, 10 e-scooters needed to be picked up from a region, those e-scooters would be removed from the region's zones in which the agents left them. The truck in this simulation was assumed to have the capacity to carry up to 40 e-scooters.

The travel behavior of the synthetic agents was simulated using a simple activity-based model (ABM). Each synthetic agent was assigned a fixed schedule that contained the agent's location (i.e., zone) for each 5-minute period of the day, including the periods necessary for the agent to travel between zones. As previously mentioned, the zonal system in which the agents operated was constituted by 219 zones. Each agent was also given a home zone; the agent's schedule began and ended in the home zone. For each trip, the ABM simulated an agent's mode choice decision using Monte Carlo simulation. This simulation depended on the mode choice probabilities generated by a logit model. Only two modes were considered: walk and e-scooter. Each agent was assigned a set of parameters to compute their mode choice probability. The deterministic utility of the walk mode was a function of a constant and the walking travel time between a trip's origin and destination. The e-scooter mode's deterministic utility was a function of the travel time and cost associated with the e-scooter closest to the agent, the number of previous e-scooter trips performed by the agent, and whether the agent was returning home. Agents were also labeled as UPRM students or non-UPRM students. A total of 3,700 agents were generated for the simulation.

The simulations were run under three scenarios: i) no rebalancing scenario, ii) scenario with rebalancing weights set to $\{\omega_1 = 0.1, \omega_2 = 2, \omega_3 = 0\}$, and iii) scenario with rebalancing weights set to

$\{\omega_1 = 0.1, \omega_2 = 0, \omega_3 = 2\}$. The computer model was run 15 times for each scenario. Note that in the second scenario the primary goal of the rebalancing model was to achieve the $\boldsymbol{\lambda}$ distribution (and therefore maximize trips), while in the third scenario the primary goal was to achieve the $\boldsymbol{\mu}$ distribution (and therefore minimize the inequality in the spatial access to the e-scooters). The results of the scenario's trials runs are summarized in the boxplots presented in Figure 22. In the simulation, on average, focusing the rebalancing operations on maximizing the number of trips increased trips by 45%, relative to the no-rebalancing scenarios, which was only slightly higher than the 43% improvement obtained when the rebalancing model was focused on achieving the equity goal. However, there was a significant difference in the equity-based performances; relative to the no-rebalancing runs, the trip-focused model resulted in a reduction of the AUC-based Atkinson inequality index of 15%, while the equity-focused model resulted in a reduction of 30%. As in the PDP tests, these tests do not provide general insights into the expected performances of real-world dockless micromobility services, but the results illustrate that there could be situations in which seeking a more spatially equitable distribution of micromobility vehicles could result in a manageable loss in performance efficiency.

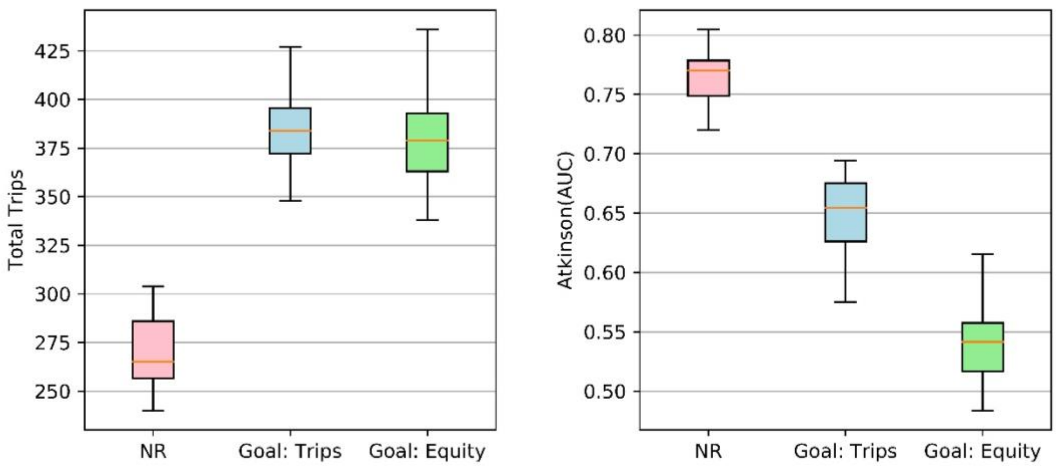

Figure 22. Total trips and average Atkinson index values for simulation under different rebalancing regimes (NR: no rebalancing scenario; "Goal: Trips": $\{\boldsymbol{\omega_1 = 0.1, \omega_2 = 2, \omega_3 = 0}\}$ scenarios; "Goal: Equity": $\{\boldsymbol{\omega_1 = 0.1, \omega_2 = 0, \omega_3 = 2}\}$ scenarios)

3.4. Closing Remarks

An optimization-based framework was proposed for vehicle rebalancing operations based on efficiency and equity objectives relevant to dockless micromobility services. This quantitative approach can be a complement to the essential community engagement work required to identify the barriers faced by people when accessing micromobility services. As discussed by Shaheen et al. (2017), there are spatial, temporal, economic, physiological, and social barriers that can hinder the full participation of individuals in shared mobility services. Naturally, the community engagement work is also fundamental for specifying the objectives and constraints of the optimization-based framework, particularly when selecting the definitions and measures of equity and access. More work is required, for example, to examine if voluntary data-sharing programs designed to enhance the equity-based operations of micromobility services are acceptable among different community groups, or, more generally, if people are interested in efforts to guide micromobility operations based on equity measures that use person-level information. More work is also required to quantify what are the costs of different equity enhancing strategies in micromobility services, including the approaches proposed in this report, and to determine how these costs can be covered to ensure that the operations of microbiology services are economically sustainable.

The algorithms used to solve the optimization problems (TDP and PDP) are likely to be the main challenge to the practical implementation of the proposed methodology. In the context of the experiments presented in Section 3.3 and of the small-scale services like MDES, the proposed heuristics are sufficient. However, given that the presented models are supposed to operate in real-time decision-making applications, more work is required to identify or develop heuristics that can be used to quickly find good solutions to the proposed models in large-scale systems.

4. Conclusions and Potential Research Opportunities

This project explored the issues of equity and access in micromobility services by studying the experience of a dockless e-scooter service in Puerto Rico and by proposing a rebalancing model that considered spatial and social inequity in the distribution of vehicles. In Chapter 2, an analysis of the dockless micromobility experience of Mayagüez, PR, was presented. In this analysis, MDES user and nonuser characteristics were identified, and the spatiotemporal patterns of e-scooter trips were explored. In addition, regression analyses were used to examine the relationship between the sociodemographic and land-use characteristics of the region and the demand levels observed in MDES. Lastly, spatial access indicators were developed and applied to the MDES case.

The key findings of the research presented in Chapter 2 are:

- MDES users tended to be young and male, and trips were concentrated in and around a university's campus.
- Costs, the built environment, and safety concerns were identified as the main reasons for not using MDES.
- Traffic congestion and lack of parking spaces were identified among the main reasons for using MDES, which suggest that the service reduced the number of auto trips in the service area. The magnitude of the auto trip reduction is unclear, but it could point to the congestion reduction potential of micromobility services, particularly in cities that lack effective public transportation services.
- Spatial access differences were observed among the neighborhoods in MDES. These differences are probably the results of the spatial distribution of MDES main users (UPRM students), the sociodemographic characteristics of the population (generally low-income and older), and the rebalancing operations aimed, naturally, at satisfying user demand.

In Chapter 3, a two-step methodology was presented for conducting rebalancing operations according to efficiency and equity objectives relevant to micromobility systems. An optimization model was proposed for identifying target micromobility vehicle distributions to achieve efficiency and equity goals in the performance of a system. Two objective function formulations were proposed to account for spatial and social equity considerations. Additionally, a pickup and delivery problem was proposed that balances efficiency and equity objectives in the search for a vehicle redistribution plan. The methods and concepts presented in this chapter contribute to the body of literature on the use of quantitative methods to design, plan, and operate systems considering equity. The simulation results suggest that, relative to efficiency-focused rebalancing, there are scenarios in which equity-focused rebalancing operations could result in minor reductions in total trips and significant improvements in spatial access to micromobility services. Additional research is necessary to reach generalizable conclusions on the likely effects of equity-focused rebalancing operations on the overall performance of dockless micromobility services.

This research contributes to real-world practice by presenting network-based methods and metrics that can be applied in the equity evaluation of dockless micromobility services. It also helps planners, engineers, and community organizers by providing additional real-world evidence that can be used to advocate for investments in infrastructure that can accommodate bicyclists, scooter users, and other travelers that do not rely on automobiles. Additionally, the research points to additional types of equity performance requirements that could be included as part of the operational goals that micromobility companies must satisfy.

There are several future research opportunities connected to the work presented in this report. Field interviews and surveys of MDES users, along with naturalistic observations, can be conducted to explore in more detail their characteristics and the reasons for using the service. In addition, this research would help to quantify the magnitude of the auto trip reductions caused by MDES and the subsequent impacts on traffic congestion, if any. Outside Mayagüez, dockless e-scooter services have begun operations in at least four other Puerto Rican cities, including the capital San Juan. The system in San Juan offers opportunities for micromobility travel demand research in urban environments with relatively large tourist populations and unreliable public transportation options. Beyond studies regarding the micromobility experience in PR, comparative studies using the proposed graph-based spatial equity measures could also be performed to gain insights on the impact of land use and transportation network infrastructure in the spatial access to micromobility.

Furthermore, research opportunities in the development of optimization models to enhance equity and access. Extensions to the models presented in Chapter 3 can be proposed to consider the presence of other modes (i.e., buses) in equity-conscious rebalancing operations. An adaptive model structure that generates the $\boldsymbol{T_r}$ would be particularly useful for operators. In addition, models can be formulated to help operators and city planners design the service areas of micromobility services. Service area requirements, including the inclusion of historically disadvantaged communities, are a common condition set by cities to allow the operation of micromobility services. Even without equity requirements, service area design is a concern for companies that want to ensure that their services generate a profit. Also, new heuristics are required to speed up the discovery of good solutions to the rebalancing optimization models, particularly for applications in large-scale micromobility systems.

Lastly, and perhaps more challenging, models can be developed to optimally make decisions on total subsidy levels for micromobility trips and on the real-time distribution of these subsidies to enhance access among low-income population groups. As the MDES experience suggests, costs are among the main barriers to access micromobility services and enhancing spatial access does not address this problem. Considerable work has been completed on algorithms to make real-time pricing decisions to maximize profits; future research can explore the potential for adapting these models to maximize equity and access.

Appendix A. Algorithm to Process Trip Start and End Coordinates

GPS errors are a fundamental challenge in the study of travel patterns in micromobility systems. In this project, GPS errors and the resulting inconsistencies in trip starting and ending locations were the main problems with the trip record data. For example, in approximately 51% of cases, an e-scooter's ending and starting locations in sequential trips were more than 150 meters apart, and these differences were not consistent with rebalancing operations (if there were no errors, the ending coordinate of trip i would be the same as the starting coordinate of the next trip, $i+1$). These coordinate differences would not be an issue in situations in which the size of the areas under analysis and the average trip lengths render them unimportant. In the current application context, however, the service area is relatively small, and the zones within this area are, naturally, even smaller. Therefore, an algorithm was developed with the objective of adjusting the recorded starting and ending trip coordinates in order to produce plausible trip sequences. The proposed procedure could be considered as a type of map-matching algorithm (Quddus, Ochieng, & Noland, 2007), but in this problem, the only data available are the starting and ending latitudes and longitudes of a sequence of trips, rather than arrays of GPS data that contain starting, intermediate, and ending coordinates (i.e., trajectories).

Next, the steps of the developed algorithm are described. Underlying this algorithm is the assumption that the e-scooter starting coordinates were more reliable than the e-scooter ending locations. This assumption was based on the observation that the starting coordinates were spatially clustered on a somewhat discrete number of regions known to be common e-scooter origins and destination, whereas the ending coordinates were spatially dispersed in travel ways (e.g., roadways), suggesting parking patterns that have not been observed in the city (particularly, in UPRM).

Step 0. Define the feasible e-scooter parking space

Define the spaces in which e-scooters can be parked. These spaces could include sidewalks, plazas, parking lots, or any space that the analyst considers as a place where it would be reasonable to expect e-scooters to be parked. Let $\boldsymbol{\Psi}$ represent the set of these feasible parking spaces. In the Mayagüez context, all sidewalks within the service area, as well as plazas and corridors within UPRM, were defined as feasible parking spaces. These spaces were identified and represented as polygon objects using GIS.

In addition, as part of Step 0, identify a set of coordinates $\boldsymbol{\Theta}$ that represents the locations in which the operator places the e-scooters at the start of the day. Again, in the current context, these locations were selected by first creating a GIS map with the starting coordinates of the first trip of all e-scooters on all days, and then identifying the centers of the resulting coordinate clusters; the center coordinates constituted the $\boldsymbol{\Theta}$ set.

The steps that follow (Steps 1-5) are part of a loop that repeats itself for each day in the data set and for each scooter. In this discussion, let $\boldsymbol{e}_{nvi}$ represent the recorded ending latitude and longitude of trip i (where i is the index of trips in an ordered sequence of trips) for e-scooter v on day n, and $\boldsymbol{s}_{nv,i+1}$ denote the recorded starting latitude and longitude of trip $i+1$. Also, let r_{max} represent the maximum distance from a recorded coordinate within which it would be reasonable to expect the true coordinate; that is, r_{max} is an error radius. This parameter could be defined based on observed GPS error ranges (Caltrans, 2020).

Step 1. Find a reasonable parking spot within the space proximal to $\boldsymbol{e}_{nvi}$ and $\boldsymbol{s}_{nv,i+1}$

For the starting coordinate $\boldsymbol{s}_{nv1}$, set the adjusted starting coordinate as the coordinate in $\boldsymbol{\Theta}$ that is closest to $\boldsymbol{s}_{nv1}$. For the subsequent coordinate pairs $(\boldsymbol{e}_{nvi}, \boldsymbol{s}_{nv,i+1})$, find the geodesic distance d_{nvi} between the pairs. If $d_{nvi} > 2 \times r_{max}$, go to Step 2. Otherwise, generate two circles of radius r_{max}, one with center at $\boldsymbol{e}_{nvi}$ and the other with center at $\boldsymbol{s}_{nv,i+1}$. Then, determine the intersection between $\boldsymbol{e}_{nvi}$'s circle, $\boldsymbol{s}_{nv,i+1}$'s circle, and $\boldsymbol{\Psi}$ (see Figure A1). If this procedure results in a null set (i.e., there is no feasible parking space), go to Step 2. Otherwise, select the parking spot closest to both $\boldsymbol{e}_{nvi}$ and $\boldsymbol{s}_{nv,i+1}$, and return the coordinate of this spot as the adjusted coordinate of $\boldsymbol{e}_{nvi}$ and $\boldsymbol{s}_{nv,i+1}$ (again, if the were no errors, $\boldsymbol{e}_{nvi}$ and $\boldsymbol{s}_{nv,i+1}$ would be the same). In this study, the implementation of this intersection procedure would result in a set of polygons from $\boldsymbol{\Psi}$. The polygons were then split using a polygon triangulation algorithm, and the centroids of the resulting triangles would be determined; these centroids constituted the candidate parking spots. The centroid coordinate with the minimum combined distance to the $(\boldsymbol{e}_{nvi}, \boldsymbol{s}_{nv,i+1})$ pair was returned as the adjusted coordinate for both $\boldsymbol{e}_{nvi}$ and $\boldsymbol{s}_{nv,i+1}$. Also, r_{max} was set to 100 meters.

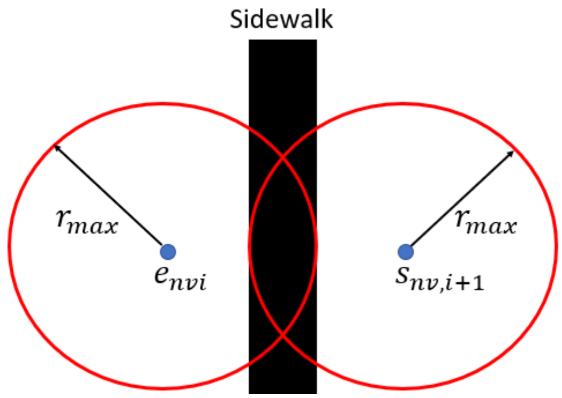

Figure A1. Illustration of the intersection procedure to identify feasible parking spots

Step 2. Anchor coordinates based on the $\boldsymbol{s}_{nv,i+1}$ or declare a rebalanced e-scooter

Generate a circle with radius r_{max} and center on $\boldsymbol{s}_{nv,i+1}$, and find the intersection between $\boldsymbol{s}_{nv,i+1}$'s circle and $\boldsymbol{\Psi}$. If this intersection procedure results in a null set, go to Step 3. Otherwise, select the parking spot closest to $\boldsymbol{s}_{nv,i+1}$, and return the coordinate of this spot as the adjusted coordinate of $\boldsymbol{s}_{nv,i+1}$ (the implementation of this procedure in this study was as explained in Step 1). Then, consider if it would be feasible for the previous trip end $\boldsymbol{e}_{nvi}$ to adopt the new adjusted coordinate given to $\boldsymbol{s}_{nv,i+1}$. To do this, find the length of the shortest path between $\boldsymbol{e}_{nvi}$ and $\boldsymbol{s}_{nv,i+1}$ on network $\boldsymbol{G}$, and divide it by the recorded travel time of trip i. If the resulting speed is less than or equal to a selected upper bound speed (e.g., the maximum e-scooter speed) and greater than or equal to a lower bound speed, then set $\boldsymbol{e}_{nvi}$ as the adjusted coordinate of $\boldsymbol{s}_{nv,i+1}$, and return this information. Otherwise, find an adjusted coordinate for $\boldsymbol{e}_{nvi}$ by applying the previous intersection-based procedure described for $\boldsymbol{s}_{nv,i+1}$. This implies that the e-scooter was moved by the operator from the adjusted $\boldsymbol{e}_{nvi}$ to the adjusted $\boldsymbol{s}_{nv,i+1}$, presumably as part of a rebalancing operation.

Step 3. Anchor coordinates based on the $\boldsymbol{e}_{nvi}$ or declare a rebalanced e-scooter

Repeat the procedure described in Step 2, but i) use $\boldsymbol{e}_{nvi}$ as the anchor, instead of $\boldsymbol{s}_{nv,i+1}$, ii) go to Step 4 if the intersection procedure results in a null set, and iii) check the implied speed from the adjusted coordinates relative to the information of trip $i+1$.

Step 4. Discard trip records

Assuming that the level of detail suggested by this algorithm is necessary, reaching this step means that no reasonable parking spot was identified for the coordinates. Therefore, the data is discarded. In this study, no data were discarded based on this criterion as the algorithm did not reach Step 4. Presumably, this would be the case in most urban areas given their roadways and sidewalk densities.

Step 5. Last $\boldsymbol{e}_{nvi}$ coordinate

Note that there is no $\boldsymbol{s}_{nv,i+1}$ for the last trip i, so the previous procedure does not work to adjust the last trip's $\boldsymbol{e}_{nvi}$ coordinate. However, the intersection procedure described in Step 2 can be applied to assign the final coordinate adjustment for the e-scooter.

Appendix B. Predictive Accuracy of Machine Learning Models in the Context of MDES Operations

Machine learning (ML) algorithms, including state-of-the-art deep learning algorithms, are at the core of the operations of emerging transportation network companies, including companies that operate dockless micromobility services. In academic research, these methods are commonly proposed as part of methods to forecast demand and, by extension, guide the rebalancing operations in micromobility systems. In this appendix, research to train machine learning-based models using MDES data is discussed. The objectives of this work were:

i. to demonstrate the predictive power of ML algorithms in situations in which there are relatively small datasets available to train the models,
ii. to support the assumption that ML algorithms could be used to generate the models required by the TDP.
iii. and to generate ML models that could be used in the numerical tests.

This appendix is divided into six sections. In the first section previous ML studies that consider micromobility services are reviewed. This is followed by a discussion of the prediction problems of interest, the available data, and the ML algorithms used in the project. The last two sections present the results of the prediction tests and discuss possible research directions.

B.1. Previous Studies

Statistical and machine learning methods, including more recent deep learning approaches, have been developed to understand and forecast demand for micromobility services. For SBS systems, extensive research has been conducted on the relationship between the sociodemographic characteristics of individuals – as discussed in Section 2.1. Findings from these studies have informed policy discussions and resulted in equity requirements for micromobility services. Beyond understanding travel behavior, models have also been developed to forecast station-level demand using primarily historical trip record data. For example, Médard de Chardon and Caruso (2015) proposed regression-based models to estimate station-level bike trips at different levels of temporal aggregation. As part of a model to solve the dynamic rebalancing problem, Regue and Recker (2014) used gradient boosting machines to produce short-term demand predictions at the SBS station level. In addition to creating variables based on the historical trip data, these researchers introduced weather-related features in the prediction models. Time series (Kaltenbrunner, Meza, Grivolla, Codina, & Banchs, 2010), Bayesian network (Froehlich, Neumann, & Oliver, 2009), and neural network (Caggiani & Ottomanelli, 2012) approaches are among other types of models that have been applied to forecast SBS trip levels.

The large-scale datasets generated by dockless micromobility services have enabled the application of deep learning methods to generate dynamic, short-term demand forecasts. For example, Xu et al. (2018) developed long short-term memory neural networks (LSTM-NN) to forecast DBS trip production and attraction at the traffic zone level. The model considered time intervals as short as 10 minutes. A convolutional LSTM-NN was used by Ai et al. (2019) to forecast SBS trips in Chengdu, China. The city was divided into a grid of equally sized cells of size 4 km × 4 km, and the temporal dimension was divided into six time periods. He and Shi (2020) proposed a graph-based neural network to predict DES flows between city zones. Using deep neural networks, Yan and Howe (2019) proposed a model that predicts the value of a linear combination of demand and a fairness metric. They propose two fairness metrics that quantify the gap between the vehicles assigned to an advantaged group and the vehicles assigned to a disadvantaged group. In a different type of application, Pan et al. (2019) developed a deep reinforcement learning framework that determines how much users should be paid to help in the rebalancing of a dockless bike-sharing system.

B.2. Data Description

Two types of data were used for the ML training and testing task: MDES observed and simulated data. The source of the observed data, as discussed in Section 2.3.1, was the MDES operator; Skootel provided records on approximately 66,000 e-scooter trips completed during the 2019-2020 academic year. These data were processed to correct GPS errors (see Appendix A). As part of the ML database preparation, for each 5-minute period and each day of operation, counts of trip arrivals and departures for each zone in the 12-zone system presented in Figure B1 were generated; the 12-zone system is called the B1 system hereafter. To complement the trip record data, information on UPRM course sections (e.g., class start and end times, classroom location), the number of students enrolled in each section, and the dates of holidays were collected. Lastly, data on rain events during the period of analysis was also collected.

B1. Zonal system used in the ML algorithm tests

The simulated data was generated using the output of the activity-based model (ABM) discussed in Section 3.3.1. The ABM was run 120 times (akin to 120 days of observation). In each trial run, an ad-hoc rebalancing plan (i.e., plan based on knowledge of selected simulation parameters) was used to adjust vehicles during the run. As with the observed MDES data, the ABM simulation generates trip information that was processed to generate zonal level counts of trip arrivals and departures for each 5-minute period in each model run (instead of each day). In addition, the AUC-based system-wide inequality indicator was computed for each 5-minute period in each model run. Simulated revenue information was also collected.

B.3. Prediction Problems and Feature Engineering

The prediction problems of interest are:

i. Prediction of trips at the service area (SA) level using observed data.
ii. Prediction of trip arrivals and departures at the B1 zonal system level using observed data.
iii. Prediction of trip arrivals and departures at the ABM zonal system level using simulated data.
iv. Prediction of AUC-Atkinson indicator (hereafter, the AUC) and revenue generation at the ABM zonal system level using simulated data.

The initial set of tests focused on the prediction of trip departures at the SA and B1 zonal levels. This initial set of tests was performed to identify the most accurate ML algorithms to use in the rest of the prediction problems and the numerical tests discussed in Section 3.3. For the initial tests, the SA models were developed to predict the hourly number of trips generated in the MDES service area. In the case of the B1 models, trip departures were estimated for each of the 12 zones in the B1 zonal system and each of the eight consecutive two-hour periods in a day of operation. The start of the first two-hour period was set at 5:45 AM. The features considered in the model development process are reported in Table B1. In general, dummies were included for days of the week, time of day, zones, rainfall events, and holidays. Historical averages, moving averages, and trip count observations from previous days were among the continuous variables incorporated in the models.

Table B1. Features used in Model Training

IDs	Feature Description	Type	Model	
			SA	B1
1	Day of Week	D	X	X
2	Time period (e.g., hour, two-hour period)	D	X	X
3	Level 1 zone	D		X
4	Month	D		X
5-7	Historical averages of trip productions for prior two-, three-, and four-hour periods	Q	X	X
8, 9	Seven-day averages of trip productions for the prior two- and three-hour periods	Q	X	X
10, 11	Moving averages of trip productions considering prior two- and three-hour periods	Q	X	X
12	Mean number of trip productions observed three days prior on same zone and period	Q		X
13	Mean number of trip productions observed two weeks prior on same zone, period, and day-of-	Q		X
14	Number of trip attraction in prior period and in the same zone	Q		X
15	Number of students enrolled in classes starting in periods after period of analysis	Q	X	X
16	Number of students enrolled in classes ending during period of analysis	Q	X	X
17	Rain event during period	D	X	X
18	Rain event in next 15 minutes (assuming weather forecasts can be used)	D		X
19	Rain event before trip on the same day	D		X
20	Rain event 15 minutes before period	D		X
21	Holiday	D	X	
22	Fall semester	D	X	X
23	Number of scooters deployed on the day	Q		X
24	Number of scooters in the zone during period	Q		X

Note: D stands for dummy (binary) variable, Q stands for continuous variable and X indicates that the variable was included in the model associated with the column.

B.4. Machine Learning Algorithms Applied

The size of the available dataset suggests, as a preliminary step, the use of standard ML algorithms, as opposed to more advanced deep learning approaches that generally require large-scale datasets (Scikit-Learn, 2020). Besides the classical linear regression model, the Bayesian ridge, ϵ-support vector machines (SVM), random forest, gradient boosting, AdaBoost, and XGBoost regression methods were applied for the prediction of trip productions at the SA and B1 models.

Next, a brief introduction is provided for the applied methods. Bayesian ridge regression is a method in which the output variable is assumed to be normally distributed, the regressor coefficients have a multivariate Gaussian prior, and the priors of the regularization parameters are Gamma distributed (Tipping, 2001). The basic objective of the applied SVM method is to attempt to find a function that produces deviations that are, at most, an ϵ value for all values being predicted in the training data (Smola and Scholkopf, 2004). The multi-layer perceptron is a simple class of feed-forward neural network that can be applied for regression (Jain, Mao, & Mohiuddin, 1996). The random forest, gradient boosting, AdaBoost, and XGBoost (T. Chen & Guestrin, 2016) methods are examples of ensemble algorithms that can be used for both classification and regression; ensemble here refers to their ability to combine the predictions of multiple, individual ML models (e.g., the random forest algorithm creates ensembles of decision trees). These methods were implemented using Python's scikit-learn library (Pedregosa et al., 2011).

B.5. Results

As standard practice suggests, the data were split into training datasets (used to train the models) and testing datasets (used to test the application and accuracy of the models). The coefficient of determination (R^2; computed by comparing observed versus predicted values), mean absolute error (MAE) and mean square error (MSE) were selected as the performance metrics for the regression models. For the first tests using the observed data, the data splits were performed by dividing the data into two discrete, time-contiguous blocks, with a single time point serving as the training/testing split boundary (i.e., the convenience sampling approach commonly used with time series (Reitermanová, 2010)). For both the SA and B1 models, a 70/30 split was used (70% training, 30% testing). Observations before August 28, 2019, were removed as they correspond to the first weeks of MDES operations, and therefore contain patterns driven by the novelty of the service. In Table B2, the performance metrics obtained by the regression models trained with the selected ML algorithms are presented. Among the SA models, the best performing model was the random forest model, in terms of R^2 (0.97), and the multi-layer perceptron model, in terms of MAE (6.18) and MSE (75.12) values. In terms of the R^2, MAE, and MSE values, the best performing B1 departure model was generated by the random forest algorithm, while XGBoost, random forest, and AdaBoost produced models with the best R^2, MAE, and MSE values.

Table B2. Performance Metrics for Models Trained with Observed Data

Algorithm	SA Model			B1 Model - Departures			B1 Model - Arrivals		
	R^2	MAE	MSE	R^2	MAE	MSE	R^2	MAE	MSE
Bayesian Ridge	0.81	6.18	78.0	0.86	1.61	7.91	0.68	2.39	16.64
Random Forest	**0.97**	6.33	88.0	**0.95**	**0.89**	**2.80**	0.85	**1.57**	7.33
Linear	0.81	6.34	79.2	0.86	1.61	7.91	0.68	2.39	16.63
Gradient Boosting	0.79	6.43	88.6	0.93	1.06	3.68	0.80	1.92	10.36
Multi-Layer Perceptron	0.82	**6.18**	**75.1**	0.86	2.03	7.92	0.58	2.78	21.45
AdaBoost	0.79	6.34	88.0	0.77	2.96	12.4	0.87	1.62	**6.25**
XGBoost	0.78	6.59	90.3	0.93	1.07	3.70	**0.87**	1.58	6.60

In Table B3, the results for additional tests performed with the XGBoost algorithms are reported. In these tests, the performance of the departure model was examined using two dynamic data splitting approaches that simulate real-life model applications. In the first approach, the training data grow as time progresses and the information of more trips is processed. With this expanding dataset, the model was trained. The model performance was tested using the data for the week following the last date in the training dataset. This is labeled as the expanding window approach. In the tests, the model was applied to make predictions over two-hour time horizons on each day of the test week. In the second approach, the training data has a fixed size, and the time window from which the training data is selected moves (or slides) as the model is trained and tested (again, using the data from the next week); this second approach is labeled the sliding window approach. The results suggest that in this system, and with the amount of data available, the explaining window approach results in better models. In Figure B1, the R^2 metric values obtained by the two departure models for the weeks in the analysis period are presented. This figure again shows that the performance of the model trained using an expanding data window was better and that the performance of both models degraded significantly during the winter break period when there were significantly fewer trips.

Table B3. Performance of XGBoost Departure Models Trained under Dynamic Data Splitting

Statistic	Expanding Window			Sliding Window		
	R^2	MAE	MSE	R^2	MAE	MSE
Mean	0.86	0.98	3.31	0.80	1.09	4.30
Standard Deviation	0.12	0.31	2.14	0.18	0.33	2.58
Min	0.49	0.36	0.31	0.14	0.43	0.45
Max	0.95	1.73	9.28	0.93	1.73	9.49

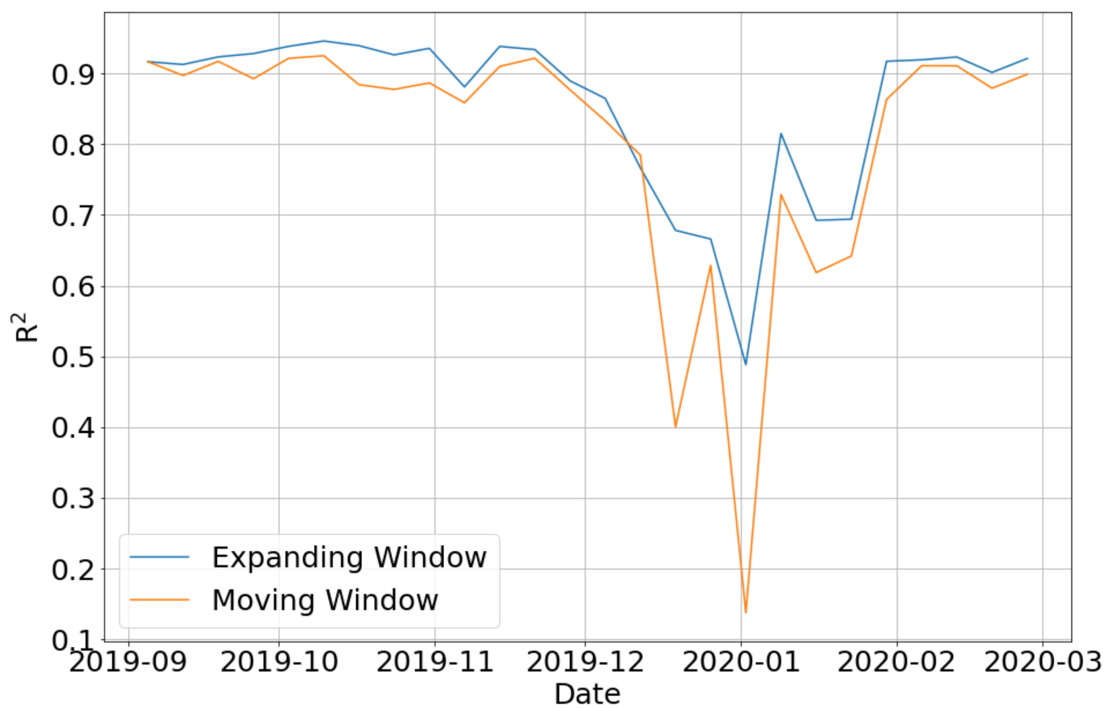

Figure B1. R^2 values (score) in time for XGBoost departure models trained under dynamic data splitting

The models trained using simulated data had even better performance metrics, which is not surprising as there are no unknown events in the simulation, and the trip behavior that is recorded is the product of an algorithm whose output, although not deterministic, is still bounded by a set of clearly defined rules. In all tests with the simulated data, the XGBoost algorithm was used given its good performance with the observed data and the fact that there are easy-to-use tools to tune the parameters of this algorithm. The R^2 values obtained for all the models trained with the simulated data (departure, attraction, AUC, and revenue models) were over 0.95. Figures B2 and B3 are presented as examples of the performance obtained for the AUC and revenue models using simulated data. Each data point is a prediction made for the value of interest (i.e., average AUC-based Atkinson index and total revenue) for a two-hour time horizon. As can be observed, there was a close relationship between the predicted and observed values in the simulation. The R^2, MSE, and MAE values for the AUC tests were 0.99, 2E-3, and 0.01, respectively, while for the revenue tests the metrics had values of 0.97, 62.7, and 5.7, respectively.

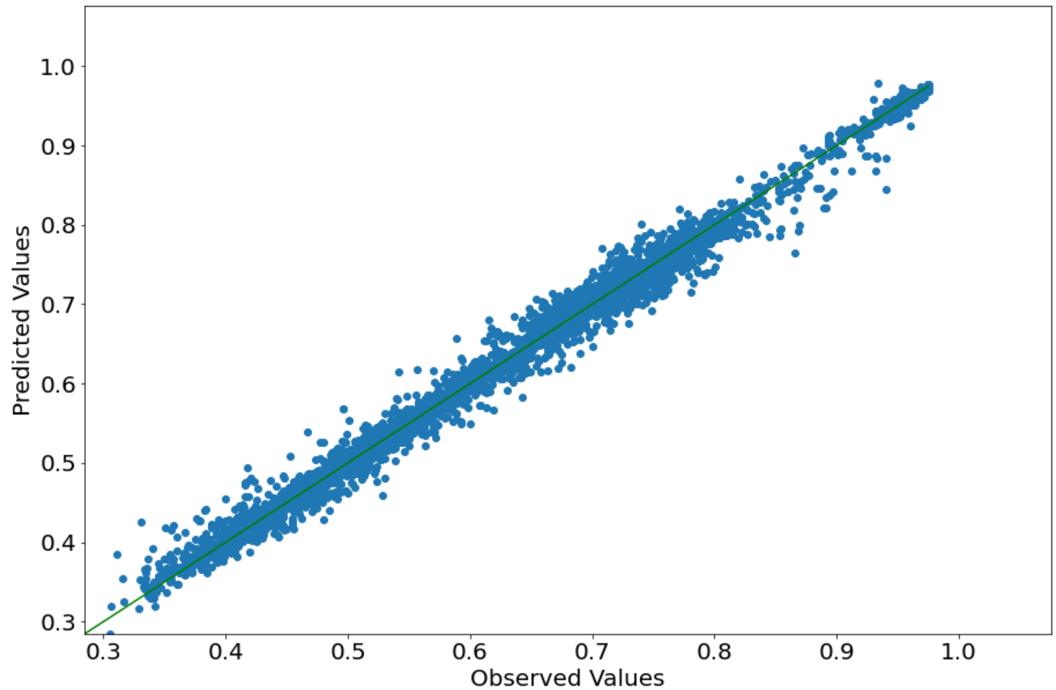

Figure B2. Predicted AUC versus observed AUC in the simulation

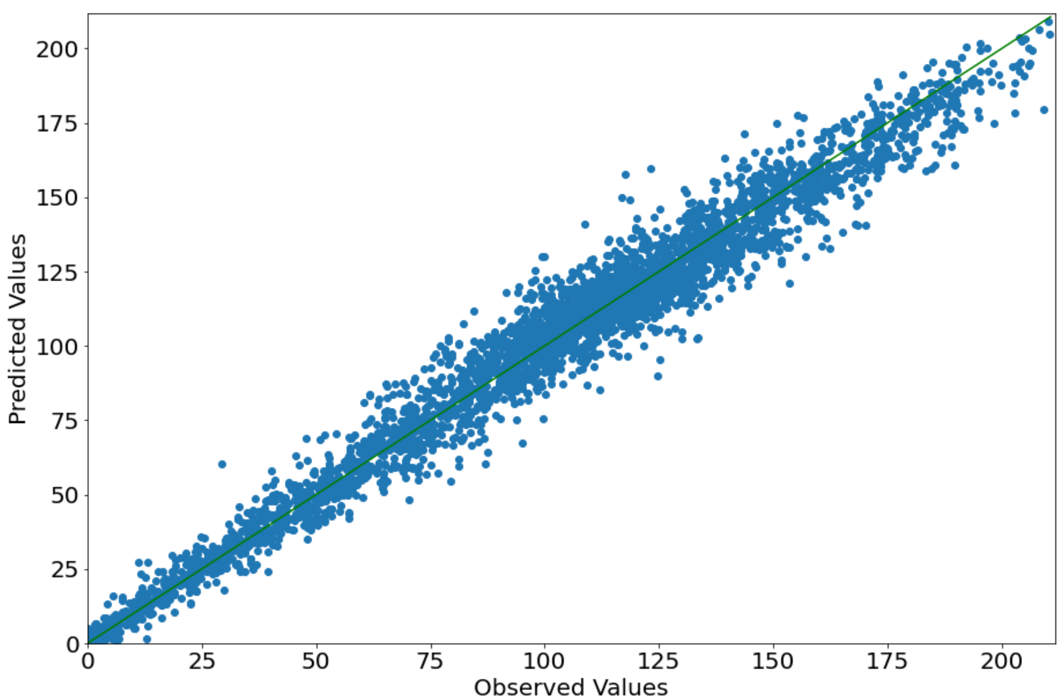

Figure B3. Predicted revenue versus observed revenue in the simulation

B.6. Closing Remarks

The results for the prediction tests reported in this appendix suggest that, even with limited data, like in the case of the MDES and simulation databases, standard ML algorithms generate accurate models. This supports the contention made in this project and in the reviewed literature that ML algorithms can be used as part of micromobility decision frameworks. This includes the use of models to predict the level of inequality in access in the system, as suggested in Chapter 3.

Appendix C. Heuristic for the Target Distribution Problem

A heuristic based on the differential evolution (DE) algorithm is proposed for the TDP. The DE algorithm is a type of evolutionary algorithm originally proposed for continuous optimization problems. Several extensions have been proposed for DE (Zhang & Sanderson, 2009), including DE approaches to solve integer programming problems. The DE-based heuristic presented here was designed with search strategies that effortlessly satisfy the TDP constraints.

As in other evolutionary algorithms, the DE algorithm generates a set of candidate solutions (offspring) by combining information from a set of previously evaluated solutions (parents). A solution in the TDP context refers to a vehicle distribution that satisfies the problem constraints. In each iteration of the proposed algorithm, an offspring solution is generated by applying two mutation operations. Let there be NP parent solutions, with the w-th parent information contained in vector $\boldsymbol{s}_w$ For each parent w in iteration n, an offspring $\boldsymbol{y}_w$ is generated by randomly selecting another parent solution v ($w \neq v$) from the pool of parent solutions and then applying the rule:

$$\boldsymbol{y}_w = \boldsymbol{s}_{nw} + f_n \cdot (\boldsymbol{s}_{nv} - \boldsymbol{s}_{nw}) \tag{C1}$$

where f_n ($f_n \in \boldsymbol{\Upsilon}$) is a combination factor whose value iteratively cycles through set $\boldsymbol{\Upsilon}$. This strategy is used to gradually shift the heuristic from exploitative search to explorative search (e.g., $\boldsymbol{\Upsilon} = \{0.1, 0.15, 0.2, 0.5\}$). Function (C1) shifts values (vehicle quantities) from the locations indicated by $\boldsymbol{s}_{nw}$ to the location indicated by $\boldsymbol{s}_{nv}$. Low values of f_n imply small changes to the distribution $\boldsymbol{s}_{nw}$ in direction to $\boldsymbol{s}_{nv}$, whereas large values of f_n create larger mutations in the direction of $\boldsymbol{s}_{nv}$. After each $\boldsymbol{y}_w$ is generated, a rounding function is used to ensure that all values in the vector are integers. If the sum of values in $\boldsymbol{y}_w$ is greater than the fleet size h, then a vehicle unit is removed from a randomly selected location (coordinate of $\boldsymbol{y}_w$) until the sum of $\boldsymbol{y}_w$ is equal to h. Alternatively, if the sum of values in $\boldsymbol{y}_w$ is less than h, then a vehicle unit is added to a randomly selected location until the sum of $\boldsymbol{y}_w$ is equal to h. Call this first set of mutation strategies the combination operation.

Having ensured that the values in $\boldsymbol{y}_w$ are integers and that they sum to the fleet capacity, the algorithm performs the second set of mutation strategies (the swapping operation) with probability p_{swap}. If the swapping operation is activated, coordinate i ($i \in \boldsymbol{J}$) is randomly selected among locations that have vehicles, coordinate j ($i \neq j, j \in \boldsymbol{J}$) is randomly selected among all possible locations, and then the vehicle quantities in each coordinate are updated according to:

$$\delta = \lceil \varphi y_{wi} \rceil \tag{C2}$$
$$y_{wi} \coloneqq y_{wi} - \delta \tag{C3}$$
$$y_{wj} \coloneqq y_{wj} + \delta \tag{C4}$$

where φ is a parameter that determines the magnitude of the vehicle swap between coordinates i and j.

Once the $\boldsymbol{y}_w$ is produced, it is evaluated using the objective function $F_z(\boldsymbol{y}_w)$ The final offspring $\boldsymbol{y}_w$ replaces the parent $\boldsymbol{s}_{nw}$ if $F_z(\boldsymbol{y}_w) > F_z(\boldsymbol{s}_{nw})$; otherwise, the solution $\boldsymbol{s}_{nw}$ remains in the pool of parent solutions that are used to generate the offspring in the next generation. To summarize, the main steps of the proposed algorithm are presented in Table C1.

Table C1. Pseudocode of DE algorithm for TDP

Procedure
Begin
Set $n = 0$.
Create a random population $\boldsymbol{s}_{nw}$ ($w = \{1,2, \ldots, NP\}$ and evaluate each solution using $F_z(\boldsymbol{s}_{nw})$
For $n = 1\ to\ N$:
For $w = 1\ to\ NP$:
Generate $\boldsymbol{y}_w$ offspring based on the combination operation.
If $rand(0,1) \leq p_{swap}$:
Mutate $\boldsymbol{y}_w$ using the swapping operation.
End If
If $F(\boldsymbol{y}_w) > F(\boldsymbol{s}_{nw})$:
$\boldsymbol{s}_{nw} \coloneqq \boldsymbol{y}_w$
End If
End For
End For
End

In the simulation tests discussed in section 3.3.2, the parameters were set as:

- $N = 40$
- $NP = 60$
- $\boldsymbol{\Upsilon} = \{0.15, 0.2, 0.2, 0.3, 0.55\}$
- $p_{swap} = 0.05$
- $\varphi = 0.5$

Appendix D. Inputs for PDP Tests

Zone	s	λ	μ
0	0	0	0
1	10	9	10
2	19	19	11
3	8	1	10
4	6	6	6
5	16	17	18
6	9	20	15
7	20	10	12
8	14	20	10
9	18	15	15
10	5	5	5
11	7	10	10
12	20	20	15
13	13	25	11
14	17	17	8
15	9	9	18
16	2	1	10
17	4	4	12
18	3	1	8
19	7	9	7
20	13	2	9

References

Ai, Y., Li, Z., Gan, M., Zhang, Y., Yu, D., Chen, W., & Ju, Y. (2019). A deep learning approach on short-term spatiotemporal distribution forecasting of dockless bike-sharing system. *Neural Computing and Applications*, *31*(5), 1665–1677. https://doi.org/10.1007/s00521-018-3470-9

Anselin, L., Syabri, I., & Kho, Y. (2006). GeoDa: An Introduction to Spatial Data Analysis. *Geographical Analysis*, *38*(1), 5–22.

Arroyo, C. (2020). *Exploring the Determinants of Student's Travel Behavior in the Face of Parking Scarcity Problems*. University of Puerto Rico at Mayagüez.

Bachand-Marleau, J., Lee, B., & El-Geneidy, A. (2012). Better understanding of factors influencing likelihood of using shared bicycle systems and frequency of use. *Transportation Research Record*, (2314), 66–71. https://doi.org/10.3141/2314-09

Bai, S., & Jiao, J. (2020). Dockless E-scooter usage patterns and urban built Environments: A comparison study of Austin, TX, and Minneapolis, MN. *Travel Behaviour and Society*, *20*(October 2019), 264–272. https://doi.org/10.1016/j.tbs.2020.04.005

Bakogiannis, E., Siti, M., Tsigdinos, S., Vassi, A., & Nikitas, A. (2019). Monitoring the first dockless bike sharing system in Greece: Understanding user perceptions, usage patterns and adoption barriers. *Research in Transportation Business and Management*, *33*. https://doi.org/10.1016/j.rtbm.2020.100432

Barabonkov, D., D'Alonzo, S., Pierre, J., Kondor, D., Zhang, X., & Tien, M. A. (2020). Simulating and evaluating rebalancing strategies for dockless bike-sharing systems. *ArXiv*, 1–12.

Barajas, J. M. (2018). How Equitable is Bikesharing? Exploring Population Characteristics and Access to Employment. *In Proceesings of the 97th Annual Meeting of the Transportation Research Board.*

Beecham, R., & Wood, J. (2014). Exploring gendered cycling behaviours within a large-scale behavioural data-set. *Transportation Planning and Technology*, *37*(1), 83–97. https://doi.org/10.1080/03081060.2013.844903

Benchimol, M., Benchimol, P., Chappert, B., De La Taille, A., Laroche, F., Meunier, F., & Robinet, L. (2011). Balancing the stations of a self service "bike hire" system. *RAIRO - Operations Research*, *45*(1), 37–61. https://doi.org/10.1051/ro/2011102

Buck, D., Buehler, R., Happ, P., Rawls, B., Chung, P., & Borecki, N. (2013). Are bikeshare users different from regular cyclists? *Transportation Research Record*, (2387), 112–119. https://doi.org/10.3141/2387-13

Caggiani, L., Camporeale, R., Dimitrijević, B., & Vidović, M. (2020). An approach to modeling bike-sharing systems based on spatial equity concept. *Transportation Research Procedia*, *45*(2019), 185–192. https://doi.org/10.1016/j.trpro.2020.03.006

Caggiani, L., Camporeale, R., & Ottomanelli, M. (2017). Facing equity in transportation Network Design Problem: A flexible constraints based model. *Transport Policy*, *55*(December 2015), 9–17. https://doi.org/10.1016/j.tranpol.2017.01.003

Caggiani, L., Camporeale, R., Ottomanelli, M., & Yuen, W. (2018). A modeling framework for the dynamic management of free- fl oating bike-sharing systems. *Transportation Research Part C*, *87*(January), 159–182. https://doi.org/10.1016/j.trc.2018.01.001

Caggiani, L., Colovic, A., & Ottomanelli, M. (2020). An equality-based model for bike-sharing stations location in bicycle-public transport multimodal mobility. *Transportation Research Part A*, *140*(December 2019), 251–265. https://doi.org/10.1016/j.tra.2020.08.015

Caggiani, L., & Ottomanelli, M. (2012). A Modular Soft Computing based Method for Vehicles Repositioning in Bike-sharing Systems. *Procedia - Social and Behavioral Sciences*, *54*, 675–684. https://doi.org/10.1016/j.sbspro.2012.09.785

Caltrans. (n.d.). *Analyzing the Potential of Geofencing for Electric Bicycles and Scooters in the Public Right of Way*. Retrieved from https://bit.ly/2Et9lOS

Cameron, A. C., & Trivedi, P. K. (2001). Essentials of count data regression. In *A companion to theoretical econometrics*.

Caspi, O., Smart, M. J., & Noland, R. B. (2020). Spatial associations of dockless shared e-scooter usage. *Transportation Research Part D: Transport and Environment*, *86*, 102396. https://doi.org/10.1016/j.trd.2020.102396

Census, U. (n.d.). 2014-2018 American Community Survey 5-Year Estimates. Retrieved from https://data.census.gov/cedsci/

Chemla, D., Meunier, F., Pradeau, T., Calvo, R. W., Chemla, D., Meunier, F., … Yahiaoui, H. (2013). Self-service bike sharing systems : simulation , repositioning , pricing.

Chen, T., & Guestrin, C. (2016). XGBoost: A scalable tree boosting system. *Proceedings of the ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, *13-17-Augu*, 785–794. https://doi.org/10.1145/2939672.2939785

Chen, Zheyan, van Lierop, D., & Ettema, D. (2020). Dockless bike-sharing systems: what are the implications? *Transport Reviews*, *40*(3), 333–353. https://doi.org/10.1080/01441647.2019.1710306

Chen, Zhiwei, Guo, Y., Stuart, A. L., Zhang, Y., & Li, X. (2019). Exploring the equity performance of bike-sharing systems with disaggregated data: A story of southern Tampa. *Transportation Research Part A: Policy and Practice*, *130*(September), 529–545. https://doi.org/10.1016/j.tra.2019.09.048

City of Chicago. (2020). *E-scooter Pilot Evaluation*.

Degele, J., Gorr, A., Hass, K., Kormann, D., Krauss, S., Lipinski, P., … Hertweck, D. (2018). Identifying E-Scooter Sharing Customer Segments using Clustering. In *2018 IEEE International Conference on Engineering, Technology and Innovation*. https://doi.org/10.1109/ICE.2018.8436288

Ferguson, E. M., Duthie, J., Unnikrishnan, A., & Waller, S. T. (2012). Incorporating equity into the transit frequency-setting problem. *Transportation Research Part A: Policy and Practice*, *46*(1), 190–199. https://doi.org/10.1016/j.tra.2011.06.002

Fishman, E. (2016). Bikeshare: A Review of Recent Literature. *Transport Reviews*, *36*(1), 92–113. https://doi.org/10.1080/01441647.2015.1033036

Froehlich, J. E., Neumann, J., & Oliver, N. (2009). Sensing and Predicting the Pulse of the City through Shared Bicycling. In *Twenty-First International Joint Conference on Artificial Intelligence* (pp. 1420–1426).

Gehrke, S. R., Sadeghinasr, B., Wang, Q., & Reardon, T. G. (2021). Patterns and predictors of dockless bikeshare trip generation and duration in Boston's suburbs. *Case Studies on Transport Policy*, (March). https://doi.org/10.1016/j.cstp.2021.03.012

Goodman, A., & Cheshire, J. (2014). Inequalities in the London bicycle sharing system revisited: Impacts of extending the scheme to poorer areas but then doubling prices. *Journal of Transport Geography*, *41*, 272–279. https://doi.org/10.1016/j.jtrangeo.2014.04.004

Gu, T., Kim, I., & Currie, G. (2019). To be or not to be dockless: Empirical analysis of dockless bikeshare development in China. *Transportation Research Part A: Policy and Practice*, *119*(November 2018), 122–147. https://doi.org/10.1016/j.tra.2018.11.007

Guo, Y., Chen, Z., Stuart, A., Li, X., & Zhang, Y. (2020). A systematic overview of transportation equity in terms of accessibility, traffic emissions, and safety outcomes: From conventional to emerging technologies. *Transportation Research Interdisciplinary Perspectives*, *4*, 100091. https://doi.org/10.1016/j.trip.2020.100091

Han, K., Liu, H., Gayah, V. V., Friesz, T. L., & Yao, T. (2015). A robust optimization approach for dynamic traffic signal control with emission considerations. *Transportation Research Part C: Emerging*

Technologies, *70*, 3–26. https://doi.org/10.1016/j.trc.2015.04.001

He, S., & Shin, K. G. (2020). Dynamic Flow Distribution Prediction for Urban Dockless E-Scooter Sharing Reconfiguration. In *WWW '20: Proceedings of The Web Conference 2020* (pp. 133–143). https://doi.org/https://doi.org/10.1145/3366423.3380101

Hirsch, J. A., Stewart, I., Ziegler, S., Richter, B., & Mooney, S. J. (2019). Residents in seattle, wa report differential use of free-floating bikeshare by age, gender, race, and location. *Frontiers in Built Environment*, *5*(March), 1–7. https://doi.org/10.3389/fbuil.2019.00017

Hirsch, J. A., Stratton-Rayner, J., Winters, M., Stehlin, J., Hosford, K., & Mooney, S. J. (2019). Roadmap for free-floating bikeshare research and practice in North America. *Transport Reviews*, *39*(6), 706–732. https://doi.org/10.1080/01441647.2019.1649318

Jain, A. K., Mao, J., & Mohiuddin, K. M. (1996). Artificial neural networks: A tutorial. *Computer*, *29*(3), 31–44. https://doi.org/10.1109/2.485891

Kaltenbrunner, A., Meza, R., Grivolla, J., Codina, J., & Banchs, R. (2010). Urban Cycles and Mobility Patterns: Exploring and Predicting Trends in a Bicycle-Based Public Transport System. *Pervasive and Mobile Computing*, *6*(4), 455–466.

Karner, A., London, J. K., Rowangould, D., & Manaugh, K. (2020). From Transportation Equity to Transportation Justice : Within , Through , and Beyond. *Journal of Planning Literature*, *35*(2), 440–459. https://doi.org/10.1177/0885412220927691

Kwan, M. (1998). Accessibility : A Comparative Anal ysis Using a Point-based Framework. *GeoGraphical Analysis*, *30*(9512451), 191–216.

Levinson, D. (2010). Equity effects of road pricing: A review. *Transport Reviews*, *30*(1), 33–57. https://doi.org/10.1080/01441640903189304

Levy, J. I., Chemerynski, S. M., & Tuchmann, J. L. (2006). Incorporating concepts of inequality and inequity into health benefits analysis. *International Journal for Equity in Health*, *5*, 1–19. https://doi.org/10.1186/1475-9276-5-2

Litman, T. (2005). Evaluating Transportation Equity: Guidance for Incorporating Distributional Impacts in Transportation Planning. *Victoria Transport Policy Institute, Victoria, British …*, *8*(2), 50–65. Retrieved from http://scholar.google.com/scholar?hl=en&btnG=Search&q=intitle:Evaluating+Transportation+Equity+Guidance+For+Incorporating+Distributional+Impacts+in+Transportation+Planning+by#2

Liu, M., & Xu, X. (2018). Dockless Bike-sharing Reallocation based on Data Analysis : Solving Complex Problem with Simple Method. *2018 IEEE Third International Conference on Data Science in Cyberspace (DSC)*, 445–450. https://doi.org/10.1109/DSC.2018.00072

Mathew, J., Liu, M., Li, H., Seeder, S., & Bullock, D. (2019). Analysis of E-Scooter Trips and Their Temporal Usage Patterns. *ITE Journal*, *89*(6), 44–49. Retrieved from http://www.nxtbook.com/ygsreprints/ITE/G107225_ITE_June2019/index.php#/44

McKenzie, G. (2019). Spatiotemporal comparative analysis of scooter-share and bike-share usage patterns in Washington, D.C. *Journal of Transport Geography*, *78*(May), 19–28. https://doi.org/10.1016/j.jtrangeo.2019.05.007

Médard de Chardon, C., & Caruso, G. (2015). Estimating bike-share trips using station level data. *Transportation Research Part B: Methodological*, *78*, 260–279. https://doi.org/10.1016/j.trb.2015.05.003

Mooney, S. J., Hosford, K., Howe, B., Yan, A., Winters, M., Bassok, A., & Hirsch, J. A. (2019). Freedom from the station: Spatial equity in access to dockless bike share. *Journal of Transport Geography*, *74*(November 2018), 91–96. https://doi.org/10.1016/j.jtrangeo.2018.11.009

Páez, A., Scott, D. M., & Morency, C. (2012). Measuring accessibility: Positive and normative implementations of various accessibility indicators. *Journal of Transport Geography*, *25*, 141–153.

https://doi.org/10.1016/j.jtrangeo.2012.03.016
Pan, L., Cai, Q., Fang, Z., Tang, P., & Huang, L. (2019). A Deep Reinforcement Learning Framework for Rebalancing Dockless Bike Sharing Systems. In *Proceedings of the AAAI Conference on Artificial Intelligence*. https://doi.org/10.1609/aaai.v33i01.33011393
Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., … Duchesnay, E. (2011). Scikit-Learn: ML in Python. *Journal of Machine Learning Research*, *12*, 2825–2830.
Perron, L., & Furnon, V. (2019). OR-Tools. Retrieved from https://developers.google.com/optimization/
Qian, X., & Jaller, M. (2021). Bikeshare destination choices and accessibility among disadvantaged communities. *Transportation Research Part D: Transport and Environment*, *91*(January), 102686. https://doi.org/10.1016/j.trd.2020.102686
Qian, X., Jaller, M., & Niemeier, D. (2020). Enhancing equitable service level: Which can address better, dockless or dock-based Bikeshare systems? *Journal of Transport Geography*, *86*(May), 102784. https://doi.org/10.1016/j.jtrangeo.2020.102784
Quddus, M. A., Ochieng, W. Y., & Noland, R. B. (2007). Current map-matching algorithms for transport applications: State-of-the art and future research directions. *Transportation Research Part C: Emerging Technologies*, *15*(5), 312–328. https://doi.org/10.1016/j.trc.2007.05.002
Radwin, D., Conzelmann, J. G., Nunnery, A., Austin Lacy, T., Wu, J., Lew, S., … Siegel, P. (2018). 2015–16 National postsecondary student aid study (NPSAS:16): Student financial aid estimates for 2015–16 (NCES 2018-466), 1–81. Retrieved from https://nces.ed.gov/pubs2018/2018466.pdf%0A%0A
Ramjerdi, F. (2006). Equity Measures and Their Performance. *In Transportation Research Board 85th Annual Meeting*, (1983), 67–74.
Raviv, T., & Kolka, O. (2013). Optimal inventory management of a bike-sharing station. *IIE Transactions (Institute of Industrial Engineers)*, *45*(10), 1077–1093. https://doi.org/10.1080/0740817X.2013.770186
Regue, R., & Recker, W. (2014). Proactive vehicle routing with inferred demand to solve the bikesharing rebalancing problem. *Transportation Research Part E: Logistics and Transportation Review*, *72*, 192–209. https://doi.org/10.1016/j.tre.2014.10.005
Reitermanov´a. (2010). Data Splitting. In *WDS'10 Proceedings of Contribued Papers, Part I* (pp. 31–36).
Schimpl, M., Moore, C., Lederer, C., Neuhaus, A., Sambrook, J., Danesh, J., … Daumer, M. (2011). Association between walking speed and age in healthy, free-living individuals using mobile accelerometry-a cross-sectional study. *PLoS ONE*, *6*(8). https://doi.org/10.1371/journal.pone.0023299
Schuijbroek, J., Hampshire, R. C., & van Hoeve, W. J. (2017). Inventory rebalancing and vehicle routing in bike sharing systems. *European Journal of Operational Research*, *257*(3), 992–1004. https://doi.org/10.1016/j.ejor.2016.08.029
Scikit-Learn. (n.d.). Choosing the Right Estimator. Retrieved from https://scikit-learn.org/stable/tutorial/machine_learning_map/index.html
Seabold, S., & Perktold, J. (2010). Statsmodels: econometric and statistical modeling with Python. In *9th Python in Science Conference*.
Shaheen, S. (2012). Public Bikesharing in North America: Early Operator and User Understanding | Mineta Transportation Institute. Retrieved from https://transweb.sjsu.edu/research/public-bikesharing-north-america-early-operator-and-user-understanding
Shaheen, S. A., & Cohen, A. P. (2019). Shared Micromoblity Policy Toolkit: Docked and Dockless Bike and Scooter Sharing. *UC Berkeley: Transportation Sustainability Research Center*, 1–9. https://doi.org/10.7922/G2TH8JW7
Shaheen, S., Bell, C., Cohen, A., & Yelchuru, B. (2017). *Travel Behavior: Shared Mobility and Transportation Equity*. *FHWA PL-18-007*. Retrieved from

https://www.fhwa.dot.gov/policy/otps/shared_use_mobility_equity_final.pdf
Shaheen, S., Cohen, A., & Martin, E. (2013). Public bikesharing in North America. *Transportation Research Record*, (2387), 83–92. https://doi.org/10.3141/2387-10
Shen, Y., Zhang, X., & Zhao, J. (2018). Understanding the usage of dockless bike sharing in Singapore. *International Journal of Sustainable Transportation*, *12*(9), 686–700. https://doi.org/10.1080/15568318.2018.1429696
Shui, C. S., & Szeto, W. Y. (2020). A review of bicycle-sharing service planning problems. *Transportation Research Part C: Emerging Technologies*, *117*(March), 102648. https://doi.org/10.1016/j.trc.2020.102648
Smith, C. S., Oh, J.-S., & Lei, C. (2015). Exploring the Equity Dimensions of US Bicycle Sharing Systems. https://doi.org/10.13140/RG.2.2.30941.72163
SMOLA, A. J., & SCHOLKOPF, B. (2004). A Tutorial on Support Vector Regression. *Statistics and Computing*, *14*, 199–222. https://doi.org/10.1023/B:STCO.0000035301.49549.88
Tipping, M. E. (2001). Sparse Bayesian Learning and the Relevance Vector Machine. *Journal of Machine Learning Research*, *1*, 211–244.
Ursaki, J., & Aultman-hall, L. (2015). *Quantifying the Equity of Bikeshare Access in US Cities*.
Wang, J., & Lindsey, G. (2019). Neighborhood socio-demographic characteristics and bike share member patterns of use. *Journal of Transport Geography*, *79*(December 2018), 102475. https://doi.org/10.1016/j.jtrangeo.2019.102475
Wang, K., & Akar, G. (2019). Gender gap generators for bike share ridership: Evidence from Citi Bike system in New York City. *Journal of Transport Geography*, *76*(December 2018), 1–9. https://doi.org/10.1016/j.jtrangeo.2019.02.003
Ward, J. H. (1963). Hierarchical Grouping to Optimize an Objective Function. *Journal of the American Statistical Association*, *58*(301), 236–244.
Washington, S. P., Karlafti, M. G., & Mannering, F. L. (2011). *Statistical and econometric methods for transportation data analysis* (2nd ed.). CRC Press.
Xu, C., Ji, J., & Liu, P. (2018). The station-free sharing bike demand forecasting with a deep learning approach and large-scale datasets. *Transportation Research Part C: Emerging Technologies*, *95*(September 2017), 47–60. https://doi.org/10.1016/j.trc.2018.07.013
Yan, A., & Howe, B. (2019). FairST : Equitable Spatial and Temporal Demand Prediction for New Mobility Systems. In *27th ACM SIGSPATIAL International Conference on Advances in Geographic Information Systems (SIGSPATIAL '19),*. Chicago. https://doi.org/https://doi.org/10.1145/3347146.3359380
Younes, H., Zou, Z., Wu, J., & Baiocchi, G. (2020). Comparing the Temporal Determinants of Dockless Scooter-share and Station-based Bike-share in Washington, D.C. *Transportation Research Part A: Policy and Practice*, *134*(December 2019), 308–320. https://doi.org/10.1016/j.tra.2020.02.021
Zhang, J., & Sanderson, A. C. (2009). *Adaptive differential evolution*. Springer-Verlag Berlin Heidelberg.
Zhu, R., Zhang, X., Kondor, D., Santi, P., & Ratti, C. (2020). Understanding spatio-temporal heterogeneity of bike-sharing and scooter-sharing mobility. *Computers, Environment and Urban Systems*, *81*(October 2019), 101483. https://doi.org/10.1016/j.compenvurbsys.2020.101483
Zou, Z., Younes, H., Erdoğan, S., & Wu, J. (2020). Exploratory Analysis of Real-Time E-Scooter Trip Data in Washington, D.C. *Transportation Research Record: Journal of the Transportation Research Board*, 036119812091976. https://doi.org/10.1177/0361198120919760