FINAL REPORT

PROJECT B3

MARCH 2022

Micromobility as a Solution to Reduce Urban Traffic Congestion

Dr. Xilei Zhao, University of Florida
Dr. Virginia Sisiopiku, University of Alabama at Birmingham
Dr. Ruth Steiner, University of Florida

STRIDE | Southeastern Transportation Research, Innovation, Development and Education Center

UF | Transportation Institute UNIVERSITY of FLORIDA

TECHNICAL REPORT DOCUMENTATION PAGE

<table>
<tr><td>1. Report No.
Project B3</td><td>2. Government Accession No.</td><td>3. Recipient's Catalog No.</td></tr>
<tr><td colspan="2" rowspan="2">4. Title and Subtitle
Micromobility as a Solution to Reduce Urban Traffic Congestion</td><td>5. Report Date
March 29, 2022</td></tr>
<tr><td>6. Performing Organization Code</td></tr>
<tr><td colspan="2">7. Author(s)
Xilei Zhao, Ph.D., University of Florida
Virginia Sisiopiku, Ph.D., University of Alabama at Birmingham
Ruth Steiner, Ph.D., University of Florida</td><td>8. Performing Organization Report No.
STRIDE Project B3</td></tr>
<tr><td colspan="2" rowspan="2">9. Performing Organization Name and Address
University of Florida/Department of Civil & Coastal Engineering,365 Weil Hall, PO Box 116580, Gainesville, FL 32611
University of Alabama, Birmingham, Hoehn Engineering Building, Rm. 311
1075 13th Street South, Birmingham, AL 35294</td><td>10. Work Unit No.</td></tr>
<tr><td>11. Contract or Grant No.
Funding Agreement Number 69A355174710</td></tr>
<tr><td colspan="2" rowspan="2">12. Sponsoring Agency Name and Address
U.S Department of Transportation/Office of Research, Development & Tech
1200 New Jersey Avenue, S.E., Washington, DC 20590
University of Florida Transportation Institute/STRIDE Center, 365 Weil Hall, P.O. Box 116580, Gainesville, FL 32611</td><td>13. Type of Report and Period Covered
11/1/2019 to 3/29/2022</td></tr>
<tr><td>14. Sponsoring Agency Code</td></tr>
<tr><td colspan="3">15. Supplementary Notes</td></tr>
<tr><td colspan="3">16. Abstract
Micromobility is an innovative transportation strategy that has demonstrated a great potential for congestion mitigation. However, the research on micromobility is very limited in the field of transportation. This project aims to conduct a comprehensive study to analyze, quantify, and understand the impacts of micromobility on congestion reduction and recommend corresponding intervention strategies for stakeholders. We firstly inferred origins and destinations of e-scooter trips in Washington, D.C. based on the General Bikeshare Feed Specification (GBFS) data and modeled the trip origin demand of e-scooter services. The Ordinary Linear Squares (OLS), Lasso, Decision Tree (DT), Random Forest (RF), and Boosting models were used to predict the trip origin demand in census block group level. The RF model had the best performance among the five models regarding root mean squared error (RMSE) and mean absolute error (MAE). Then we used feature importance (FI) and partial dependence plots (PDP) to interpret the RF model. The results showed that the most important category of variable was built environment variables. From PDPs, we also observed nonlinear relationships between the dependent variable and independent variables. After that, we developed an extended module for shared micromobility simulation in MATSim by applying modifications to its carsharing module. We also developed an effective pipeline to generate synthetic student plans by using different real data sources. The updated MATSim framework was utilized to generate realistic day plans for travelers in a case study that considered 500, 750 and 1000 e-scooters on and around the UAB campus. The case study results confirmed that the simulated traffic volumes are lower and travel speeds are higher when e-scooters are available, compared to the base case scenario. Then we discussed the policy related to shared micromobility operation and developed a decision-support tool that can collect and analyze the e-scooter-related data. Lastly, we created a new decision-support tool to assist cities to better monitor and analyze e-scooter usage and gather inputs from residents. We also presented policy recommendations on regulatory structure, general terms and conditions, operations oversight, public engagement, data, and infrastructure. This project provided rich insights of key factors associated with micromobility demand, examined the potential impact of deployment of e-scooters and other micromobility options on traffic operations, and generated new knowledge for stakeholders to facilitate planning micromobility policies and practices.</td></tr>
<tr><td colspan="2">17. Key Words
micromobility, congestion, machine learning, simulation, policy</td><td>18. Distribution Statement
No restrictions to all.</td></tr>
<tr><td>19. Security Classif. (of this report)
N/A</td><td>20. Security Classif. (of this page)
N/A</td><td>21. No. of Pages
105 pages
22. Price
N/A</td></tr>
</table>

Form DOT F 1700.7 (8-72) Reproduction of completed page authorized

DISCLAIMER

ACKNOWLEDGEMENT OF SPONSORSHIP AND STAKEHOLDERS

This work was sponsored by a contract from the Southeastern Transportation Research, Innovation, Development and Education Center (STRIDE), a Regional University Transportation Center sponsored by a grant from the U.S. Department of Transportation's University Transportation Centers Program.

We also acknowledge the contributions of Dr. Xiang Yan and Dr. Louis Merlin to this research.

Funding Agreement Number - 69A3551747104

LIST OF AUTHORS

Lead PI:

Xilei Zhao, Ph.D.
University of Florida
xilei.zhao@essie.ufl.edu
ORCID 0000-0002-7903-4806

Co-PIs:

Virginia P. Sisiopiku, Ph.D.
University of Alabama at Birmingham
vsisiopi@uab.edu
ORCID 0000-0003-4262-8990

Ruth L. Steiner, Ph.D.
University of Florida
rsteiner@dcp.ufl.edu
ORCID 0000-0001-7276-3742

Additional Researchers:

Yiming Xu, M.S.E.
University of Florida
yiming.xu@ufl.edu
ORCID 0000-0002-2983-1751

Yepeng Liu
University of Florida
yepeng.liu@ufl.edu

Da Yan, Ph.D.
University of Alabama at Birmingham
yanda@uab.edu
ORCID 0000-0002-4653-0408

Jalal Khalil
University of Alabama at Birmingham
jalalk@uab.edu

Wencui Yang, Ph.D.
University of Alabama at Birmingham
yangw@uab.edu

Mostafa Jafarzadehfadaki
University of Alabama at Birmingham
Mostaf86@uab.edu

Juan Suarez
University of Florida
juansuarez@ufl.edu

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

ABSTRACT

Micromobility is an innovative transportation strategy that has demonstrated a great potential for congestion mitigation. However, the research on micromobility is very limited in the field of transportation. This project aims to conduct a comprehensive study to analyze, quantify, and understand the impacts of micromobility on congestion reduction and recommend corresponding intervention strategies for stakeholders. We firstly inferred origins and destinations of e-scooter trips in Washington, D.C. based on the General Bikeshare Feed Specification (GBFS) data and modeled the trip origin demand of e-scooter services. The Ordinary Linear Squares (OLS), Lasso, Decision Tree (DT), Random Forest (RF), and Boosting models were used to predict the trip origin demand in census block group level. The RF model had the best performance among the five models regarding root mean squared error (RMSE) and mean absolute error (MAE). Then we used feature importance (FI) and partial dependence plots (PDP) to interpret the RF model. The results showed that the most important category of variable was built environment variables. From PDPs, we also observed nonlinear relationships between the dependent variable and independent variables. After that, we developed an extended module for shared micromobility simulation in MATSim by applying modifications to its carsharing module. We also developed an effective pipeline to generate synthetic student plans by using different real data sources. The updated MATSim framework was utilized to generate realistic day plans for travelers in a case study that considered 500, 750 and 1000 e-scooters on and around the UAB campus. The case study results confirmed that the simulated traffic volumes are lower and travel speeds are higher when e-scooters are available, compared to the base case scenario. Then we discussed the policy related to shared micromobility operation and developed a decision-support tool that can collect and analyze the e-scooter-related data. Lastly, we created a new decision-support tool to assist cities to better monitor and analyze e-scooter usage and gather inputs from residents. We also presented policy recommendations on regulatory structure, general terms and conditions, operations oversight, public engagement, data, and infrastructure. This project provided rich insights of key factors associated with micromobility demand, examined the potential impact of deployment of e-scooters and other micromobility options on traffic operations, and generated new knowledge for stakeholders to facilitate planning micromobility policies and practices.

Keywords (up to 5):
micromobility, congestion, machine learning, simulation, policy

EXECUTIVE SUMMARY

Micromobility has demonstrated a great potential to grow and become an important travel mode for short trips, but research is very limited on modeling and analyzing the impacts of micromobility on the existing transportation system and exploring its impacts on congestion mitigation. This project aims to conduct a comprehensive study to analyze the impacts of micromobility on urban mobility to understand the potential of micromobility to serve as a solution to mitigate congestion and recommend corresponding intervention strategies for stakeholders. It consists of three major tasks: 1) Task 1: e-scooter big data analytics and travel demand modeling; 2) Task 2: traffic simulation of e-scooters - a pilot study for an urban university campus; 3) Task 3: shared micromobility policy analysis and decision-support tool. The main findings are summarized as follows.

Task 1 inferred origins and destinations of e-scooter trips in Washington, D.C. based on GBFS data and modeled the e-scooter trip origin demand using socioeconomic and demographic variables, built environment variables, and transit supply variables. The results showed that the RF model achieved the best performance among the five models (i.e., Ordinary Linear Squares, Lasso, Decision Tree, Random Forest, and Boosting). The RF model was further interpreted using used feature importance and partial dependence plots. The most important category of variable was built environment variables. We also observed nonlinear relationships between the demand and key factors such as WalkScore and parking density.

Task 2 used MATSim to simulate traffic for a base case scenario at an urban university campus (i.e., UAB campus) and developed an extended module that allowed the consideration of e-scooter use for shared micromobility simulation by applying proper modifications to MATSim' carsharing. We also developed an effective pipeline to generate synthetic student plans by using different real data sources. The case study results confirmed that the simulated traffic volumes are lower and travel speeds are higher when e-scooters are available, compared to the base case scenario.

Task 3 discussed the policy related to shared micromobility operation and developed a decision-support tool, called SERMOS. We presented policy recommendations on regulatory structure, general terms and conditions, operations oversight, public engagement, data, and infrastructure. We also developed a decision-support system that can collect and analyze the e-scooter-related data to help local stakeholders to facilitate better-targeted decision making and policy intervention.

This project provided insights of key factors associated with micromobility demand, examined the potential impact of deployment of e-scooters on traffic operations, and generated new insights for key stakeholders to facilitate planning micromobility policies and practices. Several issues require future research. For example, more features may be needed to develop a more comprehensive demand forecasting model. In addition, more work is needed to refine the Birmingham MATSim model. We also plan to add more analytics modules in the decision-support tool in future work.

1.0 INTRODUCTION

In the recent years, the urban transportation system is experiencing a rapid change with the rise of micromobility, i.e., a variety of small, lightweight transportation devices such as e-scooters and dockless bikes (USA Today, 2019). A recent study by Populus (2018) has found that around 70% people view e-scooters positively as they believe that e-scooters can expand transportation options by replacing short trips in automobile and complementing public transit. The first perception was also validated empirically by a study conducted in Chicago, which showed that for trips between 0.5 and 2 miles, e-scooters present a strong alternative to private vehicles (Smith and Schwieterman, 2018). Similarly, after analyzing half a million e-scooter trips (during a three-month period) in the Indianapolis region, Mathew et al. (2019) found that the median duration and distance of these trips were 8 minutes and 0.7 miles respectively. As most of the short car or ridehailing trips take place in the downtown and its surrounding areas, e-scooters are presenting an opportunity to relieve traffic congestion by replacing automobile trips.

As micromobility continues to grow in size and importance, public entities should start to consider its broader social impacts and its potential to address some of the transportation problems that cities face. In particular, some transportation experts have suggested e-scooters to be part of a solution to reduce congestion and to mitigate the environmental problems brought by automobile use (USA Today, 2019). Moreover, some studies have indicated that micromobility has the potential to account for 8 to 15 percent of all the trips under five miles and grow to a market that is worth $200B to $300B in the U.S. (Shaheen and Cohen, 2019). Furthermore, an integrated transportation system that includes public transit and micromobility also has a great potential to reduce all the car trips by offering a solution to the infamous first/last-mile problem, which may lead to congestion mitigation and emission reduction.

Micromobility has demonstrated a great potential to grow and become an important travel mode for short trips, but research is very limited on modeling and analyzing the impacts of micromobility on the existing transportation system and exploring its impacts on congestion mitigation. In an attempt to fill some of these research gaps, this project aims to use the state-of-the-art machine-learning techniques and activity-based traffic simulations to understand the potential of micromobility to serve as a solution to mitigate congestion.

1.1 OBJECTIVE

In this project, we will examine how micromobility will impact traffic operations and its potential to ease traffic congestion in the U.S. More specifically, this research aims at making the following contributions to the scientific knowledge and practice.

1) Providing a comprehensive assessment for micromobility as a solution to congestion mitigation by integrating big data analytics, travel demand modeling, activity-based traffic simulation, and policy analysis.
2) Leveraging historical e-scooter travel demand data, socio-demographic data, employment data, land-use data, and other relevant data to explore travelers' usage patterns.
3) Applying interpretable machine learning techniques to model and explain the relationships between e-scooter travel demand and other important features, including availability of bike lanes, connectivity to transit, among many others, in order to forecast the travel demand for e-scooters.
4) Identifying under which scenarios micromobility can help reduce urban congestion by applying state-of-the-art traffic simulation models. Furthermore, sensitivity analyses will be conducted to analyze the level of congestion reduction under various market penetration rates of e-scooters.
5) Improving existing activity-based traffic simulation models to account for new modes, such as e-scooters, and provide more accurate simulations for future scenarios.
6) Identifying needs, opportunities, and potential obstacles for policy and operational cooperation between municipal governments and micromobility service providers and proposing effective urban policy and intervention strategies for promotion of e-scooter usage.

1.2 SCOPE

This project conducts a comprehensive study to analyze, quantify, and understand the impacts of micromobility on urban mobility and recommend corresponding intervention strategies for stakeholders by integrating transportation big data analytics, travel demand modeling, activity-based traffic simulation, and policy analysis. It can be divided into three major tasks: 1) Task 1: e-scooter big data analytics and travel demand modeling; 2) Task 2: traffic simulation of e-scooters - a pilot study for an urban university campus; 3) Task 3: shared micromobility policy analysis and decision-support tool.

In Task 1, we firstly inferred origins and destinations of e-scooter trips in Washington, D.C. based on GBFS data. Then we modeled the trip origin demand of e-scooter services in Washington, D.C. using socioeconomic and demographic variables, built environment variables, and transit supply variables. The OLS, Lasso, DT, RF, and Boosting models were used to predict the trip origin demand in census block group level. The in-sample and out-of-sample performance of these five models were compared using MAE and RMSE.

The results of the best performed model, the RF model, were further interpreted using FI and PDPs.

In Task 2, we used MATSim to simulate traffic for a base case scenario at an urban university campus (i.e., UAB campus). Then, we developed an extended module that allowed the consideration of e-scooter use for shared micromobility simulation. This was done by successfully applying proper modifications to MATSim' carsharing module to enable the simulation with the mode of dockless e-scooters. We also changed the scoring function for cars and e-scooters, so both modes can work together in a way that realistic plans get better scores. In addition, we developed an effective pipeline to generate synthetic student plans by using different real data sources. The updated MATSim framework was utilized to generate realistic day plans for travelers in a case study that considered 500, 750 and 1000 e-scooters on and around the UAB campus.

In Task 3, we discussed the policy related to shared micromobility operation and developed a decision-support tool named SERMOS. Firstly, we discussed considerations of scooter shares as environmentally beneficial when considering the mode substitution, rebalancing, and life cycle costs of scooters. Then we presented policy recommendations on regulatory structure, general terms and conditions, operations oversight, public engagement, data, and infrastructure. We also developed a decision-support system that can collect and analyze the e-scooter-related data.

2.0 LITERATURE REVIEW

2.1 Micromobility in the US and travel behavior of e-scooter users

Micromobility refers to small, single-passenger transportation modes rented for short-term use. Existing micromobility studies mainly focus on two micromobility modes: e-scooters and shared bicycles. In this section, studies on usage of e-scooters and bike-sharing are reviewed.

2.1.1 Micromobility safety and regulatory challenges

The rapid emergence of e-scooter without adequate regulatory apparatus or infrastructure in many cases has resulted in public concerns and raised problems such as usage regulation and safety issues. The research community has been dedicated to exploring these topics for years, and studies on e-scooters are still in growth. One topic is crashes and safety of the e-scooter usage since the e-scooter may travel with pedestrians and motor vehicles in some areas. Factors that related to the crashes of e-scooters were examined by researchers, and the results showed that illegal riding (i.e., not wearing a helmet, carrying a passenger) and riding under influence were important factors that influence e-scooter safety (Haworth et al., 2021; Yang et al., 2020). The quick spread of e-scooters as a transportation mode in cities also leads to a regulatory challenge for cities. As discussed in existing studies, the regulations can be spatial and non-spatial. Non-spatial regulations include permitting, insurance, scooter caps, vehicle requirements, rider requirements, app requirements, pricing regulation, marketing, and public education, helmet laws, equipment maintenance, safety regulations such as speed limitations, and data sharing (Verkehrswende, 2019; Merlin et al., 2021). Spatial regulations concern the proper areas for riding and parking e-scooters, and the provision of appropriate infrastructure. These regulations are usually related to safety, efficiency, and equity of the urban transportation system. For example, several cities have identified equity areas and require companies to offer a minimum supply of e-scooters in these areas for equity purposes (Clewlow et al., 2018; Arnell et al., 2020), and some other cities may want to discourage an oversupply of scooters in congested areas (Verkehrswende, 2019). However, it is difficult for cities to make proper regulatory policy to ensure the safety of e-scooter operation and take more advantages of this new transportation mode without a comprehensive and in-depth understanding of travel behavior of e-scooter users. Therefore, cities may wish to know about the users' willingness to use e-scooters, as well as the usage patterns of e-scooters.

2.1.2 Factors associated with micromobility usage

Many factors have been examined by the researchers using the survey data to understand the influence of factors on users' willingness to use e-scooters. These factors can be categorized roughly as socio-demographic factors, trip characteristics, and built environment factors. The socio-demographic factors include age, gender, income, education level, race, resident status, and so on. The literature suggested that the young people, the male, the people with high income, and highly educated people were more willing to use the e-scooters (Lee et al., 2021; Cao et al., 2021; Mitra and Hess, 2021; Christoforou et al., 2021; Laa and Leth, 2020; Sanders et al., 2020). Sanders et al. (2020) also found that non-white people were significantly more likely to intend to try e-scooters. In addition, Mitra and Hess (2021) suggested that the living situation also affected the adoption of e-scooters: single people were more open to using e-scooter services. Trip characteristic factors include travel time, monetary cost, transfer, and travel purpose (Lee et al., 2021; Cao et al., 2021; Mitra and Hess, 2021; Christoforou et al., 2021). Traveler with lower satisfaction with existing alternative modes, which caused by transfer and travel time, were more likely to switch to e-scooters (Lee et al., 2021), and e-scooters mostly replaced walking and public transport modes (Laa and Leth, 2020). Christoforou et al. (2021) indicated that the most common travel purposes of e-scooters were leisure, strolling, and visits, and playfulness was also an important motivation for travelers to choose e-scooters. The most significant built environment factor related to users' adoption of e-scooters was the quality of riding environment, especially the street safety (Mitra and Hess, 2021; Sander et al., 2020; Hosseinzadeh et al., 2021). Areas with higher Walk Score and Bike Score usually had more e-scooter trips (Hosseinzadeh et al., 2021). In addition, the access-egress walking distance also influenced users' willingness to user e-scooters (Cao et al., 2021). That means an optimized distribution of available e-scooters is important to promote e-scooter use, thus an exploration of e-scooter travel behavior was needed.

2.1.3 The usage patterns of micromobility

The researchers have taken to the advent of e-scooters with alacrity, likely because of the abundance of available location data like General Bikeshare Feed Specifications (GBFS) and the Mobility Data Specification (MDS) data (citation), and e-scooter trip data published by the civic authorities (NABSA, 2020). Based on these data, researchers have examined diverse questions about e-scooter travel behavior such as trip distance, trip duration, spatial patterns, temporal patterns, and equity. Trip distance and trip duration vary across studies. The average trip distance in Austin was 0.77 miles with an average travel time of 7.55 minutes (Jiao and Bai, 2020), while a different study in Washington DC found a shorter mean trip distance of 0.6 km and a mean trip length of just 5 min (McKenzie, 2019). Temporal usage patterns were found different from the conventional traffic peaks in some studies. For example, Liu et al. (2019) found that the e-scooter usage peak in Indianapolis was 4 pm to 9 pm on weekdays and 2 pm to 7 pm on weekends. Spatial patterns of e-scooter usage were also examined by the several studies. The findings suggested that downtown areas and university areas usually had a density of scooter trips, while suburban areas had a much lower trip density (McKenzie, 2019; McKenzie, 2020; Bai and Jiao, 2020; Hosseinzadeh et al., 2021). Researchers also have examined spatial associations of e-scooter usage. E-scooter usage were found to be positively associated with some demographic characteristics of areas including employment rate, proportion of young population, and proportion of high educated population (Merlin et al., 2021; Caspi et al., 2021). Studies also found that built environment were correlated to spatial distribution of e-scooter trips. Areas with better riding environment (i.e., higher Walk Score and Bike Score, and better bicycle infrastructure) often had a high density of e-scooter trips (Hosseinzadeh et al., 2021; Caspi et al., 2021). Another important built environment factor was access to transit stations. Areas with higher transit station density usually had more e-scooter trips (Bai and Jiao, 2020; Jiao and Bai, 2020; Merlin et al., 2021). Land use factors were also associated with the spatial usage patterns of e-scooters: greater land use diversity and higher proportion of commercial land use were positively correlated with higher e-scooter use (Hosseinzadeh et al., 2021; Bai and Jiao, 2020; Merlin et al., 2021; McKenzie, 2019).

As an emerging travel mode, e-scooter were usually compared with other existing micromobility options, especially the bike-sharing. Studies suggested that the usage patterns of e-scooter and bike-sharing were different temporally and spatially. McKenzie (2019) found that the temporal distribution of bike-sharing trips clearly reflected stand commuting patterns in Washington D.C. while e-scooter trips did not. Reck et al. (2021) suggested that morning peak positively influences mode choice for shared e-bikes and bikes and negatively for e-scooters. For spatial usage patterns, e-scooters trips were found to have spatially compact and quantitatively denser distribution compared with shared bikes in Singapore (Zhu et al., 2020). On the contrary, membership-based bike-sharing trips appeared to be more concentrated in the downtown core of Washington D.C. than e-scooter and non-membership-based bike-sharing trips (McKenzie, 2019). The trip purposes preference of e-scooter and bike-sharing trips were also different. Shared bikes were more preferred for commuting, whereas e-scooters were more often used for leisure rides (Reck et al., 2021; Bieliński and Ważna, 2020). Studies also found that e-scooter users were on average younger than e-bike users, more sensitive to gas price changing, and less sensitive to the weather (Younes et al., 2020; Bieliński and Ważna, 2020).

2.1.4 The impact of micromobility on the existing transportation system

The rapid growth of micromobility creates the need to explore modal shift practices toward micromobility option. In markets where micromobility options are available, it is important to understand how micromobility affects other modes of transportation such as walking, biking, bike-share, and public transportation in terms of travel time, cost, and convenience.

Some survey-based studies have shown that the emerging micromobility transportation modes replaced some traditional transportation options including automobile travel. Between October 2018 and September 2019, 2.7 million journeys were performed during a trial of e-scooters and e-bikes in Santa Monica, California, 49% of which replaced journeys that would have been undertaken by cars (City of Santa Monica, 2019). According to a global study of Lime customers conducted by Barclays, 30% of Lime customers have substituted an automobile journey with an e-bike or e-scooter trip (Barclays, 2019). In addition, e-scooter rides replaced 28% of the private automobile or ride-sharing journeys, and 57% of walking, biking, e-biking, and skateboarding trips (Fitt & Curl, 2019). Chang et al. (2019) claimed that shared e-scooters are substantially replacing walking and cycling trips in Denver and Portland. According to results from online surveys, e-scooter journeys replaced walking (43%) and bike (14%) journeys in Denver, Colorado. If a shared e-scooter had not been available for their previous trip, 46 percent of respondents in Portland, Oregon said they would have walked (37%) or cycled (9%) instead (Chang et al., 2019). In a 2019 online study of Auckland and Christchurch citizens, 14% said they would replace a private car trip with an e-scooter, and 10% reported that they would replace an Uber/taxi trip with an e-scooter (Kantar, 2019). Cao et al. (2021) conducted a survey study to examine the possibility of substituting short-distance transit trips with shared e-scooter trips users in Singapore and showed that people tend to favor shared e-scooters if transit service has long walking distances and more transfers.

Other researchers use statistical models to estimate the impact of micromobility. Smith & Schwieterman (2018) applied Chaddick Institute's multimodal travel model to assess 30,000 trips around the city and concluded that people would prefer e-scooters over their private cars for short trips between 0.5 and 2 miles while e-scooters would not probably be an appropriate option in comparison with public transportation for trips over 3 miles. Lee et al. (2019) conducted a study to forecast the percentages of e-scooters substitute for carpool, taxi, and bike trips in four US cities, namely Portland, Austin, Chicago, and New York. The results showed that e-scooter can substitute short-distance trips of up to 32%, 13%, 7.2% of carpool, bike, and taxi trips, respectively (Lee et al., 2021). Reck et al. (2021) analyzed shared micromobility competition and mode choice among four types of shared micromobility modes and found that docked modes are preferred for commuting. Their results also revealed a fundamental relationship between the density of micromobility fleet size and the volume of usage (Reck et al., 2021). (Ciociola et al., 2020) proposed a simulation model to forecast the demand for e-scooter usage in Minneapolis and Louisville by using the Poisson method for temporal estimation and Kernel Density Estimation (KDE) for spatial estimation. Their demand model provides a data-driven approach to compare and improve the design of e-scooter sharing systems in smart cities.

2.2 The potential application of microscopic simulation to micromobility

To date, microscopic traffic simulation models have focused on automobile traffic. To the author's knowledge, there has not been a comprehensive model to simulate micromobility traffic, such as e-scooter and e-bike. However, there is research that studies bicycles as a transportation mode, which has some similar motion and maneuverability as e-scooter and e-bike. Therefore, the review of literature on bicycle microsimulation can contribute to the development of micromobility microsimulation applications.

2.2.1 Bicycle microsimulation models

Most microsimulation models are built upon the car-following and lane change formulae, where vehicle actions are highly influenced by lead vehicles. We cannot assume the same stimulus-response mechanism in bicycle traffic given the independent movement of the bicycle. While the principles and algorithms of microsimulation for an automobile can apply to bicycles, bicycle traffic has more flexibility and maneuverability affected by bicyclist's attributes and environmental factors, such as individual's age, physical ability, weather, topology, and the availability of designated lanes. The heterogeneity and stochasticity are more noticeable in cyclists' behavior (Taylor & Mahmssani, 1998).

2.2.1.1 Cellular Automata Model

One of the most common approaches to modeling microscopic cyclist behavior is cellular automata (CA) (Wolfram, 1983). CA uses a discrete space algorithm in which cells occupied by cyclists change as cyclists move by following a set of behavior and interaction rules (Mohammed et al., 2021). CA has been extensively used to simulate cyclists' behavior. Jiang et al. (2004) modeled the cyclists' behavior by developing a stochastic CA model that used the discretized cells containing multiple bicycles. However, this model cannot differentiate the heterogeneity and stochasticity of the cyclists. To overcome the limitation, many CA models have been developed to account for different cyclists' characteristics, motion, and behavioral patterns (Jia et al., 2007; Gould & Karner, 2009; Xue et al.,2017; Tang et al., 2018). Despite its wide adoption, the CA model has its limitations. First, it cannot produce the continuous state-space representation. Second, the CA model only allows a pre-defined number of agent groups in cells with its own set of rules. These limitations restrict the ability to model multiple levels of heterogeneity and environment dynamics (Mohammed et al., 2021).

2.2.1.2 Psychopysical Model

Another method of modeling microscopic cyclist behavior adopts the concept of the psychophysical car-following model. Liang et al. (2012) developed a psychophysical model that stimulates the cyclists' acceleration, deceleration, and turns by assuming cyclists follow the rules of collision avoidance. Zhao and Zhang (2017) extended the model application to motor vehicles, bicycles, and pedestrians with parameter values estimated from experimental data. These methods have a similar limitation as the CA model and only model generalized behavior.

2.2.1.3 Agent-Based Model

Unlike the conventional modeling methods, the Agent-Based Model (ABM) is a powerful tool that can capture the complexity of the real-world environment and the variability of human behavior. In ABM, each traveler or vehicle represents an agent that moves according to the interactions between agents and the environment.

In recent years, a small number of studies use the ABM to simulate bicycle traffic patterns. Loidl et al. (2016) simulated the single trips of employees, students, and leisure cyclists. Ziemke et al. (2017) developed the extension of agent-based modeling framework MATSim with infrastructure attributes that affect cyclists' routing choices. Ziemke et al. (2019) extended previous work to consider the interaction of bicycles with the automobile. Veldhuis (2018) simulated the single bicycles trips of employees, shoppers, and tourists in Amersterdam. These studies proposed agent-based models that simulated simplified daily bicycling trips, however, they fall short of simulating a complete day of a total regional population in detail (Kaziava et. al., 2021). Kaziyeva et al. (2021) simulated the spatial-temporal patterns of bicycle flows of 186,000 inhabitants in the Salzburg region, Austria. Eventhough the study simulated only bicycle movements; the simulation of mode choice included all major available transportation modes.

A key step to develop ABM is to decide the agent's interaction rules. One approach is to use a structural analytical model with parameters that guide agent behavior in different situations (Baster et al., 2013; Hussein & Sayed, 2017). This method has wide transportation modeling applications, including modeling the effect of "Mobility as a Service (MaaS)" trends (Djavadian & Chow, 2017), solving transit network design problems (Liu & Zhou, 2016), and even modeling the interactions between autonomous vehicles (de Oliveira, 2017). One major limitation of this method is the preset rules: agents do not learn from the interactions, nor do they evolve through the learning process (Abdou et al., 2012).

To overcome this limitation, the Reinforcement Learning approach can be applied in ABM. In Reinforcement Learning, intelligent adaptive agents can learn from expert demonstrations and evolve their goals and strategies over time (Plekhanova, 2003). The framework of a finite-state Markov Decision Process is used to guide agents' behavior. Agents execute sequential decision processes based on a reward function that represents the attractiveness of potential future states (Sutton & Barto, 1999). Mohammed et al. (2021) applied reinforcement learning in an agent-based model to simulate both longitudinal and lateral motion dynamics of cyclists. Their approach imposed fewer restrictions and assumptions and demonstrated better accuracy in predicting cyclists' behavior than other cyclist simulation models.

2.2.2 Bike-sharing microsimulation

Compared to conventional bicycles, the emergence of micromobility transportation is quite new and there are fewer studies of microsimulation for micromobility. Soriguera et al. (2018) developed an ABM in Matlab to emulate bike-sharing in the city of Barcelona, Spain. The model includes the model choice between biking and walking and the deployment of bikes by the truck fleet. However, this model is built upon a simplified system that focuses on rebalancing the shared bikes among stations. Hebenstreit and Fellendorf (2019) developed an ABM that incorporates public transit, bicycle, and bike-sharing into a multimodal transportation system in the city of Vienna, Austria. The MATSim-based model focused on the station-based bike-sharing system. Some unique features associated with bike-sharing, such as station, battery charging status, and relocation, were considered for modal choice and routing choice decisions. This is one of the pioneer works that incorporates the micromobility model into multimodel transportation microsimulation. The MATSim based framework provides great flexibility for the extension and future research.

2.2.3 Summary

In recent years, researchers have studied the different aspects of micromobility transportation modes, including the patterns of usage, the factors influencing the usage, and the impact on the existing transportation system. However, there has not been a comprehensive microscopic traffic model to simulate micromobility traffic options (e.g., shared e-scooter and e-bike). The microscopic traffic simulation models have been developed and improved for over 50 years. However, most models share some major drawbacks including the inability to model the heterogeneity and stochasticity of human behavior. The introduction of new technologies, such as artificial intelligence and machine learning techniques, helps to overcome these limitations. Still, the conventional microscopic traffic simulation models that are currently available mainly imitate the automobile traffic flow and cannot properly capture the flexibility and complexity of micromobility traffic, such as e-scooter and e-bike.

To address these concerns, Agent-Based Models may be considered as an option for integrating micromobility with automobile and other modes. Agent-Based Models have the ability to incorporate the complexity and interactions of real-life traffic situations and has shown promising results in simulating bicycle traffic and mixed traffic. Since bicycle traffic shares similarities with micromobility, the techniques used for stimulating bicycle traffic can be applied for modeling micromobility. Some pioneer studies on bike-sharing have been conducted around European cities, and their methods can be adopted to develop micromobility models in U.S. cities. MATSim, an open-sourced ABM holds great promise as a potential platform that can allow integration of micromobility options with automobile and transit options. Such an integration will enable to study the impacts from mode choice shifts between automobile and micromobility modes, in networks that offer such options.

2.3 Literature review on existing policy and practice for planning and managing micromobility

The literature on micromobility is growing and providing more clarity into various aspects of dockless shared scooters (DSS). Most of the literature addresses four major topics regarding dockless shared scooters: benefits of DSS usage, concerns surrounding their use and implementation in the city, observations resulting from the implementation of DSS, and patterns resulting from DSS interaction. These discussions are summarized in Figure 1.

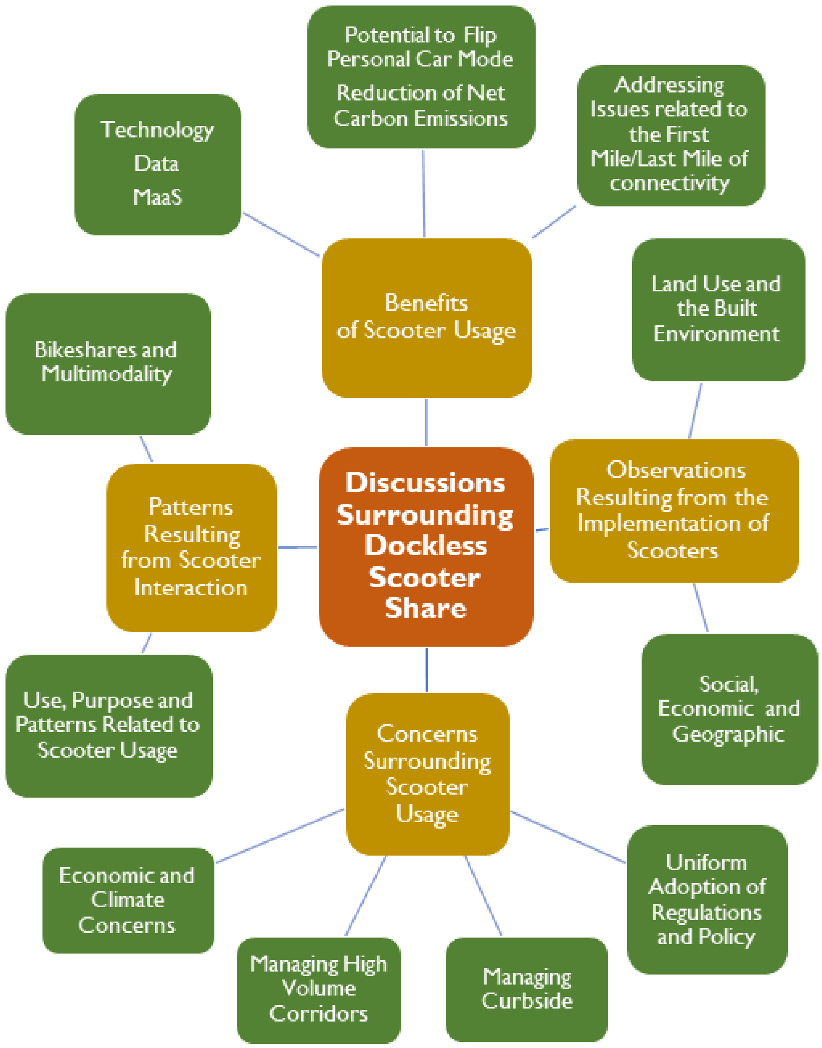


Figure 1. Discussions surrounding dockless scooter share

2.3.1 Benefits of dockless shared scooter usage

Three major benefits of DSS are often cited in the literature: their capacity to address first mile/last mile problems, the perception of them as a sustainable mobility alternative, and their application of technology.

One of the more promising aspects of scooter share is its capacity to address or to bridge the often discussed first mile/last mile mobility problem. McKenzie (2019) asserts that scooters are presented as a first mile/last mile solution, however, scooter operators flooded cities with scooters so quickly that municipalities have not yet properly evaluated if this assertion is true. In the research article that developed scenarios for scooters in Chicago, Smith and

Schweiterman (2018) contend that scooters have the capacity to replace private car trips that are between 0.5 miles and 2 miles. Button, Frye and Reaves (2020) suggest that scooters address a demand that the mobility alternatives do not address, citing the average scooter trip distance of 2 miles, which is too short for car trips or taxi hailing, but too long a walk. Estimating the potential demand of scooters in New York City, Lee, Chow, Yoon, and He (2021) models suggest that many short trips (perhaps within this 0.5-2-mile range) could be replaced by scooters, including 32% of carpool, 13% of bike, 7.2% of taxi, 1.9% of walking, and 1.8% of auto trips. Additionally, their model results suggest that nearly 24% of access/egress trips to public transit could be complemented by scooters. In Portland, nearly 34% of scooter riders would have chosen a motorized alternative had a scooter not been available, suggesting that scooters are capable of addressing not only a first mile last mile gap, but also supporting a positive modal shift (PBOT, 2018).

Numerous sources find that DSS's potential to create a substantial modal shift from personal cars is part of their appeal as a sustainable mobility alternative, mainly because there is a perception that scooters are responsible for significantly less carbon emissions than cars. These two sustainability goals go hand in hand in most of this discussion. Bai and Jiao (2020) suggest that a reduction in automobile usage is evident and further suggest that there are key built environment and socioeconomic factors that are responsible for this shift. Based on their forecasts, Smith and Schweiterman (2018) speculate that the introduction of scooters in parking-constrained areas in the 0.5-to-2-mile range could increase the number of trips from 47% to 75% for non-auto, time competitive, mobility options. Acknowledging that scooters are a new mobility mode, Caspi, Smart and Noland (2020) suggests that if they are successful, they could be an environmentally friendly mode that can address many mobility needs. At the moment, however, Caspi, Smart and Noland (2020) assert that commuting is not one of those mobility needs for scooter users, citing a more recreational usage. de Bortoli and Christoforou (2020) concur and add that, in order to reduce the DSS carbon footprint, municipalities must deploy them carefully, using regulations that consider local characteristics.

Less discussed but equally important is the most evident characteristic of scooters: their efficient application of technology. Hosseinzadeh, Algomaiah, Kluger, and Li (2021) suggest that what has led to such a rapid adoption of scooters in "Smart Cities" is the convenience of finding, using and dropping-off scooters through a mobile application. This technological advancement is consistent with Smart City concepts, demonstrating a successful integration of technology in transportation. The convenience of a shared, dockless system is extended to any given geographical point, but of particular consideration, to public transportation access and egress points. Lee, et al. (2021) suggests that integrating scooters into a Mobility as a Service (MaaS) system would allow users to experience a seamless, multimodal trip. The literature often suggests that the scooter's potential to improve accessibility is best realized through its inclusion in a MaaS system that can connect users to their nearest public transportation system and effectively transfer them within the same technological interface (in this case, a mobile app), using the same payment methods, and perhaps scheduling the connection more efficiently than the user could have on their own.

2.3.2 Concerns surrounding dockless shared scooter usage

For as many benefits of DSS usage being cited, many concerns have emerged in the literature. Four major concerns are often discussed in the literature about DSS: the uniform implementation of policy and regulation, economic and environmental concerns, curbside management, and corridor management.

Policy makers are generally considering three key factors when discussing policy related to the implementation of scooters: safety, equitable access to vehicles and impact on traffic and sustainability (Populus, 2018). However, this may be reducing the circumstances surrounding policy adoption very narrowly since the main concern surrounding policy implementation has been a lack of uniformity throughout the places they are adopted. This is hardly of policy makers own doing; scooters arrived in cities during ongoing discussion on the proper implementation of rideshares and transportation network companies (TNC) (Button, Frye & Reaves, 2020). Additionally, the novelty of DSS coupled with a lack of data about them made it challenging to develop regulatory structures, especially when compromises were being reached to balance enhanced personal mobility, protections for scooter users and protections for the scooter operators. (Button, Frye & Reaves, 2020) As a result, Button, Frye, and Reaves (2020) asserts that policy makers have been "chasing events" rather than leading with proactive policies. This has led to numerous shifts and a constantly evolving regulatory environment for scooter operators.

Their impact on safety is well known since a lack of dedicated areas in the right of way for scooters has led to their use on sidewalks and elsewhere, causing issues in dense areas (Smith and Schwieterman, 2018). Lee et al. (2021) recognize that the legality of scooters varies across the U.S. but demand for them has pushed for their rapid implementation, leaving policymakers to primarily focus on equitable access to them. Policy makers need to address the more glaring issues of safety, and the scooters impact on traffic and sustainability. Button, Frye and Reaves (2020) cite the seven fatalities, a relatively high number, related to scooter usage reported in a brief period in

2019. The Insurance Institute for Highway Safety (IIHS) completed that "compared e-scooter injury statistics with interviews with emergency room patients who had sustained injuries while riding bicycles." (Preston, 2020 quoting a study by IIHS). "IIHS researchers found that e-scooter riders sustained more injuries per mile than bicyclists and were twice as likely to be injured because of potholes, pavement cracks, lampposts, and signposts, although bicyclists were three times as likely to be hit by a motor vehicle" (Preston, 2020). No national studies have focused on scooter safety, leaving policy makers without any data on safety to make informed regulations.

Furthermore, there is also a lack of solid sustainability benchmarks for regulators to evaluate. Eccarius and Lu (2020) suggest that the success of public policy and the continuation of scooters as a mobility service is to identify the users of micromobility to determine how the service aligns with local mobility goals and strategies. Eccarius and Lu (2020) emphasize a particular point that is assumed by many researchers: people shifting from public transit and from active modes to scooters will not necessarily make them the sustainable mobility alternative that they presumed to be.

If sustainability is a goal for regulators, then Eccarius and Lu's (2020) consideration needs to be central to decision making about scooters, since any, "claims of environmental benefits from their use should be met with skepticism unless longer product lifetimes, reduced materials burdens, and reduced e-scooter collection and distribution impacts are achieved." The circumstances under which scooters are a sustainable alternative are directly tied to their environmental benefits. de Bortoli and Christoforou (2020) suggests that the facilitation of longer trips with e scooters should be a policy priority as they tend to replace non-active modes of transport. This negative modal shift could lead to serious environmental issues. Hollingsworth (2019) suggests that while e-scooters may be an effective solution to urban congestion and last-mile problem, they do not necessarily reduce environmental impacts of the transportation system. de Bortoli and Christoforou (2020), who also modelled the impact of scooters on the environment suggests that scooters can be a positive in certain circumstances. Regulators of scooters need to be aware of these circumstances because recommendations are vary depending upon context, suggesting the importance of localized strategies.

The economics of scooter shares are also concerning. According to Button, Frye and Reaves (2020), scooter share companies are struggling to break even, just like rideshare companies such as Uber and Lyft have in the recent past. Numerous articles evaluate the market penetration and the potential income generated by these companies. Noland (2019) estimates that users in Louisville incurred an average cost of $3.40 per trip, which was more expensive than a local bus trip estimated between $1.50 and $1.75. However, the income generated by the companies seems to be around $501.85 per day, which comes down to a meager $183,000 per year in revenues. Lee et al.'s (2021) model, seems more optimistic, with an estimated $77 million in annual revenues and a market penetration of about 75,000 potential daily trips in New York City. Button, Frye and Reaves (2020) found that Lime (a DSS operator) lost an average of $6 million a month the first half of 2018 and by October 2018 accounted for a net loss of nearly $23 million as it sought to compete with Bird, another DSS operator.

Survey results from de Bortoli and Christoforou (2020) show that most users are males aged 18 to 29; de Bortoli and Christoforou (2020) further asserts that the market penetration for people over 30 remains significantly low, suggesting that scooters are addressing a niche in the market. But if scooters are to be capable of being a sustainable alternative, should they be so only for a niche? McKenzie (2020) raises this concern when observing that the predominantly African American neighborhoods of Wards 7 and 8 in Southeast Washington, D.C. have low ridership for both scooter shares and bikeshares. McKenzie (2020) says, "this suggests one of two things, either these mobility services only appeal to a small socio-economic subset of the population, or these new services are contributing to a further socio-economic divide fueled by technology-based transportation." With little data to design robust economic models, the question of DSS's economic stability remains unclear, but at the moment, the current income generation does not suggest endurance, and the current market penetration suggests a small subset of people being served.

Finally, in the domain of the built environment, scooters raise two major concerns involving scooter users interacting with the local environment: curbside and corridor management. Zou, Younes, Erdoğan and Wu (2020) suggest that in high-demand areas that are busy with traffic and pedestrians, designated parking areas are necessary to prevent clutter from disorderly parked dockless e-scooters, especially on pedestrian sidewalks. This suggestion comes in a long list of complaints associated with the dockless nature of scooters, which have cluttered streets and sidewalks to the amazement of residents. While generally convenient, Button, Frye, and Reaves (2020) contends that scooters have contributed to an aesthetically unappealing look in cities, especially in areas attractive to scooter users themselves, since scooters are generally scattered along sidewalks and other locations where previous customer have left them. Additionally, they clutter an already limited public space despite their relatively smaller size (Eccarius and Lu, 2020). Lee et al. (2019) considers that while curbside management is a hot (and often negative) discussion surrounding scooters, understanding the demand for scooters can promote the proper development of policy and infrastructure.

More challenging to address is their use along corridors. Operating on bike infrastructure that has been historically underinvested Zou et al. (2020) highlight the importance of managing e-scooter traffic at critical street corridors, especially principal arterial roads with high volumes of traffic, a high number of historical bike crash incidents, and a lack of bikeway design. This is mainly because there is a lack of safe infrastructure for scooters to navigate, and often scooters operate either on open roads alongside cars and heavy vehicles or on sidewalks alongside slower moving pedestrians. Zou et al. (2020) found that local streets with heavy traffic are popular corridors for scooter users, adding to a safety challenge that policy makers have not been able to wrestle with appropriately due to lack of data. Moreover, Zou et al. (2020) warn that an analysis on scooter trajectory data is challenging due to the amount of API (Application Programming Interface) scrapping involved to get one trajectory, suggesting a highly technical process achievable with open data and technologically capable researchers. But it remains an important part of research since Button, Frye, and Reaves (2020) find that in Southern California, between September 2017 and August 2018, 249 people required medical care after e-scooter related incidents. Over the same period of time 195 received medical attention for bicycle injuries and 181 for pedestrian injuries. The incidents include falls, collisions and getting struck by a moving vehicle, suggesting that many incidents while the scooter is in use.

2.3.3 Understanding the implementation of dockless shared scooter

While questions remain about the concerns and benefits of DSS, a growing number of research papers have evaluated the implementation of DSS, helping to bridge a gap of knowledge that has existed since DSS gained popularity in 2017. The implementation of DSS has revealed much about where in particular they are most successfully implemented, namely by being associated to land uses and built environment attributes or to particular socioeconomic and demographic characteristics.

Numerous research papers associate scooter usage to downtown areas and universities (Mathew, Liu, Seeder, & Li, 2019; Bai and Jiao, 2020; Noland, 2019; Caspi, Smart & Noland, 2020). However, Bai and Jiao (2020) found that areas with higher indexes of walkability, bikeability and compact land uses are equally opportune areas. Similarly, they found that a diversity of land use types generates more demand than the actual mix of specific land uses. They attributed this to larger numbers of points of interest in diverse land uses. Hosseinzadeh et al. (2021) found positive correlations with commercial land uses, industrial land uses and areas with high employment. Caspi, Smart, and Noland (2020) found that the proportion of residential, commercial, educational, and industrial land uses have an impact on the number of trips generated in an area.

Caspi, Smart, and Noland (2020) also found that scooter usage is not related to neighborhood affluence but did recognize that usage by low-income populations is associated most with high student presence near university areas. Caspi, Smart, and Noland (2020) suggest that it is possible that areas with low socioeconomic statuses may be served disproportionately by operators, however there is no data on the supply side of scooter deployment to confirm this. Eccarius and Lu (2018) found that values related to environmental concern was an indirect intention for many scooter users in their research. Several researchers cite low-parking availability and the short-distance trips as important elements in generating trips from scooters (Noland, 2019; Smith and Schwieterman, 2018; Lee et al., 2021). Lee et al. (2021) suggests that scooters would be less competitive on walking trips in New York City compared to other modes. Noland (2019), McKenzie (2019), and other researchers have found that DSS are not a significant mode for commuters, citing low morning usage and high afternoon plateaus suggest other non-commuters use scooters more frequently.

2.3.4 Understanding patterns of dockless shared scooter interaction

Some of the most important findings in the literature discuss the patterns of usage and purpose of DSS, as well as their interaction with other modes, namely bikeshares. Noland (2019) found that nearly 400 trips per day are made in Louisville. Trips lasted on average 15 minutes and travelled on average about 1.25 miles (within the often mentioned 0.5-to-2-mile scooter range). On average, speeds were found to be around 5 miles per hour. McKenzie (2019) by comparison found that nearly 7,050 trips per day were made on Lime scooters in Washington between June and October 2018. Trips lasted 5 minutes and travelled about 0.4 miles with speeds of about 4.8 miles per hour. Between December 2018 and June 2019, Younes, Zou, Wu, and Baiocchi (2020) found that 4,300 trips were made daily in Washington with distances of .6 miles, 11 minutes and nearly 3.5 mph speeds. The effects of winter weather can be the cause of the variance in trips between the research put forth by McKenzie (2019) and Younes et al. (2020). On the other hand, Caspi, Smart, and Noland (2020) found 11,000 trips per day in Austin with an average of 0.6 mi and a trip duration of close to 6.5 minutes, travelling at about 5.5 miles per hour. Mathew et al. (2019) found 4,700 daily scooter trips in Indianapolis travelled 1.1 miles on average, lasted 13.9 minutes and travelled at nearly 5.5 miles per hour.

Mathew et al. (2019) found that only 15% of individual scooters in Indianapolis were used for more than an hour a day, meaning most scooters remained parked most of the day. Younes et al. (2020) found that changing gasoline prices were positively correlated with trips and found that large events in Washington were influential in trip generation. Smith and Schwieterman (2018) find that, "e-scooters would make about 16% more jobs reachable within 30 minutes compared to the number of employment opportunities currently accessible by public transit and walking alone." Bai and Jiao (2020) found that proximity to the city center in Austin and Minneapolis was associated to higher scooter accessibility.

Most importantly, previous research indicates correlations between various mobility modes and scooters. Smith and Schwiterman (2018) find that the benefits of scooters can vary significantly in only a few blocks due to public transit accessibility. Various sources (Smith and Schwieterman, 2018; Lee et al., 2021) suggest that short distance trips can help scooter users begin or complete long-distance transit trips with great success. Achieving multi-modality is broadly suggested to be the highest potential for scooters. de Bortoli and Christoforou (2020) finds that multi-modality is important and that there were few, rare trips that did not start without accessing some other mode.

The most discussed comparison exists between scooter shares and bike shares since they belong to the micromobility family. A gender gap in the adoption of bike-sharing for women is discussed in Populus Technologies (2018), which indicates a 12% adoption rate compared to 21% for males. This is due to the lack of safe infrastructure that is important to female adoption of services. This gap found in bikeshares is almost closed in scooter share programs, where 3% of females have adopted them compared to 4% of males. The League of American Bicyclists (2019), for example, found 33% of Lime riders identify as female compared to just over 25% of bicycle commuters. Additionally, 36% of riders identify themselves as people of color compared to 27% who cycle.

Hosseinzadeh et al. (2021) and Bai and Jiao (2020) concur that high bikeability index scores are needed for scooter shares to be successful. Hosseinzadeh et al. (2021) further points out that bikeability index does not also mean that there will be a high density of scooters suggesting that bikes may not be entirely replaced by scooters. In their research, McKenzie (2019) compares member bikeshare, non-member bikeshare and scootershare. Member bikeshare are associated with commuting, however, non-member bikeshare closely resembled trip purposes of scootershare users such as recreational use, leisure or tourism. Younes et al. (2020) concurs with this notion and suggests that non-member bikeshare and scootershare are in competition, where scootershare complements membership bikeshare. While no membership scootershare system exists, there could be potential ramifications for multi-modality, MaaS and improved economics in the idea of a membership scootershare. Moreover, McKenzie (2019) suggests that scooters and non-members seem to have competing attributes, yet member bikeshares seem to be complimented by dockless scooters. Noland (2019) and McKenzie (2019) find that adverse weather, such as rain or cold weather, has more negative effects on bikeshare when compared to scootershares. This is attributed to the ease of ending the scooter trip wherever, while bikeshares need to be ended in locations perhaps relatively distant from sheltered areas. Similar to bikeshares, scooter shares are also associated with areas of high employment rates (Caspi, Smart and Noland, 2020). de Bortoli and Christoforou (2020) finds that in Paris, the user demographics of scooter shares and bikeshares are similar as well as similar modal substitutions. Smith and Schwieterman (2018) finds that scooter users will often choose scooters over docked bicycles because of their coverage and ease of access. Smith and Schwieterman (2018) also finds that compared to bikeshares, scooters could not compete on long distance trips. This is due to their marginal costs, which are significantly lower on bikeshares than on scooters. Within a range of $4 and $5 per trip, scooters are less convenient to bikeshare member users who after paying the annual membership and using the membership for at least 200 one-way trips, they only pay about $0.50 per trip. Zou et al. (2020) use bike statistics to indicate corridors where safety for scooters indicate concern, demonstrating an important link between bicycles and scooters that is not shared by other modes of transportation. Similarly, the use of historical bike data can be used to discuss safety or bike lane design for scooters. This could suggest an interchangeability between bikes and scooters in the long run. Tables 1 and Table 2 below summarize the literature on DSS.

Table 1. Summary of Literature Review on Scooters

Author(s)	**Case Study Area(s)**	**Dependent Variable**	**Independent Variables**		**Modelling Approach(es)**
			Internal	**External**	
Bai & Jiao 2020	Austin, TX and Minneapolis, MN	Spatiotemporal Patterns and relationships of scooter usage	-Ridership -Trip Characteristics	-Population Demographics -Land Use	-GIS Hotspot -Negative Binomial Regression Model
de Bortoli & Christoforou 2020	Paris	Scooter Impact on Climate Change	-Carbon Emissions from scooter production, rebalancing, recycling, etc.	-Modal Shifts from various modes -Respective Carbon Emissions	-Life Cycle Assessment
Button Frye & Reaves 2020	United States (multiple cities)	Regulations and Policies regarding Scooters	Comparative Literature Review		
Caspi, Smart & Noland 2020	Austin, TX	Spatiotemporal Patterns and relationships of scooter usage	-Ridership -Trip Characteristics	-Population Demographics -Land Use	-Spatial Lag Model -Spatial Durbin Model -Generalized Weighted Regression

Christoforou, Gioldasis, de Bortoli & Seidowsky 2020	Paris	Propensity to use Scooters and characteristics of scooter users	-Scooter General Usage -Trip Characteristics	-General Travel Habits -Survey-takers socioeconomic characteristics	-Road Survey -Multinomial Logit Model
Eccarius & Lu 2018	Taiwan	Propensity to use Scooters and characteristics of scooter users	-Trip Purpose -Reasons for Scooter usage	-Experience with various forms of shared mobility	-Survey -Exploratory Factor Analysis -Binary Logit Model
Eccarius & Lu 2020	Taiwan	Intentions for Scooter Usage	-Intentions -Global Motives -Personal Values	-Demographics	-Survey -Theory of Planned Behavior
Hollingsworth, Copeland & Johnson 2019	Raleigh, NC	Scooter Impact on Climate Change	-Carbon Emissions from scooter production, rebalancing, recycling, etc.	-Modal Shifts from various modes -Respective Carbon Emissions	-Life Cycle Assessment
Hosseinzadeh, Algomaiah, & 2021	Louisville, KY	Spatial Factors associated with Scooter Trips	-Ridership -Trip Characteristics	-Population Demographics -Land Use -Urbanism Scores -Built Env. Characteristics	-Generalized Additive Model
Lee,, Chow, Yoon & He 2021	New York City (Manhattan)	Forecasting e-scooter substitution of direct and access trips	-Ridership	Sociodemographic Data -Traffic Data	-Log-Log Regression Model
Mathew, Liu, Seeder & Li 19	Indianapolis, IN	Spatiotemporal Patterns of Scooter Usage	-Trip counts -Trip Characteristics	None	-Descriptive Statistics
McKenzie 2019	Washington, DC	Comparison of Spatiotemporal Patterns and relationships of scooters and bikeshare	-Trip counts -Trip Characteristics	-Population Demographics -Land Use	-Watson's U2 two sample test for homogeneity
Noland 2019	Louisville, KY	Spatiotemporal Patterns and relationships of scooter usage	-Trip counts -Trip Characteristics	-Weather Data -Land Use -Transit Data	-Ordinary Least Square Regressions
Smith and Schwieterman 2018	Chicago	Scooter Impact on other mobility modes	-Trip Data -Scooter Availability Scenarios	-Employment Data -Built Env. Characteristics -Multimodal Characteristics	-Multimodal Network Analysis
Younes, Zou, Wu & Baiocchi 2020	Washington, DC	Hourly number of trips and Median Hourly Trip Duration for Scooter, Bikeshare Member and Non-Member users	For Scooters and Bikeshares: -Trip counts -Trip Characteristics	-Weather -Gas Prices -Special Events	-Negative Binomial Regression Model
Zou, Younes, Erdoğan & Wu2020	Washington, DC	Correlation between trips and crash data to evaluate safety at the route level	-Trip Counts -Trip Trajectories	-Bike Crash Data -Time of Day -Day of Week	-Bivariate Correlation

Table 2. Topics Addressed in Literature on Scooters

Reference Articles →		Bai and Jiao 2020	de Bortoli and Christoforou 2020	Button et al. 2020	Caspi et al. 2020	Christoforou et al. 2021	Eccarius and Lu 2018	Eccarius and Lu 2020	Hollingsworth 2019	Hosseinzadeh et al. 2021	Lee et al. 2021	Mathew et al. 2019	McKenzie 2019	Noland 2019	Smith and Schwieterman 2018	Younes et al. 2020	Zou et al. 2020	
Number of Themes Addressed		6	6	3	8	4	3	5	6	7	8	2	6	2	7	4	3	
Benefits of Scooter Usage	Potential to Flip Personal Car Mode	■	■		■	■			■						■			6
	Reduction of Net Carbon Emissions	■	■		■				■						■			5
	First Mile/Last Mile of connectivity			■							■		■		■			4
	Technology/MaaS/Data									■	■							2
Concerns Surrounding Scooter Usage	Economic and Climate Concerns		■						■		■		■	■				5
	Managing High Volume Corridors							■							■		■	3
	Curbside Management			■				■			■				■		■	5
	Regulations and Policy		■		■		■	■	■	■	■				■	■		9
Implementation of Scooters	Spatial Socioeconomic Correlations	■			■	■		■		■	■	■	■		■			9
	Land Use & Built Environment Corr.	■			■					■							■	4
Scooter Interaction	Bikeshares	■	■		■		■		■	■	■		■			■		9
	Multimodality	■	■		■	■	■		■	■	■		■			■		10
	Scooter Usage Purpose and Patterns			■	■	■		■		■		■	■	■		■		9

Micromobility as a Solution to

3.0 TASK 1: E-SCOOTER BIG DATA ANALYTICS AND TRAVEL DEMAND MODELING

3.1 Introduction

We are experiencing a dramatic shift in transportation system induced by technology and social awareness of sustainability. For example, smart-phone-based ride-hailing services and micromobility are providing new forms of shared mobility services to travelers and bring a significant impact on their travel mode choices. Micromobility refers to small, single-passenger transportation modes rented for short-term use, such as e-scooters, docked bikes, and dockless bikes. Micromobility is flexible, convenient, affordable, accessible, environmentally friendly, and fun to use, and it is especially attractive for serving short-distance trips. Therefore, micromobility is a good solution to the "first mile/last mile" problem (the problem of public transit being unable to get passengers to their doorstep) that has long troubled public transit. Micromobility can connect the neighborhoods to the transit stops, which makes it more convenient for residents to use public transit. However, although the integration of micromobility and public transit is conceptually appealing, to make it work in practice requires knowledge on patterns of micromobility usage.

Among all the micromobility options, e-scooters are growing at the fastest pace (NACTO, 2019). E-scooters can be seen traveling on streets or parking asides bicycle lanes in many cities of the United States such as Washington D.C., Los Angeles, Chicago, and Atlanta. Based on smartphone applications, the e-scooter sharing services are easy to use. The user first accesses a map of available scooters via the application in a smartphone. After locating and navigating to the target scooter, the user unlocks it by scanning the Quick Response (QR) code on the vehicle or entering the plate number and then starts the trip. When reaching the destination, the user parks the scooter in designated areas and completes the trip on the application. The fee will be charged to the credit card registered by the user.

While e-scooter as a travel mode can greatly enhance urban mobility, it has two key limitations. Firstly, the demands of e-scooter trip origin and destination are unbalanced spatially and temporally. For example, the spatial distribution of commuter trip origins and destinations are significantly unbalanced during the morning peak and the afternoon peak. The operators need to rebalance the vehicles to meet the high demand of e-scooter use in some areas. Therefore, accurate spatiotemporal e-scooter demand predictions are needed to help the operators generate optimal rebalancing strategies. Secondly, the e-scooters need to be recharged when they are at a low power level. Under normal usage conditions, a typical e-scooter must be recharged at least once within 24 hours. To accomplish this, the operators pay citizens to recharge e-scooters on their private property. Participants are instructed to pick up scooters with low power levels and drop them off at specific locations when finishing recharging. Accurate e-scooter demand predictions are needed to determine the optimal scooter drop-off locations. However, few existing studies focus on the e-scooter demand prediction problem.

The Machine Learning (ML) methods have provided researchers powerful tools to predict the demands of different travel modes including public transit, ride-splitting, bike-sharing, and so on. Unlike conventional statistical methods that often assume a pre-determined functional form, ML allows the model structure to freely vary and thus can readily capture the nonlinear patterns underlying the data to generate more accurate prediction results. In addition, ML methods allows researchers to explore nonlinear and threshold effects conveniently. Therefore, in this task, we use ML methods to model the e-scooter demand and explore effects of different factors.

Public application programming interfaces (APIs) are the main data source for the public to understand micromobility. As a part of the micromobility permit requirement, cities often require micromobility providers to share data through APIs prescribed by standard formats, including the General Bikeshare Feed Specification (GBFS) and the Mobility Data Specification (MDS). GBFS was initially developed as the open data standard for bike share system availability back in 2015, but now it is applicable for nearly all shared micromobility systems in the North America (NABSA, 2020). GBFS APIs report real-time information about available vehicles, which typically includes vehicle location, vehicle type (bike or scooter), and battery level. Created by the Los Angeles Department of Transportation (LADOT) in 2018, MDS extends GBFS to require additional information from mobility providers. The additional information may include data on unavailable vehicles in the network, trip characteristics, and trip trajectories (MobilityData, 2020). However, the MDS has received limited adoption so far, and the MDS APIs are usually not made available to the public. Accordingly, we focus on GBFS in this task.

In this task, we firstly infer origins and destinations of e-scooter trips in Washington, D.C. based on GBFS data. Secondly, we develop several ML models to predict the e-scooter demand in block group level. Then, the in-sample and out-of-sample performance of ML models are compared. Finally, we interpret the ML models using some interpretation tools (i.e., feature importance and partial dependence plot) to explore the effects of factors.

3.2 Data

3.2.1 Data collection and processing

The data used in this research consist of two parts: scooter trip OD data in Washington D.C., and independent variables including socioeconomic and demographic data, built environment data, and transit supply data in block group level.

3.2.1.1 Trip origin and destination data

The scooter trip OD data are inferred from the GBFS data in Washington D.C. from Feb 24, 2020 to Mar 01, 2020. Data of six operators in Washington, D.C., including Bird, Jump, Lime, Lyft, Skip, and Spin, are collected. Data feeds are scraped using APIs provided by the vendors. Since the data update frequency varies among vendors, we use different scraping intervals. Specifically, scraping interval for Bird, Jump, Skip, and Spin data is 60 seconds, and scraping interval for Lime and Lyft is 300 seconds. The attributes of the data include scooter ID, latitude and longitude of scooter location, battery level, etc. However, GBFS data only report real-time information about available vehicles. We need to infer trip origins and destinations from the raw data.

Some recent studies have extracted trip information from the GBFS data to examine the spatiotemporal patterns of scooter usage (McKenzie, 2019; Younes et al., 2020). The trip inference method used in these studies usually assumes that GBFS APIs report *Static Vehicle ID*s, that is, the ID of a given micromobility device does not change over time. However, this assumption no longer holds for many circumstances because the GBFS data standards are updated frequently. For example, in GBFS v2.0, it is required to randomly rotate vehicle IDs after each rental, in order to reduce the potential exposure of private data (MobilityData, 2020). To enhance rider privacy, many micromobility providers currently operate GBFS APIs that report *Resetting Vehicle ID*s (i.e., the vehicle ID will be randomly rotated once the vehicle is unlocked for a new trip) or *Dynamic Vehicle ID*s (i.e., the vehicle ID of a given scooter randomly changes every several minutes). Among the six scooter operators, Jump, Skip, and Spin were using *Static Vehicle ID*s, Bird was using *Resetting Vehicle ID*s, and Lime and Lyft were using *Dynamic Vehicle ID*s. The trip origins and destinations are inferred using algorithms developed by Xu et al. (2020).

3.2.1.2 Independent variables

The independent variables include socioeconomic and demographic data, built environment data, and transit supply data. Table 3 presents the description and the descriptive statistics of the independent variables used in this study.

All the variables are aggregated to block groups level. Block groups are defined as clusters of blocks within the same census tract that have the same first digit of their 4-character census block number. A block group usually covers a contiguous area. The socioeconomic and demographic properties are homogeneous in a block group; thus this kind of segmentation can well represent the functional and administrative properties of the zones (Ke et al., 2021). The block groups in downtown area of Washington, D.C. are relatively small, where the micromobility usage intensity is high. Although the block groups in rural and sub-rural areas are relatively large, micromobility trip demand is small in these areas. To better understand the usage pattern and regulate the operation of micromobility, we want the segmentation to be detailed in areas with high demand, which consistent with the segmentation of block groups. The average area of the block groups in Washington, D.C. is $0.4\ km^2$, and the median is $0.2\ km^2$. The size is suitable for management and research for micromobility.

We constructed a list of input features by merging data from various sources. We obtained a list of socioeconomic and demographic variables from the American Community Survey 2014-2018 5-year estimates data. Furthermore, we used General Transit Feed Specification data to estimate some transit-related variables, applied geographic information system (GIS) techniques to calculate several built environment variables, and used the Walkscore.com API to obtain the Walk Score of a census tracts centroid.

Table 3. Descriptive profile of Input variables

Variable	Description	Mean	SD	Min	Max
Ori	Trip origin demand	125.60	380.85	0	5167
Totpop	Total population	1521	740	60	6019
Popden	Population density	21029.1	16423.6	23.7	115858.7
Pctmale	Proportion of male population	0.47	0.07	0.21	0.91
Hhsize	Average household size	2.4	0.59	1.2	4.6
Pcthighschool	Proportion of population with high school degree and above	1	0	1	1
Pctsomecollege	Proportion of population with some college degree and above	0.73	0.21	0.15	1
Young1	Proportion of population aged 18-34	0.32	0.16	0.01	1
Pctbachelor	Proportion of population with bachelor’s degree and above	0.56	0.3	0	1
Medage	Population in median age	36	7.64	15	66
Pctwhite	Proportion of white population	0.41	0.33	0	1
Pctblack	Proportion of black population	0.47	0.35	0	1
Pcthisp	Proportion of hispanic population	0.1	0.1	0	0.57
Pctasian	Proportion of asian population	0.04	0.05	0	0.31
Carown	Proportion of households with at least one car	0.68	0.19	0.13	1
Pct2car	Proportion of households with at least two cars	0.23	0.16	0	0.75
Pcttransit	Proportion of workers taking transit to work	0.35	0.15	0	0.9
Pctdrialone	Proportion of workers driving alone to work	0.38	0.17	0	0.75
Numworker	Number of workers	814	484	60	4896
Unemploy	Proportion of unemployment	0.08	0.08	0	0.57
Incpercap	Income per capita ($)	55100	31868	3743	182111
Medhhinc	Median household income ($)	96519	55202	12229	250001
Pctmodinc	Proportion of moderate income	0.14	0.099	0	0.55

Pctlowinc	Proportion of low-income ($25k less)	0.13	0.16	0	0.37
Pctupmidinc	Proportion of up middle-income households	0.1	0.06	0	0.37
Pctlowmidinc	Proportion of low middle-income households	0.13	0.07	0	0.37
Pctmidinc	Proportion of middle-income households ($50k to $75k)	0.23	0.1	0	0.73
Pcthighinc	Proportion of high-income ($75k more)	0.42	0.24	0	1
Pctrentocc	Proportion of renter-occupied housing units	0.53	0.27	0	1
Pctsinfam2	Proportion of single-family homes	0.45	0.32	0	1
WalkScore	WalkScore of centroid of census tract	73.47	21.84	4	99
Attraction_Den	Attraction density (per mile square)	2.03	10.64	0	144.54
Bikelane_Den	Bike lane density (miles per mile square)	11.47	13.34	0	90.31
Biketrail_Den	Bike trial density (miles per mile square)	6.97	19.34	0	260.10
Hotel_Den	Hotel density (per mile square)	57.74	182.00	0	1911.92
Parking_Meters_Den	Parking meter density (per mile square)	316.3	702.94	0	4736.4
Parking_Valets_Den	Parking valet density (per mile square)	504.96	1094.57	0	10214.23
RdNtwk_Den	Road network density (miles per mile square)	53.24	20.28	10.19	134.29
Interst_Den	Intersection density (per mile square)	706.08	314.23	43.49	2005.37
MetroStop_Den	Metro stop density (per mile square)	6.65	11.63	0	72.14
BusStop_Den	Bus stop density (per mile square)	564.43	404.41	25.43	3158.83
PctMetroBuf	Percentage of tract within 1/4 mile of a metro stop	0.14	0.26	0	1
PctBusBuf	Percentage of tract within 1/4 mile of a bus stop	0.95	0.14	0.15	1

3.2.1.3 Reducing multicollinearity

Multicollinearity has adverse effects on the reliability of estimates of model parameters and the model's predictive power. Therefore, highly correlated independent variables should be excluded to mitigate the multicollinearity concern before the modeling process. We use variance inflation factor (VIF) to check multicollinearity among independent variables, and then remove variables with VIF greater than 10, a common threshold applied to determine multicollinearity (Sheather, 2009). The VIF values of the remaining independent variables are listed in Table 4.

Table 4. Final independent variable list with VIF values

Variable	VIF	Variable	VIF
Popden	4.54	Pctmidinc	1.87
Pctmale	1.24	Pctrentocc	6.07
Hhsize	3.72	Pctsinfam1	2.86
Medage	3.45	Pctsinfam2	5.97
Young1	4.28	WalkScore	3.61
Pctwhite	6.29	Attraction_Den	1.41
Pcthisp	1.47	Bikelane_Den	2.05

Pctasian	1.94	Biketrail_Den	1.25
Carown	4.82	Hotel_Den	5.14
Pct2car	5.43	Parking_Meters_Den	2.21
Pcttransit	2.11	Parking_Valets_Den	6.25
Pctdrialone	3.55	RdNtwk_Den	3.41
Numworker	2.14	Interst_Den	3.44
Unemploy	2.26	MetroStop_Den	2.64
Medhhinc	8.71	BusStop_Den	5.65
Incpercap	7.34	PctMetroBuf	2.21
Pctlowinc	6.63	PctBusBuf	1.77
Pctmodinc	2.84		

3.2.2 Temporal distribution of trip origins and destinations

We have 65,601 observed trip origins and 65,600 observed trip destinations in total. Figure 2 presents the total number of origins and destinations in different time periods. We can see that there are two peaks of demand every day. Peak hours include 8-9 a.m. and 5-7 p.m. Besides, the temporal usage pattern of weekday and weekend is different.

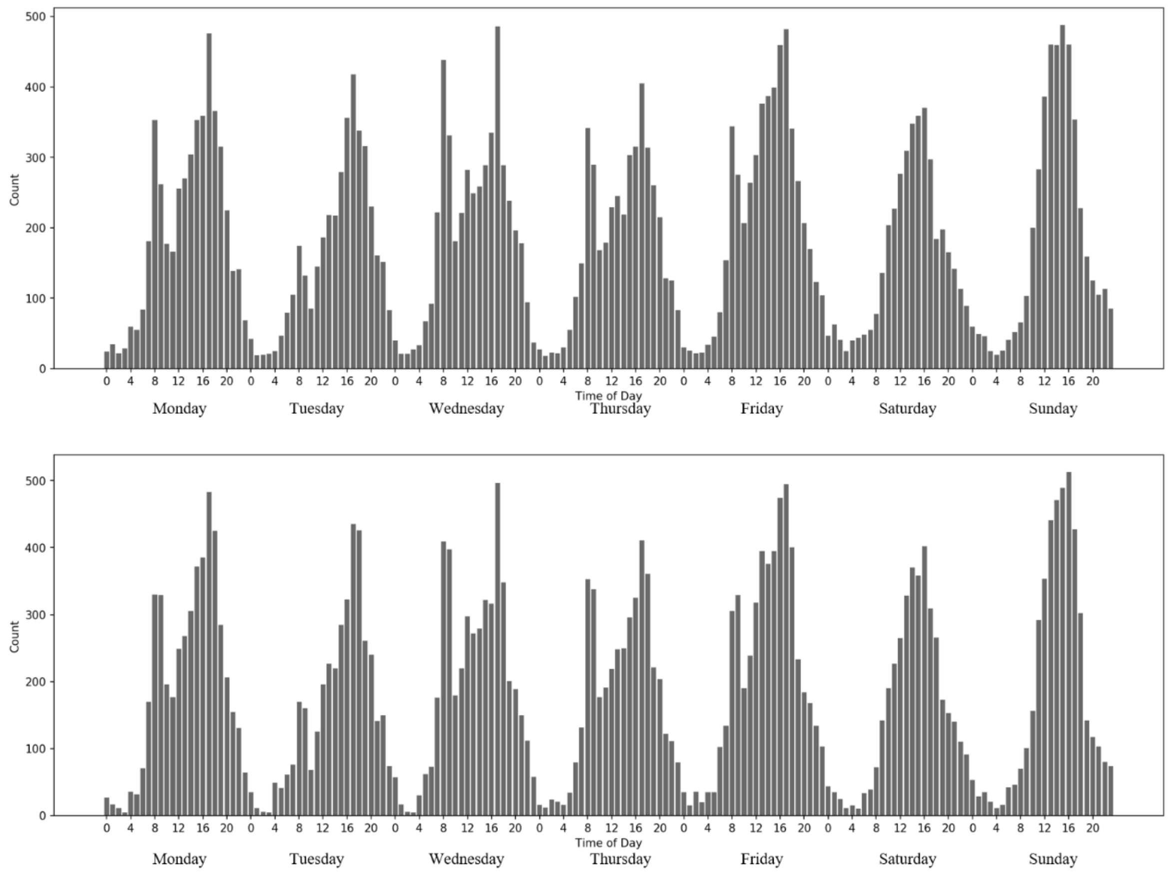


Figure 2. Temporal distributions of total trip origins and destinations (top: trip origin; bottom: trip destination)

We also examined temporal distribution of trip origins and destinations for different vendors in Washington, D.C. (Figure 3.2 and Figure 3.3). Note that the y-axis of the subfigures is different as the trip volume differs among vendors. According to Figure 3 and Figure 4, while the six vendors have different trip volumes, the temporal usage patterns are generally similar, which is consistent with conclusions in McKenzie (2020). All of the six vendors have two significant trip peaks during weekdays. The morning peak hours are 8:00 a.m. to 10:00 a.m., and the afternoon peak hours are 5:00 p.m. to 7:00 p.m. But on weekends, there is only one peak, during the afternoon from 3:00 p.m. to 5:00 p.m. This difference between weekdays and weekends presumably results from commuting trips, which are more common during weekdays. Compared with the study by McKenzie (2020), where data from December 2018 to March 2019 were used, the morning and afternoon peaks in weekdays are more prominent in our study. This result suggests that the proportion of commuting trips in e-scooter trips might be increasing over time, but more analysis is needed in the future to assess the e-scooter trip purpose.

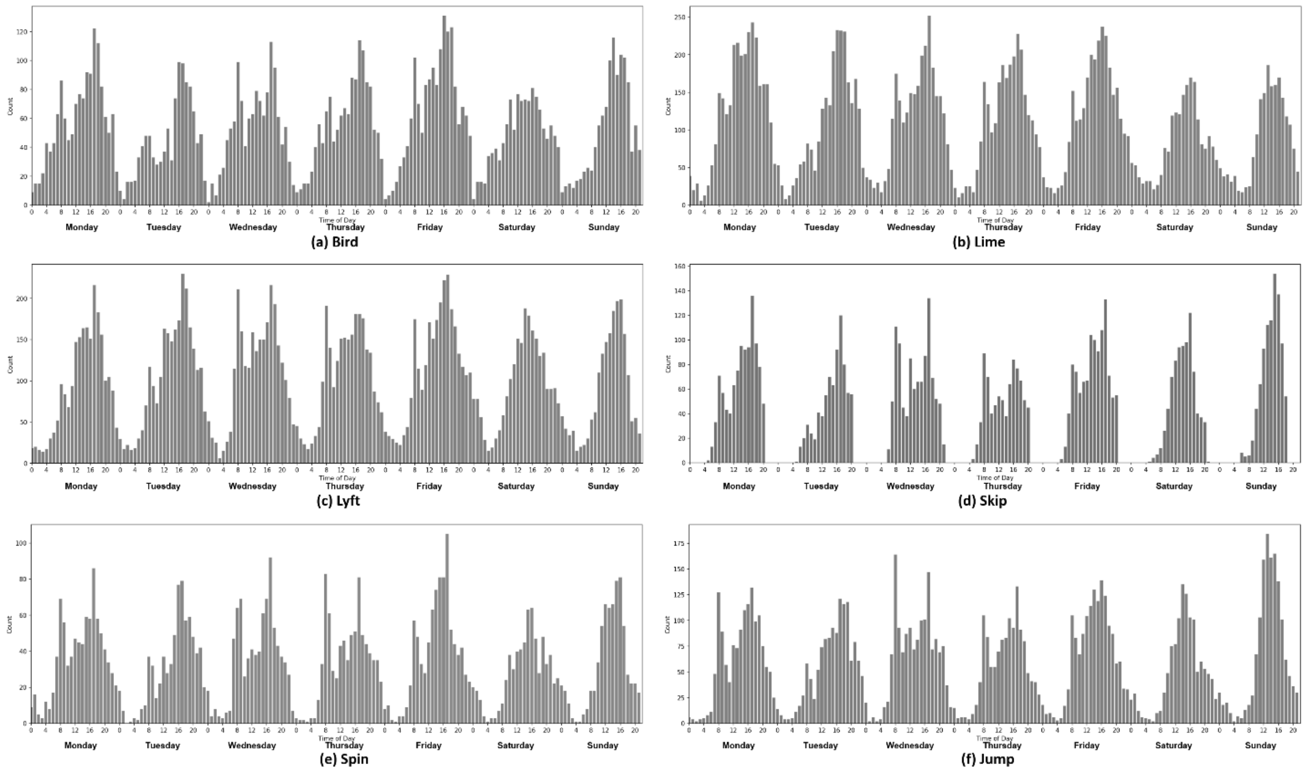


Figure 3. Temporal distribution of Trip Origins for different vendors

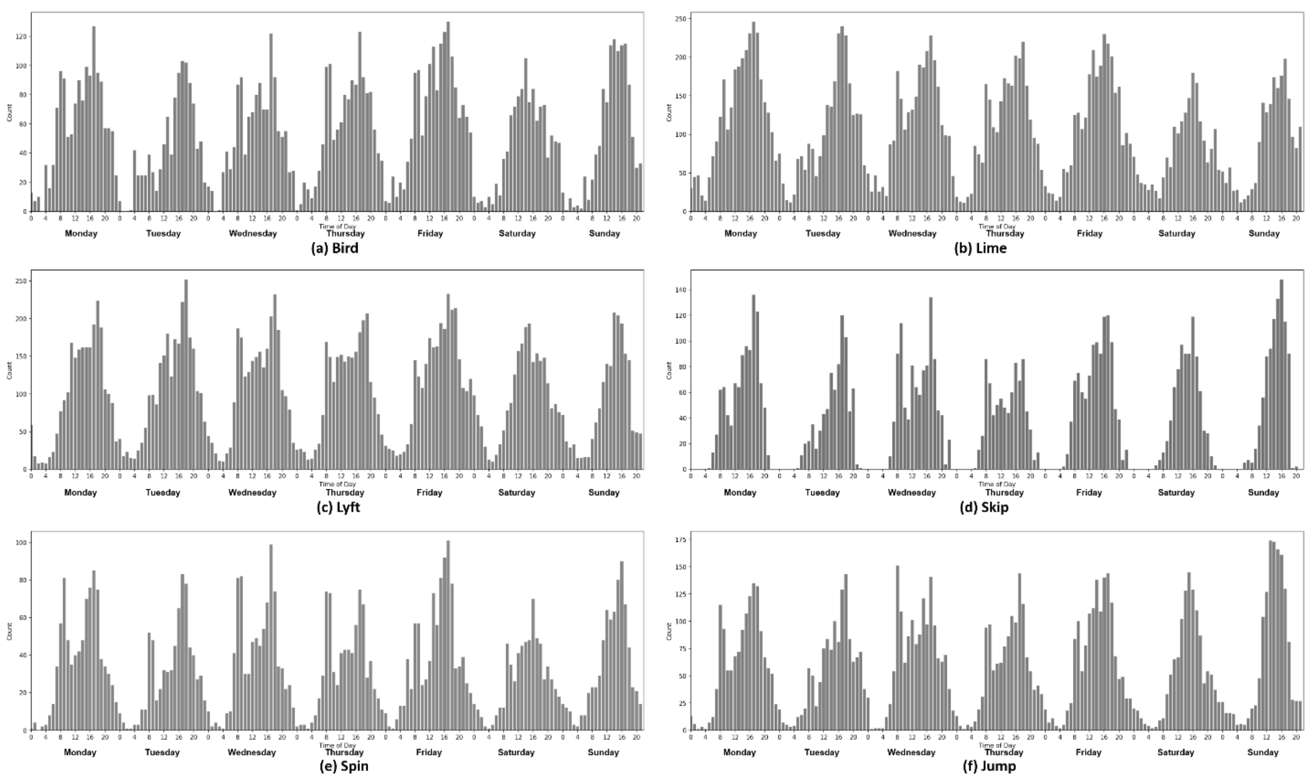


Figure 4. Temporal distribution of Trip Destinations for different vendors

3.2.3 Spatial distribution of trip origins and destinations

We then generate heat maps based on number of trips in each block group to explore spatial e-scooter usage patterns. The results are presented in Figure 5. The general spatial pattern is that the e-scooter demand is significantly high in downtown areas. Interestingly, the block group with the largest e-scooter demand has the greatest number of tourist attractions. This indicate that the e-scooter services are playing an important role in tourism trips.

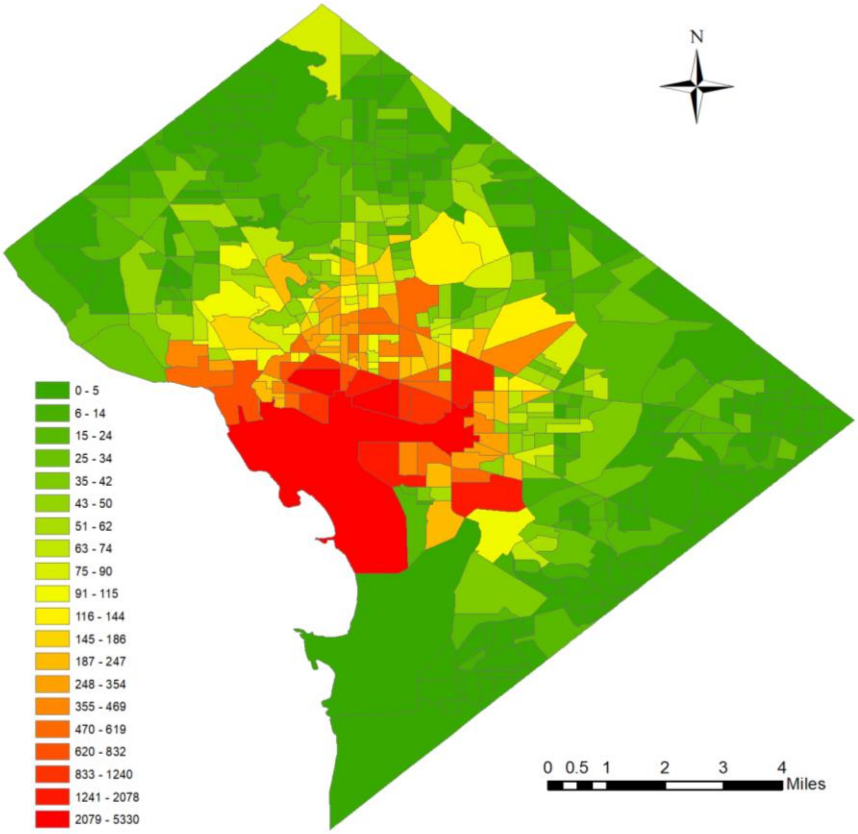


Figure 5. Spatial distributions of e-scooter trips (origins) in Block Group level

3.3 Methodology

In this study, we use 5 methods, including Ordinary Linear Squares (OLS), Lasso, Decision Tree (DT), Random Forest (RF), and Boosting to develop models to predict the e-scooter trip origin demands of block groups in Washington, D.C. The model performance is evaluated by in-sample and out-of-sample root mean squared error (RMSE) and mean absolute error (MAE). The model that has the best performance in our data is further interpreted by feature importance (FI) and partial dependence plots (PDP).

3.3.1 Modeling methods

Ordinary least squares model is a type of method for estimating the unknown parameters in a linear regression model. The principle of OLS is to minimize the sum of squares of the differences between the observed dependent variable y_i in the dataset and the dependent variable predicted by the model $\hat{y}_i$. The function form is:

$$y_i = \beta_1 x_{i1} + \beta_2 x_{i2} + \cdots + \beta_n x_{in} + \varepsilon_i$$

where β_1, β_2, …, β_n are parameters, x_{i1}, x_{i2}, …, x_{in} are independent variables, ε_i is error term.

OLS model is a widely used linear model. The structure of OLS model is simple, and it is easy to use, while the OLS model has disadvantages such as limitations of the linear shape, possibly poor extrapolation properties, and sensitivity to outliers.

Lasso (Least Absolute Shrinkage and Selection Operator) is a linear regression method that performs both variable selection and regularization in order to enhance the prediction accuracy and interpretability (Tibshirani, 1996). Lasso forces the sum of the absolute value of the regression coefficients to be less than a fixed value, which sets certain coefficients to be zero, effectively shrinking the size of coefficients. The lasso coefficients minimize the quantity:

$$\sum_{i=1}^{n} (y_i - \beta_0 - \sum_{j=1}^{p} \beta_j x_{ij})^2 + \lambda \sum_{j=1}^{p} |\beta_j| = RSS + \lambda \sum_{j=1}^{p} |\beta_j|$$

where y_i is the observed dependent variable, β_0, β_j, λ are parameters, RSS is the residual sum of squares.

Lasso produces simpler and more interpretable models that involve only a subset of the predictors compared with OLS model. However, Lasso may generate a slightly higher variance in some cases.

Decision tree model is a tree-based nonparametric ML model. It can be applied to both regression and classification problems (Breiman, 1984). In this study, we use a decision tree for regression. A regression tree can be built by two steps. First, divide the predictor space by the set of possible values into distinct and non-overlapping regions. Then, for every observation that falls into the same region, take the mean of the response values as prediction. This kind of regression tree may produce good predictions on the training set but is likely to overfit the data, leading to poor test set performance. That is because of the high complexity of the resulting tree. We further prune the original large tree to a subtree, aiming to minimize the test error rate.

Decision tree model is easy to interpret because of its tree structure. Trees can be displayed graphically and easily interpreted. In addition, trees can easily handle qualitative predictors without the need to create dummy variables. However, the decision tree model is sensitive to noise and susceptible to overfit (Hastie et al., 2009).

Random Forest (RF) is a tree-based ensemble ML model (Breiman, 2001). RF generates a set of decision trees where the training set for each decision tree is selected using bootstrap sampling from the original sample set, and the optimal node splitting feature for each node is selected from a random subset of the original set of features. The bootstrap sampling and the random selection of features could reduce the correlation between the generated decision trees and thus the average prediction response of multiple trees is expected to overcome the overfitting problems and has lower variance than individual decision trees.

In general, RF tends to have high accuracy prediction and can handle large numbers of features due to the embedded feature selection in the model generation process. RF is also sufficiently robust: the predictor features for RF can be of any type (numerical, categorical, continuous, or discrete) and RF is insensitive to skewed distributions, outliers, and missing values (Breiman, 2001). More importantly, as a tree-based ensemble learning model, RF can model both linear and nonlinear relationships between the input features and the response variable as well as capture interactions among features because of its flexible modeling structure (Breiman, 2001).

Boosting is another tree-based ensemble approach for improving the predictions resulting from a decision tree (Friedman, 2001). The first step of boosting is to create multiple copies of the training data using bootstrap. The trees of boosting are growing sequentially, which means each tree is grown using information from previously grown trees. And boosting does not involve bootstrap sampling, instead, each tree is fit on a modified version of the original data set.

The boosting approach has a strong predictive. Compared with multiple regression, boosting does not require the output variable to be normally distributed. It can also deal with missing values and multicollinearity issues. Meanwhile, it can fit nonlinear relationships between variables. However, it cannot conduct significance tests or produce confidence intervals for coefficients.

3.3.2 Interpretation methods

To identify key factors associated with the e-scooter trip demand and examine their relationships, we further interpret the models using feature importance (FI) and partial dependence plots (PDP). The FI is used to identify key determinants, and PDP is used to explore the relationships between the factors and the e-scooter trip demand.

The importance of each feature in predicting the outcome variable can be quantified by the FI metrics. For linear regression models such as OLS and Lasso, the FI can be evaluated by the coefficients of the independent variables using a standardized input. For tree-based model such as random forest, and boosting, the most commonly used feature importance measures is *Mean Decrease Impurity*. Mean Decrease Impurity evaluates the importance of variable x_i by averaging the weighted reduction in cost for all nodes where feature x_i is selected over all trees. Gini impurity is usually used as the cost function to evaluate variable importance. In this study, we will report the relative importance of features, with the total relative importance of all features scaled to 100%. Note that relative feature importance represents the relative contribution of a variable to the predictive power of a model, and it does not indicate the direction to which a variable is associated with the outcome variable.

The PDP shows the marginal effect that a variable has on the predicted outcome of a machine learning model. PDP works by marginalizing the model output over the distribution of the variables in the complement set of the selected variable(s), so PDP shows the relationship between the selected variable(s) we want to evaluate and the predicted outcome (Molnar, 2019). Consider the sub-vector X_S of the input predictor variables $X^T = (X_1, X_2, \dots, X_p)$, indexed by $S \subset \{1,2,\dots,p\}$. Let C be the complement set, with $S \cup C = \{1,2,\dots,p\}$. A general function $f(X)$ will in principle depend on all of the input variables: $f(X) = f(X_S, X_C)$. The partial dependence of $f(X)$ on X_S is:

$$f_S(X_S) = E_{X_C} f(X_S, X_C)$$

and partial dependence of $f(X)$ on X_S can be estimated by:

$$\bar{f}_S(X_S) = \frac{1}{N}\sum_{i=1}^{N} f(X_S, x_{iC})$$

where x_{iC} is the value of X_C occurring in the training data, N is the sample size.

As a popular machine-learning interpretation method, PDPs are easy to implement. Besides, the computation of partial dependence plots is intuitive, and under uncorrelated cases, the interpretation is clear: PDP shows how the average prediction changes when the corresponding feature is changed. However, the PDP method assumes that the feature(s) under evaluation is not correlated with the other features. If they are correlated, PDP creates new data points in the areas of the feature distribution where the actual probability is very low, often leading to biases in results (Molnar, 2019).

3.4 Results

3.4.1 Model comparison

The model fit (measured using samples in training set) and predictive accuracy (measured using samples in test set) of the OLS, Lasso, DT, RF and Boosting models are evaluated by root mean squared error (RMSE) and mean absolute error (MAE). These metrics can be calculated by:

$$RMSE = \sqrt{\frac{\sum_{k=1}^{N}(\hat{y}_k - y_k)^2}{N}}$$

$$MAE = \frac{1}{N}\sum_{k=1}^{N}|\hat{y}_k - y_k|$$

where N is the total number of observations, y_k is the kth observed value for the outcome variable, and $\hat{y}_k$ is the kth predicted value for the outcome variable.

The five models are evaluated by 10-fold cross-validation. That is, we randomly split the dataset into 10 equal sized subsets. For a single subset, we used it as the test set, and use the remaining 9 subset as training set. The process is repeated 10 times, with each of the 10 subsets used once as the test set. Then we averaged the 10 results to produce the final estimate. The performance metrics of the five models are presented in Table 5.

Table 5. Model Performance of OLS, Lasso, DT, RF and Boosting

Model	In-Sample Performance		Out-of-Sample Performance	
	MAE	RMSE	MAE	RMSE
OLS	141.33 ± 16.24	267.83 ± 27.33	161.26 ± 39.38	300.56 ± 204.38
RF	36.00 ± 2.12	122.27 ± 11.75	85.69 ± 41.04	238.56 ± 203.50
Boosting	9.94 ± 1.56	15.24 ± 2.90	102.48 ± 39.01	250.64 ± 182.83
DT	82.15 ± 7.85	232.15 ± 28.56	128.08 ± 39.42	365.10 ± 190.36
Lasso	140.66 ± 16.15	267.84 ± 27.33	160.39 ± 39.37	299.63 ± 204.78

According to Table 5, the Boosting model and the RF model have better in-sample performance than other models. However, the out-of-sample performance of the Boosting model and the RF model is significantly worse than their in-sample performance. This result indicates that the Boosting model and the RF model have an overfitting problem. The linear models (i.e., OLS and Lasso) have larger fit and predictive errors than DT, RF, and Boosting. This is because the distribution of the outcome variable denoting a positive skew, with the majority of trip origin demand presenting a small value. The linear models are not good at dealing with this kind of data. The RF model has the best out-of-sample performance, and it also has good in-sample performance. To conclude, the RF model has the best performance on our data set among the five models. Therefore, the further interpretation will focus on the RF model.

3.4.2 Modeling interpretation

The relative importance of each feature is presented in Figure 6, with the total relative importance of all features scaled to 100%.

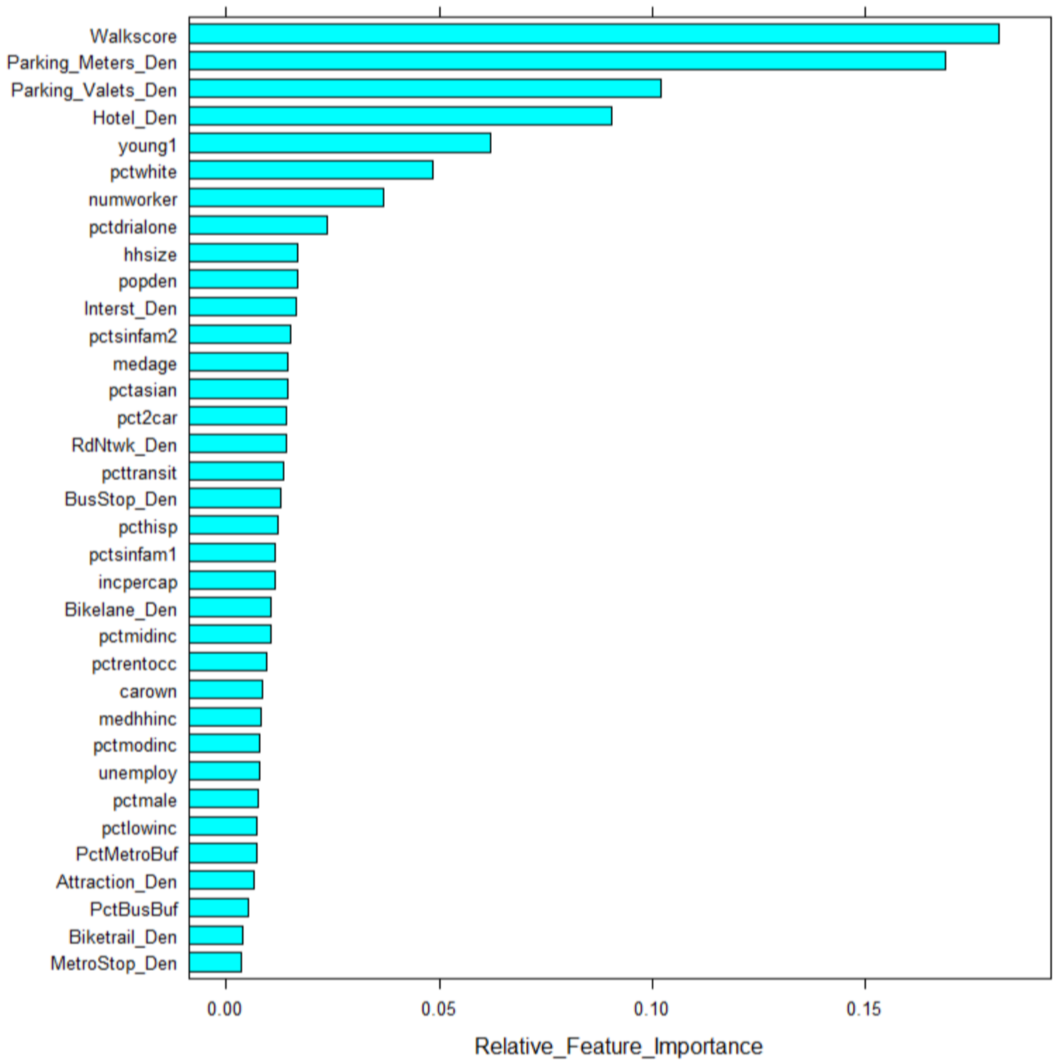
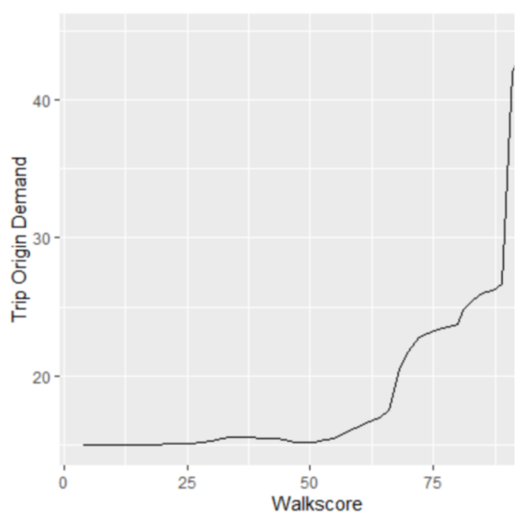


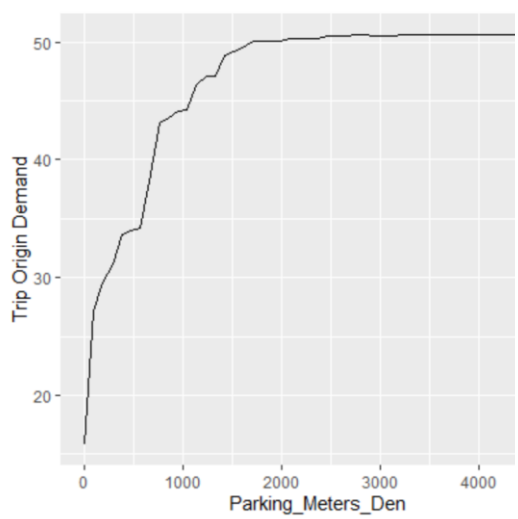
Figure 6. Relative Feature Importance of RF Model

According to Figure 6, the most important feature in the RF model is WalkScore with 18.1% relative importance. This result makes sense because the block groups with high WalkScore usually have a good environment to ride e-scooters. This result consistent with previous studies indicating that areas with better transportation facilities and environment (e.g., the street safety) are more likely to have more e-scooter trips (Mitra and Hess, 2021; Sander et al., 2020; Hosseinzadeh et al., 2021). The second and third important features are parking related variables, Parking_Meters_Den (i.e., parking meter density), and Parking_Valtes_Den (i.e., parking valet density). This may result from the "P+R" (park and ride) trips: travelers use the e-scooters near their parking spots. Hotel_Den (i.e, hotel density) ranks 4th in relative feature importance. This is probably caused by the tourism trips. Previous studies showed that a large proportion of e-scooter trips were leisure, strolling, and visiting trips, and playfulness was also an important motivation for travelers to choose e-scooters (Christoforou et al., 2021). Tourists may ride e-scooters to metro stops, attractions, restaurants, and so on. The proportion of young (aged 18-34) population and the proportion of white population are ranks 5th and 6th, respectively. This means the census block groups with different demographic compositions have different e-scooter trip demands, probably because of the residents' travel preferences. These results are consistent with some survey-based studies. The young people and the non-white people were significantly more likely to use the e-scooters (Lee et al., 2021; Cao et al., 2021; Mitra and Hess, 2021; Christoforou et al., 2021; Laa and Leth, 2020; Sanders et al., 2020). The work-related variable, number of workers, ranks 7th among the variables. This may result from the commute trips. For different variable groups, the built environment variables have the highest relative feature importance, and the transit-supply-related variables have the lowest relative feature importance. That indicates the riding environment and infrastructure are important for promoting e-scooter use.

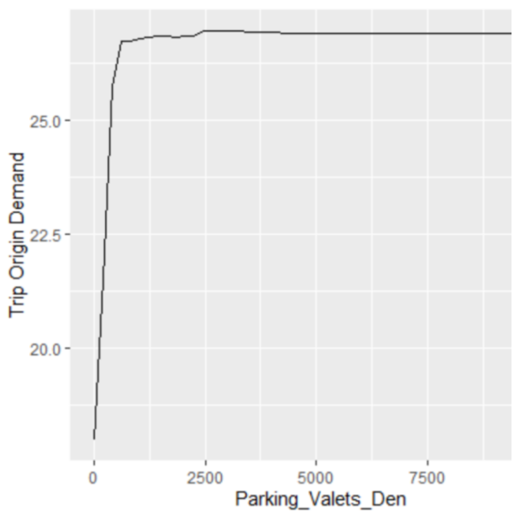
The partial dependence plots of top six important features are presented in Figure 7.



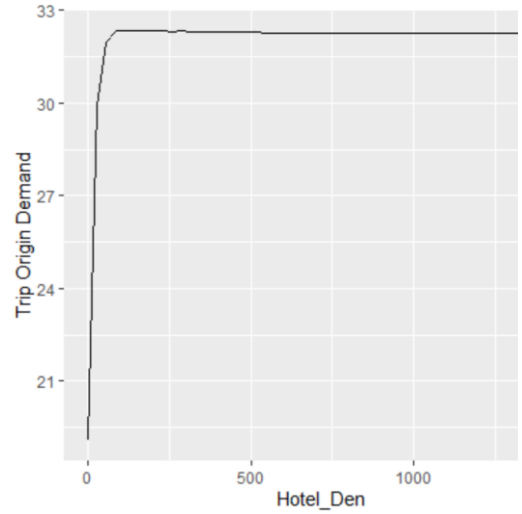
(a) Walkscore



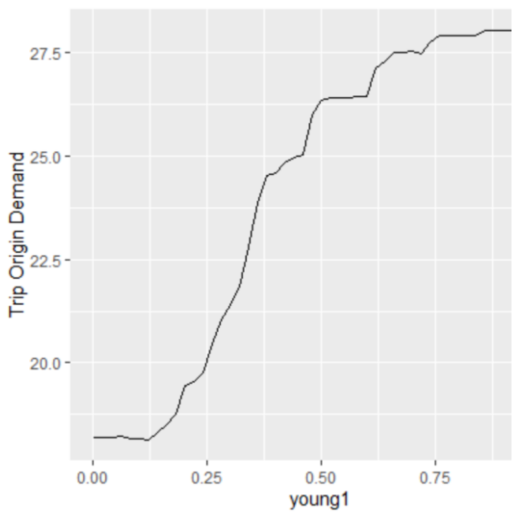
(b) Parking meter density



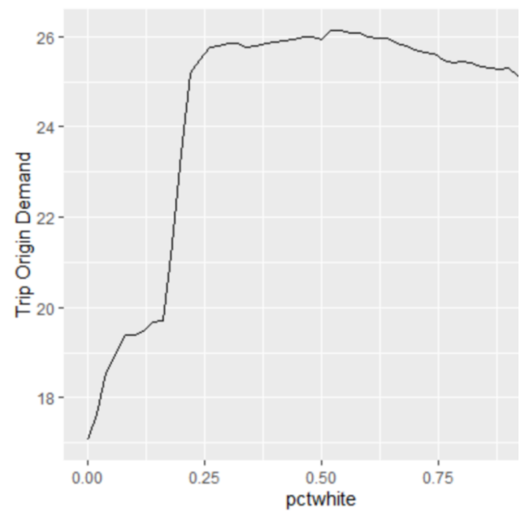
(c) Parking valet density



(d) Hotel density



(e) Proportion of population aged 18-34



(f) Proportion of white population

Figure 7. Partial Dependence Plots of Top Six Important Features

According to Figure 7, all of the six features have positive relationships with trip origin demand. We can also observe nonlinear relationships from these figures. For example, when WalkScore increases from 85 to 95, there is a sharp increase in trip origin demand while the trip origin demand remains flat when WalkScore increases from 25 to 50. For parking meter density, parking valet density, and hotel density, the demand increases with the density increases. After the density exceeding a threshold, the demand does not change much. For proportion of population aged 18-34, the demand increases rapidly when the proportion increases from 0.12 to 0.38. Then the increase slows down. For proportion of white population, the demand increases sharply when the proportion increases from 0 to 0.25. Then the demand remains steady. After the proportion exceeding 0.5, the demand begins to decrease.

3.5 Conclusion

In this task, we model the trip origin demand of e-scooter services in Washington, D.C. The independent variables include socioeconomic and demographic variables, built environment variables, and transit supply variables. The OLS, Lasso, DT, RF, and Boosting models are used to predict the trip origin demand in census block group level. The in-sample and out-of-sample performance of these five models are compared using MAE and RMSE. The results of the best performed model, the RF model, are further interpreted using FI and PDPs. The most important variable is WalkScore and the most important category of variable is built environment variables. From PDPs, we can observe nonlinear relationships between the dependent variable and independent variables. For example, when WalkScore increases from 85 to 95, there is a sharp increase in trip origin demand.

Several issues require future research. Firstly, a detailed and in-depth analysis of modeling results need to be addressed. More literatures should be included to support the conclusions we draw from FI and PDPs. Secondly, more features may be needed to develop a more comprehensive model and to generate richer insights. Thirdly, the results and insights regarding the trip origin demand found in Washington D.C may not be directly transferable to other cities with different characteristics. Therefore, transferability requires further research in the future.

4.0 TASK 2: TRAFFIC SIMULATION OF E-SCOOTERS- A PILOT STUDY FOR AN URBAN UNIVERSITY CAMPUS

4.1 Introduction

In the recent years, there has been a growing interest in introducing shared micromobility options in urban areas and university campuses. These initiatives are driven from a desire to provide alternative options to automobile travel, especially for short distance trips, in hopes of reducing automobile vehicle miles traveled and the associated environmental footprint, easing traffic congestion and improving traffic circulation and user satisfaction. Also, increased use of e-scooters, e-bikes and similar micromobility options can lead to the increase of future bike lanes, greenways, and other facilities that make streets and neighborhoods user-friendly for residents and transportation users alike (Powell, 2020). The recent popularity in shared micromobility services is enabled by advances in technology (such as smartphones, GPS systems and mobile payment options) and supported by social and environmental concerns related to vehicle ownership and urban living (Shaheen et al, 2017). According to the National Association of Transportation Officials (NACTO), people in the United States took 136 million trips on shared bikes, e-bikes, and scooters in 2019, which is 60% more than 2018 (NACTO, 2019).

A prime example of a city that embraced e-scooters and e-bike sharing systems is Portland, Oregon where city officials worked diligently to create an environment in which these services thrive (Powell, 2020). In 2018, e-scooter ridesharing popped up in cities like San Francisco, Santa Monica, St. Louis and 63 other cities. Some of these cities experienced e-scooter services only for a limited amount of time. For example, Bird, a private e-scooter company, introduced e-scooters in Birmingham, AL in August 2018 (Egmon, 2018). However, this move became controversial and short lived in Birmingham as well as in the many other cities across the country. This was due to the fact that ridesharing companies like Bird deployed their e-scooters without seeking the proper permits first. In fact, many cities did not have proper regulations in place for e-scooter services to ensure the safety of riders and pedestrians, and the efficient storage of scooters. The City of Birmingham impounded the scooters and charged the company daily storage fees until e-scooters were removed, and in the meanwhile, proceeded with the drafting of legislation to regulate e-scooters.

The approval of a 2020 ordinance that allows the operation of motorized scooters in the city [City of Birmingham, 2021a] enabled the City of Birmingham to embark on a new pilot program to reintroduce e-scooters in 2021 by partnering with two companies, Veo and Gotcha, [Birmingham Department of Transportation, 2021]. As part of this 90-days pilot program, Veo began operations in late April 2021 offering shared e-scooters and e-bikes in downtown Birmingham and other neighborhoods [City of Birmingham, 2021b]. The City of Birmingham Department of Transportation established 94 parking corrals throughout the service area and each vendor was approved to deploy 500 devices in the first 90 days of operations. Figure 8 shows Veo e-scooters from the 2021 Birmingham pilot deployment.

Veo e-Scooters at UAB (Photo credit: Virginia Sisiopiku)

Veo Scooters at Caldwell Park (Photo credit: Pat Byington)

Veo e-Scooters Parked at Pepper Place (Photo credit: pepperplace.com)

Figure 8. E-scooters in Birmingham, AL

Many cities took a cautious approach by introducing pilot programs in order to gauge the interest in micromobility amongst their residents and the potential impacts on travel patterns prior to a broader deployment. To quantify such impacts and guide local urban and regional planners and transportation authorities as they try to determine the merits of micromobility deployment in their areas, simulation studies are desirable. However, the literature (see Chapter 2) identified a current lack of off-the-shelve micro-simulators that allow the incorporation of shared micromobility options into traffic simulation modeling.

In an effort to bridge this gap, our team performed a study to determine the feasibility of using simulation modeling to incorporate micromobility options into a traffic network in order to measure the impact of such modes on local congestion. Building on our earlier research related to STRIDE B and STRIDE I2 projects (Sisiopiku et al, 2019; Sisiopiku and Salman, 2019; Guo et al., 2019a and Guo et al., 2019b), we used the MATSim platform (www.matsim.org), an open-source agent-based transportation simulation software to implement our traffic simulation around the University of Alabama at Birmingham (UAB) campus located in Birmingham, AL. The Birmingham pilot study integrated e-scooters as an option of travel in the vicinity of UAB, through the development of new Python programs that generated realistic plans of UAB employees and students, which were then fed into MATSim and used to compare traffic impacts with and without e-scooters on the basis of speed and traffic volume. In addition to the baseline (no scooters), three scenarios were considered with gradually increased e-scooter availability to study the likely benefits of increased micromobility services on traffic operations.

It should be noted that the development of the Birmingham MATSim simulation model that considered e-scooters in our study was limited in scope and aimed at demonstrating feasibility rather than measuring actual traffic impacts from the 2021 pilot e-scooter deployment. Still, our team made sincere effort to use any information available about the Birmingham 2021 pilot deployment in the design of the simulation scenario in order to increase the realism of the experiment (e.g., size of the e-scooter fleet, location of parking corals, max speed, e-scooter dimensions, etc.).

When using MATSim to run traffic simulation, MATSim takes the travel day-plans of a population and executes them on the underlying road network (obtained from OpenStreetMap) to generate simulated traffic. So, a key challenge is to generate a realistic synthetic population along with their travel day-plans to be fed to MATSim. In our prior work (Guo et al., 2019a; Guo et al., 2019b], we have proposed a solution to generate realistic population day-plans for an entire city like Birmingham using user surveys plus publicly available data sources. However, that approach does not take micromobility into account. Moreover, micromobility is unique in that the travel distance is usually much shorter than that of car trips, and the service deployment is often in a localized area. As a result, more fine-grained travel behavior modeling is essential.

Therefore, in this project we decided to conduct an area-specific survey of travelers designed based on the travel demands and types in the target area (e.g., based on zone functionality (Guo et al., 2019b) such as commercial, entertainment, or educational). Since the scale of a survey is limited compared with an entire population in the area, we needed to further enrich the day-plan details using openly available data sources.

We choose the campus of the University of Alabama at Birmingham (UAB) as the target area to study. Since UAB is a functional zone for education, research, and healthcare, we surveyed both students and employees in our case study. For this purpose, we conducted a mobility survey of 4,137 students and 4,920 employees of UAB, which was used to obtain the location distributions of their home/apartments, classrooms, and workplaces. To enable the generation of rich information of day-plans, we also resorted to various openly available data sources, and internal data from UAB (e.g., parking lots/decks). Our workflow can be used in other scenarios when area-specific surveys and data sources are provided.

The remaining of Chapter 4 is organized as follows. Section 4.2 describes how we modified MATSim's carsharing module to support the mode of e-scooters, and how we designed a proper scoring function to allow MATSim to select the proper travel mode among alternative modes. Section 4.3 introduces our approach to generate realistic travel plans of UAB students and employees, including the various data sources that we used to achieve this goal. In Section 4.4, we report our experimental findings and conclude this report in Section 4.5.

4.2 MATSim Adaptation for Shared Micromobility Simulation

The following paragraphs provide a brief introduction to the MATSim model and discuss the approach that we used to customize the model in order to meet the needs and objectives of this study. Additional details on MATSim capabilities and applications to date are summarized in the literature and in particular in Horni et al. (2016). Details on our earlier work on the development of the MATSim base model for the Birmingham area are available in Sisiopiku et al. (2019) and Guo et al. (2019a).

4.2.1 MATSim model background

The Multi-Agent Transport Simulation (MATSim) model was used as a simulation tool in this study. MATSim (www.matsim.org) is an extendable, multi-agent simulation framework implemented in Java that offers considerable customizability through its modular design structure. The software is designed to simulate large-scale scenarios by adopting a computationally efficient queue-based approach (Horni et al., 2016). MATSim uses a microscopic simulation methodology for travel demand forecasting that traces the daily travel diary and agents synthetic travel decisions. It is designed to model a single day, similar to other activity-based models.

MATSim is an open-source software that requires its input files to be as XML files. Minimum *input files required* to run the software are: a) Configuration file; b) Network file, and c) Population/plans file. The *configuration file* builds the connection between MATSim tool and all other Extensible Markup Language (XML) files (e.g., network, population, etc.), and contains a list of settings that influence how the simulation behaves. MATSim's *network file* consists of nodes and links and describes the infrastructure that agents can use to move around. Nodes are defined by coordinates while the link requires definition of several attributes including the length of the link, capacity, speed, and the number of available lanes. The *population file* provides information about travel demand, e.g., a list of agents and their travel diaries. The travel demand is described by the daily plans of each agent. The *population file* contains a list of transportation users and their daily plans, activities, and legs.

Each simulation job executes in iterations where each iteration executes the selected plans of all agents over an underlying road network (Horni et al., 2016). Figure 9 shows the execution flow of a MATSim job (Horni et al., 2016). It starts with an initial population demand (a.k.a. plans) in the studied area which is application dependent. Each agent in the population maintains a pool of up to 5 day-plans. In each iteration, (i) MATSim's "mobsim" simulation executor runs the selected plans of the agents in the synthetic road network environment; (ii) then, a scoring function assigns a score to each plan based on the corresponding agent's experiences with the selected day plans (e.g., if congestion happens or not); (iii) afterwards, the replanning step selects a candidate plan based on the plan scores in each agent's day-plan pool, and may modify this plan for execution in the next iteration.

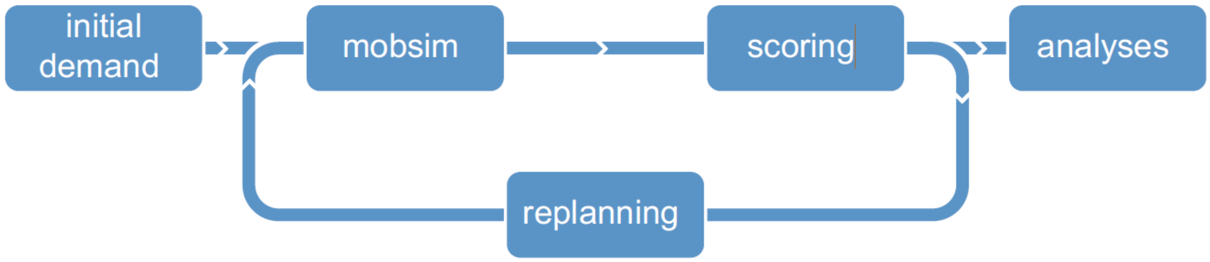


Figure 9. The Execution Flow of MATSim (Horni et al., 2016)

Specific details on our approach for computing the mobility demand around the UAB campus for the Birmingham case study are available in Section 4.3.

As far as *simulation outputs* are concerned, MATSim creates output data that can be used to monitor the current simulation setup progress as well as to analyze results. In each iteration, a *linkstats* file containing hourly count values and travel times on every network link is generated by the model. The user can specify the output interval of simulation statistics for individual links. MATSim provides overview summaries of counts and other statistics for the whole network, but also analyzes for individual links. Also, a google maps-based visualization is available, showing simulation output results for each station in a pop-up window.

For the purposes of our study that is interested in the introduction of shared micromobility travel options, we need to use an extended module of MATSim. However, MATSim does not have an extended module that directly supports e-scooters for shared micromobility simulation. The closest available extended module to our needs is MATSim's carsharing module, which only supports shared cars (e.g., Zipcar) rather than other vehicle types such as e-scooters.

To address this issue, we revised MATSim's carsharing module to simultaneously support multiple types of shared micromobility modes, including e-scooters. Specifically, we studied the code of MATSim's carsharing module, and made the necessary revisions so that it to allow users to add more shared- mobility modes. Each shared-mobility mode needs to specify a few attributes for proper simulation by MATSim. For the e-scooter mode, we adopted the following attributes which are consistent with those documented in published studies and/or reported by manufacturers:

- Passenger Car Unit (PCU): 0.2;
- Maxi Speed: 18.6 miles per hour (or 8.314944 meters per second);
- Length: 3.8 feet (or 1.15824 meters);
- Width: 1 foot (or 0.3048 meters);

Also, for each travel leg, MATSim has a parameterized score function that computes the cost of the leg based on the executed travel time and distance. For different modes, we configured the parameters of the score function to be different so that the most proper mode will be selected based on the leg length. For example, a 5-minute walk is preferred than driving, but driving is preferred over a 30-minute walk. E-scooters are preferred for a middle-ranged travel distance that is too long for walking but relatively short for driving so driving a car is not preferred, especially in downtown areas and University campuses where finding a parking spot is often a challenge. For example, the literature suggests that the typical trip distance for shared e-scooters is 1.6-1.9 km or 1-1.2 miles (Ensor et al., 2021). We modified the score function of the original carsharing module to integrate the e-scooter mode and allow it to be properly chosen among alternative modes.

In any MATSim model, each MATSim job is associated with a configuration file, which defines job parameters. In our study, in order to allow proper simulation with e-scooters as a mode of shared mobility, we made the following changes to the configuration file:

(1) MATSim uses a FIFO queue on each road link for vehicle traffic simulation by default, but this will cause the slower e-scooters to block the faster cars following them. We configured the job to use MATSim's "PassingQueue" instead, which allows e-scooters to use a designated lane so that they will not block the cars. This aligns better with the real situation around UAB where dedicated bike lanes are available, and

(2) MATSim's "subtourmodechoice" strategy for replanning was enabled, so that MATSim can change modes between iterations. This would allow a more reasonable mode setting to be explored by the co-evolutionary framework of MATSim, so that the best plan (using e-scooters or not) can be retained and selected due to its better score.

The MATSim's carsharing module also uses an input file called "CarsharingStations.xml" where users need to specify the parking locations of the shared vehicles. In our case, e-scooters are dockless (a.k.a. freefloating in MATSim's carsharing terminology), so the stations are just used to specify the initial locations of e-scooters when a day begins (for simulation). Even though e-scooters are dockless, the companies will redistribute their e-scooters at the end of the day to a set of pre-designated stations to meet the demands for the next day, which aligns well with our use of "CarsharingStations.xml" for simulation. To increase the realism of our simulation scenarios, we configured CarsharingStations.xml to match the Veo's stations in the Birmingham micromobility pilot deployment. Figure 10 shows the e-scooter stations for the 2021 pilot deployment in Birmingham.

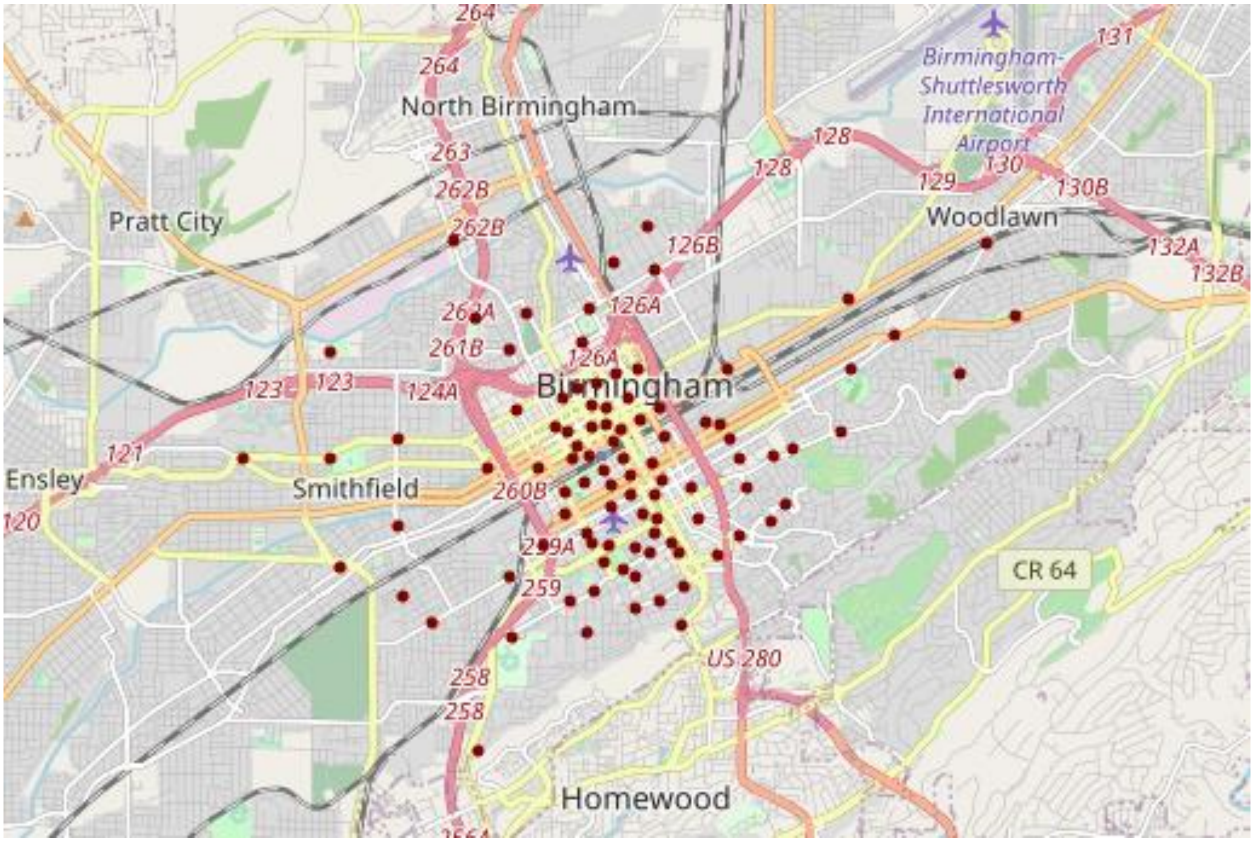


Figure 10. Veo e-Scooter Stations in the 2021 Birmingham Pilot Deployment

4.3 MATSim plan generation for the Birmingham study

MATSim requires users to provide a file named “plans.xml” containing the day-plans of each agent (traveler) to run a simulation job. So, the key to a successful simulation is to generate realistic population day-plans in the simulated area. We considered three types of traffic that constitute the majority of UAB traffic:

(1) Traffic generated by students attending classes.
(2) Traffic generated by employees going to work.
(3) Background traffic where travelers are just passing through UAB, where UAB is neither origin nor destination.

4.3.1 Data background traffic generation

In Guo et al. (2019a), we used MATSim to generate a transportation simulation for the entire City of Birmingham with the help of a small seed mobility survey of 451 participants, where day-plan details are enriched with the help of various open data sources that we completed as part of STRIDE B project (Sisiopiku et al., 2019; Sisiopiku et al., 2021; Sarjana et al., 2020). There, we partitioned the city into ZIP Code Tabulation Areas (ZCTAs), and for each OD (origin-destination) pair (where origin refers to "Home"), we determined the number of trips by iterative proportional fitting (IPF) seeded from the mobility survey data and scaled up with the help of ZCTA aggregates (obtained from open data sources) serving as IPF marginals. Since the seed mobility survey was small, the OD matrix obtained was sparse with many zeros (see Figure 12 of Guo et al, 2019a). To overcome this issue, we smoothed this OD matrix with an OD-marginal-product matrix by taking a weighted sum and tuned the weight parameter so that the simulated traffic on a validation set of roads align best with the real road monitoring data from National Performance Management Research Data Set (NPMRDS). The resulting OD matrix A is shown in Figure 11.

	35215	35173	35206	35209	35226	35205	35242	35223	35207	35233	...	35080
35215	1709.557371	283.697728	370.190938	422.086864	567.395457	141.848864	1065.596346	366.731210	290.617185	1377.423445	...	859.968278
35173	272.777556	823.857093	237.294296	270.559852	363.703407	90.925852	683.052741	235.076593	186.287111	59.878000	...	139.715333
35206	1429.891815	97.254025	126.904642	144.695012	194.508049	48.627012	365.295605	125.718617	99.626074	674.028056	...	74.719556
35209	343.422074	228.948049	298.749284	982.635414	457.896099	114.474025	1501.956599	937.962624	876.537537	75.385333	...	175.899111
35226	388.045259	258.696840	337.567827	1668.901198	517.393679	129.348420	2255.701346	334.412988	907.011908	727.186056	...	198.754889
35205	878.238278	799.493982	205.503407	234.312296	314.977185	2004.760464	591.542519	203.582815	161.329778	1977.872168	...	120.997333
35242	594.090000	1038.065389	516.810000	1873.270778	792.120000	840.035389	3413.656168	511.980000	405.720000	2056.426168	...	304.290000
35223	104.425481	69.616988	90.841679	103.576494	139.233975	676.813883	261.488198	731.998081	71.314963	664.928056	...	53.486222
35207	56.847259	37.898173	49.452494	56.385086	75.796346	18.949086	142.349235	48.990321	38.822519	12.478667	...	29.116889
35233	18.197926	12.131951	15.830716	18.049975	24.263901	6.065975	45.568790	657.688155	12.427852	3.994667	...	9.320889
35211	212.231185	141.487457	826.629266	210.505728	282.974914	70.743728	531.440691	182.898420	144.938370	46.587333	...	108.703778
35235	205.109333	778.744945	178.428444	203.441778	273.479111	68.369778	1155.612500	176.760889	140.074667	45.024000	...	105.056000
35222	92.945481	61.963654	80.855012	92.189827	123.927309	30.981827	232.741531	80.099358	63.474963	20.402667	...	47.606222
35243	188.484593	125.656395	163.966272	186.952198	251.312790	1346.838976	1113.983068	162.433877	128.721185	683.380056	...	96.540889
35203	20.132519	655.427068	659.519044	19.968840	26.843358	648.716229	50.413136	17.349975	13.749037	4.419333	...	10.311778
35216	416.766519	277.844346	362.552988	1697.388951	555.688691	780.927562	1685.615858	1001.170031	284.621037	1375.496112	...	213.465778
35217	106.062444	70.708296	92.265704	105.200148	141.416593	35.354148	265.587259	91.403407	72.432889	665.287389	...	54.324667
35244	397.016667	264.677778	345.372222	1035.794278	529.355556	132.338889	1636.160945	2268.160612	271.133333	87.150000	...	1487.360778
35208	112.461481	74.974321	97.832346	111.547160	149.948642	37.487160	281.610864	96.918025	76.802963	666.692056	...	57.602222

Figure 11. ZCTA OD Matrix A for Background Traffic

UAB spans ZCTAs 35233 and 35205. To generate the background traffic, we further set A[i] [35233] and A[i][35205] to 0 for all origin ZCTAs i since we have counted their traffic in the first two types for UAB students and employees. For the remaining OD pairs, we generate their day-plans similarly as in [6] to count traffic that merely pass through UAB.

4.3.2 Student traffic generation

Figure 12 shows the three types of student plans that we adopted. Each plan consists of a sequence of activities connected by legs. For example, Plan 1 has an activity sequence "Home → Parking → Class → Parking → Home" describing a student's day plan where he/she drives from home to a parking lot/deck on the campus, parks his/her car and then walks to the classroom to attend the class; after the class ends, the student then walks back to the parking lot/deck, and then drives home. In Plans 2 and 3, the student goes directly to the classroom from home, either on foot (e.g., he/she lives in an on-campus apartment near the classroom) or by car (e.g., roadside parking).

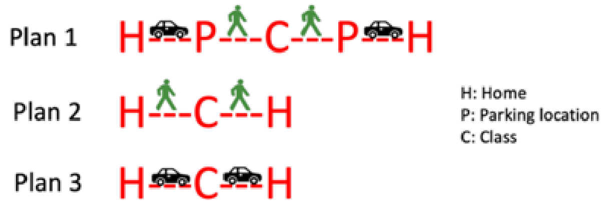


Figure 12. Types of Student Plans Considered

In a plan, each activity has an associated location, and each leg has a starting and end time for the travel, as well as a trip mode (e.g., car, walk or e-scooter). Figure 13 shows how to specify the day-plans of an agent in MATSim.

```
<person id="p_1">
    <plan selected="yes">
        <activity type="Home" x="518513.90340469655" y="3706169.0093547124" start_time="00:00:00" end_time="07:43:55">
        </activity>
        <leg mode="car">
        </leg>
        <activity type="Class" x="517955.12758688163" y="3706776.006161157" start_time="07:56:29" end_time="10:50:29">
        </activity>
        <leg mode="car">
        </leg>
        <activity type="Home" x="518513.90340469655" y="3706169.0093547124" start_time="11:03:03" end_time="24:00:00">
        </activity>
    </plan>
```

Figure 13. A Type-3 Student Plan in “plans.xml” for MATSim

Simplification Assumption. Figure 4-5 assumes that a student always attends one class and then goes back home. This may not always hold true since a student may, for example, take two classes before going back home, or have lunch in a cafeteria after class before going back home. However, since we do not have data on the diversified possibility of day-plans, but we know UAB’s class information, this simplified assumption is the best we can reach based on the principle of Occam's razor.

The plans in Figure 4-5 are for the baseline scenario where e-scooters are not used. If e-scooters are allowed, the “walking” legs in Plans 1 and 2 can be replaced by “e-scooter” legs. To generate realistic plans, we need to build our plan generation process upon real data. Therefore, we first describe our data sources used to generate realistic locations for (1) home/apartments, (2) parking lots/decks, and (3) classrooms.

Home Locations. Building on earlier work by Sisiopiku, 2018 and Sisiopiku and Ramadan, 2017 we had access to a mobility survey of 4,137 UAB students and 4,920 employees of UAB, which contained information about their daily trips commuting between home and the UAB campus. There were two challenges in directly using the survey data. First, to protect the privacy of the participants, the exact home locations (e.g., building numbers) were not disclosed, but rather information was entered in the form of "nearest intersection" + "zip code" + "city." Figure 14 shows an illustration of the relevant columns in the data table.

How do you enter the UAB campus? Response	Where do you live?City	Where do you live? Nearest intersection (e.g. Hickory Trc and Magnolia Dr)	Where do you live? Zip Code
From US 280	Birmingham	NaN	35242
I live on campus	Birmigham	280	35205
From I-65 Northbound (traveling from the south)	McCalla	Hwy 459/20/59	35022
From US 280	Birmingham	Highland Lakes	35242
From an arterial street	Birmingham	Gravlee Lane	35206

Figure 14. A Home-Related Columns from UAB Mobility Survey

To tackle this problem, we used positionstack (https://positionstack.com/), an address geocoding API, to get an approximate home location for each participant in GPS coordinates. Some records may miss the "nearest intersection" information, in which case we use a random address from the zip code area, sampled from the addresses obtained from OpenAddresses (https://openaddresses.io/), so that the sampled location is populated (e.g., not on a lake).

It should be noted, that in our study and as per Figure 12 we considered 3 types of student-plans: Plan 1 is for students living off campus, Plan 2 is for students living on campus, and Plan 3 can be for both. We, therefore, use a UAB campus polygon (from OpenStreetMap, see Figure 15) to separate the (approximate) "Home" locations in the survey into two sets: an on-campus set S_{on} and an off-campus set S_{off}.

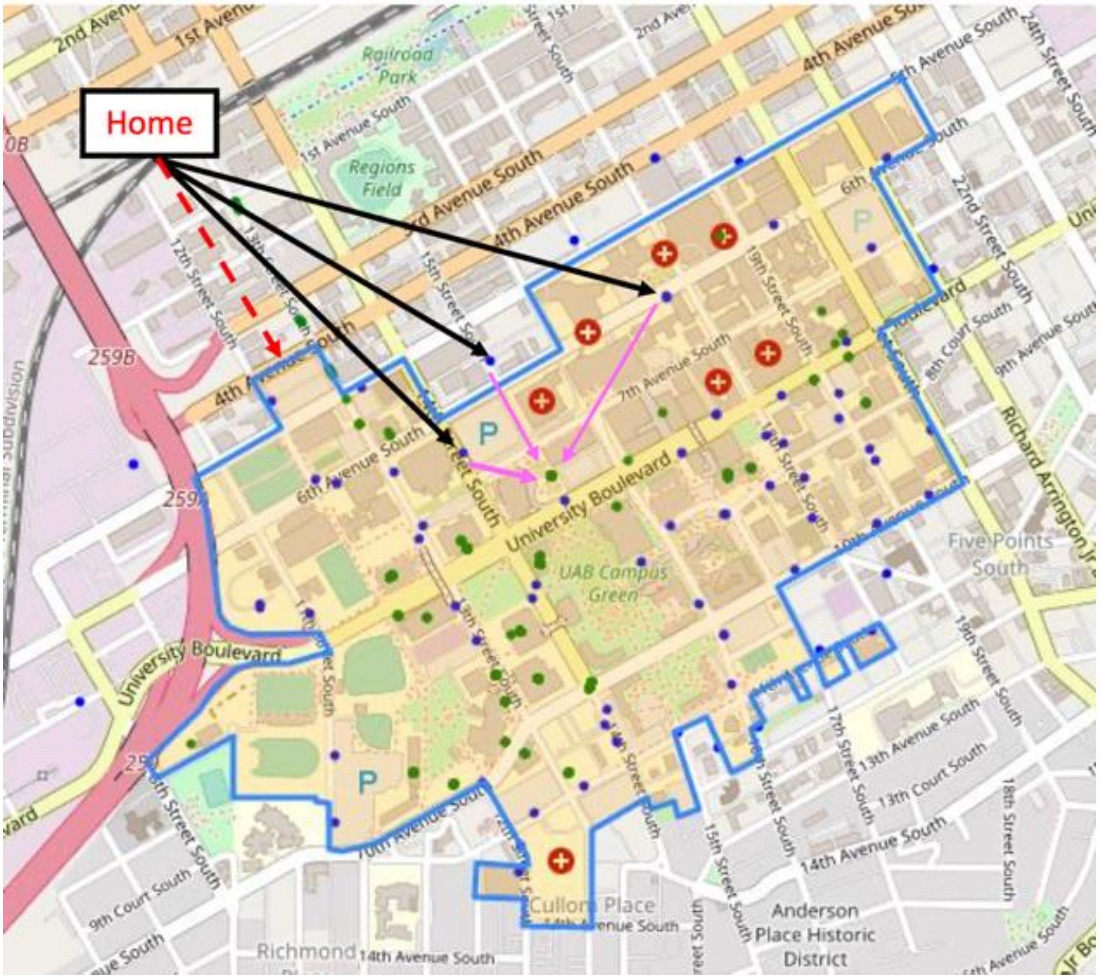


Figure 15. The UAB Campus Polygon from OpenStreetMap, and Parking Score Computation

The second challenge that we faced is that the survey only included a certain portion of all the UAB employees and students, but MATSim simulation requires the day-plans of an entire population. Therefore, we used Kernel Density Estimation (KDE) to fit the 2D location distribution of the on-campus set Son (resp. off-campus set S_{off}), denoted by P_{on} (resp. P_{off}). The tuned KDE bandwidth parameter is 0.001 (resp. 0.01) for fitting P_{on} (resp. P_{off}). Later, when we generate home locations for students, we sample from this distribution on demand.

However, a sampled home location may not correspond to an actual home/apartment location. We, therefore, needed to align the sampled location to a nearby apartment/address location for realistic simulation. For this purpose, we collected and used data on the locations of apartments inside and round the City of Birmingham from online sources (Figure 16). Such sources included apartments.com, rent.com, etc., and on-campus apartments from https://www.uab.edu/students/housing/residence-halls.

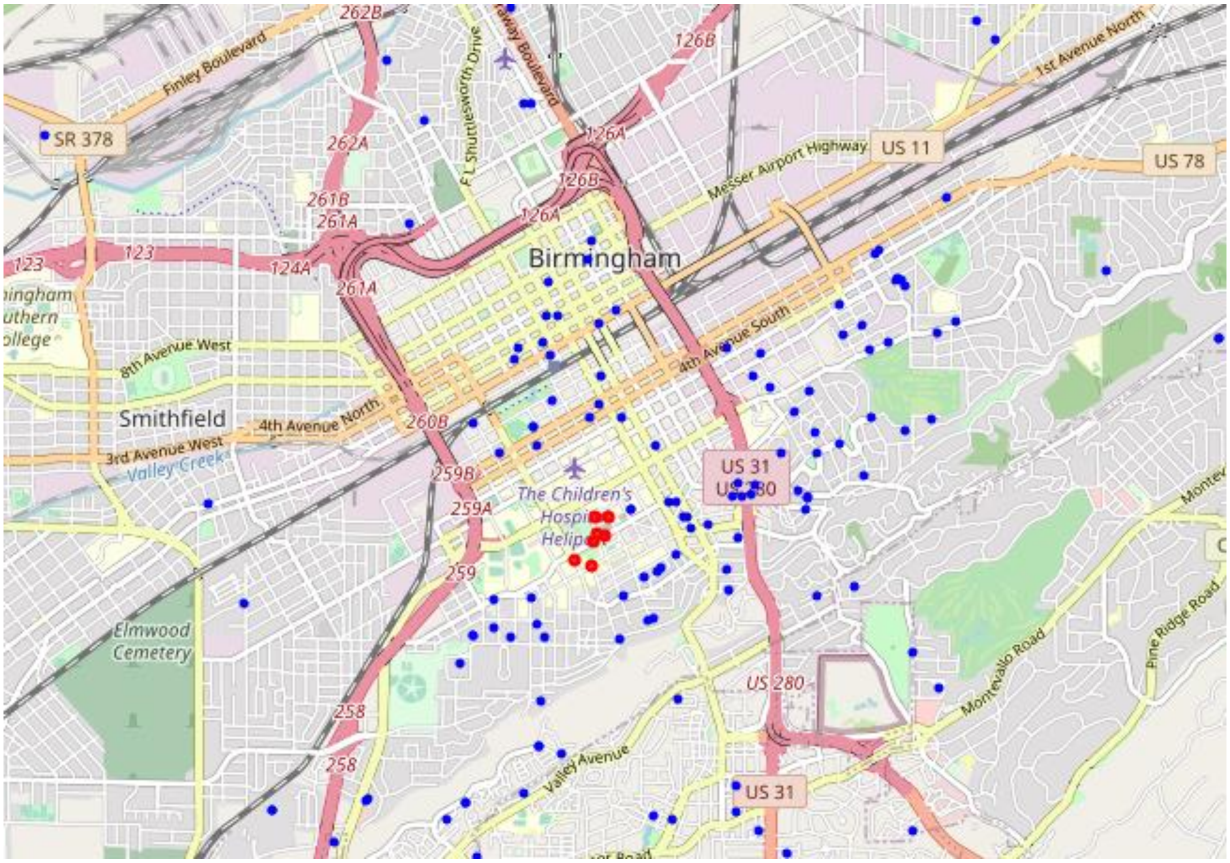


Figure 16. UAB On-Campus (Red) and Off-Campus (Blue) Apartments

For a student living on campus, we sampled a location p_{on} from the on-campus distribution P_{on}, and then align it to an on-campus apartment as follows:

- We first tried to align the sampled location p_{on} to a random apartment within 300 meters (by a range query over an R-tree indexing all apartment locations);
- If we could not find such an apartment, we then sampled a random on-campus apartment.

For a student living off campus, we sampled a location p_{off} from the off-campus distribution P_{off}, and then aligned it to an off-campus apartment as follows:

- We first tried to align the sampled location p_{off} to a random apartment within 300 meters;
- If we could not find such an apartment, we then sampled a random address (from OpenAddresses) within 300 meters;
- If such an address cannot be found, we then found the nearest address to p_{off} for alignment (using a nearest neighbor query over an R-tree indexing all addresses).

Classes. As per Figure 12, we had to determine students' day-plans for attending classes, so we also needed to obtain classroom locations/buildings. We used UAB Class Schedule Listing (UAB, 2021) to obtain the class information of a particular day for simulation purpose. The day we chose as a typical class day was a Wednesday in Fall 2020, which was before the guidance for UAB remoting teaching due to the COVID-19 pandemic went into effect.

Notably, besides the location of a class c (denoted by $p(c)$), we also had access to its starting time $t_s(c)$ and end time $t_e(c)$, and the number of students $n(c)$ in the class. Such data were useful in our student plan generation process to be described later.

Parking Locations. As per Figure 12, Plan 1 for off-campus students involves parking locations, which can be either parking lots or parking decks. We obtained and used in our MATSim model detailed data on the parking locations from UAB Transportation, including the parking type (lot or deck), address, and parking capacity. We also obtained their GPS coordinates by the positionstack API.

Parking Location Determination. Assuming that we have already determined the home location p_h and classroom location p_c of a student plan, the next task is to determine the parking location p_p that the student uses to park his/her car after arriving on campus. Intuitively, students would prefer parking locations closer to their classrooms. However, in busy hours where many classes are ongoing, the parking locations nearby a classroom may be full, forcing students to find a farther-away parking location with available space. Moreover, for two such parking locations that have similar distance to the destination classroom, a student would prefer the one on the side of the campus that is closer to his/her home.

We regard finding a parking location as the following process: a student has a list of on-campus parking locations in mind ordered by preference based on home and classroom locations; he/she tries the parking locations down the list one by one until a parking location with available space is found.

We assigned the preference score to each parking location p_p, denoted by $s(p_p)$, based on two distances (1) $d_{hp} = distance(p_h, p_p)$ and (2) $d_{pc} = distance(p_p, p_c)$:

$$s(p_p) = \alpha \cdot d_{hp} + (1 - \alpha) \cdot d_{pc}, \qquad (1)$$

where α balances the importance of d_{hp} and d_{pc}. We adopted $\alpha = 0.3$, which gives more weight to d_{pc} so that a student would prefer a parking location closer to the classroom to one that is closer to his/her home. In Figure 15, the black arrows correspond to d_{hp}, and the pink arrows correspond to d_{pc}.

One problem remains: since we are considering off-campus trips, the distance d_{hp} is usually much longer than d_{pc}, so the first term in Eq (1) is overrated. To avoid this issue, we deducted from d_{hp} a constant distance that the student has to travel anyway no matter which on-campus location (including parking locations) he/she arrives. This constant distance is denoted by $d_{h\text{-}UAB}$, computed as the distance from home to the nearest point of the UAB campus polygon (see the dashed red arrow in Figure 15).

Thus, we actually used the following formula (Eq. 2):

$$s(p_p) = \alpha \cdot (d_{hp} - d_{h\text{-}UAB}) + (1 - \alpha) \cdot d_{pc}, \qquad (2)$$

We computed d_{hp}, $d_{h\text{-}UAB}$ and d_{pc} for on-campus parking locations p_p using Dijkstra's algorithm. Notably, we can compute d_{hp} and d_{pc} for all on-campus parking locations (125 in total) together using just 2 calls of Dijkstra's algorithm (single-source multi-destination), and the steps are listed as follows:

- Find d_{hp} for all parking locations p_p by running one Dijkstra's algorithm from p_h on the underlying road network, until all the 125 parking locations have been visited.
- Find $d_{h\text{-}UAB}$ by running one Dijkstra's algorithm from p_h on the underlying road network, until all the border points on the UAB campus polygon are visited; then $d_{h\text{-}UAB}$ takes the distance value between p_h and the closest UAB border point.
- Find d_{pc} for all parking locations p_p by running one Dijkstra's algorithm from p_c on the reverse graph of the underlying road network (i.e., the directions of all edges are reversed), until all the 125 parking locations have been visited.
- Compute $s(p_p)$ for all parking locations p_p by Eq (2).

Now that we have scores for all parking locations p_p, we can sort the parking lots in non-decreasing order of their scores $s(p_p)$ to obtain the preference list L_p. Recall that we will have to go down the list to find the first parking location with available space. We used a distribution to determine if a parking location has space. Poisson distribution is a discrete probability distribution that expresses the probability of a given number of events (denoted by X) occurring in a fixed time interval, if these events occur with a known constant mean rate λ and independently of the time of past events. The probability mass function (PMF) is:

$$f(k;\lambda) = \Pr(X{=}k) = \frac{\lambda^k e^{-\lambda}}{k!}, \qquad (3)$$

where k is the number of occurrences (k = 0,1, 2, ...), and λ is equal to the expected value of X (i.e., $\lambda = \mathrm{E}(X)$). Through the use of the Poisson distribution, we could answer the question: what is the probability of finding k vehicles in a parking location p_p at a specific time t (e.g., right before the class beginning time), such that $k <$ the parking capacity of p_p, denoted by $c(p_p)$. If this does not hold, i.e., $k \geqslant c(p_p)$, then p_p is not available, and we need to continue checking the next parking location in the preference list L_p.

Let the Poisson cumulative distribution function (CDF) be:

$$CDF(K;\lambda) = \sum_{k=1}^{K} f(k;\lambda), \qquad (4)$$

In other words, p_p is available only with probability $CDF(c(p_p) - 1; \lambda)$. In the preference list L_p, we choose the current parking location p_p with probability $\Pr(X < c(p_p)) = CDF(c(p_p) - 1; \lambda)$ by sampling a random number in [0, 1] and check if it is less than $\Pr(X < c(p_p))$; otherwise, we move on to check the next parking location in L_p. One problem remains: λ should be the expected number of vehicles in parking location p_p at time t when the student arrived at p_p, and we next explain how to obtain $\lambda \triangleq \lambda(p_p, t)$, which is a function of location p_p and time t.

Given a time t when a student arrives at a parking location p_p, we would like to know the ongoing classes c (i.e., $t_s(c) \leqslant t \leqslant t_e(c)$) happening near p_p so that we can compute $\lambda(p_p, t)$ to be used for availability check. Let us denote the set of ongoing classes near p_p by $C_t(p_p)$. We assumed that $C_t(p_p)$ includes all ongoing classes happening within a distance of 150 meters from the parking location p_p, though this distance threshold is tunable. Then intuitively, the larger $|C_t(p_p)|$ is and the larger the sizes of classes $c \in C_t(p_p)$ (i.e., $n(c)$) are, then the larger the value of $\lambda(p_p, t)$ should be since $\lambda(p_p, t)$ is the expected number of vehicles in p_p at time t. Based on the above intuition, a straightforward way to set $\lambda(p_p, t)$ is to compute:

$$\lambda(p_p, t) = \sum_{c \in C_t(p_p)} \gamma \cdot n(c), \qquad (5)$$

where for each nearby class c, we assume a certain fraction γ of all the $n(c)$ students will park at p_p, where γ is a tunable parameter. However, for different classes, their γ's could be different. For example, a class c_A may have many nearby parking locations while another class c_B may only have one nearby parking location; in this case, γ should be smaller for c_A than c_B since students of class c_A has more parking options. We, therefore, defined γ to be class-dependent as $\gamma(c) = 1/|P(c)|$, where $P(c)$ is the set of parking locations near c.

We assumed that $P(c)$ includes all parking locations within a distance of 150 meters from class c, but this distance threshold is tunable. Intuitively, we here split students of a class c evenly among its nearby parking locations:

$$\lambda(p_p, t) = \sum_{c \in C_t(p_p)} \gamma(c) \cdot n(c) = \sum_{c \in C_t(p_p)} \frac{n(c)}{|P(c)|}. \qquad (6)$$

To compute $\lambda(p_p, t)$ efficiently, we built an interval tree $T(p_p)$ for each parking location p_p beforehand, so that given a timestamp t when a student arrives at a parking location, we can look up $C_t(p_p)$ to compute $\lambda(p_p, t)$ using the above formula. Here, $T(p_p)$ keeps all the classes that are near p_p (within 150 meters), indexed by their time intervals $[t_s(c), t_e(c)]$ as the key, and stores a weight value $w(c) = n(c)/|P(c)|$. Given a time t as a point-query, $T(p_p)$ returns all ongoing classes $c \in C_t(p_p)$ since $t \in [t_s(c), t_e(c)]$.

We built the interval trees $T(p_p)$ for all the 125 parking locations p_p together as follows: for each class c, we obtained its nearby parking locations $P(c)$, and for each $p_p \in P(c)$, we inserted entry <key = $[t_s(c), t_e(c)]$, value = $n(c)/|P(c)|$> into the interval tree $T(p_p)$.

Plan Time Generation. Recall that we need to determine the timestamp t to arrive at a parking location for Plan 1 in Figure 16. Additionally, we need to decide when a student should leave home in order not to be late for class. We next discuss how we determine these timestamps for different types of plans.

Figure 17 shows a plan where a student travels directly from home to class, and then back home. The travel mode could be either car, walk, or e-scooter. For an activity (event) e, we denote its starting (resp. end) time by $t_s(e)$ (resp. $t_e(e)$). For example, the time to leave home is denoted by $t_e(e_h)$, i.e., the end time of the "staying home" event e_h; and the time to attend a class c is denoted by $t_s(e_c)$, i.e., the starting time of the "attending class c" event e_c.

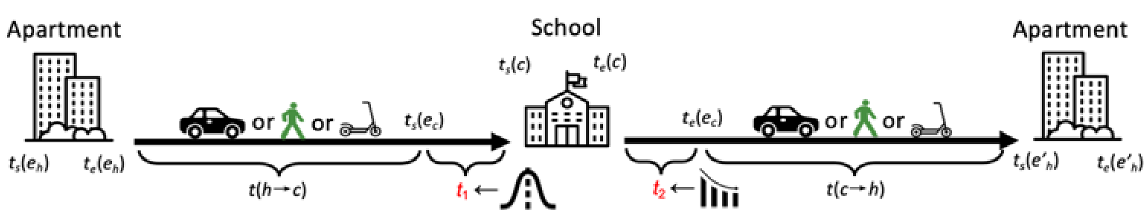


Figure 17. Time Sampling for a Direct Home-to-Class Plan

Additionally, we denote the time taken by the leg "Home → Class" by t(h→c), and the time taken by the leg "Class → Home" by t(c→h), both of which can be computed using Dijkstra's (or A*) algorithm on the underlying road network (depending on the vehicle speed for the selected travel mode, as well as the road freeflow speeds).

To generate a realistic leaving-home time $t_e(e_h)$ for a student plan, we needed to make sure that the student is able to arrive at the classroom before the class c begins, i.e., $t_s(e_c) < t_s(c)$. In particular, we had to count the travel time t(h→c) and an additional amount of time t_1 describing how much earlier a student typically tends to arrive at the classroom before a class begins as shown in Eq. 7.

$$t_e(e_h) = t_s(e_c) - t(h \rightarrow c) = t_s(c) - t_1 - t(h \rightarrow c) \quad (7)$$

To ensure the diversity of classroom arrival time by different students, we sample t_1 from a normal distribution with 10 minutes as the mean, and a standard deviation of 5 minutes, modeling the fact that students tend to arrive at the classroom 10 ± 5 minutes earlier than the class beginning time $t_s(c)$.

We also needed to generate the leaving-class time $t_e(e_c)$ and the back-home time $t_s(e'_h)$, where we denote the event of "staying home" after the class by e'_h to differentiate from the one before the class. Students do not leave before a class ends, and many students leave the classroom a few minutes after $t_e(c)$, e.g., to discuss among peers, to ask the instructor questions, or just because the classroom door has a limited passing rate. We denote the length of the time from when class c ends to when a student leaves the classroom by t_2, and to impose diversity, we sample t_2 from an exponential distribution with rate parameter λ = 2 minutes. Note that since E(t_2) = λ in the exponential distribution, we assume that on average a student leaves the classroom 2 minutes after the class ends. Exponential distribution is reasonable since most students leave the classroom quickly after the class, and only some students stay for a longer time. One may also use the gamma distribution instead with a proper shape parameter, to account for the fact that students need some time to move to the classroom door so that the density peak (aka. mode) is not at $t_2 = 0$. Figure 18 below illustrates the probability density functions (PDFs) of exponential and gamma distributions:

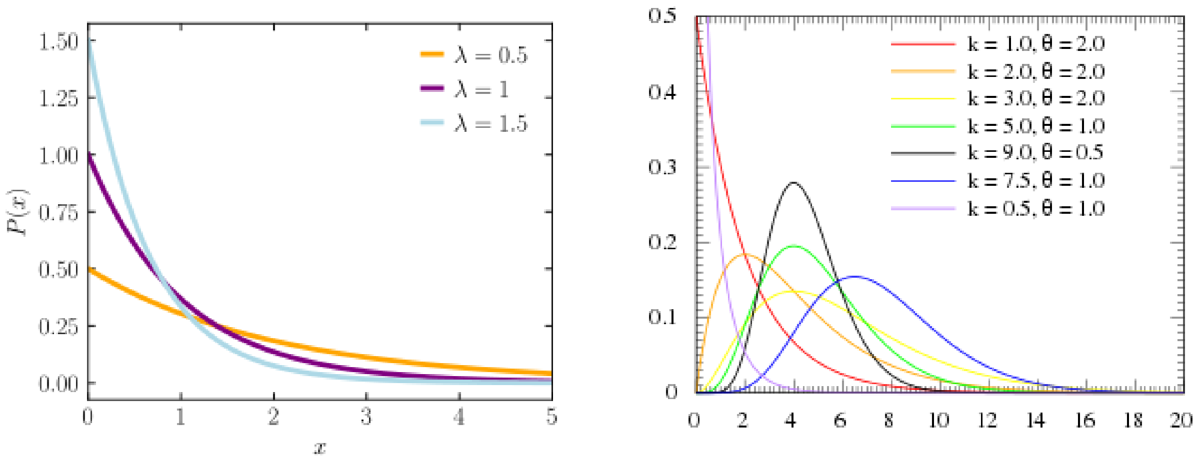


Figure 18. PDFs of Exponential Distributions (Left) and Gamma Distributions (Right)

According to the above discussion, we should set the leaving-class time $t_e(e_c)$ and the back-home time $t_s(e'_h)$ as follows:

$$t_e(e_c) \; = \; t_e(c) + t_2 \qquad (8)$$

$$t_s(e'_h) \; = \; t_e(e_c) + t(c \rightarrow h) \; = \; t_e(c) + t_2 + t(c \rightarrow h) \qquad (9)$$

So far, we only considered plans without a parking location, i.e., Plans 2 and 3 in Figure 12. For Plan 1 in Figure 12, we also need to generate the plan timestamps related to the parking events (assuming that the parking location has been determined as described previously). Figure 19 summarizes time sampling for a plan with parking location.

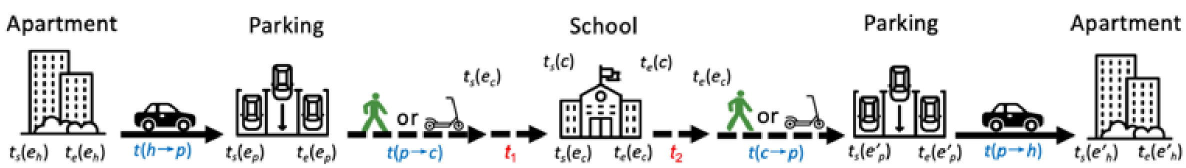


Figure 19. Time Sampling for a Plan with Parking Location

For "Home → Parking → Class," we denote that parking event by ep, and denote the travel times of the two legs by t(h→p) and t(p→c), respectively. Similarly, for "Class → Parking → Home," we define parking event e'_p, and travel times t(c→p) and t(p→h). Additionally, we assumed a typical time spent in the parking location to be 2 minutes. By default, MATSim increases the score of a plan if the activity (resp. leg) duration is longer (resp. shorter) in simulation. To avoid long waiting time in parking lots, we revised the plan score calculation code in MATSim to treat "Parking" also as a leg.

In summary, for Plan 1 in Figure 12, we compute for "Home → Parking → Class", where $t_{\text{park}} \sim N(\mu = 2 \text{ min}, \sigma = 1 \text{ min})$:

$$t_e(e_p) = t_s(c) - t_1 - t(p \rightarrow c), \quad t_s(e_p) = t_e(e_p) - t_{\text{park}} \qquad (10)$$

$$t_e(e_h) = t_s(e_p) - t(h \rightarrow p) \qquad (11)$$

and we compute for "Class → Parking → Home":

$$t_e(e_c) = t_e(c) + t_2 \qquad (12)$$

$$t_s(e'_p) = t_e(e_c) + t(c \rightarrow p), \quad t_e(e'_p) = t_s(e'_p) + t_{\text{park}} \qquad (13)$$

$$t_s(e'_h) = t_e(e'_p) + t(p \rightarrow h) \qquad (14)$$

Finally, for Plan 1 in Figure 12, we computed travel times $t(p \rightarrow c)$ and $t(c \rightarrow p)$ using mode "walk" rather than "e-scooter" since MATSim may select either one during simulation when mutating the leg modes in its co-evolutionary algorithm, so calculating time using the slower "walk" mode ensures that students would not attend class late or leave class early. Similarly, for Plan 2 in Figure 12, we compute travel times t(h→c) and t(c→h) using mode "walk" rather than "e-scooter."

Putting Things Together. Now that we have explained how we chose a parking location for an off-campus student plan, and how we determined the activity starting/end time for a plan, we are ready to describe our plan generation procedure below:

Algorithm 1: Generation of Student Plans

(Note: Walk Modes May Later be Mutated into E-scooter Modes)

```
for each class c ∈ ℂ:
    repeat for n(c) times:
        create a student plan object o
        o's on-campus flag ← sampling with prior probability P_UAB obtained from survey
        if o's on-campus flag is set:
            o's home location ← sampling from P_on = KDE(S_on)
            align o's home location to an actual on-campus apartment location
        else:
            o's home location ← sampling from P_off = KDE(S_off)
            align o's home location to an actual apartment or OpenAddresses address
            set o's parking flag with a prior probability P_park
        if o's on-campus flag is not set and o's parking flag is set:
            determine o's parking location
            determine o's activity starting and end timestamps, with modes car and walk
        else:
```

```
16:        if walking distance > 30 minutes:   o's leg mode ⟵ car
17:        else:   o's leg mode ⟵ walk
18:        determine o's activity starting and end timestamps
19:     write the created student plan object o to "plan.xml"
```

Specifically, let $\mathbb{C}$ be the set of all UAB classes on Wednesday in Fall 2020, the day we chose to simulate. For each class $c \in \mathbb{C}$, there are $n(c)$ students taking the class, so Line 2 generates $n(c)$ student plans for class c.

For each student plan o that we created in Line 3 for class c, we set the activity locations and timestamps in Lines 4-18, and then write the generated plan o to "plan.xml" in Line 19 to be inputted to MATSim for simulation later. Our earlier survey of 4,137 UAB students, can provide information to help us obtain a prior probability P_{UAB} that a student lives on campus. In Line 4, we therefore set o to be an on-campus plan with probability P_{UAB}, and off-campus otherwise. Then, Lines 5-18 generate the plan's location and timestamp information as follows.

First, consider an off-campus plan. Line 9 samples the home location from the distribution P_{off} that we obtained by KDE fitting over S_{off}, which is the set of off-campus home/apartment locations reported by students in survey. Line 10 then aligns the sampled location to an actual apartment or home address. Line 11 decides if the student parks at a parking location or just parks roadside near the classroom, using a prior probability P_{park} (set as 10%) which acts as a tunable hyperparameter.

If a student parks, Lines 13 and 14 will determine the parking locations and activity timestamps as we discussed previously. While if a student parks roadside, Lines 16-18 will set the travel mode and activity timestamps accordingly as we discussed previously. Note that even if a student lives off campus, he/she may still walk or ride an e-scooter to the classroom, if the student lives near UAB with a walking distance $\leqslant$ 30 minutes.

Next, consider an on-campus plan. Then Line 6 samples the apartment location of the student from the distribution P_{on} that we obtained by KDE fitting over S_{on}, which is the set of on-campus apartment locations reported by students in the survey. Line 7 then aligns the sampled location to an actual on-campus apartment. Subsequently, Lines 16-18 will set the travel mode and activity timestamps accordingly as we discussed previously. Note that, according to our design, students are most likely to walk or ride an e-scooter to the classrooms since they live on campus, and the condition in the if-branch in Line 16 is most likely *false*, but we do not rule out the possibility that a student may drive to the classroom if his/her apartment is far from the classroom (e.g., the student has to travel across the entire campus). The threshold of 30 minutes is a tunable hyperparameter and can be further adjusted in future work.

4.3.3 Employee traffic generation

We next introduce how employee plans were generated in our study. Since most steps are the same as in student plan generation, we focus only on the differences. The first difference is that instead of classrooms, employees go to their workplaces. In our survey all the 4,920 employees report the building codes of their working places (Sisiopiku and Ramadan, 2017). Thus, so we can fit a working location distribution from these reported locations using KDE, denoted by P_{work}. Later, when we generate employee plans, we sample the work locations from P_{work}, and then align them to their nearest UAB buildings.

The second difference is that the student plans have class starting and end time, but the working hours for employees are unknown. We sample the arrival time at the workplace randomly from [8 am, 10 am] (i.e., uniform distribution), and sample the departure time randomly from [4 pm, 6 pm], since most UAB employees commute during peak hours.

The last difference is that in order to compute Poisson parameter $\lambda(p_p, t)$ for parking location sampling, we need to know the workplaces and arrival/departure time of all employees to be generated. So we sample them for all plans first, and for each workplace w, we obtain its nearby parking locations $P(w)$, and for each $p_p \in P(w)$, we insert entry <key = [$t_s(w)$, $t_e(w)$], value = $1/|P(c)|$> into the interval tree $T(p_p)$. Note that the interval trees should aggregate the class and workplace intervals from all student and employee plans before being used to compute $\lambda(p_p, t)$ for parking location availability sampling. That also means that for all the partially generated employee plans mentioned above, we have to then determine their parking locations, the departure time from home and the arrival time back home.

4.4 Simulation experiments and results

4.4.1 Experimental setup

Starting with the Birmingham MATSim model development in STRIDE Project B (Sisiopiku et al, 2019) and through updates and model refinements detailed in Guo et al. (2019a, and 2019 b) we have conducted a MATSim simulation by generating realistic day-plans for the entire population in Birmingham, AL. There, we obtained an origin-destination (OD) matrix M by iterative proportional fitting (IPF), where M[i][j] denotes the number of people traveling from region i to region j. Regions here correspond to the different ZIP Code Tabulation Areas (ZCTAs) from the US Census Bureau. Since UAB is in the zip code area 35233, we are essentially generating “Home → Class” plans that fall in M[i][ZCTA-35233] from different regions i in Birmingham.

We note that the background traffic is still important since traffic through the UAB campus may just be passing by, with neither source nor destination being UAB. Such background traffic can be obtained by generating day-plans similarly as in Guo et al. (2019a), but each M[i][ZCTA-35233] should deduct the number of “Home → Class” trips generated for UAB micromobility simulation. The background traffic generated by these plans, combined with the UAB student and employee plans generated for micromobility simulation, would produce a complete simulation.

The network, population and demand data were prepared for use with MATSim. For efficiency reasons, a 10 % sample of the population was used for our simulations. This is a common practice in activity-based models. In MATSim this is done by the *countsScaleFactor* parameter, which was set to 10 in our study. Details are available at Horni et al. (2016).

We considered four scenarios. Scenario 1 refers to baseline conditions with no e-scooter availability (baseline) and Scenarios 2, 3, and 4 consider availability of 500, 750, and 1,000 e-scooters, respectively. In the baseline scenario, we only enabled the car and walking modes. In the e-scooters experiment, we enabled the e-scooter mode, so MATSim can change the leg modes among car, walking and e-scooter.

In Scenario 2 we conducted our Birmingham micromobility simulation considering 500 e-scooters to match the number of scooters deployed as part of the 2021 Birmingham micromobility pilot program [City of Birmingham, 2021b]. Then in Scenarios 3, and 4 we increased the number of scooters to 750 and 1,000, respectively.

As mentioned earlier, MATSim runs in iterations, where each iteration executes the selected plans of all agents over an underlying road network. We chose the number of iterations to be 50. In addition to the background traffic for Birmingham as previously described, we also used the methodology proposed in Section 4.3 to generate the student and employee plans specific to the UAB campus. The number of students was directly available from the course enrollment numbers reported by UAB Class Schedule. We generated 20k employee plans. The number of UAB students and employees used in the MATSim simulation scenarios is comparable to the number of total students (22.85k) and the UAB employee number (27k) reported by Birmingham Business Alliance [BBA, 2021]. We assumed 80% students (resp. 90% employees) who drive a car would use parking lots/decks, and the remaining 20% (resp. 10%) park roadside.

MATSim's "*subtourmodechoice*" strategy for replanning was enabled, so that after each iteration, the leg modes may change among car, walk and e-scooter. However, MATSim does not provide a mechanism to lock the mode of a particular leg in a plan. More advanced customization on the allowed mutation restrictions would require changing the Java code of MATSim, which we avoid for now. However, a bad mutation (e.g., where a student walks from a home far from UAB) would get a very bad score and thus being dropped during the co-evolutionary algorithm of MATSim ultimately, and it would be selected with a very small probability since MATSim uses a logit model for plan selection (i.e., the selection probabilities are computed by softmax over the plan scores). We did observe a very small number of plans that are bad after the last round of simulation, but since the vast majority of the plans are good, the overall trend is informative and sheds light on the real traffic.

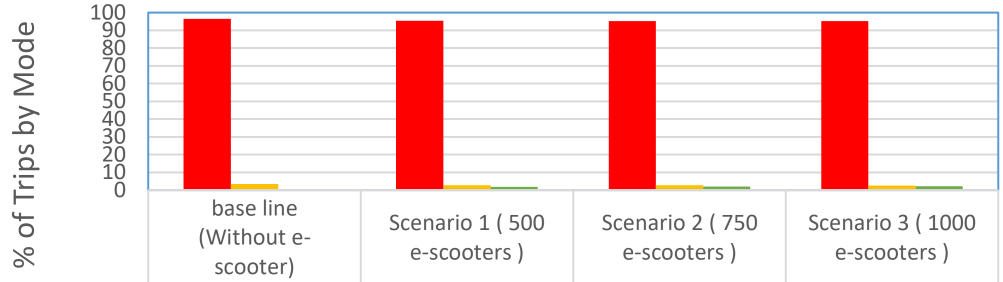
A demo video of our simulation with e-scooters (Scenario 2) can be found at https://youtu.be/zh_mHQ6ck4U.

4.4.2 MATSim simulation model outputs

In our study, we used Via (https://simunto.com/via/) to visualize the simulated traffic. Via is a software designed to visualize MATSim's output traffic. In Via, each vehicle is displayed using a customizable icon such as a green triangle ( ), where the color can be changed based on the vehicle speed and mode. Via also contains a time slider that can be used to choose a specific time for display. For example, one can move the slider to 8:00 AM to check the traffic at that time. The simulation video is replayed by sliding this time slider at a certain speed that can be adjusted.

The expectation from the deployment of shared e-scooter services at university campuses, such as UAB, is that they will contribute to the improvement of the on-campus traffic operations, especially during peak hours. Some students may choose to ride an e-scooter to the classrooms instead of driving thus reducing the traffic demand and the need of parking. Also, students and employees who live off campus may park their vehicle at a remote parking lot/deck and then travel to their trip destination by e-scooter rather than drive around campus until they can find roadside parking.

In order to quantify the impacts on traffic congestion on or around university campuses from mode choice shifts toward micromobility options, we can compare MATSim outputs from the simulation of a) baseline conditions (no micromobility options considered) and b) conditions with the presence of micromobility options (i.e, Scenarios 1 through 4 in our study). Figure 20 shows the distribution of trips by mode generated by MATSim for the Birmingham case study (i.e., car, walk and e-scooter) for all 4 simulation scenarios considered.



	base line (Without e-scooter)	Scenario 1 (500 e-scooters)	Scenario 2 (750 e-scooters)	Scenario 3 (1000 e-scooters)
Cars	96.49	95.45	95.26	95.15
Walk	3.51	2.73	2.72	2.63
Escooters	0	1.82	2.02	2.22

Figure 20. MATSim Simulated Mode Choices for Birmingham Case Study Scenarios

In our study, we obtained and compared two performance metrics (i.e., link hourly volumes and average speeds) from a sample of network links under four study scenarios (i.e., with and without e-scooter availability). For demonstration purposes, in this report, we share a sample of our findings considering two major road segments at the center of UAB (around UAB Green) where many classroom buildings and office/lab buildings locate along University Boulevard and 14th St S, as shown in Figure 21.

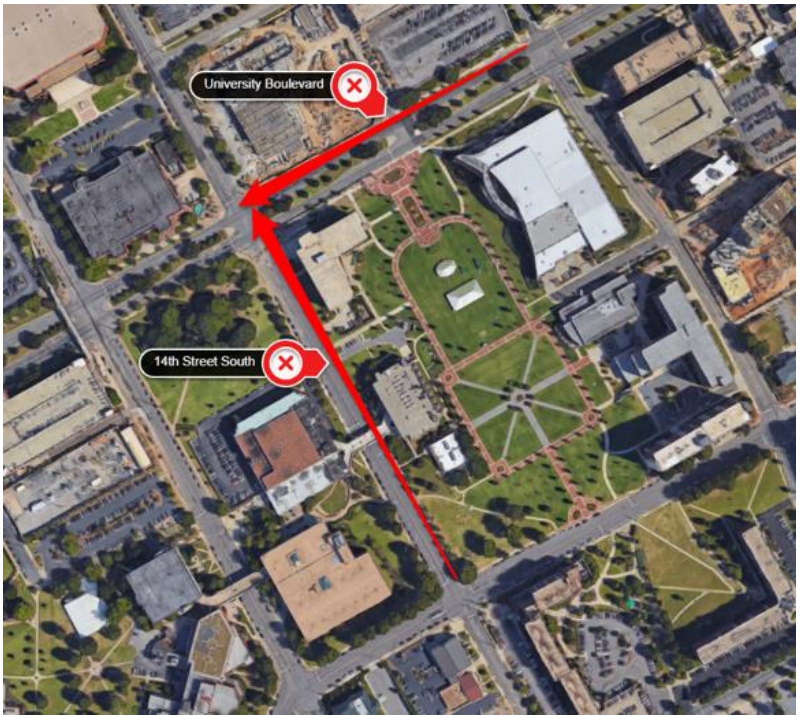


Figure 21. Sample Road Segments Used for Evaluation

4.5 Sample Birmingham study results

Given the planned pilot deployment of 500 e-scooters in the Birmingham region, we first considered how the option of having 500 e-scooters available on and around the UAB campus impacts traffic volumes (vehicles per hour). It should be noted that the road volume data output was scaled according to the sample used (10%) and aggregated to an hourly count per link. Also, we limited comparisons during times that e-scooters are allowed to operate on Birmingham streets (6AM-11PM) and classes and other activities are scheduled on campus (till 7PM).

4.5.1 Comparison of scenarios 1 and 2

Link Volumes: Results from the comparison of Scenario 1 (baseline) and Scenario 2 (500 e-scooters available) are displayed in Figure 22 and show that using e-scooters clearly reduces the link traffic volumes of both road segments in almost all hours of the day. It is further observed that the 14th Street South sees much more improvement in traffic volume than the University Boulevard. This is expected as the traffic on 14th Street South is mainly contributed by students taking classes or employees going to offices in the various buildings round UAB Green whereas University Boulevard is a major arterial providing access to the UAB campus from/to major highways where traffic is more diversified and includes background traffic.

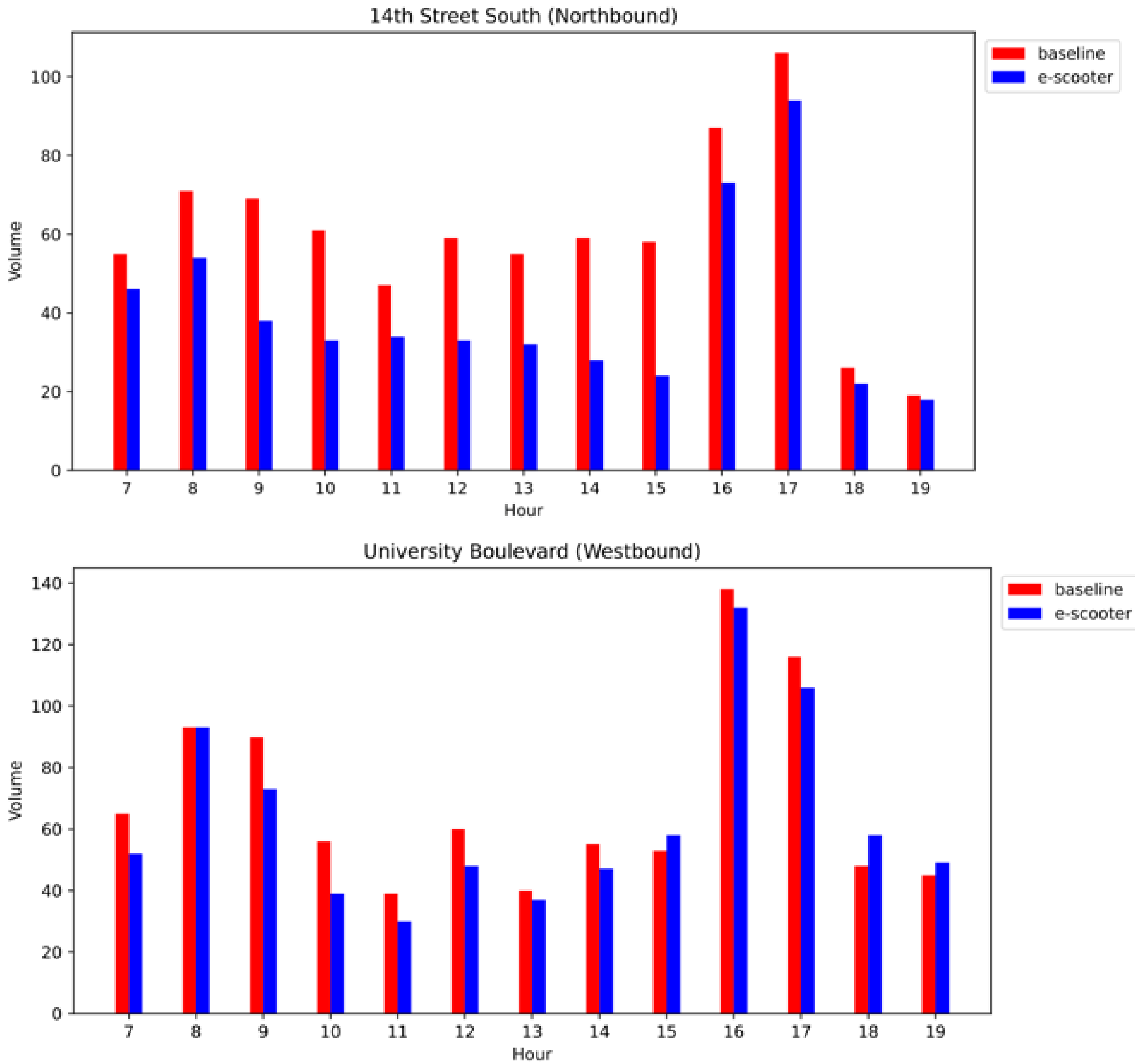
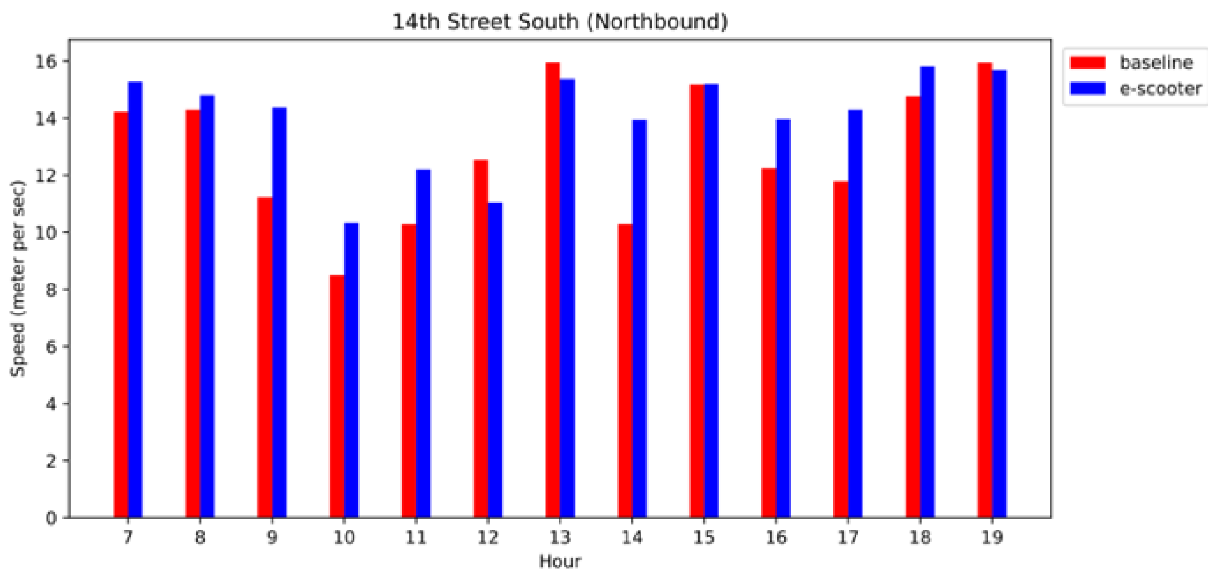


Figure 22. Sample Hourly Link Traffic Volumes (vph) – Cars Only

Average Speeds: Figure 23 shows the hourly average car speed (meter/sec) comparison on two road segments between Scenario 1 (baseline) and Scenario 2 (500 e-scooters available). We can see that allowing 500 e-scooters to operate on and around campus generally increases the average hourly speeds, especially during peak hours. These findings indicating that traffic flow is improved as a result of modal shifts towards the e-scooter mode. Thus, the availability of shared micromobility options shows promise toward easing traffic congestion and improving mobility of UAB students and employees alike.



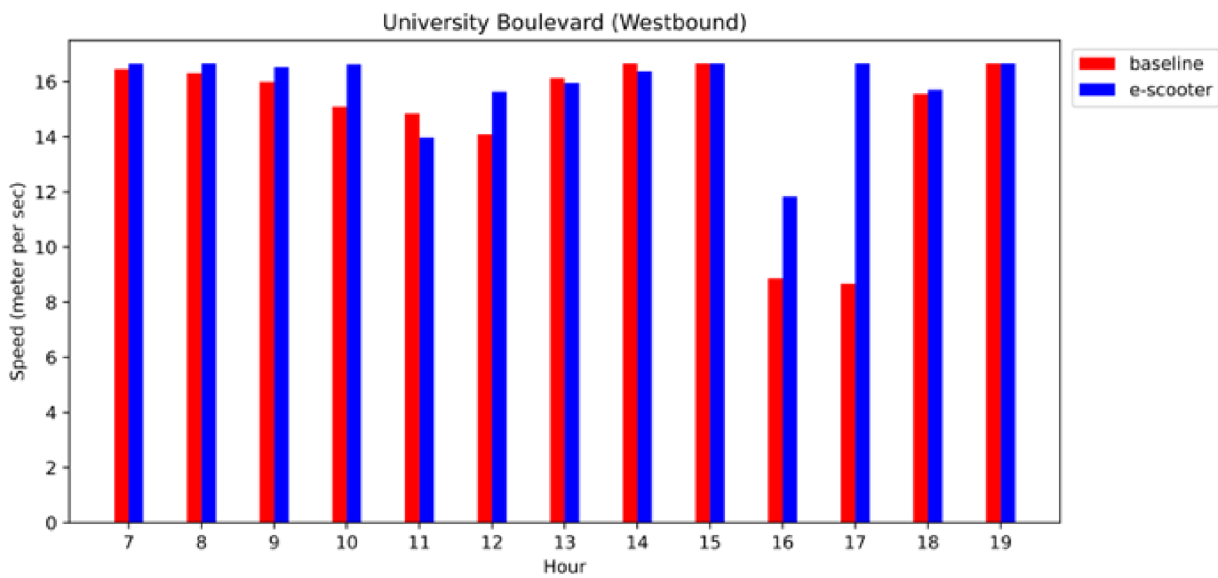


Figure 23. Sample Hourly Link Hourly Average Car Speed (meter/sec)

4.5.2 Sensitivity analysis

City Planning officials and e-scooter operators often find it hard to answer the question of how many shared e-scooters should be deployed. From the network operational performance perspective one can consider potential benefits of e-scooters on congestion mitigation for various levels of e-scooter deployment. For demonstration purposes we performed a sensitivity analysis by simulating the Birmingham network in MATSim for 4 levels of e-scooter deployment and comparing performance on the basis of volumes and speeds. The four levels correspond to our four scenarios (Scenario 1: 0 e-scooters available; Scenario 2:

500 e-scooters; Scenario 3: 750 e-scooters, and Scenario 4: 1,000 e-scooters available).

Link Volumes: Results from the comparison of Scenario 1 (baseline) and e-scooter Scenarios 2, 3 and 4 are displayed in Figure 24 and are fairly consistent. The findings show that e-scooters' availability reduces the link traffic volumes of both study road segments in almost all hours of the day. Close inspection of the graphs further confirms that the more e-scooters are available, the higher is the reduction of the link volumes (compared to the baseline). However, the differences resulting from modal shifts from automobile toward e-scooter use are small. This is due to the dominance of the automobile presence on the UAB campus and is properly captured by the Birmingham MATSim model which assigns more than 95% of the total trips to automobiles, even in the presence of micromobility options (see Figure 20).

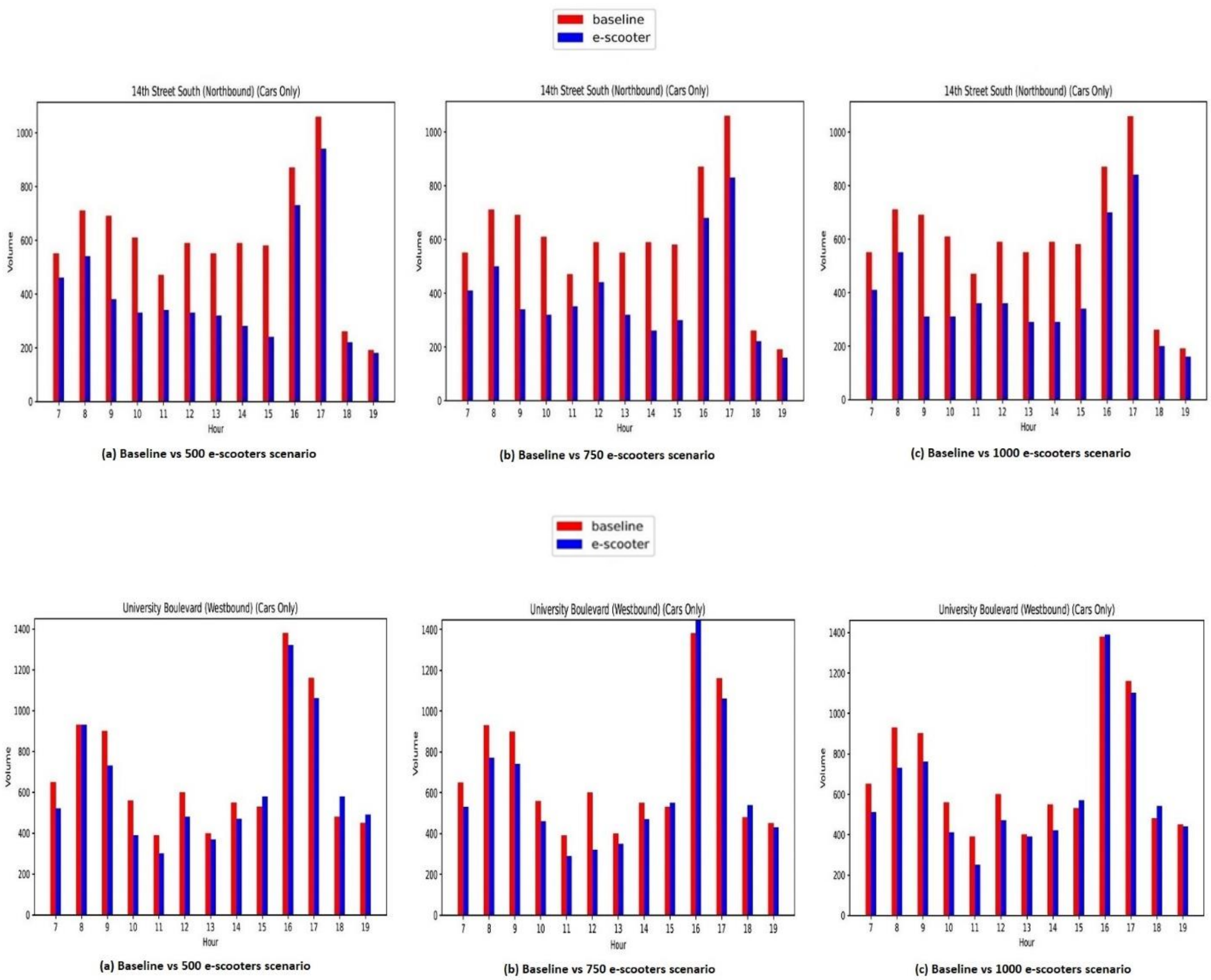


Figure 24. Comparison of Volumes on Two Study Links (Baseline versus E-scooters Scenarios)

Average Speeds: Figure 25 displays changes in average hourly car speeds (meter/sec) for baseline and the three e-scooter scenarios considered in this study and for two sample study links (14th Street South and University Boulevard).

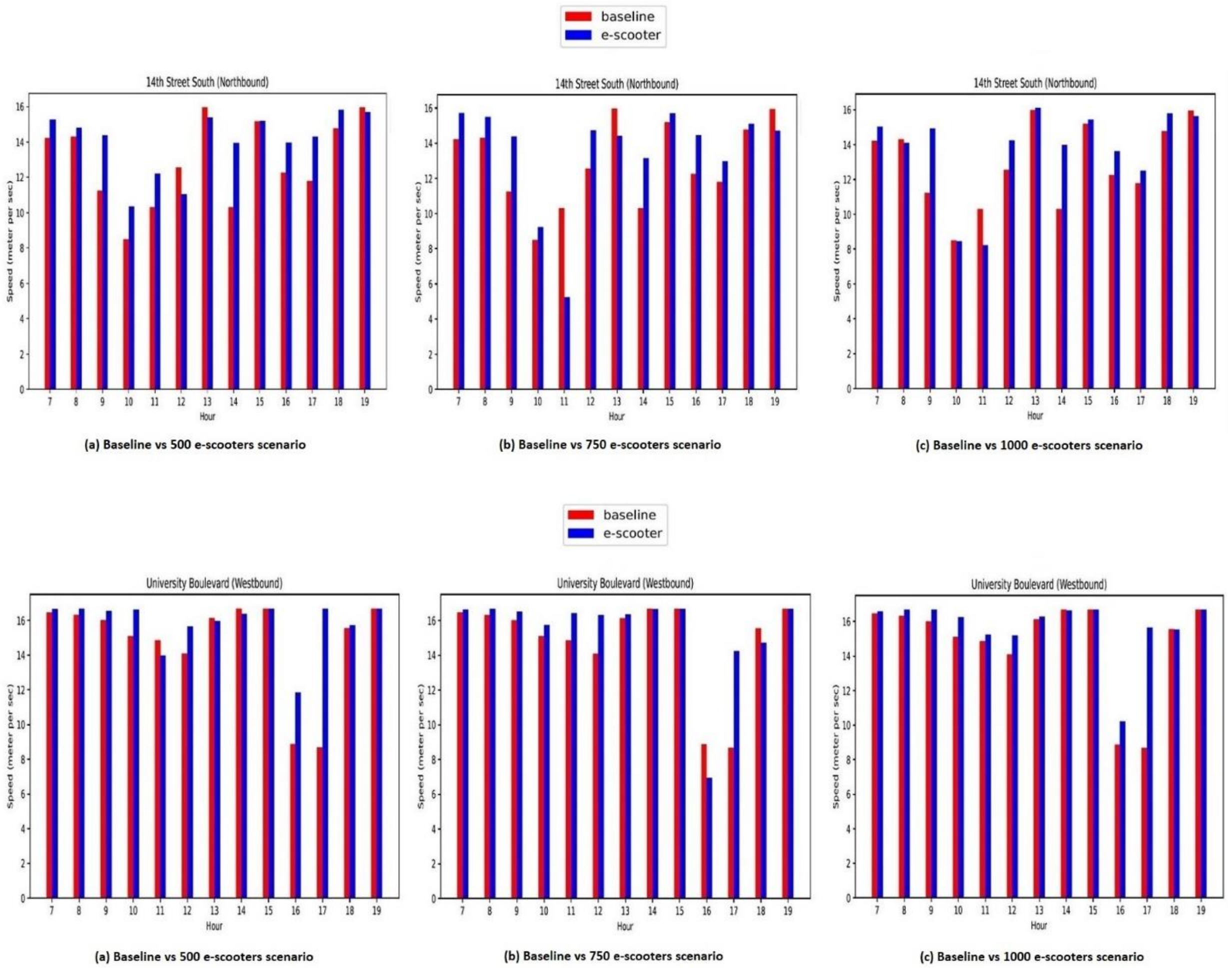


Figure 25. Comparison of Speeds on Two Study Links (Baseline versus E-scooters Scenarios)

Figure 25 confirms that, compared to baseline conditions, average automobile speeds improve when e-scooters operate on and around campus. Increases in average speeds are more pronounced during peak hours and when larger shared e-scooter fleets are available.

The sample findings from the two representative network links (i.e., 14th St S and University Blvd) presented in Figures 22 through Figure 25 are consistent across other simulated network links on the UAB campus where e-scooters are available for use. The results of the study are also consistent with expectations and suggest that the replacement of short trips in personal vehicles by e-scooter trips not only offers convenience to the users but also has a positive impact on traffic operations, helping to ease traffic congestion and related adverse impacts on university campuses and similar settings.

4.6 Conclusions and study contributions

In this study, we used MATSim to simulate traffic for a base case scenario at an urban university campus. Then, we developed an extended module that allowed the consideration of e-scooter use for shared micromobility simulation. This was done by successfully applying proper modifications to MATSim' carsharing module to enable the simulation with the mode of dockless e-scooters. We also changed the scoring function for cars and e-scooters, so both modes can work together in a way that realistic plans get better scores. In addition, we developed an effective pipeline to generate synthetic student plans by using different real data sources.

The updated MATSim framework was utilized to generate realistic day plans for travelers in a case study that considered 500, 750 and 1000 e-scooters on and around the UAB campus. The case study results confirmed that the simulated traffic volumes are lower and travel speeds are higher when e-scooters are available, compared to the base case scenario.

A major contribution of this study is the demonstration of the feasibility of integrating micromobility options in simulation modeling. Our pioneering work on the modification of the MATSim carsharing module now allows for the consideration of e-scooters as a travel mode option in the generation of travel plans when using MATSim. This is a notable achievement, given the recent increase in the popularity of micromobility modes and the current lack of simulation models that allow the incorporation of such options into their structures. Our study also offers examples of resources that can provide data inputs for making the traffic simulation model of a university campus a realistic one. Moreover, it shares valuables details on how to address issues arising from the use of a complex simulation platform like MATSim for conducting large scale simulations with (or without) e-scooters. Enabling the modeling of e-scooters in the MATSim framework gives researchers a new tool that they can use to study and quantify impacts on traffic congestion due to modal shifts from private automobile to e-scooters.

The study fulfilled its stated goal of improving activity-based traffic simulation models to account for new travel modes in different kinds of scenarios and improving the accuracy of traffic simulations for future scenarios. Given that case studies of this nature are very limited, the study is expected to be of great interest to researchers, e-scooter providers and decision makers interested in the deployment of e-scooters and other micromobility options and its potential impact on traffic operations.

Future research will focus on refinement and enhancement of the Birmingham MATSim model. Our plan is to run more iterations (at least 100) in MATSim to ensure the plans outputted are of good quality. In this demonstration study, we incorporated some assumptions. For example, we assumed that on average students leave the classroom 2 min after the class ends; drivers spent an average of 2 min at the parking location, and parking locations sought off are within a distance of 150 meters from class. However, these assumed parameter values may vary from location to location, thus calibration is recommended in future work. We also recommend conducting a more comprehensive study on the impact of using e-scooters on UAB campus traffic, by observing the traffic heatmap on various roads in UAB, and further varying the number of deployed e-scooters until the optimal number of available e-scooters can be determined.

Future plans include a) the incorporation of actual e-scooter ridership data gathered from the Veo and Gotcha e-scooters during their pilot deployment in Birmingham, b) comprehensive analysis of actual and simulated e-scooter usage patterns and related network performance impacts, and c) development of a set of recommendations to assist decision makers in effectively integrating e-scooters into the urban mobility ecosystem in their communities.

5.0 TASK 3-1: POLICY ANALYSIS

Many of the concerns of scooter-shares involve the implementation of services in a regulatory environment. Alternatively, many of the concerns surrounding scooter share, as well as their perceived benefits can be matured to full potential within the regulatory environment through thoughtful and educated discussion as well as considerations based on data analysis. At the heart of this discussion is an appropriate combination of these elements that correspond to engineers and planners, often in separate discussions, but arriving at similar targets.

Three crucial points that should be at the forefront of the discussion on scooter-share programs since they are fundamental in the regulatory approach that local governments and operators should have on these programs in the coming years. These include considerations of scooter shares as environmentally beneficial when considering the mode substitution, rebalancing, and life cycle costs of scooters.

One of these points is related to a general perception that assumes that scooter-shares are environmentally beneficial. Chrispoforou et al., (2021) and Hollingsworth, Copeland and Johnson (2019) reach similar conclusions when they claim that this is not entirely true. The suggestions come as the result of a perception that scooters – because they are electric – do not contribute to a significant amount of carbon emissions. Relative to the average car, this may be true, however, the collective impact of scooters must be evaluated according to various factors that make scooters less sustainable.

Rebalancing is the process by which scooters are picked-up by operators, recharged, and redistributed. This process is especially taxing on the environment when scooters are distributed in areas of low demand, at a distance, and far from each other, as well as far from the locations where they will be recharged. The issue surrounding rebalancing has less to do with the scooters themselves, and more to do with the vans used for this process. Both Christoforou et al., (2021) and Hollingsworth, Copeland and Johnson (2019) attempt to quantify this process and conclude that they attach a significant amount of carbon emissions on a per scooter trip basis.

Additional factors contributing to the environmental concerns have to with the scooter production process compared to the life cycle. During production, many elements, namely the lithium battery, are acquired in processes that produce significant levels of carbon emissions for the lifespan of the shared vehicles. It is estimated that shared scooters last around 6 months. The only positive aspect is that nearly 95% of the scooters are recyclable. This situation is marginal when compared to the highly used personal vehicle, which are responsible for much larger carbon emissions than a scooter even when life cycle costs are considered.

However, the main concern is the modal shift to scooter-share. As discussed in the literature review, various articles estimate that less than 10 percent of individual car drivers are shifting from car trips to scooter share. The more alarming statistic is that more than 60 percent of trips made on scooter-shares could have been feasibly done by walking. As such, when talking about the environmental impact of scooters, the issue needs to be framed as a question of modal shift. If the statistics were inversely true, it could be said that scooters are in fact a sustainable alternative since many vehicle owners would be replacing car trips with scooter trips. This is not true at the current moment in most places, and therefore, the suggestion that scooter-shares are unquestionably a sustainable alternative should be restated with caution and framed within the considerations mentioned here.

5.1 Modal Shift

Following from the point made on sustainability, the second crucial discussion point is the modal shift. Certainly, it is conceivable that under the appropriate circumstances, scooter-share services may be sustainable. The measure by which this is possible is by influencing the modal shift from personal car to scooter share. We can now begin referring to the growing literature that is giving operators and local governments an idea on how to do this. As is the case with modes of public transportation, there are places in which scooter-shares are more suitable given other considerations. Given the research so far, it is possible to recommend that scooters be used around certain land uses such as commercial and multifamily, especially if these coincide with the downtown area or a local university. Areas of interest that may open at odd times or hours, as well as large events that are generally less served by other transportation modes can also benefit from scooter-shares. While a broad city coverage can be desired, the sustainable benefits of scooters are curtailed when and where the demand for them is low. As such, local governments would be highly advised to assess areas of demand and reduce the coverage areas adequately. Moreover, a good measure to keep in mind when planning for scooter-shares is the sustainable trip range of 0.5-to-2 miles, which most researchers in this area agree provide the greatest level of benefits. Areas with high levels of walkability may contribute to a negative pedestrian to scooter-share modal shift, but areas with high levels of bikeability may help promote an accessible alternative especially when they exist within the 0.5-to-2-mile range. At the current moment, providing scooter shares in low-income, single-family areas may not be advisable since this generates low rates of usage. While an equitable approach is desired, it is possible that the current pricing structure is a deterrent in achieving an equitable scooter-share system. Equitable concerns should not be at all discarded however, and perhaps modifications to the existing relationship between government and operator may be key in addressing this concern.

A point that is hardly discussed but that would be beneficial to consider, especially if equity goals are of great concern, is the establishment of a membership scooter-share program. In their research, Younes et al. (2020) and Bai and Jiao (2020) found that scooter-shares and bikeshares have similar attributes that group them in the same micromobility family. When Younes et al. (2020) found that scooter shares and non-member bikeshares in Washington D.C. shared similar characteristics when compared to member bikeshares, the resulting question would be to consider whether a membership type scooter-share service could operate similar to membership type bikeshare systems. A simple evaluation of membership-type systems demonstrate that commuters do in fact use membership bikeshares given how they peak during morning and afternoon peak hours. While this may cause competition between both micromobility types, it is possible to further understand their respective strengths and adequately serve the needs they address. It would be in the best interest of planners to begin thinking of micromobility as a toolbox capable of addressing small scale mobility solutions instead of individual tools that operate independently or in competition with each other. In addition, a service in this direction would be capable of addressing equity concerns previously noted, if local governments would provide fare-reduction programs that would help make scooter trips cheaper in low-income communities. Creating a membership type scooter program could help facilitate a goal in multi-modality where riding scooters to bus stops or trains could be completed seamlessly. Establishing MaaS systems that include scooters could achieve the highest potential of a dockless scooter share system and could help to improve alternative mobility systems enough to produce a significant shift from personal cars in such a way that we could state that micromobility is a sustainable transportation alternative. In order for scooters to reach their maximum potential as a mobility service, an evaluation of the current policy recommendations should be considered.

5.2 NACTO's Guidelines for Regulating Shared Micromobility

In 2019, the National Association of City Transportation Officials developed a set of Guidelines for Regulating Shared Micromobility that included two broad sections – Best Practice Recommendations and Current State of the Practice. The Best Practices section includes discussions of regulating shared mobility, general terms and conditions, scope and operations insight, public engagement, mobility data and user privacy, and infrastructure. Instead of repeating their information on bets practices, we will present policy recommendations on regulatory structure, general terms and conditions, operations oversight, public engagement, data, and infrastructure.

5.3 Regulatory Structure

Multiple scooter-share programs already exist around the world and continue to expand to more cities. However, the implementation of regulations that adequately respond to the nature of these programs are not mature enough because of the program's relative novelty and a lack of consistency in regulatory structure. It is likely that in the coming years, a more structured regulatory framework should be proposed. Already NACTO (2019) has published its first set of regulatory considerations, especially focused on micromobility implementation in North American cities. Concern for public safety in all stages of operation is necessary for the operation of scooter-share systems. There are two ways to think of public safety regarding these programs: safety with regards to the scooter's interaction with the public and safety with regards to the scooter users; in this section we discuss the former. When discussing the regulatory framework, it is important that municipalities firmly discuss expectations of operators in favor of dockless, scooter-share services. The main concern is the scooter's interaction with the right-of-way, which local governments can address. One of the larger issues within this area is the clutter of scooters on sidewalks and within the right-of-way. In favor of public safety, and for the long-term operability of micromobility, it will be important to mitigate issues in whichever way possible. In later sections, we identify several infrastructural recommendations that may help to consider what kinds of investments would help to address safety on the right-of-way.

It is highly recommended that any micromobility program be first observed as a pilot program. However, when the pilot program is developed, it should be for at least a year to account for seasonal variabilities. Moreover, the purpose of the pilot program must be to evaluate mobility trends and hotspots that could eventually lead to adopting a solid coverage zone. Pilots should never try to adjust usage patterns or adopt policies as the program moves along. Some of the first cities that allowed scooters have had issues ensuring that scooter users stay on roadways and not sidewalks after they were initially used on sidewalks. There was great confusion for users and operators whenever policies were implemented on where to use them, and then making other adjustments as was seen fit. This has led to the improper use of scooters from the beginning of their implementation in many of the first cities that adopted them leading to safety issues and reducing the life cycle of the scooters themselves. Rather, all policies and modes of use need to be in place from the beginning to mitigate confusion and challenging circumstances with operators and users.

5.4 General Terms and Conditions

Many cities had multiple operators providing scooter-share services, which led to multiple issues and possibly excess competition. Now many cities are only allowing three operators maximum to operate in their cities. When this happens, operators need to be committed to addressing the various sustainable considerations recommended and stipulated by the local governments that are allowing the permits. Fewer operators ensure ease of data sharing and makes the interaction between the local governments easier since the governments only need to correspond with three operators, at most.

When discussing terms and conditions, it is highly recommended that cities request more detailed information on the supply-side of micromobility. Operators are often concerned with how their operations may be affected if the supply-side is shared with competitors, however, in a public-private partnership, it may be easier to manage the supply side in a discrete manner. Being able to access information on supply-side data will allow for more appropriate analysis to be made. One of the glaring shortcomings of current research is the lack of discussion on this matter and it is one of the challenges of modelling since there is no way of knowing if low usage is due to demographics and the built environment or to the lack of availability of scooters on the supply-side.

It is becoming more common to see the establishment of consolidated mobility departments that join public transit, bike/pedestrian and micromobility programs in an effort to tackle mobility objectives collaboratively. An office of mobility would help to be a liaison between local governments and operators. This would help to manage compliance of operators and to clearly define performance measures that are similar to those in other programs, while at the same time considering the local needs. This same office could help provide data analytics internally that would help determine how to collaborate with other mobility modes in an effort to achieve greater multimodality. Additionally, when pilot periods are over, a mobility department can focus on the data collected over the year and strategize on a concerted effort to address hidden mobility patterns that could be addressed by other mobility modes such as mass transit, buses or trams.

5.5 Operations Oversight

The largest discussion and the greatest change in operations needs to occur in rebalancing methods. Local governments need to discourage rebalancing and not simply on the basis of usage but by establishing alternative methods to remediate rebalancing. In a later section, we go into more detail about corrals and docking that may help to reduce the amount of rebalancing needed. Adjusting the coverage area after the pilot period should be directly linked to efforts in reducing rebalancing and aiming for modal shift.

While NACTO (2019) considers that equity in distribution should be of high priority, especially in low-income communities, it is highly advised to proceed with caution in this regard. Equity should be a concern but only when pricing schemes use incentives to not only use scooters, but to adopt them as a feasible mobility mode. Additionally, rebalancing techniques need to be well strategized to reduce further environmental impacts, considering that low demand and long driving distances may make rebalancing in these areas less sustainable overall. This is perhaps more achievable when some degree of multi-modality is achieved by the combined efforts of a mobility department, the operators, and the use of a MaaS platform.

There are further discussions on the impact of the wheel size on the scooter itself when compared to bicycles. This is because scooters have been regulated as a personal consumer product rather than a shared-use fleet vehicle (NACTO, 2019). The result is that scooter users experience greater deformities on the pavement more severely than on bicycles for example. NACTO (2019) estimates that on average bicycle wheel sizes measure more than 26 inches, where scooter wheel sizes range between 8 and 10 inches in diameter. Other standards that are not addressed include center of gravity, platform size, acceleration and braking interface, and lights (NACTO, 2019). A lack of precise standards, coupled with geometric design features of the paths used by scooters have led to numerous safety issues. Reports that safety concerns, injuries and fatalities exceed those of bicycles are alarming, and justify a need to pay attention to the design of the vehicles and to design for the vehicles, a task that may interest some factions of transportation engineering. In fact, NACTO (2019) suggests that bicycle infrastructure planned by local governments be expedited since the demand to use these lanes seem to continue increasing, creating a pressing need for engineers to respond.

5.6 Public Engagement

To engage the public more broadly, local governments should provide incentives to participate in a user survey as a part of the terms and conditions for usage of DSS. Using the in-app experience creates the advantage of being able to reach users very broadly. Public engagement can also manifest itself in providing educational materials that promote the use of scooters as well as adequate understanding as to the benefits of scooter usage. Additionally, other forms of public engagement could be to incentivize independent contractors to recharge scooters, especially in low-income communities. Moreover, these communities, as NACTO (2019) highly suggests, should be involved in the planning process, as opposed to planning micromobility for these communities.

5.7 Data

One of the initial challenges (and a continued challenge) of scooter-share analytics was the lack of data accessibility. Louisville is one of few cities that has a consistent and streamlined open data service that has allowed for a clean analysis of the scooter share services since they were first adopted in August 2018. Two factors led to this, primarily that operators did not wish to share data with other competitors and secondly that the data was kept anonymous for the scooter user's sake. It is essential that local governments require data from operators for the purposes of analysis. This data should be as detailed as Louisville's Open Data source (LOJIC, 2019). Many operators are only required to have an API; however, this kind of data structure requires detailed computer science techniques that are often not used by local governments, neither are they easily accessible to the general public. It is essential nonetheless to maintain user anonymity. In many cases, this is done best by aggregating trip data to the third latitude and longitude points and by assigning an anonymized data name to each trip.

5.8 Infrastructure

There are two main areas where infrastructural investment could help micromobility develop into full potential: on the curbside and on corridors.

Curbside management has emerged as an important area of concern in planning in recent years. In recent years, scooter-share programs have added to the challenges of managing the curbside as a part of their operations. One of the main qualms about dockless micromobility is the propensity to clutter on the curb, especially in high demand areas. At the same time, what has made dockless scooters so popular have been their dockless attribute, therefore, attending to clutter would mean to address how and where they are placed at the end of trips. Additionally, this would require higher investments on behalf of operators for a service that they perhaps would not feel inclined to provide. A suggested approach would be to wait until the end of the pilot program to identify areas of high demand and areas of constant demand, in addition to hotspot areas. Considering that scooters are most beneficial when they are ridden in a .5 mile to 2-mile area, the first .5-mile areas around hotspots and high demand areas should allow for corrals and docking if possible. Placing corrals at adequate distances may also deter too many pedestrian trips from becoming scooter trips while at the same being accessible enough to make a .5 mile to 2-mile trip. Moreover, considerations on rechargeable docks would help to reduce rebalancing if scooters in these zones were required to be docked at the end of use. Areas of constant demand, where usage is significantly lower than in areas of high demand could allow for dockless trips, allowing for the service to retain its hallmark style.

Corridor management may be the area that most concerns transportation engineers. While some recommendations are made by AASHTO, there is very little consistent push to establish design requirements for micromobility lanes when compared to the uniform requirements stipulated for roadways. One of the challenges of corridors or bike lanes where scooters (and bikes for that matter) operate on is that they travel right next to heavy cars without a buffer between them. Additionally, a lack of buffering does not deter vehicles from disregarding these lanes and using them to pull aside from the road. Therefore, it is important for AASHTO to provide more consistent updates on roadway design that is inclusive of micromobility lanes. This could be determined as a function of demand, and land use. Suggested considerations could include the requirement for 1-foot buffers with plastic bollards on roads in downtown areas or university areas. Perhaps the .5 mile to 2-mile area could be used as a reference for the implementation of such design characteristics. If requirements were made in this direction, safer scooter-usage could be the results. Similarly curb cuts with cross-slopes adequate for the wheel-size could help prolong the life cycle of scooters. To avoid right hook hits from cars for the safety of micromobility users, a two second delayed red-light signal should be considered in highly trafficked mobility lanes.

6.0 TASK 3-2: DECISION-SUPPORT TOOL: SERMOS

In this project, we developed a new decision-support tool, called SERMOS, which is partially based on the products of Task 1. SERMOS collects and analyzes the e-scooter-related data (e.g., GBFS data, e-scooter equity zone data, e-scooter parking zone data, among others), which is expected to benefit various stakeholders. For example, cities can use the tool to monitor and regulate micromobility operations. MPOs can use the tool to assist long-term planning of multimodal transportation systems.

For the SERMOS system, there are three modules, including reporting module, mapping module, and analytics module. The overall structure of the system is shown in Figure 26.

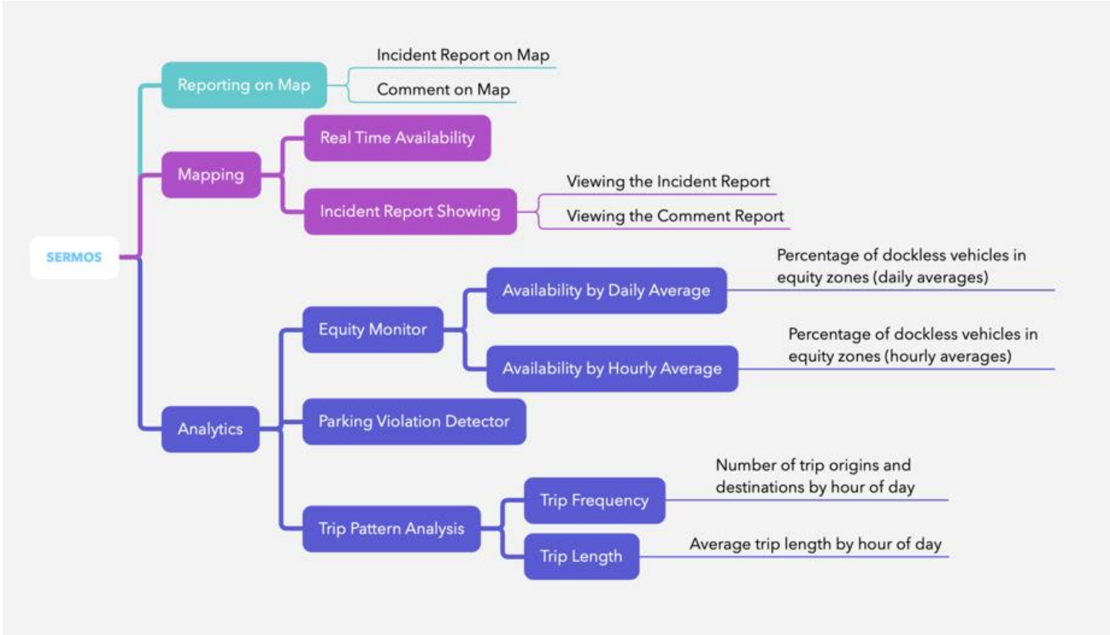


Figure 26. Overall structure of SERMOS system

The reporting module provides the functionality that users can click any location on the map to report an e-scooter incident or just leave a comment. Figure 27 shows the user interface of the reporting module. After the report is submitted, the reported location and the reporting time will also be recorded. All these information will be then stored and maintained in the database to serve as the input for the mapping module.

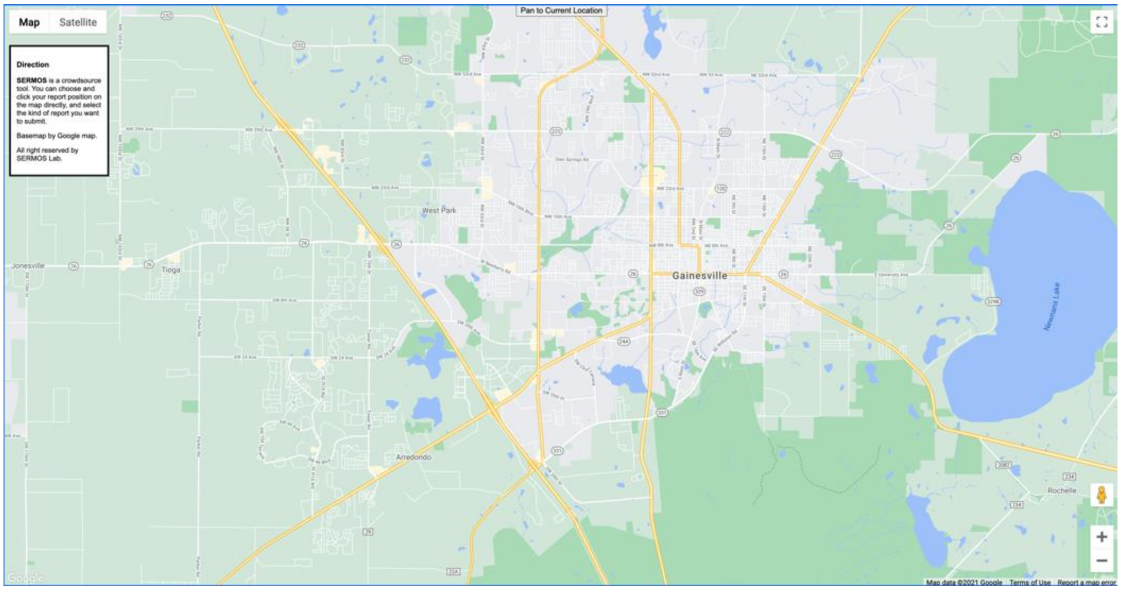


Figure 27. Overall structure of SERMOS system

All the reports that have been gathered in the reporting module are illustrated at the corresponding locations in the mapping module. The user interface is shown in Figure 28. Users like a city's mobility manager can click the marker to view the report details.

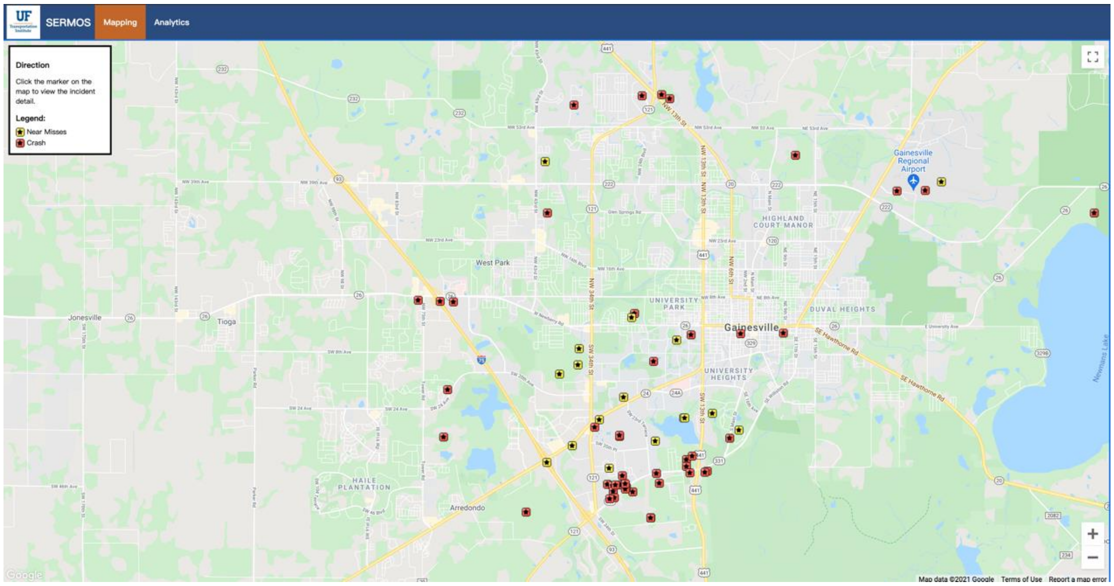


Figure 28. User interface of mapping module

The analytics module includes equity monitor, parking violation detector, and trip pattern analysis. In Figure 29, we present an example of trip pattern analysis. The analytics module will be fully developed if additional funding is available.

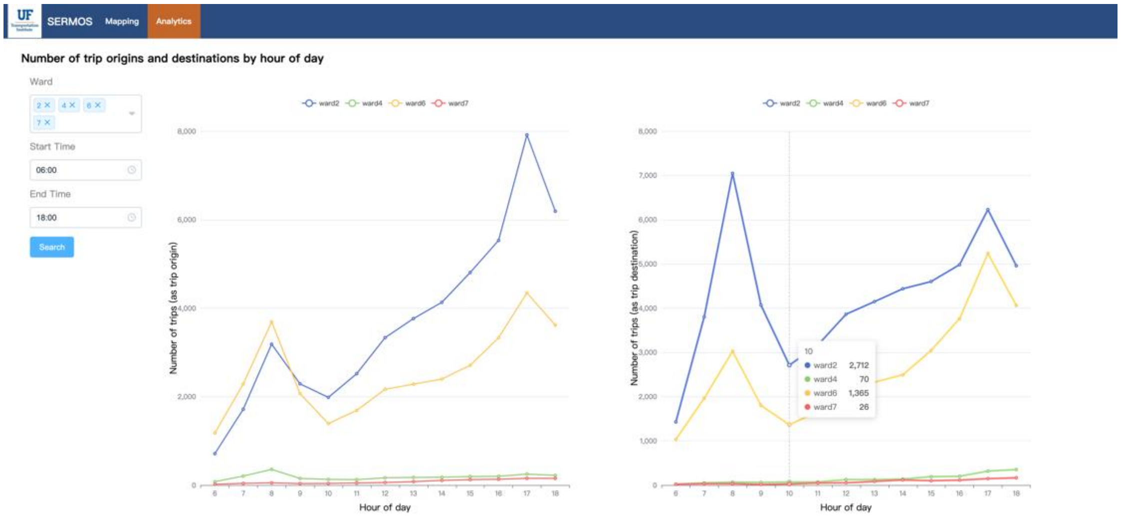


Figure 29. Number of trip origins and destinations by hour of day

7.0 CONCLUSION

This project conducted a comprehensive study to analyze, quantify, and understand the impacts of micromobility on urban mobility and recommend corresponding intervention strategies for stakeholders. E-scooters have been welcomed as a transportation option that can expand transportation options by replacing short trips by automobile and provide first-mile-last mile access to public transit. Some transportation experts suggest that e-scooters can be a part of the solution to reducing congestion and to mitigate the environmental impacts associated with automobile usage. This project included three tasks: 1) Task 1: e-scooter big data analytics and travel demand modeling; 2) Task 2: traffic simulation of e-scooters - a pilot study for an urban university campus; 3) Task 3: shared micromobility policy analysis and decision-support tool – that were designed to understand policy issues and model the demand for e-scooters. E-scooters are still in the early stages of adoption and will continue to evolve and adapt to the urban environments in which they are placed.

In Task 1, we firstly inferred origins and destinations of e-scooter trips in Washington, D.C. based on GBFS data. Then we modeled the trip origin demand of e-scooter services in Washington, D.C. The independent variables included socioeconomic and demographic variables, built environment variables, and transit supply variables. The OLS, Lasso, DT, RF, and Boosting models were used to predict the trip origin demand in census block group level. Comparisons of in-sample and out-of-sample performance of these five models were performed using MAE and RMSE. The results of the best performed model, the RF model, were further interpreted using FI and PDPs. The results showed that the most important variable was WalkScore and the most important category of variable was built environment variables. From PDPs, we observed nonlinear relationships between the dependent variable and independent variables such as WalkScore, parking density, hotel density, proportion of young population, proportion of white population. For example, when WalkScore increases from 85 to 95, there is a sharp increase in trip origin demand. This task improved the predictive accuracy as well as provide rich insights of how black-box machine-learning models make such predictions to facilitate the policy analysis.

In Task 2, we used MATSim to simulate traffic for a base case scenario at an urban university campus (i.e., UAB campus). Then, we developed an extended module that allowed the consideration of e-scooter use for shared micromobility simulation. This was done by successfully applying proper modifications to MATSim' carsharing module to enable the simulation with the mode of dockless e-scooters. We also changed the scoring function for cars and e-scooters, so both modes can work together in a way that realistic plans get better scores. In addition, we developed an effective pipeline to generate synthetic student plans by using different real data sources. The updated MATSim framework was utilized to generate realistic day plans for travelers in a case study that considered 500, 750 and 1000 e-scooters on and around the UAB campus. The case study results confirmed that the simulated traffic volumes are lower and travel speeds are higher when e-scooters are available, compared to the base case scenario. This task improved the existing activity-based traffic simulation models by considering new travel modes as well as generating more accurate simulation outcomes.

In Task 3, we discussed the policy related to shared micromobility operation and developed a decision-support tool named SERMOS. Firstly, we discussed considerations of scooter shares as environmentally beneficial when considering the mode substitution, rebalancing, and life cycle costs of scooters. Then we presented policy recommendations on regulatory structure, general terms and conditions, operations oversight, public engagement, data, and infrastructure. We also developed a decision-support system, named SERMOS, that can collect and analyze the e-scooter-related data to facilitate e-scooter planning (e.g., where to plan bike lanes and how to best deploy e-scooters) and real-time management (e.g., where parking violations occur and when and where equity requirements of e-scooter availability are not met). SERMOS provides information for decision makers to overcome two of the major concerns of policy makers about the implementation of e-scooters – safety and equitable access to vehicles in low-income neighborhoods. This task can help local stakeholders to facilitate better-targeted decision making and policy interventions when deciding where to expand, change or reduce service throughout an urban area.

In conclusion, this project integrated big data analytics, demand modeling, traffic simulation, and policy analysis to provide a comprehensive assessment of the impacts of micromobility on congestion mitigation. It developed machine learning methodologies to model and interpret travel demand of micromobility, improved activity-based traffic simulation models, suggested micromobility related policy and developed an interactive decision-support tool for local stakeholders. This project provided rich insights of key factors associated with micromobility demand, examined the potential impact of deployment of e-scooters and other micromobility options on traffic operations, and generated new insights for key stakeholders to facilitate planning micromobility policies and practices.

8.0 RECOMMENDATIONS

Several issues require future research. For Task 1, more features may be needed to develop a more comprehensive demand forecasting model and to generate richer insights. For example, since the usage of micromobility can be significantly influenced by the weather conditions (Noland, 2021), taking the weather information as additional model input may enhance the predicting accuracy. In addition, the results and insights regarding the trip origin demand found in Washington D.C may not be directly transferable to other cities with different characteristics. Transferability of the models requires further research.

For Task 2, future research will focus on refinement and enhancement of the Birmingham MATSim model. Our plan is to run more iterations (at least 100) in MATSim to ensure the plans outputted are of good quality. We also recommend conducting a more comprehensive study on the impact of using e-scooters on UAB campus traffic, by observing the traffic heatmap on various roads in UAB, and further increasing the number of deployed e-scooters (>1,000) until the optimal number of available e-scooters can be determined. Future plans include a) the incorporation of actual e-scooter ridership data gathered from the Veo and Gotcha e-scooters during their pilot deployment in Birmingham, b) comprehensive analysis of actual and simulated e-scooter usage patterns and related network performance impacts, and c) development of a set of recommendations to assist decision makers in effectively integrating e-scooters into the urban mobility ecosystem in their communities.

For Task 3, we plan to include more analytics modules in the decision-support tool in future work. For example, we will add a hot spot identification module to identify areas with high micromobility demand.

9.0 REFERENCE LIST

1. Abdou, M., Hamill, L., & Gilbert, N. (2012). Designing and building an agent-based model. In Agent-based models of geographical systems (pp. 141-165). Springer, Dordrecht.
2. Arnell, B., Noursalehi, P., Huntley, E., & Zhao, J. (2020). Shared electric scooters and transportation equity: A cross-city analysis. In Transportation Research Board 99th Annual Meeting (pp. 12-16).
3. 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, 264-272.
4. Barclays. (2019). Micromobility: Fast, cheap and good solution for 'smart cities.' https://www.investmentbank.barclays.com/our-insights/micromobility-fast-cheap-and-good-solution-for-smart-cities.html.
5. Baster, B., Duda, J., Maciol, A., & Rębiasz, B. (2013). Rule-based approach to human-like decision simulating in agent-based modeling and simulation. In 2013 17th International Conference on System Theory, Control and Computing (ICSTCC) (pp. 739-743). IEEE.
6. Bieliński, T., & Ważna, A. (2020). Electric Scooter Sharing and Bike Sharing User Behaviour and Characteristics. Sustainability, 12(22), 9640.
7. Birmingham Business Allience -BBA (2021). Major Employer List. Available at https://www.birminghambusinessalliance.com/major-employers.
8. Birmingham Department of Transportation (2021). Shared Micromobility has Launched! https://www.birminghamal.gov/transportation/shared-micromobility/ (accessed September 6, 2021).
9. Breiman, L., Friedman, J., Stone, C. J., & Olshen, R. A. (1984). Classification and regression trees. CRC press.
10. Breiman, L., (2001). Random forests. Machine learning, 45(1), pp.5-32.
11. Brown, A., Klein, N. J., Thigpen, C., & Williams, N. (2020). Impeding access: The frequency and characteristics of improper scooter, bike, and car parking. Transportation Research Interdisciplinary Perspectives, 4, 100099.
12. Button, K., Frye, H., & Reaves, D. (2020). Economic regulation and E-scooter networks in the USA. Research in transportation economics, 100973. https://doi.org/10.1016/j.retrec.2020.100973
13. Cao, Z., Zhang, X., Chua, K., Yu, H., & Zhao, J. (2021). E-scooter sharing to serve short-distance transit trips: A Singapore case. Transportation Research Part A: Policy and Practice, 147, 177-196.
14. 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.
15. Chang, A. Y. J., Miranda-Moreno, L., Clewlow, R., & Sun, L. (2019). Trend or Fad? Deciphering the Enablers of Micromobility in the US. SAE International.

16. Christoforou, Z., Gioldasis, C., de Bortoli, A., & Seidowsky, R. (2021). Who is using e-scooters and how? Evidence from Paris. Transportation Research Part D: Transport and Environment, 92, 102708.
17. Ciociola, A., Cocca, M., Giordano, D., Vassio, L., & Mellia, M. (2020). E-Scooter Sharing: Leveraging Open Data for System Design. 2020 IEEE/ACM 24th International Symposium on Distributed Simulation and Real Time Applications (DS-RT), 1–8.
18. City of Birmingham (2019a). City of Birmingham Shared Micromobility Ordinance No. 20-21. https://www.birminghamal.gov/wp-content/uploads/2020/02/COB-Shared-Micromobility_Ord.-No.-20-24.pdf (accessed October 3, 2021)
19. City of Birmingham (2019b). Shared bikes and scooters now available in Birmingham. https://www.birminghamal.gov/2021/04/18/shared-bikes-and-scooters-now-available-in-birmingham/ (accessed September 6, 2021).
20. City of Santa Monica. (2019). Shared micromobility pilot program summary report. https://www.smgov.net/uploadedFiles/Departments/PCD/Transportation/SantaMonicaSharedMobilityEval uation_Final_110419.pdf
21. Clewlow, R., Foti, F., & Shepard-Ohta, T. (2018). Measuring equitable access to new mobility: A case study of shared bikes and electric scooters. In Transportation Research Board 97th Annual Meeting.
22. de Bortoli, A., & Christoforou, Z. (2020). Consequential LCA for territorial and multimodal transportation policies: method and application to the free-floating e-scooter disruption in Paris. Journal of Cleaner Production, 273, 122898.
23. de Oliveira, Í. R. (2017). Analyzing the performance of distributed conflict resolution among autonomous vehicles. Transportation Research Part B: Methodological, 96, 92-112.
24. Djavadian, S., & Chow, J. Y. (2017). An agent-based day-to-day adjustment process for modeling 'Mobility as a Service' with a two-sided flexible transport market. Transportation research part B: methodological, 104, 36-57.
25. DuPuis, N., Griess, J., & Klein, C. (2019). Micromobility in cities: A history and policy overview.
26. Eccarius, T., & Lu, C. C. (2018). Exploring consumer reasoning in usage intention for electric scooter sharing. Transp. Plan. J. 運輸計劃季刊, 47(4), 271-295
27. Eccarius, T., & Lu, C. C. (2020). Adoption intentions for micro-mobility–Insights from electric scooter sharing in Taiwan. Transportation research part D: transport and environment, 84, 102327.
28. Edgemon, Erin (2018). Bird Electric Scooters Have Landed in Birmingham. Al, 28 Aug. 2018, https://www.al.com/news/birmingham/2018/08/electric_scooters_have_landed.html.
29. Ensor, M., Maxwell, O., & Bruce, O. (2021). Mode shift to micromobility. Waka Kotahi NZ Transport Agency Research Report 674.
30. Fitt, H., & Curl, A. (2019). E-scooter use in New Zealand: Insights around some frequently asked questions. https://ir.canterbury.ac.nz/handle/10092/16336

31. Friedman, J. H. (2001). Greedy function approximation: a gradient boosting machine. Annals of statistics, 1189-1232.
32. Gould, G., & Karner, A. (2009). Modeling bicycle facility operation: Cellular automaton approach. Transportation research record, 2140(1), 157-164.
33. Guo, G., Khalil, J.M., Yan, D., and Sisiopiku V. (2019a). Realistic Transport Simulation: Tackling the Small Data Challenge with Open Data, Proceedings of International Workshop on Big Data Tools, Methods, and Use Cases for Innovative Scientific Discovery, 2019 IEEE International Conference on Big Data, pp. 4512-4519, Los Angeles, CA.
34. Guo, G., Khalil, J.M., Yan, D., and Sisiopiku V. (2019b). Realistic Transport Simulation with Open Data, BigData 2019: 6066-6068, Los Angeles, CA.
35. Hastie, T., Tibshirani, R. and Friedman, J., (2009). The elements of statistical learning: data mining, inference, and prediction. Springer Science & Business Media.
36. Haworth, N., Schramm, A., & Twisk, D. (2021). Comparing the risky behaviours of shared and private e-scooter and bicycle riders in downtown Brisbane, Australia. Accident Analysis & Prevention, 152, 105981.
37. Hollingsworth, J., Copeland, B., & Johnson, J. X. (2019). Are e-scooters polluters? The environmental impacts of shared dockless electric scooters. Environmental Research Letters, 14(8), 084031.
38. Horni, A., Nagel, K. and Axhausen, K.W. (eds.) 2016. The Multi-Agent Transport Simulation MATSim. London: Ubiquity Press. DOI: http://dx.doi.org/10.5334/baw. License: CC-BY 4.0.
39. Hosseinzadeh, A., Algomaiah, M., Kluger, R., & Li, Z. (2021). E-scooters and sustainability: Investigating the relationship between the density of E-scooter trips and characteristics of sustainable urban development. Sustainable cities and society, 66, 102624.
40. Hosseinzadeh, A., Algomaiah, M., Kluger, R., & Li, Z. (2021b). Spatial analysis of shared e-scooter trips. Journal of Transport Geography, 92, 103016. https://doi.org/10.1016/j.jtrangeo.2021.103016
41. Hussein, M., & Sayed, T. (2017). A bi-directional agent-based pedestrian microscopic Jia model. Transportmetrica A: Transport Science, 13(4), 326-355.
42. Jia, B., Li, X. G., Jiang, R., & Gao, Z. Y. (2007). Multi-value cellular automata model for mixed bicycle flow. The European Physical Journal B, 56(3), 247-252.
43. Jiang, R., Jia, B., & Wu, Q. S. (2004). Stochastic multi-value cellular automata models for bicycle flow. Journal of Physics A: Mathematical and General, 37(6), 2063.
44. Jiao, J., & Bai, S. (2020). Understanding the shared E-scooter travels in Austin, TX. ISPRS International Journal of Geo-Information, 9(2), 135.
45. Kantar, T. N. S. (2019). Public response to shared e-scooters in Auckland and Christchurch. Report prepared for NZTA, Auckland City Council and Christchurch City Council.
46. Kaziyeva, D., Loidl, M., & Wallentin, G. (2021). Simulating Spatio-Temporal Patterns of Bicycle Flows with an Agent-Based Model. ISPRS International Journal of Geo-Information, 10(2), 88.
47. Ke, J., Qin, X., Yang, H., Zheng, Z., Zhu, Z., & Ye, J. (2021). Predicting origin-destination ride-sourcing demand with a spatio-temporal encoder-decoder residual multi-graph

convolutional network. Transportation Research Part C: Emerging Technologies, 122, 102858.
48. Laa, B., & Leth, U. (2020). Survey of E-scooter users in Vienna: Who they are and how they ride. Journal of transport geography, 89, 102874.
49. League of American Bicyclists (LAB). (2019). Bicycling & Walking in the United States: 2018 Benchmarking Report. Washington D.C.: League of American Bicyclists. Retrieved from https://data.bikeleague.org/
50. Lee, H., Baek, K., Chung, J. H., & Kim, J. (2021). Factors affecting heterogeneity in willingness to use e-scooter sharing services. Transportation Research Part D: Transport and Environment, 92, 102751.
51. Lee, M., Chow, J. Y. J., Yoon, G., & He, B. Y. (2019). Forecasting e-scooter substitution with direct and access trips by mode and distance in New York City. ArXiv Preprint ArXiv:1908.08127.
52. Lee, M., Chow, J. Y. J., Yoon, G., & He, B. Y. (2021). Forecasting e-scooter substitution of direct and access trips by mode and distance. Transportation Research Part D: Transport and Environment, 96, 102892. https://doi.org/https://doi.org/10.1016/j.trd.2021.102892
53. Liang, X., Baohua, M. A. O., & Qi, X. U. (2012). Psychological-physical force model for bicycle dynamics. Journal of Transportation Systems Engineering and Information Technology, 12(2), 91-97.
54. Liu, J., & Zhou, X. (2016). Capacitated transit service network design with boundedly rational agents. Transportation Research Part B: Methodological, 93, 225-250.
55. Liu, M., Seeder, S., & Li, H. (2019). Analysis of E-scooter trips and their temporal usage patterns. Institute of Transportation Engineers. ITE Journal, 89(6), 44-49.
56. Loidl, M., Wallentin, G., Cyganski, R., Graser, A., Scholz, J., & Haslauer, E. (2016). GIS and transport modeling—Strengthening the spatial perspective. ISPRS International Journal of Geo-Information, 5(6), 84.
57. McKenzie, G. (2019). Spatiotemporal comparative analysis of scooter-share and bike-share usage patterns in Washington, DC. Journal of transport geography, *78*, 19-28.
58. McKenzie, G. (2020). Urban mobility in the sharing economy: A spatiotemporal comparison of shared mobility services. Computers, Environment and Urban Systems, 79, 101418.
59. Merlin, L. A., Yan, X., Xu, Y., & Zhao, X. (2021). A segment-level model of shared, electric scooter origins and destinations. Transportation Research Part D: Transport and Environment, 92, 102709.
60. Mitra, R., & Hess, P. M. (2021). Who are the potential users of shared e-scooters? An examination of socio-demographic, attitudinal and environmental factors. Travel Behaviour and Society, 23, 100-107.
61. MobilityData, What's new in GBFS v2.0. https://mobilitydata.medium.com/whats-new-in-gbfs-v2-0-63eb46e6bdc4, Medium, 2020
62. Mohammed, H., Sayed, T., & Bigazzi, A. (2021). Microscopic modeling of cyclists on off-street paths: a stochastic imitation learning approach. Transportmetrica A: Transport Science, 1-22.

63. Molnar, C., (2019). Interpretable machine learning. https://christophm.github.io/interpretable-ml-book/
64. NACTO, Shared Micromobility in the U.S., 2018.New York, NY, 2019.
65. National Association of City Transportation Officials –NACTO (2018). Shared Micromobility in the U.S.: 2019. https://nacto.org/shared-micromobility-2019 (accessed on September 6, 2021).
66. National Association of City Transportation Officials (NACTO). (2019). Guidelines for regulating shared micromobility. New York. Retrieved from: https://nacto.org/wp-content/uploads/2019/09/NACTO_Shared_Micromobility_Guidelines_Web.pdf.
67. National Association of City Transportation Officials. (2020). https://nacto.org/program/bike-share-initiative/
68. Nickkar, A., Banerjee, S., Chavis, C., Bhuyan, I. A., & Barnes, P. (2019). A spatial-temporal gender and land use analysis of bikeshare ridership: The case study of Baltimore City. City, Culture and Society, 18, 100291.
69. Nikiforiadis, A., Chrysostomou, K., & Aifadopoulou, G. (2019). Exploring Travelers' Characteristics Affecting their Intention to Shift to Bike-Sharing Systems due to a Sophisticated Mobile App. Algorithms, 12(12), 264.
70. Noland, R. B. (2019). Trip patterns and revenue of shared e-scooters in Louisville, Kentucky. Findings, 7747.
71. Noland, R. B. (2021). Scootin'in the rain: Does weather affect micromobility? Transportation Research Part A: Policy and Practice, 149, 114-123.
72. North American Bikeshare Association, General Bikeshare Feed Specification. https://github.com/NABSA/gbfs, NABSA, 2020.
73. Plekhanova, V. (Ed.). (2003). Intelligent agent software engineering. IGI Global.
74. Populus (2018). The Micro-Mobility Revolution: The Introduction and Adoption of Electric Scooters in the United States. San Francisco, CA.
75. Populus Technologies, Inc. (2018). The Micro-Mobility Revolution. Retrieved from: https://static1.squarespace.com/static/5fc6dab681da8a590dace76d/t/5ffb8ddbf752e30a15facc7d/1610321380415/Populus_MicroMobility_2018-July.pdf
76. Portland Bureau of Transportation (PBOT). (2018). 2018 E-Scooter Findings Report. The City of Portland, Oregon. Available at: https://www.portlandoregon.gov/transportation/78431.
77. Powell, Meerah (2020). Portland Extends E-Scooter Pilot Program Until End Of 2020. Oregon Public Broadcasting, OPB. www.opb.org/news/article/portland-extends-e-scooter-pilot-program-2020/ (accessed on December 20, 2019).
78. Preston, B. (2020, October 15) New Study Shows Safety Risks of Riding e-Scooters on the Sidewalk: IIHS finds that riding on sidewalks is dangerous for riders and pedestrians. Consumer Reports. Retrieved from https://www.consumerreports.org/electric-scooters/safety-risks-of-riding-e-scooters-on-the-sidewalk-iihs-study/
79. Reck, D. J., Haitao, H., Guidon, S., & Axhausen, K. W. (2021). Explaining shared micromobility usage, competition, and mode choice by modelling empirical data from Zurich, Switzerland. Transportation Research Part C: Emerging Technologies, 124, 102947.

80. Reck, D. J., Guidon, S., Haitao, H., & Axhausen, K. W. (2020). Shared micromobility in Zurich, Switzerland: Analysing usage, competition, and mode choice. Paper Presented at the 20th Swiss 34 Transport Research Conference (STRC), Ascona, May, 0–18. https://doi.org/https://doi.org/10.3929/ethz-b-000414863
81. Sanders, R. L., Branion-Calles, M., & Nelson, T. A. (2020). To scoot or not to scoot: Findings from a recent survey about the benefits and barriers of using E-scooters for riders and non-riders. Transportation Research Part A: Policy and Practice, 139, 217-227.
82. Sarjana, S., Ramadan, O.E., and Sisiopiku, V.P. (2020). Analysis of Transportation Users' Preferences and Attitudes for Identifying Micro-Level Determinants of Transportation Network Companies' (TNCs) Growth. Journal of Transportation Technologies, Vol. 10, No. 3. Pp. 251-264, June 19, 2020.
83. Shaheen, S., C. Beill, A. Cohen, and B. Yelchuru (2017). Travel Behavior: Shared Mobility and Transportation Equity. Publication PL-18-007, FHSW, Washington, DC
84. Sheather, S. (2009). A modern approach to regression with R. Springer Science & Business Media.
85. Sisiopiku, V. (2018). "Travel Patterns and Preferences of Urban University Students. Athens Journal of Technology & Engineering, Vol. 5(1), March 2018, pp. 19-31.
86. Sisiopiku, V.P., Hadi, M., McDonald, N., Steiner, R., and Ramadan, O.E. (2019). Technology Influence on Travel Demand and Behaviors (Project B), Final Report to the Southeastern Transportation Research, Innovation, Development and Education Center (STRIDE).
87. Sisiopiku, V.P., Morshed S. A., Sarjana, S., and Hadi, M. (2021). Transportation Users' Attitudes and Choices of Ride-Hailing Services in Two Cities with Different Attributes. Journal of Transportation Technologies, Vol. 11 (2), pp. 196-212.
88. Sisiopiku, V. P., and Ramadan, O. E. (2017). Understanding Travel Behavior and Mode Choice of Urban University Campus Employees. Proceedings of the 58th Transportation Research Forum Annual Conference, Chicago, IL.
89. Sisiopiku, V.P., and Salman, F. (2019). Simulation Options for Modeling Shared Mobility, Proceedings of the 2019 AlaSim International Conference and Exhibition, Huntsville, AL.
90. Smith, C. S., & Schwieterman, P. J. (2018). E-Scooter Scenarios: Evaluating the Potential Mobility Benefits of Shared Dockless Scooters in Chicago. Chaddick Institute Policy Series, 1, 6–8.
91. Sutton, R. S., & Barto, A. G. (1999). Reinforcement learning. Journal of Cognitive Neuroscience, 11(1), 126-134.
92. Tang, T. Q., Rui, Y. X., Zhang, J., & Shang, H. Y. (2018). A cellular automation model accounting for bicycle's group behavior. Physica A: Statistical Mechanics and Its Applications, 492, 1782-1797.
93. Taylor, D. B., & Mahmassani, H. S. (1998). Behavioral models and characteristics of bicycle-automobile mixed-traffic: planning and engineering implications (No. SWUTC-98-60056-1). University of Texas at Austin. Center for Transportation Research.
94. Tibshirani, R. (1996). Regression shrinkage and selection via the lasso. Journal of the Royal Statistical Society: Series B (Methodological), 58(1), 267-288.

95. UAB Class Schedule (2021). Available at: https://ssb.it.uab.edu/pls/sctprod/z1449_class_lookup.enter_parms
96. Veldhuis, L. O. (2018). Applications of Agent Based Modelling: Analysis and Simulation of Bicycle Traffic in Urban Environments. Report. Wageningen University: Wageningen, The Netherlands.
97. Verkehrswende, A. (2019). Shared E-Scooters: Paving the Road Ahead. Policy Recommendations for Local Government. Consulted, 17, 2020.
98. Wolfram, S. (1983). Statistical mechanics of cellular automata. Reviews of modern physics, 55(3), 601.
99. Xu, Y., Yan, X., Sisiopiku, V. P., Merlin, L. A., Xing, F., & Zhao, X. (2020). Micromobility Trip Origin and Destination Inference Using General Bikeshare Feed Specification (GBFS) Data. arXiv preprint arXiv:2010.12006.
100. Xue, S., Jia, B., Jiang, R., Li, X., & Shan, J. (2017). An improved Burgers cellular automaton model for bicycle flow. Physica A: Statistical Mechanics and its Applications, 487, 164-177.
101. Yang, H., Ma, Q., Wang, Z., Cai, Q., Xie, K., & Yang, D. (2020). Safety of micro-mobility: analysis of E-Scooter crashes by mining news reports. Accident Analysis & Prevention, 143, 105608.
102. Yang, Y., Heppenstall, A., Turner, A., & Comber, A. (2019). A spatiotemporal and graph-based analysis of dockless bike sharing patterns to understand urban flows over the last mile. Computers, Environment and Urban Systems, 77, 101361.
103. Younes, H., Zou, Z., Wu, J., & Baiocchi, G. (2020). Comparing the temporal determinants of dockless scooter-share and station-based bike-share in Washington, DC. Transportation Research Part A: Policy and Practice, 134, 308-320.
104. Zhang, W., Buehler, R., Broaddus, A., & Sweeney, T. (2021). What type of infrastructures do e-scooter riders prefer? A route choice model. Transportation Research Part D: Transport and Environment, 94 (March), 102761. https://doi.org/10.1016/j.trd.2021.102761
105. Zhao, Y., & Zhang, H. M. (2017). A unified follow-the-leader model for vehicle, bicycle, and pedestrian traffic. Transportation research part B: methodological, 105, 315-327.
106. 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, 101483.
107. Ziemke, D., Metzler, S., & Nagel, K. (2017). Modeling bicycle traffic in an agent-based transport simulation. Procedia Computer Science, 109, 923-928.
108. Ziemke, D., Metzler, S., & Nagel, K. (2019). Bicycle traffic and its interaction with motorized traffic in an agent-based transport simulation framework. Future Generation Computer Systems, 97, 30–40. https://doi.org/10.1016/j.future.2018.11.005
109. Zou, Z., Younes, H., Erdoğan, S., & Wu, J. (2020). Exploratory analysis of real-time e-scooter trip data in Washington, DC. Transportation research record, 2674(8), 285-299.
110. Zuniga-Garcia, N., & Machemehl, R. (2020). Dockless electric scooters and transit use in an urban/university environment. 99th Annual Meeting of the Transportation Research Board, Washington, DC.

10.1 APPENDIX A - Acronyms, abbreviations, etc.

Acronyms	Definition
ABM	Agent-Based Model
API	Application Programming Interface
CA	Cellular Automata
CDF	Cumulative Distribution Function
DSS	Dockless Shared Scooters
DT	Decision Tree
FI	Feature Importance
GBFS	General Bikeshare Feed Specification
GIS	Geographic Information System
IIHS	Insurance Institute for Highway Safety
IPF	Iterative Proportional Fitting
KDE	Kernel Density Estimation
LADOT	Los Angeles Department of Transportation
MaaS	Mobility as a Service
MAE	Mean Absolute Error
MATSim	Multi-Agent Transport Simulation
MDS	Mobility Data Specification
ML	Machine Learning
NPMRDS	National Performance Management Research Data Set
OD	Origin-Destination
OLS	Ordinary Linear Squares
PDF	Probability Density Function
PDP	Partial Dependence Plots
PMF	Probability Mass Function
QR	Quick Response
RF	Random Forest
RMSE	Root Mean Squared Error
TNC	Transportation Network Companies
UAB	University of Alabama at Birmingham
VIF	Variance Inflation Factor
XML	Extensible Markup Language
ZCTA	ZIP Code Tabulation Area

10.2 APPENDIX B - Associated websites, data, etc., produced

Website: https://faculty.eng.ufl.edu/sermos-lab/transit-innovative-mobility

10.3 APPENDIX C - Summary of Accomplishments

Date	Type of Accomplishment	Detailed Description
01/2020	Conference Paper	Zhao, X., Liu, X., Yan, X. (2020). Modeling demand for ridesourcing services in the City of Chicago: A direct demand machine learning approach. Proceedings of Transportation Research Board 99th Annual Meeting, Washington, DC.
01/2020	Conference Presentation	Xu, Y., Yan, X., Liu, X., Zhao, X. (2020) Applying Interpretable Machine Learning to Identify Key Factors Associated with Ride-Splitting Adoption Rate and to Model Their Nonlinear Relationships. Transportation Research Board 99th Annual Meeting, Washington, DC.
02/2020	Publication	Yan, X., Liu, X., Zhao, X. (2020). Using machine learning for direct demand modeling of ridesourcing services in Chicago. Journal of Transport Geography, 83, 102661.
02/2020	Publication	Zhao, X., Yan, X., Yu, A., Van Hentenryck, P. (2020). Prediction and behavioral analysis of travel mode choice: A comparison of machine learning and logit models. Travel Behaviour and Society, 20, 22-35.
06/2020	Conference Presentation	Elefteriadou, L., Du, L., Zhao, X. (2020). Autonomous vehicles and micromobility in a disruptive society and transportation system. The 5th Conference on Sustainable Urban Mobility.
08/2020	Conference Presentation	Sisiopiku, V., Zhao, X., Xu, Y., Yan, D., Steiner, R. (2020) Can Micro-mobility Reduce Urban Traffic Congestion? ITE Annual Meeting.
12/2020	Conference Presentation	Zhao, X. (2021). Micromobility for smart cities: Planning, design, and operations. Interstate Transit Research Symposium.
01/2021	Conference Paper	Zhang, X., Zhao, X. (2020). A Clustering-aided Ensemble Method for Predicting Ridesourcing Demand in Chicago. Proceedings of Transportation Research Board 100th Annual Meeting.
01/2021	Conference Paper	Noei, S., Zhao, X. (2020). Longitudinal Dynamics in Traffic Microsimulation. Proceedings of Transportation Research Board 100th Annual Meeting.

01/2021	Conference Paper	Xu, Y., Yan, X., Sisiopiku, V. P., Merlin, L. A., Xing, F., Zhao, X. (2020). Micromobility Trip Origin and Destination Inference using General Bikeshare Feed Specification (GBFS) data. Proceedings of Transportation Research Board 100th Annual Meeting.
01/2021	Educational Product	Xu, Y., Paliwal, M., Zhao, X. (2020). Real-time forecasting of micromobility demand: A context-aware recurrent multi-graph convolutional neural network approach. 2021 TRB workshop sponsored by AED50.
01/2021	Publication	Xu, Y., Yan, X., Liu, X., Zhao, X. (2021). Identifying key factors associated with ride-splitting adoption rate and modeling their nonlinear relationships. Transportation Research Part A: Policy and Practice. https://doi.org/10.1016/j.tra.2020.12.005
01/2021	Publication	Merlin, L. A., Yan, X., Xu, Y., Zhao, X. (2021). A segment-level model of shared, electric scooter origins and destinations. Transportation Research Part D: Transport and Environment.
03/2021	Media (article, etc.)	https://www.sciencedaily.com/releases/2021/03/210301091147.htm
03/2021	Media (article, etc.)	https://nyc.streetsblog.org/2021/03/04/e-scooters-are-best-for-short-trips-to-transit-shops-study/
03/2021	Educational Product	Zhao, X. (2021). Planning micromobility for future smart cities. University of Miami School of Architecture.
05/2021	Faculty Accomplishment or Award	Our lab's paper led by Dr. Xilei Zhao won the Travel Behaviour & Society (TBS) Outstanding Paper Award 2020. Find more details here: https://www.journals.elsevier.com/travel-behaviour-and-society/news/tbs-outstanding-paper-award-2020.
06/2021	Media (article, etc.)	https://www.alligator.org/article/2021/06/eletric-scooters-are-coming-to-uf
06/2021	Media (article, etc.)	https://joyride.city/3-ways-micromobility-services-can-socially-impact-cities/
07/2021	Educational Product	Zhao, X. (2021). Planning innovative mobility systems with machine learning. Transportation Data Science Seminar Series, Texas A&M University. (Invited talk)
09/2021	Educational Product	Xu, Y. (2021). Real-time forecasting of dockless scooter-sharing demand: A spatio-temporal multi-graph convolutional network approach. The UF AI Research Catalyst Fund Seminar Series. (Invited talk)