FINAL REPORT

PROJECT I2

NOVEMBER 2022

Mitigating Network Congestion by Integrating Transportation Network Companies & Urban Transit

Virginia Sisiopiku, Ph.D., University of Alabama at Birmingham
Mohammed Hadi, Ph.D., Florida International University

STRIDE | Southeastern Transportation Research, Innovation, Development and Education Center
UF | Transportation Institute UNIVERSITY of FLORIDA

TECHNICAL REPORT DOCUMENTATION PAGE

1. Report No.	2. Government Accession No.	3. Recipient's Catalog No.
Project I2		

4. Title and Subtitle
Mitigating Network Congestion by Integrating Transportation Network Companies and Urban Transit

5. Report Date
10/20/2022

6. Performing Organization Code

7. Author(s)
Virginia Sisiopiku, Ph.D., University of Alabama at Birmingham
Mohammed Hadi, Ph.D., Florida International University

8. Performing Organization Report No.
STRIDE Project I2

9. Performing Organization Name and Address
University of Alabama at Birmingham, Hoehn Engineering Building, Room Number 140, 1075 13th Street South, Birmingham, AL 35294-4440

Florida International University, Civil & Environmental Engineering, 10555 W Flagler Street, EC 3680, Miami, FL 33174

10. Work Unit No.

11. Contract or Grant No.
Funding Agreement Number 69A3551747104

12. Sponsoring Agency Name and Address
University of Florida Transportation Institute/ Southeastern Transportation Research, Innovation, Development and Education Center (STRIDE) 365 Weil Hall, P.O. Box 116580 Gainesville, FL 32611

U.S Department of Transportation/Office of Research, Development & Tech
1200 New Jersey Avenue, SE
Washington, DC 20590

13. Type of Report and Period Covered
8/1/2018 to 10/20/2022

14. Sponsoring Agency Code

15. Supplementary Notes N/A

16. Abstract - Transportation Network Companies (TNCs) like Uber and Lyft, provide a transportation option that offers a higher level of availability, reliability, and convenience than traditional taxi and transit services. However, there are widespread concerns about their impacts on urban congestion and their threat to public transit and taxi services, some of which are affirmed by recent case studies. The research team developed and demonstrated novel methods for calibrating MATSim models using a regionally approved mode split behavioral model and real-world traffic counts; collecting and processing Uber trip-level data using crowdsourcing to address the lack of publicly available TNC data; and modeling ride-hailing, in addition to automobile and transit trips, in the same simulation testbed. Products of this research include: (a) a questionnaire survey for documenting awareness and use of TNC services in the Southeastern US; (b) a rigorously calibrated MATSim model of the Miami Beach area; and (c) a comprehensive digital twin model of the Birmingham region. The latter MATSim model successfully incorporates public transit and ride-hailing services into the Birmingham transportation network, in addition to private automobiles. Overall, this research work provides valuable contributions to the current body of knowledge related to multimodal modeling using an open-source large-scale agent-based transportation simulation platform. The findings of the case studies reported herein provide evidence on the benefits of adopting transit, TNC, and road pricing strategies in small- and medium-size urban settings and can assist transportation decision makers, urban planners, transit agencies, and TNC providers in their efforts to optimize their operations and serve the needs of the traveling public.

17. Key Words	18. Distribution Statement
Transportation Network Companies (TNCs); MATSim Simulation Modeling; Model Calibration; Mode Integration; Digital Twin.	No restrictions

19. Security Classif. (of this report)	20. Security Classif. (of this page)	21. No. of Pages	22. Price
N/A	N/A	129 pages	N/A

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

DISCLAIMER

ACKNOWLEDGEMENT OF SPONSORSHIP AND STAKEHOLDERS

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

Funding Agreement Number - 69A3551747104

LIST OF AUTHORS

Lead PI:

Virginia P. Sisiopiku, PhD
Professor, Department of Civil, Construction and Environmental Engineering
University of Alabama at Birmingham
vsisiopi@uab.edu | ORCID 0000-0003-4262-8990

Co-PI:

Mohammed Hadi, PhD
Professor, Department of Civil & Environmental Engineering
Florida International University
hadim@fiu.edu | ORCID 0000-0003-2233-8283

Additional Researchers:

***Da Yan**, PhD*
Assistant Professor, Dept. of Computer Science
University of Alabama at Birmingham
yanda@uab.edu

Jalal Khalil
Graduate Student, Dept. of Computer Science
University of Alabama at Birmingham
jalalk@uab.edu

Syed Ahnaf Morshed
Graduate Student, Dept. of Civil & Environmental Engineering
Florida International University
smors005@fiu.edu

Taniya Sultana
Graduate Student, Dept. of Civil, Construction, and Environmental Engineering
University of Alabama at Birmingham
tans834@uab.edu

***Seyedmostafa Jafarzadehfadaki**, MS*
Graduate Student, Dept. of Civil, Construction, and Environmental Engineering
University of Alabama at Birmingham
mostaf86@uab.edu

***Furat Salman**, MS*
Graduate Student, Dept. of Civil, Construction, and Environmental Engineering
University of Alabama at Birmingham
furat@uab.edu

***Ossama E. Ramadan**, PhD, PE*
Post-Doc, Dept. of Civil, Construction, and Environmental Engineering
University of Alabama at Birmingham
oramadan@uab.edu

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

ABSTRACT

Transportation Network Companies (TNCs) like Uber and Lyft, provide a transportation option that offers a higher level of availability, reliability, and convenience than traditional taxi and transit services. However, there are widespread concerns about their impacts on urban congestion and their threat to public transit and taxi services, some of which are affirmed by recent case studies. Other studies, including one in Los Angeles, CA, reported benefits from integrating transit and TNCs. These contradictory findings motivate further research and investigation to examine and document effects of transit and TNC operations on transportation network performance. Moreover, they call for the development of new methods and models for assessing the mobility impacts of TNC presence and multimodal integration. Using the MATSim agent-based simulation platform and two cities in the Southeast (Birmingham, AL and Miami Beach, FL) as test sites, this study closely examined: (a) perceptions and predictors of TNC use; (b) mode choice shifts in the presence of transit, TNC, and road pricing options; and (c) impacts of various levels of public transit and/or TNC presence on transportation network performance. In doing so, the research team developed and demonstrated novel methods for calibrating MATSim models using a regionally approved mode split behavioral model and real-world traffic counts; collecting and processing Uber trip-level data using crowdsourcing to address the lack of publicly available TNC data; and modeling ride-hailing, in addition to automobile and transit trips, in the same simulation testbed. Products of this research include: (a) a questionnaire survey for documenting awareness and use of TNC services in the Southeastern US; (b) a rigorously calibrated MATSim model of the Miami Beach area; and (c) a comprehensive digital twin model of the Birmingham region. The latter MATSim model successfully incorporates public transit and ride-hailing services into the Birmingham transportation network, in addition to private automobiles. Overall, this research work provides valuable contributions to the current body of knowledge related to multimodal modeling using an open-source large-scale agent-based transportation simulation platform. The findings of the case studies reported herein provide evidence on the benefits of adopting transit, TNC, and road pricing strategies in small- and medium-size urban settings and can assist transportation decision makers, urban planners, transit agencies, and TNC providers in their efforts to optimize their operations and serve the needs of the traveling public.

Keywords: Transportation Network Companies (TNCs); MATSim Simulation Modeling; Model Calibration; Mode Integration; Digital Twin.

EXECUTIVE SUMMARY

Transportation Network Companies (TNCs) such as Uber and Lyft offer additional choices to travelers in their service area. However, to date, the impacts of these services on transportation network performance and transit use are not clear. One reason is the difficulty of modeling TNCs using traditional simulation models, which often lack the ability to simulate TNC trip characteristics. Privacy concerns, cited by TNC operations, also hinder the access to TNC trip records in most markets where such companies operate. The objective of this project was twofold: (a) study the influence of transit- and TNC operations on individual travelers' mode choices and (b) assess the impact of such choices on transportation network performance. To meet the study objectives, the research team performed four interrelated case studies that used Birmingham, AL and/or Miami Beach, FL as test beds. This report is organized as a compilation of four reports based on these four case studies.

Case Study 1 analyzed 790 questionnaire surveys of transportation system users across the Birmingham and Miami Beach regions with the purpose of establishing links between travel behaviors and TNC use in the Southeastern US. The survey responses confirmed that TNC service coverage and user characteristics influence user perceptions and adoption of such services. In the Birmingham region, TNC use was strongly correlated with vehicle availability and waiting time. In the Miami Beach region, variables identified as significant predictors of TNC use included vehicle ownership, vehicle availability, availability of Uber/Lyft, age range (18-29), high income (>$75K), and residency.

Case Study 2 developed a base model of the Miami Beach network using the multi-agent simulation model (MATSim). The model was then calibrated extensively using a regionally approved mode split behavioral model and real-world traffic counts. The calibrated Miami Beach MATSim model was used to estimate modal shifts between the passenger car and transit services due to the introduction of (a) an enhanced transit alternative and (b) road pricing, i.e., a toll fee for using a nearby highway facility. The case study findings confirmed that the modal shift towards transit is greater when the new transit option is introduced in combination with a road pricing strategy (4.1% shift towards transit), rather than alone (0.5% shift towards transit).

Case Study 3 examined potential impacts on traffic volumes, speeds, and travel times from expanding public transit options in the Birmingham region. This case study used a comprehensive activity-based simulation model of Birmingham to simulate traffic operations under various transit ridership scenarios ranging from 0% (base) to 1.1% (scenario 1-current) to 5.7% (scenario 2-future) to 10.1% (scenario 3-future). The findings identified ridership scenarios, demand levels, and time periods when such effects on traffic operations were significant.

Case Study 4 demonstrated the feasibility of modeling ride-hailing services using an agent-based simulation platform and evaluated the impact of such services on traffic operations in the Birmingham region. The findings of the study helped to determine the optimal number of Uber vehicles in the network on an hour-by-hour basis, given the varying demand for TNC service. Moreover, examination of speed and volume data confirmed that the availability of Uber services did not cause additional congestion in Birmingham, compared to the base case scenario (no TNC service).

Products of this research include: (a) a questionnaire survey for documenting awareness and use of TNC services in the Southeastern US; (b) a rigorously calibrated MATSim model of the Miami Beach area; and (c) a comprehensive MATSim model of the Birmingham region that models automobile, public transit and ride-hailing trips in the same network.

The methods proposed and tested in this study can be used in other medium-sized cities to understand travelers' mode choices in the presence of public transit and TNC services, and access the impacts of such preferences on their travel patterns and, in turn, the performance of the transportation network.

1.0 INTRODUCTION

1.1 BACKGROUND

Transportation Network Companies (TNCs) are companies that “provide prearranged transportation services for compensation using an online-enabled application or platform (such as smart phone apps) to connect drivers using their personal vehicles with passengers” (California Public Utilities Commission as cited by Cooper et al., 2018). TNCs, like Uber and Lyft, were initially perceived as a solution for urban congestion. However, in several cities in the US where these companies operate, TNCs were not able to deliver on their promise of fewer cars on urban networks. In fact, recent studies from heavily congested cities in the US have reported that TNCs were taking over the transit ridership and inducing a new line of business that is the Uber- or Lyft-driver, which in turn increased the levels of congestion represented by higher VMT. In addition, TNCs have introduced regulatory and policy challenges, mainly because of the controversial aggressive models of market entry and the push-back from regulated for-hire transport industry (Taylor et al., 2016).

In San Francisco, CA, the San Francisco County Transportation Authority partnered with researchers from Northeastern University who developed a methodology for collecting data through TNCs Application Programming Interfaces (APIs) with high spatial and temporal resolution (Cooper et al., 2018). Despite not having an independent data source to validate against, they were able to quantify the market penetration rate of TNCs in the study area. They estimated that TNCs serve over 170,000 trips on a typical weekday compared to 40,000 passengers served by public transit. Furthermore, they concluded that TNC trips followed traditional time-of-day distributions and were mostly transit substitution trips.

In Boston, MA, a recent report by the Boston Metropolitan Planning Organization (Gehrke et al., 2018) surveyed 1,000 travelers who frequently use Uber and/or Lyft. The finding showed that introducing TNCs in Boston, MA resulted in transit substitution at a rate of 54% with 12% occurring during the morning or afternoon commute periods. In addition, the survey concluded that transit substitution was more frequent among riders with a weekly or monthly transit pass. Thus, those who ride the transit more often are more likely to drop it for TNC services.

In New York, NY, Schaller (2017) raised concerns about the effects of TNCs on traffic congestion, emissions, and their potential to undermine public transit and taxi services that are essential components of urban transportation networks. Such concerns were based upon the fast-growing market share of TNCs. In 2016, TNCs transported 15 million passengers per month, and the ridership tripled between June 2015 and the fall of 2016. In addition, Shaller’s analysis indicated that TNCs added 600 million miles of vehicular travel to the city. Furthermore, he proposed a type of road pricing scheme to counter the rapid growth of TNCs (Schaller, 2017).

In Austin, TX, Lavieri et al. (2018) analyzed a dataset that contained trip-level information; including the location of trip origins and destinations, total trip length, and corresponding fare, of 1,494,125 trips that occurred between June 4, 2016 and April 13, 2017. Their analysis was performed using a spatial multivariate count model that consisted of recasting of the basic count model and spatial dependency formulation. In addition, they used a quasi-likelihood estimation approach to estimate the fractional split model. The first model was used to describe how many trips are generated in a specific zone on weekdays and weekends, and the second model helped in identifying the characteristics of zones that attract TNC-serviced trips. The study results showed spatial dependence in TNC-serviced trips among proximally located zones, as well as correlation between weekday and weekend trips originating in a zone. Furthermore, the results indicated that bus frequencies had a negative impact on the generation of TNC-serviced trips during the week, suggesting a substitution effect between TNC services and transit use for weekday trips. Moreover, the analysis suggested that travelers with different income brackets may use TNC services for different activity purposes.

On the global scale, several researchers investigated the issue of transit substitution, and the factors that influence travelers' choice and potential for reviving transit through integrating novel modes as feeder systems. One example is in Copenhagen, Denmark, where Anderson et al. (2017) analyzed the responses of 5,641 public transit users using a two-step traditional approach to route choice modeling. The first step employed a doubly stochastic approach for choice set generation to the public transport context to create plausible alternatives to the observed choices. The second step used a mixed path size correction logit to account for similarity across routes and heterogeneity across travelers. The utility function specified by Anderson et al. (2017) considered the entire door-to-door experience in the multimodal network in terms of access time, egress time, waiting time, in-vehicle time, and transfer time. In addition, close attention was paid to trip length, trip purpose and travelers' characteristics to gain insight into the preference structures of different travelers. The study findings indicated that the primary factors for choosing public transit are punctuality, fixed timetables, and free access to WiFi. Additionally, the results indicated that the most problematic factors are transfers, egress/access times, and walking times (especially during inclement weather). Their study can be regarded as comprehensive; however, they did not consider specifically TNCs as a mode or a feeder to the public transit.

More recently, Nguyen-Phuoc et al. (2018) conducted a qualitative survey of public transit users in Melbourne, Australia, to find out how would public transit users travel to their destinations if the entire public transit system was no longer available, either in the short term or long term. Their main focus was on identifying the main factors affecting people's mode choice in that context. Their results confirmed that many of the factors affecting mode shift from public transit to travelling by car in the event of public transit removal are similar to factors influencing car mode shift choice in normal conditions. These include car ownership, driver's license availability, travel distance, or travel cost. In addition, accessibility to public transit stations was found to be an influencing factor which had an effect on public transit substitution.

Despite of stated challenges and controversial views on TNCs' impacts on congestion and transit trips subsitution, the case of Los Angeles, CA, provides some evidence of beneficial integration of TNCs and transit. This was realized by the use of Uber by transit riders to fill-in the last/first mile gap (Uber, 2015). Analysis of one month of trip data that started or ended within ¼th of a mile of a metro station, indicated that trips near metro stations accounted for more than 16% of Uber trips that started or ended in Los Angeles.

The literature review confirms that case studies focusing on TNCs impacts on traveler's mode choices are limited. This identifies a present need to investigate the impact of TNCs presence on transportation operations. Furthermore, a need exists to study TNCs, urban transit, and other modes as an ecosystem in integration rather than in competition and understand how they can work in harmony to provide more choices and reduce urban congestion. While such studies are highly desirable, they are currently limited due to the lack of (a) available data and (b) simulation models that allow for consideration of passenger car, TNC, and urban transit trips to be modeled simultaneously using commonly used simulation platforms.

1.2 OBJECTIVE

This study examined the important issue of multi-modal trip integration in urban settings in the presences of TNCs. The objective was to explore, identify, and model methods for integrating TNCs and public transit along with automobile traffic in order to study associated impacts on operational performance. To achieve the stated objective, the study:

- Documented transportation users' choice preferences on the selection of ride-hailing services as a mobility option in the Southeast US (Chapter 2)
- Demonstratef methods to calibrate a multi-agent simulation model (MATSim) and usef it to estimate the mode split between the passenger car and transit services in the Maimi Beach region (Chapter 3).
- Developed a comprehensive activity-based simulation model of the Birmingham area and used it to simulate traffic operations under various transit ridership scenarios (Chapter 4)
- Demonstrated novel methods for collecting and processing TNC trip data through surveys of Uber drivers and the use of crowdworkers (Chapter 5), and
- Developed a digital twin of Birmingham's transportation network that accounts for various modes of transportation including TNCs, public transit, and private vehicles (Chapter 5).

1.3 SCOPE

This is one of the first studies in the US to attempt to model Uber, transit, and passenger car trips in an integrated manner using the agent-based platform MATSim. Using the Birmingham, AL and Miami Beach, FL as modeling test beds, the study team collected data pertinent to transportation users' mode choice preferences, transportation network characteristics, existing transit routes and schedules, and TNC trips. Population data for Birmingham were obtained from the STRIDE project 2017-B while similar data were collected in Miami through a questionnaire survey of transportation system users using the Qualtrics online platform. Novel approaches were used to expand the survey data using open-source databases in order to enhance the quality and realism of the simulation modeling. Due to the lack of available TNC trip records, the study team collected app screenshots of Uber rides from Uber drivers in the Birmingham region and applied a crowdsourcing approach to extract ride information from the app screenshots. Another novelty of the study was to demonstrate the calibration of the multi-agent simulation based on the regionally approved mode split behavioral model and real-world traffic counts obtained for the Miami Beach network.

The methods used in the case studies documented in this report and related findings are expected to benefit various transportation agencies, transit providers, MPOs, municipalities, and the general public. The aim is to assist transportation stakeholders in planning and managing transportation network operations where car/ridesharing platforms are active with the ultimate goal of relieving congestion off urban and regional networks and improving transportation operations.

1.4 REFERENCES

1. Anderson, M. K., Nielsen, O. A., & Prato, C. G. (2017). Multimodal route choice models of public transport passengers in the Greater Copenhagen Area. *EURO Journal on Transportation and Logistics,* 6(3), 221-245. DOI: 10.1007/s13676-014-0063-3
2. Cooper, D., Castiglione, J., Mislove, A., & Wilson, C. (2018). Profiling TNC Activity Using Big Data*. TRB 97th Annual Meeting Compendium of Papers*, Paper № 18-05899.
3. Gehrke, S. R., Felix, A., & Reardon, T. (2018). Fare Choices: A Survey of Ride-Hailing Passengers in Metro Boston. Boston, MA: Boston Metropolitan Planning Organization.
4. Lavieri, P. S., Dias, F. F., Juri, N. R., Kuhr, J., & Bhat, C. R. (2018). A Model of Ridesourcing Demand Generation and Distribution. *Transportation Research Record: Journal of the Transportation Research Board*. DOI: 10.1177/0361198118756628
5. Nguyen-Phuoc, D. Q., Currie, G., De Gruyter, C., & Young, W. (2018). How do public transport users adjust their travel behaviour if public transport ceases? A qualitative study. *Transportation Research Part F: Traffic Psychology and Behaviour,* 54, 1-14. DOI: 10.1016/j.trf.2018.01.009
6. Schaller, B. (2017). UNSUSTAINABLE? The Growth of App-Based Ride Services and Traffic, Travel and the Future of New York City. Brooklyn, NY: Schaller Consulting.

7. Taylor, B. D., Chin, R., Crotty, M., Dill, J., Hoel, L. A., Manville, M., . . . Kortum, K. (2016). *Between Public and Private Mobility: Examining the Rise of Technology-Enabled Transportation Services*. Washington, DC: Transportation Research Board.
8. Uber. (2015, March 11, 2016). Uber and LA's Public Transportation Working Together. Retrieved May 21, 2018, from https://newsroom.uber.com/us-california/uber-and-las-public-transportation-working-together/

CHAPTER 2: CASE STUDY 1 - TRANSPORTATION USERS' ATTITUDES AND CHOICES OF RIDE-HAILING SERVICES IN TWO CITIES WITH DIFFERENT ATTRIBUTES

2.1 INTRODUCTION

Transportation Network Companies (TNCs) such as Uber and Lyft are smartphone app-based ride-hailing services that have grown rapidly over the past decade. Such services match passengers with drivers using online enabled platforms. The launch of TNC services took place in 2009, when Uber (formerly known as Uber Cab) introduced the service in the San Francisco area (Hartmans, 2018). Soon after, TNCs made their appearance in various other markets across the US, thus adding transportation options that competed or complemented available transportation services. The promise to save time, increase affordability and convenience, reduce stress, and the lack of need to own and use a personal automobile has been appealing to many customers who embraced TNC services, especially in large metropolitan areas. Among available TNCs in the US market, Uber is the market leader with 65% market share.

In addition to providing user benefits, TNCs were initially perceived as a solution for urban congestion. However, recent studies from heavily congested cities in the US have reported that TNCs took over part of the transit ridership rather than promoting ridesharing among solo drivers. To make things worse, Uber- or Lyft vehicles waiting for rides contributed to increased Vehicle Miles Traveled (VMT) and urban pollution. In addition, TNCs have been involved in regulatory and policy challenges, mainly because of the controversial aggressive models of market entry and the pushback from regulated for-hire transport industry (National Research Council Committee for Review of Innovative Urban Mobility Services, 2015).

A number of recent studies explored the emerging trend of TNC services as a mode of transportation. A concise summary is available by Sisiopiku et al. (2019). Shaheen (2018) discussed the recent cultural shift from the auto-dependency to shared mobility and the impact of such shift on the growth of ride-hailing services such as Uber and Lyft. Several studies attempted to define TNC market characteristics using surveys. These studies showed great variations in their findings depending on the geographical locations and the surveyed user demographics. For instance, studies conducted in large metropolitan areas like Boston, Chicago, Los Angeles, New York, San Francisco, Seattle and Washington D.C. showed that the typical TNC user is 18-29 years of age and possesses an advanced degree (Clewlow & Mishra, 2017). However, TNC users in cities like Pittsburgh and Puget where predominantly 34-44 years old and holding Bachelors' degrees (Vinayak et al., 2018 & Chen, 2015). Circella et al. (2017) examined the differences in travel mode choice between Millennials and Generation Xers in California using inputs from 2,155 individuals. When compared to Gen Xers, Millennials were three times more likely to use Uber or Lyft. With respect to older transportation users, Freund et al. (2020) suggested that door-to-door assistance service could increase the use of TNC service among 65+ years old population. A survey of 380 TNC users in San Francisco conducted by Rayle et al. (2014)

reported that 67% of responders used ride sourcing for social/leisure trips (bars, restaurants, concerts, friends/family visits) while only 16% used it for commuting purposes. Responders reported that their leaning towards TNCs was driven by the availability of a secure payment system with short wait time. Approximately 40% of TNC users in the San Francisco region reported using their private vehicle less due to the availability of on-demand mobility sharing services (Rayle et al., 2014). A national Pew Research Center survey of 4,787 American adults in 2015 found only 15% of Americans had used ride-hailing apps, whereas one third had yet to even hear of them (Smith, 2016). While the author did not find race or gender as influential factors in the use of these apps; age, education, income level and type of locale (i.e., urban, suburban, or rural) were all found to be strongly explanatory.

Overall, the literature review confirms that knowledge and utility of TNC services among travelers vary greatly in accordance with a vast array of socio-demographic variables, as with most new technology (Smith, 2016). Moreover, systematic documentation of actual impacts of TNCs presence on the preferences and daily travel patterns of the transportation system users is still limited. This is due to the lack of availability of detailed data resulting from privacy concerns and resistance of TNC companies to share company data (Sarjana et al., 2020). Thus, localized studies are of great value in order to document transportation users' attitudes and preferences and identify socio-demographic variables that influence the use of TNC services.

In light of this need, the objective of this case study was to examine the impact of transportation users' choice preferences on the selection of ride-hailing services as a mobility option in the Southeast US. Using questionnaire responses from two different geo-locations in the Southeast (i.e., Birmingham, Alabama and Miami Beach, Florida) the study documented and compared preferences and attitudes toward TNC use as a travel mode of choice. The study considered demographic data in the analysis and interpretation of the survey findings and the identification of indicators that affect the use of TNCs at the study locations.

2.2 METHODOLOGY AND DATA COLLECTION

This study compared results from two surveys that collected and documented public perceptions related to ride-hailing services in two TNC markets: namely Birmingham, AL and Miami Beach, FL. The study builds on the authors' earlier work (Sarjana et al., 2020) that used a questionnaire survey to understand the leading reasons and conditions that drive people towards the use of TNCs services in the Birmingham Metro Area.

Both surveys were developed using the Qualtrics Research Core tool in accordance with the Institute of Transportation Engineers Manual on Transportation Engineering (ITE) Studies guidelines (Institute of Transportation Engineers, 2011). Qualtrics LLC facilitated the identification and recruitment of survey participants and automated the data entry and management process. The research team obtained the survey responses from Qualtrics LLC and performed validation checks, data processing, and data analysis. All necessary approvals were obtained from the Institutional Review Board (IRB) for Human Use prior to conducting the surveys. For quality assurance, both questionnaires were pretested and refined prior to distribution.

The surveys sought to get information about users' attitudes towards using TNCs along with detailed socio-demographic such as age, gender, education level, and employment type. The demographic data were categorized based on the US Census criteria. The survey also requested participants to report detailed trip information for a typical day (i.e., 24-hr travel diary) during a typical weekday including origin and destination of each trip, travel time, trip purpose and the travel mode used. Additionally, information related to vehicle ownership, alternate mode choices, and recommendations for future transportation improvements (including expansion of the TNC services) was solicited. The survey tool was initially developed as part of STRIDE B project and was used to collect data for the Birmingham area. Detailed information on the survey tool available in Sisiopiku et al., 2019. The survey tool was adopted and slightly revised in preparation of use in STRIDE I2 for collecting data from the Miami Beach area. Copies of both the Birmingham and Miami Beach surveys are available in Appendix 7.2

Participants were presented with simple multiple-choice questions with specified context and were asked to answer each question categorically based on the context. While some survey questions were identical or similar between the two study sites, others solicited inputs on issues of unique importance to each study site. For example, since Miami Beach Area is a popular tourist spot that is busy during the weekends, Miami Beach survey participants were asked to provide trip information for a typical weekend, in addition to a typical weekday. Also, residing within the geographical area of interest was a requirement for participation in the Birmingham study, but not in the Miami Beach survey, in order to allow for documentation of responses from tourists that visited the Miami Beach area for recreational purposes.

The collected responses were carefully checked and validated. After eliminating any surveys that included incomplete, duplicate, or irregular answers, 451 responses from the Birmingham area and 339 from the Miami Beach area were analyzed for a total of 790 surveys. It is important to note that out of the 339 respondents in Miami Beach Area, 71 (21%) were Miami Beach residents and the rest (79%) were visitors from the greater Miami area or out-of-city tourists. This allowed for examination of potential differences in the preferences and attitudes toward TNCs between residents and tourists in the Miami Beach case study. For a quick reference, Table 2.1 summarizes characteristics of both study locations along with information relevant to the two surveys.

Table 2. 1 Summary Characteristics of Study Sites and Survey Responses

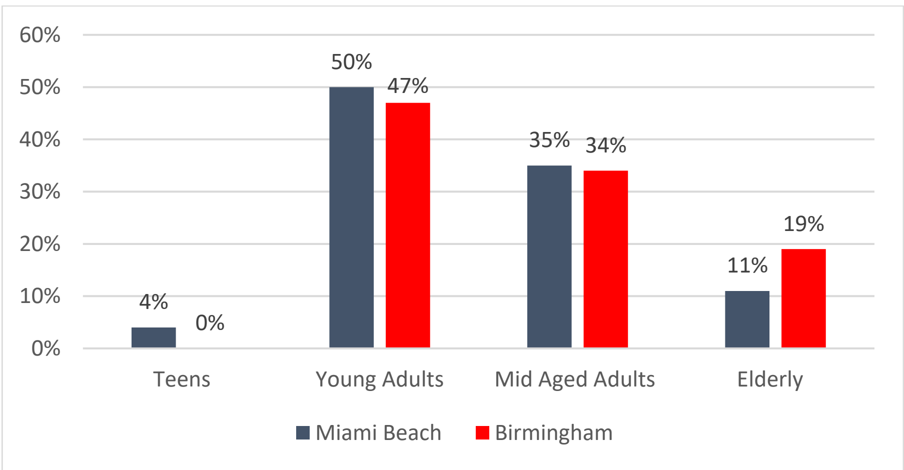
	Birmingham, AL	**Miami Beach, FL**
Population	212,297 capita (2019 Census)	88,885 capita (2019 Census)
Available Travel Modes	Private vehicle, bus, taxi, TNCs, bikes, carpool, vanpool	Private vehicle, bus, taxi, TNCs, bikes, carpool, vanpool, subway
Total Responses	451	339
Number of Self-reported Trips	1,130	878
Type of Respondents	100% residents	21% residents, 79% tourists

2.3 DATA ANALYSIS AND RESULTS

2.3.1. DESCRIPTIVE ANALYSIS

Population segmentation through demographic characteristics illustrates the size of potential TNC market in the selected study regions. Among the 451 responders from Birmingham and 339 from Miami Beach considered in the analysis, 342 and 204 respectively were women. Based on the responses provided in the Birmingham and Miami Beach surveys, more female than male travelers are TNC users where the female to male ratio is 74:26 and 55:44 respectively.

When considering the age of the survey participants, the largest percentage of participants in both the Miami Beach and Birmingham surveys represented the young adult age group. Inspection of the survey results confirmed that the peak age group for the overall survey correlated with the TNC users. Figure 2.1 displays the distribution of the TNC users by age group. Teens include participants under 18 years of age; young adults include those between 18 and 34 years of age; mid aged adults include participants 35-54 years of age; and elderly include those that were over 55 years old.



Note: Teens: <18 yrs; Young Adults: 18-34 yrs; Mid Aged Adults: 35-54 yrs; Elderly: >55 yrs

Figure 2. 1 Age group of survey participants (TNC users).

Figure 2.2 shows the usage of modes other than private automobile in the past year for the survey participants. It can be observed that ride-hailing services were more popular than public transit service and organized ride sharing programs among the users in both study locations (73% FL and 45% AL).

As shown in Figure 2.3, the analysis of self-reported trip data over a 24-hr period showed that approximately 6.4% of the reported trips in the Birmingham region were conducted by TNCs whereas in Miami Beach the reported TNC market share was higher (20% of the weekday trips and 17% of the weekend trips). There are many possible explanations for the difference. First, Uber was introduced in the Miami market in June 2014; more than a year and a half before coming into the Birmingham market thus users are more familiar with TNC presence. More importantly, there is the severe shortage of parking availability in the Miami Beach area, which makes ride-hailing services an attractive alternative to automobile use, especially for tourists and visitors within the Miami area. The Miami Beach area has a large number of restaurants, bars,

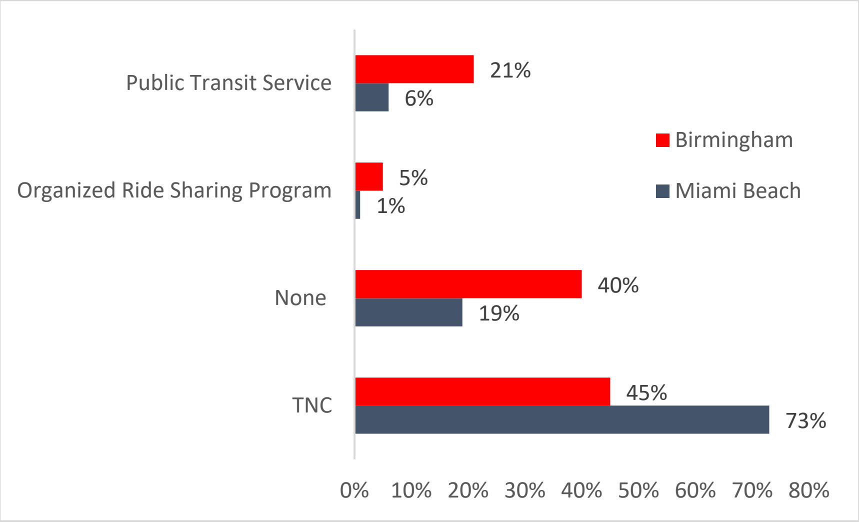


Figure 2. 2 Modes other than private car used in the past year by the survey participants.

and late-night entertainment venues, which may make TNC use attractive, as users try to avoid driving after alcohol consumption. It is also worth noting that personal automobile usage in Birmingham is significantly higher (85%) than that at the Miami Beach (56%) area. This is consistent with earlier studies in the Birmingham area (Sisiopiku, 2018; Sisiopiku & Ramadan, 2018) which reported automobile use in the Birmingham metro between 83% and 86%. It is also evident that use of transit services for both regions is very low (2%).

Additional analysis was performed to compare mode choices between weekday and weekend trips in the Miami Beach region (Figure 2.4). As previously mentioned, weekends are typically busy in this region due to high attraction of tourists and visitors. It can be observed that there are no major differences in the preference of mode selection in Miami Beach region based on the day of the week consideration (weekday versus weekend).

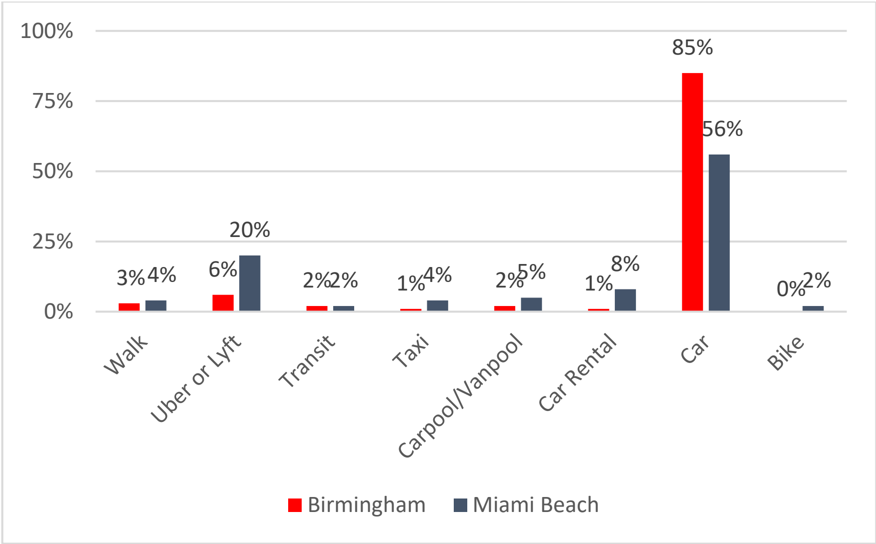


Figure 2. 3 Mode Choices of Birmingham and Miami survey respondents.

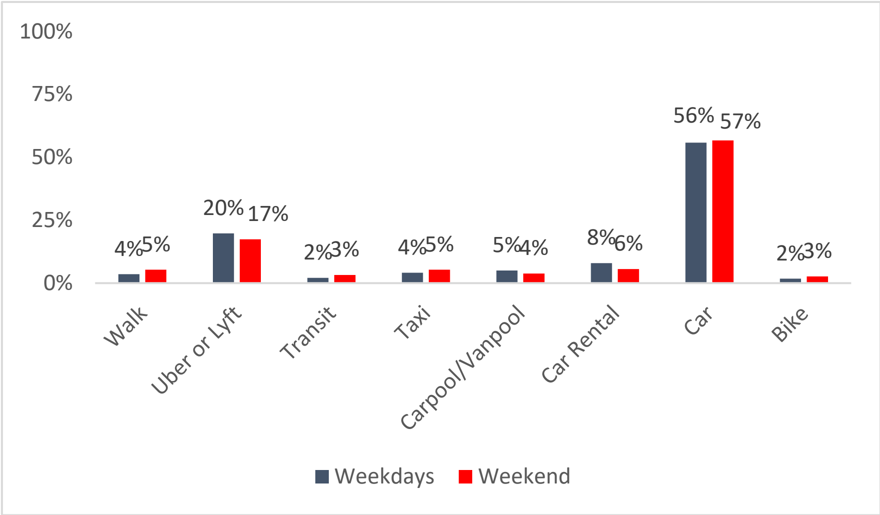


Figure 2. 4 Mode choices during typical weekdays and weekends (Miami Beach)

Figure 2.5 illustrates trip purposes of trips performed by survey respondents using TNC as their mode of transportation during typical weekdays in the Birmingham and Miami Beach regions. Additionally, the weekend trip purpose versus TNC trip data was considered to assess any difference in behaviors during the weekend scenario in the Miami Beach region.

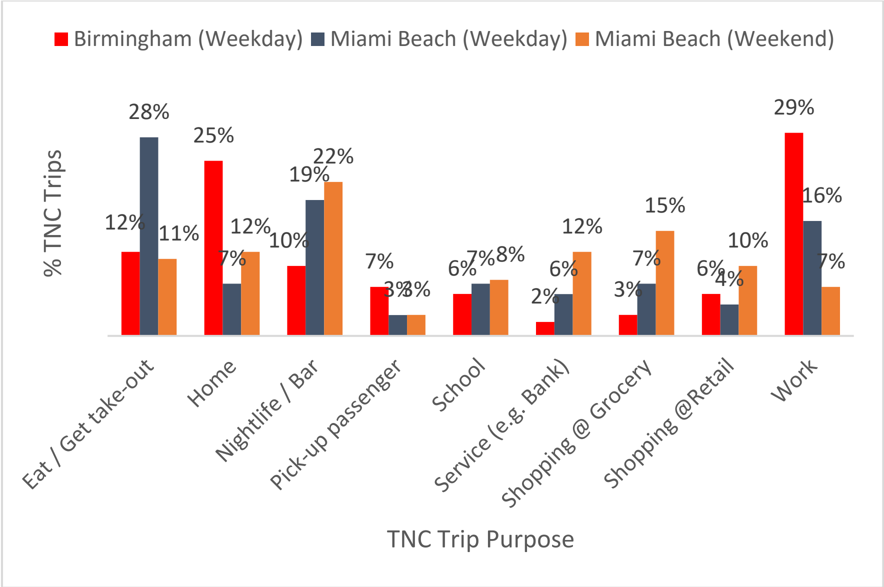


Figure 2. 5 Trip purposes vs TNC trips.

It can be observed from Figure 2.5 that the majority of Birmingham respondents use TNCs to travel to work (29%) or home (25%). During weekdays, the TNC trips to work in Miami Beach are just 16%, noticeably lower to those in the Birmingham Metro area. This is expected, as the majority of the survey respondents in the Miami Beach area are either tourists or commute to Miami downtown with personal automobile. It can also be seen that the peak destinations for weekday TNC trips in the Miami Beach are restaurants (28%) and nightlife/bar (19%). As expected, during weekends in Miami Beach, the majority of the TNC trips are geared toward entertainment and shopping with trips to nightlife/bar (22%) being the dominating trip purpose category.

The relationships between vehicle ownership, TNC usage, and TNC trip purpose are illustrated in Figures 2.6 and 2.7 for Birmingham and Miami Beach, respectively. In the Birmingham metro area, it is evident that individuals having vehicle ownership tend not to use TNC as a transportation mode of choice. Survey respondents having regular access to personal automobiles use TNC service occasionally to commute to work (17%) or home (10%) and for recreational purposes. In the Miami Beach area, the scenario is opposite to that of Birmingham. Despite personal vehicle ownership, survey respondents opted to use TNC service, primarily to access restaurants (21%) and bars (13%) during nighttime. When non-resident trips were excluded from the analysis, similar results were obtained with the majority of TNC trips by Miami Beach residents being destined to restaurants (40%) and bars (13%).

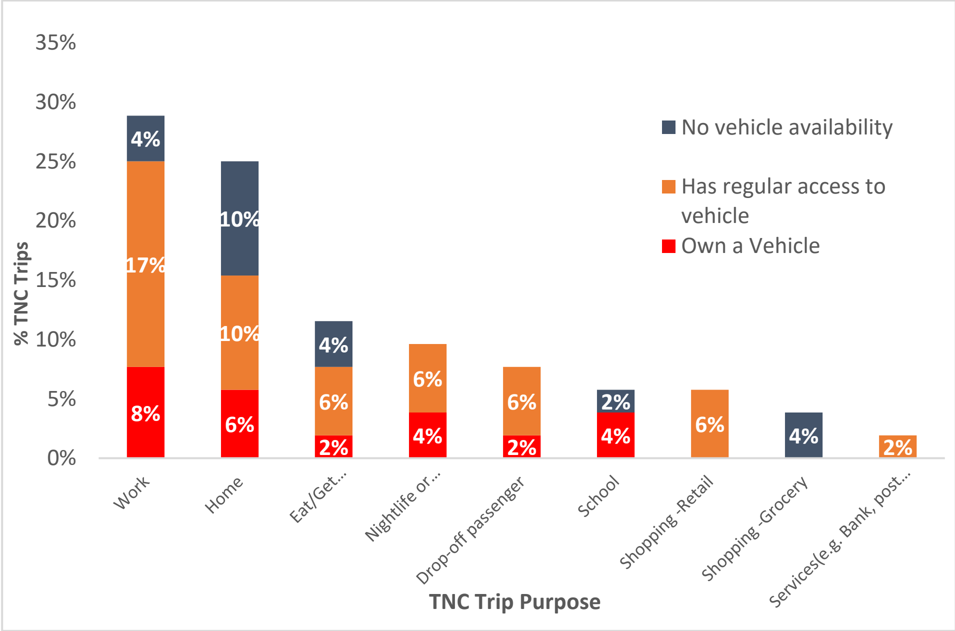


Figure 2. 6 Car availability of TNC users (Birmingham). Source: [Sarjana et al., 2020].

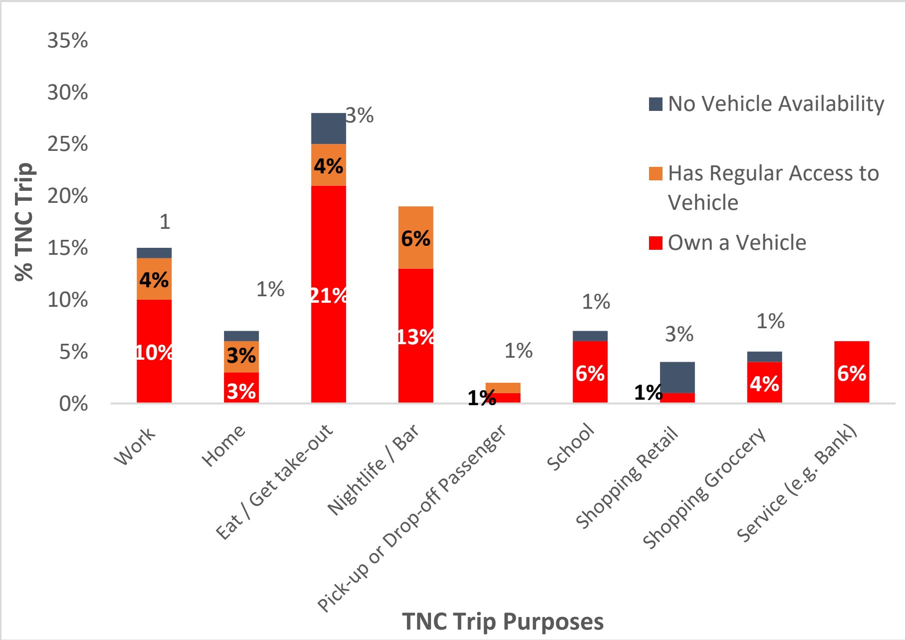


Figure 2. 7 Car availability of TNC users (Miami Beach).

Table 2.2 and Table 2.3 show the correlation between TNC trip time, waiting time and vehicle availability on mobile apps for the Birmingham and Miami Beach regions, respectively. TNC trip times are categorized based on the time of the day (i.e., daytime and nighttime). It can be observed that the range of waiting times is different for the two regions under consideration reflecting local conditions. The maximum waiting time for the Miami Beach region is more than 30 minutes (especially during weekends), whereas in Birmingham it is much shorter, ranging between 11 and 15 minutes. Also, the findings confirm that the range of available vehicles in the app differs by region. In contrast to the maximum range of vehicles in Birmingham (3-5), the range of maximum vehicles in Miami Beach is more than 5.

Table 2. 2 Correlation of TNC trip time with waiting time and vehicle availability on app (Birmingham)

Available vehicles on App	1-2	3-5	Total
Day-time trips	**38%**	**12%**	**50%**
0-5 minute	19%	6%	25%
11-15 minutes	4%	4%	8%
6-10 minutes	15%	2%	17%
Night-time trips	**50%**	**0%**	**50%**
0-5 minute	33%	0%	33%
6-10 minutes	17%	0%	17%
Total	**88%**	**12%**	**100%**

Table 2. 3 Correlation of TNC trip time with waiting time and vehicle availability on app (Miami Beach)

Available vehicles on App	None	1-5	5+	Total
Day-time trips	**4%**	**31%**	**16%**	**51%**
<5 minute	0%	7%	2%	9%
5-15 minutes	0%	15%	13%	28%
15-30 minutes	4%	9%	2%	15%
More than 30 minutes	0%	0%	0%	0%
Night-time trips	**7%**	**33%**	**9%**	**49%**
<5 minute	2%	7%	4%	13%
5-15 minutes	4%	18%	4%	25%
15-30 minutes	2%	4%	0%	5%
More than 30 minutes	0%	4%	2%	5%
Total	**11%**	**64%**	**25%**	**100%**

The trips are similarly allocated during day and night in both regions. The majority of the waiting times in Birmingham are between 0-5 minutes, both in daytime (25%) and nighttime (33%). However, in Miami Beach, the majority of survey responders report waiting times that range between 5-15 minutes both during daytime (27%) and nighttime (25%), possibly reflecting the nature of the transportation network in Miami Beach that provides accessibility challenges to the drivers, including Uber drivers.

The survey participants were asked about their preference with respect to future improvements related to transportations infrastructure and services. Figure 2.8 illustrates such preferences based on survey responses in the Birmingham and Miami Beach regions. According to their responses, 26% and 19% of survey participants recommended an expansion of TNC services in the Birmingham and Miami Beach regions respectively. Furthermore, survey participants from both regions equally prioritized improving public transit services (43%).

Survey respondents in Birmingham were also asked to note the reason(s) for using TNCs in the past. To identify the most influential reasons for selecting TNC services as a mode of transportation, the reasons stated by the respondents were factorized in terms of binary values where 1 is assigned to 'selected values' and 0 is assigned to values that were 'not selected'. Table 2.4 documents the mean and standard deviation according to the survey responses.

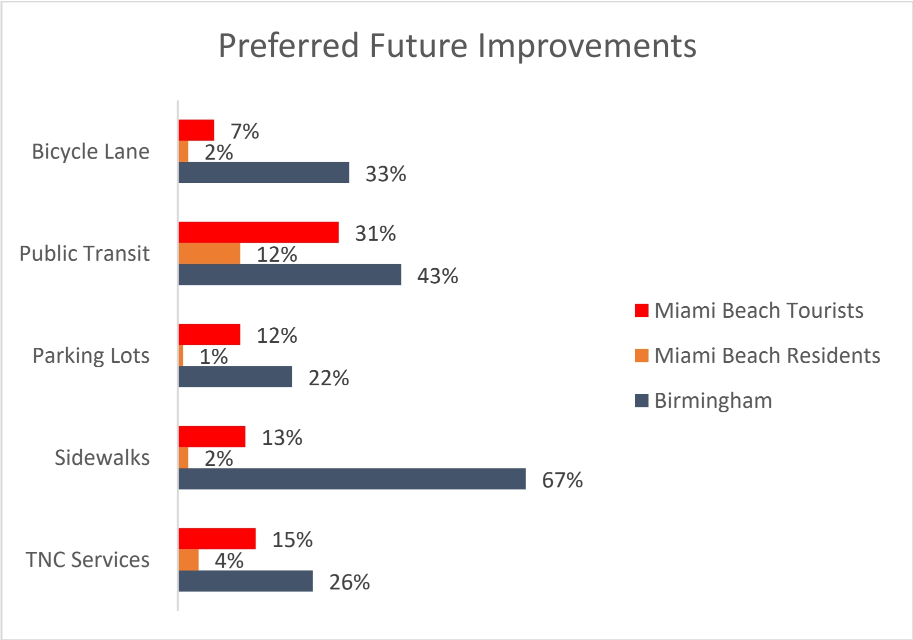


Figure 2. 8 Preference for future improvements of transportation infrastructure and services.

The results in Table 2.4 clearly show that convenience was reported by transportation users in the Birmingham area as the main driving force for selection of TNCs as a travel mode. Safety/avoiding driving when intoxicated, and lack of automobile availability were cited as the second and third most important reasons for use of TNCs in the survey of Birmingham users.

Table 2. 4 Summary Characteristics of Study Sites and Survey Responses

Reason	Mean	Standard Deviation
Convenience	0.56	0.50
Safety	0.30	0.46
No car availability	0.27	0.44
Destination has limited/no parking	0.24	0.43
Cheaper than alternatives	0.21	0.41
Parking at destination is expensive	0.19	0.39
Transit is not accessible	0.06	0.23
Transit is not reliable	0.03	0.17
Other reason	0.03	0.18
Other mode not available	0.02	0.15

The survey conducted in Birmingham also asked respondents with no previous TNC experience within the past year to mark the main reason for not considering TNCs as a mode of transportation. From Figure 2.9, it can be observed that nearly 30% survey respondents reported that the use of TNCs was not convenient for them, while another 20% noted that they do not use TNCs due to associated cost. The “other reasons” that were cited by the survey respondents include personal preference to use other transportation modes, distance to destination making TNC unattractive, concerns about riding in another person’s car and the lack of accommodations for individuals with limited mobility.

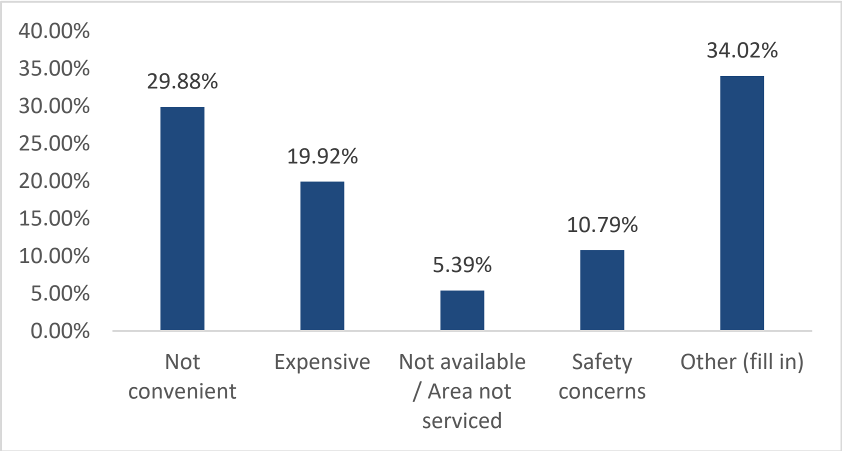


Figure 2. 9 Reasons for not Using TNCs (Birmingham).

The study also confirmed that trip distance plays an important role as a potential determinant of TNC use. According to the characteristics of the TNC trips reported in the Birmingham case study, TNC users use TNC services for trips under 10 miles. A comparison between TNC and non-TNC trips revealed that the average trip length performed by TNC was 5.19 miles, far lower than the average trip length of automobile trips (9.28 miles) in the region.

Further analysis indicated that the longest TNC trips involved drop-off of a passenger or trips to work or home. The average trip length per trip purpose for TNC trips is shown in Figure 2.10.

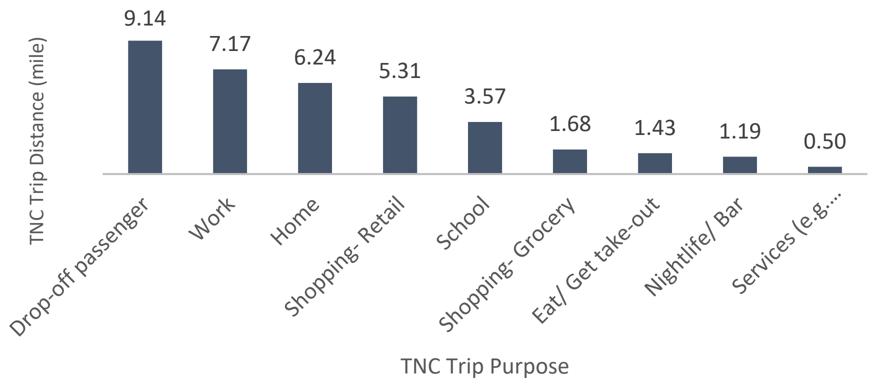


Figure 2. 10 TNC trip distance (miles) for various trip purposes (Birmingham).
SOURCE: [Sarjana et al., 2020]

The survey conducted in Miami Beach also requested feedback from respondents about their willingness to use ride-hailing services to commute, if incentives or special services were provided. The four questions that were designed to portray the hypothetical scenarios are as follows:

- Willingness to use ride-hailing services for the first or last mile if incentives (discounts) are provided
- Willingness to use the ride-hailing services for commuting if the public transit service is made free
- Willingness to use the ride-hailing services for the first or last mile to reach home or to the nearby public transit stop or station, and
- Willingness to use the ride-hailing services for commuting if a rewards point system is introduced by their workplace.

The majority of the respondents expressed their willingness to use the TNC services when incentives or special offers are provided (Figure 2.11). The type of incentive offered appeared to make little difference in their response. It is worth noting that only 20% of the survey participants reported a complete lack of interest in considering TNC services for commuting purposes, even if incentives were offered. This is consistent with other studies that suggest that automobile-dependent users show resistance in embracing alternative modes (Morshed et al., 2020).

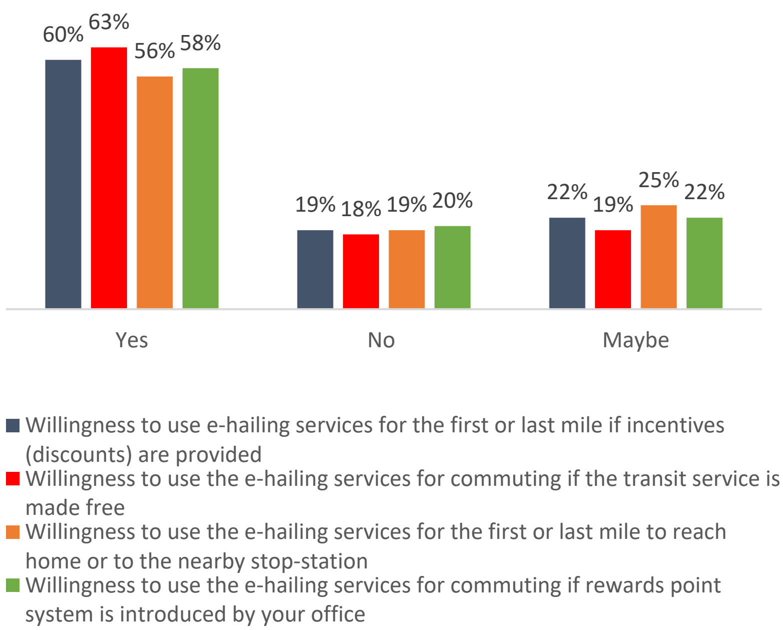


Figure 2. 11 Willingness to use ride-hailing services in hypothetical scenarios (Miami Beach).

2.3.2. REGRESSION ANALYSIS

In addition to the descriptive analysis, this case study employed the Least Absolute Shrinkage and Selection Operator (Lasso), a very popular machine learning algorithm to perform regression analysis (Tibshirani, 1996). The purpose of this effort was to identify parameters that can be used to predict TNC use. Lasso regression provides two major advantages over linear regression, which made it an attractive alternative to linear regression, namely (a) clear variable or feature selection and (b) better prediction accuracy. Lasso regression, through shrinkage or regularization of the coefficients, increases the prediction accuracy and decreases the variance of the model interpretability. This regression technique selects strong variables in high dimension data for clearer interpretations of the results since models with too many variables are hard to interpret (Hastie et al., 2015). Additionally, Lasso regression eliminates over-fitting, i.e., large variance and unbiased estimates, which increases prediction accuracy (Steele et al., 2018).

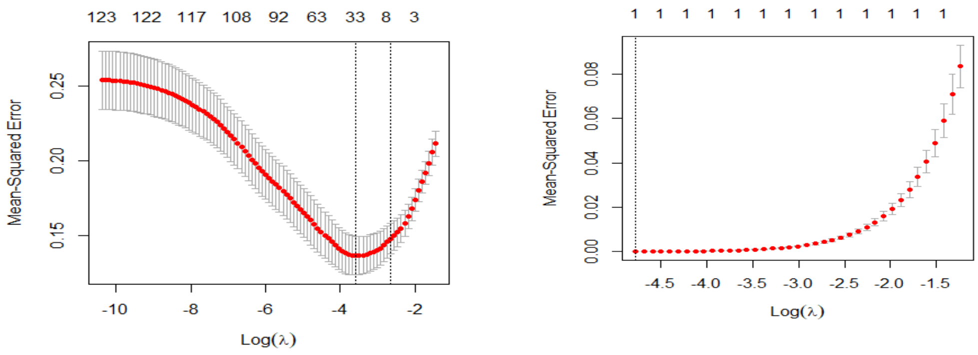
In Lasso regression, categorical variables are encoded into a set of indicators by transforming the variables into factors. Consequently, a dummy variable matrix of predictors is created, along with continuous predictors to serve as inputs to the model. Dummy coding includes binary attributes to indicate category membership. The reference category is indicated as ‘0’ and corresponding category is coded as ‘1’ in the dummy coding. As shown in Equation (2.1), Lasso adds a penalty term i.e., product of a bias parameter λ with the absolute value of the slope to regulate the size of the coefficients (β_lasso) which can affect the number of predictors included in the model.

β_lasso=Min (Sum of squared residuals)+ λ*|slope|…………………………………………………………..(2.1)

The tuning parameter (λ) is chosen by cross validation, i.e. when λ= 0 mean square error is 0. As λ increases, shrinkage occurs so that the variables that are insignificant ('0' value) are eliminated.

In this study, the Lasso method was applied separately on the Miami Beach Survey and Birmingham Survey Data. There were 155 baseline variables for the Miami Beach case study and 103 base line variables for the Birmingham case study. 80% of the data was used to train the model and the remaining 20% was used for model predictions. The dependent variable 'y' represented the TNC usage among the respondents in both study regions. The estimated coefficients and the variables selected under the Lasso method considered only features that were significant for the model fitting.

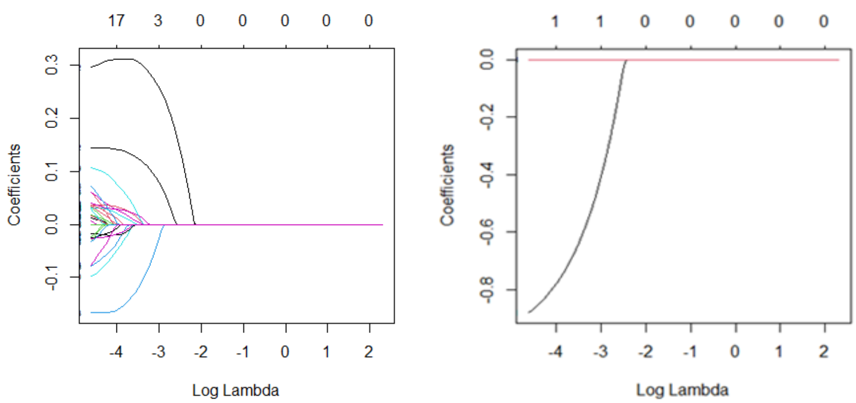
Lasso fits the most significantly contributed variables by making the insignificant variables exactly to zero. Figure 2.12 shows the shrinkage of the coefficients towards zero to eliminate the insignificant variables present in the two models developed in this study (i.e., Miami Beach and Birmingham models).



(a) (b)

Figure 2. 12 Coefficient Shrinkage for (a) Miami Beach Model; and (b) Birmingham Model.

Figure 2.13 represents selection of the optimum value of λ using cross validation. The optimum λ value for the Miami Beach and Birmingham models was 0.009294 and 0.00088 respectively. The graphs indicate that the unregularized models are a good fit. The best mean squared error (MSE) values were 13.75% and 0% for the Miami Beach and Birmingham models, respectively.

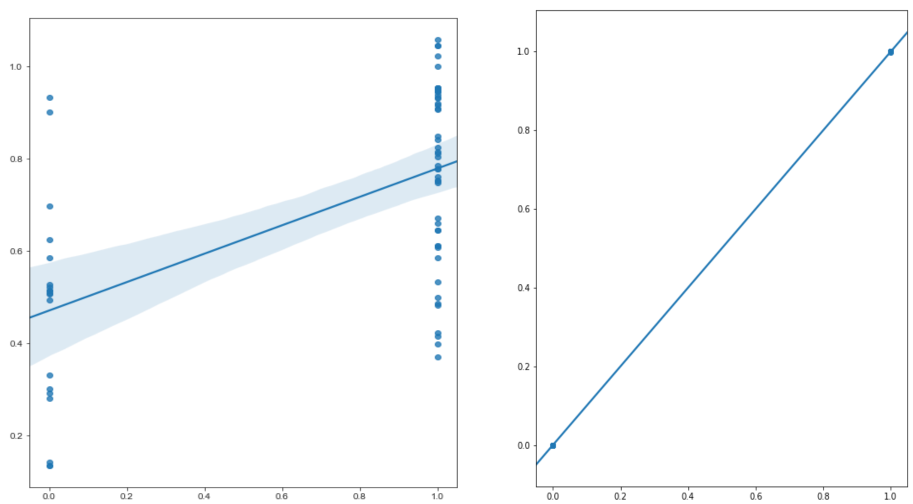


(a) (b)

Figure 2. 13 Optimum Tuning Parameter for (a) Miami Beach Model; and (b) Birmingham Model

The application of the Lasso procedure led to the elimination of 123 variables and 101 variables from Miami Beach and Birmingham survey respectively. This resulted in the selection of 32 variables for the Miami Beach model and 2 variables for the Birmingham model.

Figure 2.14 showcases the results from the regression analysis by fitting the model with significant independent variables identified by the Lasso method. For the Miami Beach model, notable features such as car availability, vehicle ownership, availability of Uber/Lyft, age range (18-29), high income (>$75K), Miami residency etc., possess positive significance for using TNC services in the Miami Beach region. Additionally, features such as lack of prior use of public transit, lack of use of TNC services and interest in future expansion of sidewalks possess a negative significance with respect to using TNC services in the Miami Beach area. In the Birmingham region, only the car availability and waiting time were found to be significant predictors to choosing TNC as a mode of transportation. A list of all significant variables is provided in the Appendix.



(a) (b)

Figure 2. 14: Regression Model-Fit for (a) Miami Beach Model; and b) Birmingham Model.

2.4 CONCLUSIONS

The analysis of 790 questionnaire surveys of transportation system users in the Birmingham and Miami Beach regions shed light on users' awareness and use of TNC services in the Southeastern US. Examination of the survey participants' responses confirmed that TNC service coverage and other geographic considerations and user characteristics impact user perceptions and adoption of such services. Thus, local transportation users' surveys are an important tool to document travelers' preferences and guide planning of TNC services accordingly.

The study also confirmed that, even in small- and medium-size urban areas, transportation users are aware of ride-hailing services and are taking advantage of them. It was also found that the trip purpose for using TNC services varied according to the composition of the survey participants (i.e., local residents versus visitors/tourists). Residents used TNC trips more for trips to home or work while visitors chose ride-hailing trips mostly to access entertainment establishments including restaurants and bars.

An interesting difference between the findings from the two surveys was the relationship between vehicle ownership and TNC use. In the highly automobile-dependent Birmingham market, TNC use was more closely linked to need (i.e., lack of vehicle availability) than choice, while the opposite was the case in the Miami Beach case study.

The findings of the surveys also helped us to define the profile of the typical TNC user in the study regions. In the Birmingham metro area, the typical TNC users are 25-34 years of age that use the ride-hailing services for commuting or entertainment purposes for short to medium range distances (or average of 5 miles). The typical profile of a Miami Beach TNC user is that of younger traveler (18-29 years of age) that uses the ride-hailing service primarily for entertainment purposes, especially during weekends in order to get to the tourist-attraction locations.

Finally, the model fitting exercise identified predictors for TNC use. For the Birmingham region, TNC use was strongly correlated with vehicle availability and waiting time. For the Miami Beach region, several independent variables were identified as significant predictors of TNC use i including vehicle availability, vehicle ownership, availability of Uber/Lyft, age range (18-29), high income (>$75K), and residency.

Overall, the study findings established valuable links between travel behaviors and TNC use. These can inform transportation agencies about the needs and opportunities for ride-hailing services in the local markets. The findings can also be used to create targeted marketing plans and incentives to encourage mode switching to shared modes, including TNCs. Overall, the study highlights the importance of understanding the user characteristics of the local market when planning for TNC and other ride-sharing services in the future.

2.5 REFERENCES

1. Hartmans, A., & Leskin, P. (2018). The History of how Uber went from the Most Feared Startup in the World to its Massive IPO. Business Insider. Retrieved from https://www.businessinsider.com/ubers-history.
2. National Research Council Committee for Review of Innovative Urban Mobility Services (2015). Between Public and Private Mobility Examining the Rise of Technology-Enabled Transportation Services. Transportation Research Board, Washington, DC.
3. Sisiopiku, V.P., Hadi, M., McDonald, N., Steiner, R., and Ramadan, O.E. (2019). Technology Influence on Travel Demand and Behaviors (STRIDE B), Final Report to the Southeastern Transportation Research, Innovation, Development and Education Center (STRIDE).
4. Shaheen, S. (2018). Shared mobility: The potential of ride hailing and pooling. In D. Sperling (Ed.), Three Revolutions: Steering Automated, Shared, and Electric Vehicles to a Better Future (2nd Ed.), (pp. 55-76). Washington D.C.: Island Press. Retrieved from https://escholarship.org/uc/item/46p6n2sk.

5. Clewlow, R. R., & Mishra, G. S. (2017). Disruptive Transportation: The Adoption, Utilization, and Impacts of Ride-Hailing in the United States. Institute of Transportation Studies, University of California, Davis.
6. Vinayak, P. F., Dias, F. F., Astroza, S. F., Pendyala, R. F., & Garikapati, V. M. (2018). Accounting for Multidimensional Dependents among Decision-Makers within a Generalized Model Framework: An Application to Understanding Shared Mobility Service Usage Levels. Transport Policy, 77(2018), 129–137. Retrieved from https://doi.org/10.1016/j.tranpol.2018.09.013.
7. Chen, Z. (2015). Impact of Ride-Sourcing Services on Travel Habits and Transportation Planning. Retrieved from http://d-scholarship.pitt.edu/25827/.
8. Circella, G., Alemi, F., Berliner, R., Tiedeman, K., Lee, Y., Fulton, L., Handy, S. & Mokhtarian, P. L. (2017). The Multimodal Behavior of Millennials: Exploring Differences in Travel Choices between 1 Young Adults and Gen Xers in California 2 (University of California, Davis Institute of Transportation Studies Research Report UCD_ITS_RR_17-54) Retrieved from https://itspubs.ucdavis.edu/wp-content/themes/ucdavis/pubs/download_pdf.php?id=2872.
9. Freund, K., Bayne, A., Beck, L., Siegfried, A., Warren, J., Nadel, T., & Natarajan, A. (2020). Characteristics of Ride Share Services for Older Adults in the United States. Journal of Safety Research, 72(2020), 9–19. Retrieved from https://doi.org/10.1016/j.jsr.2019.12.008.
10. Rayle L, Shaheen S, Chan N, Dai D, Cervero R (2014) App-based, on-demand ride services: Comparing taxi and ridesourcing trips and user characteristics in San Francisco University of California Transportation Center (UCTC). Technical report, UCTC-FR-2014-08.
11. Smith, A. (2016). 2. On-demand: Ride-hailing apps. Chapter 2 in Shared, Collaborative and On Demand. Pew Research Center: Internet, Science & Tech. Retrieved from https://www.pewinternet.org/2016/05/19/on-demand-ride-hailing-apps/.
12. 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, Available at https://www.scirp.org/pdf/jtts_2020061810033242.pdf
13. Institute of Transportation Engineers (ITE). (2011). Manual of Transportation Engineering Studies 2nd ed., Vol. 12.
14. Sisiopiku, V. (2018). “Travel Patterns and Preferences of Urban University Students”. Athens Journal of Technology & Engineering, Vol. 5(1), March 2018, pp. 19-31.
15. Sisiopiku, V.P., and Ramadan, O.M. (2018). “Understanding Women’s Needs as Determinants of Mode Choice: A Case Study of the University of Alabama at Birmingham”, In Annual Meeting Online Proceedings of the TRB 97th Annual Meeting, Washington, D.C.
16. Morshed, S. A., Arafat, M., Ashraf Ahmed, M., & Saha, R. (2020, August). Discovering the Commuters’ Assessments on Disaster Resilience of Transportation Infrastructure. In International Conference on Transportation and Development 2020 (pp. 23-34). Reston, VA: American Society of Civil Engineers.

17. Tibshirani, R. (1996). Regression shrinkage and selection via the Lasso. Journal of the Royal Statistical Society, 58(1), 267 – 28.
18. Hastie, T., Robert, T., & Wainwright, M. (2015). Statistical Learning with Sparsity: The Lasso and Generalizations.
19. Steele, A. J., Denaxas, S. C., Shah, A. D., Hemingway, H., & Luscombe, N. M. (2018b, 08). Machine Learning Models in Electronic Health Records can Outperform Conventional Survival Models for Predicting Patient Mortality in Coronary Artery Disease. PLOS ONE, 13, 1-20.

CHAPTER 3: CASE STUDY 2 - CALIBRATION OF A MULTI-AGENT BASED SIMULATION MODEL FOR THE MIAMI BEACH NETWORK

3.1 INTRODUCTION

Modeling the impact of transportation network attributes on traveler's choice behaviors is critical in order to understand the impacts of changes in traffic and transit systems and various confounding factors. The reliability of travel modes, travel time, and fares are important forecasters for mode choice and route choice among travelers. For example, transit ridership was found to be more sensitive to changes in travel time than to the cost of transit fares, with wait time being considered as much more "costly" than in-vehicle time (Frank et al., 2008). Furthermore, travelers' socioeconomic and demographic attributes such as income, education, car availability, age, gender, and education are also important for consideration in mode choices forecasting (Bhat, 1997). Among several characteristics of traveler related variables, income has been the most widely used in such models. Travelers of higher income are generally assumed to choose an alternative mode that provides fast and convenient service, despite it being more expensive (Koppelman, 1989). Additionally, the purpose of trip has significant impacts on mode choice (Morrison & Winston, 1985).

Existing travel demand forecasting software uses static demand in which a fixed set of trips or origin-destination demand is assumed. These models also use simplistic traffic flow models to calculate travel times as part of the process, which are generally based on equations that relate the travel time to the volume over capacity ratio. The forecasting models use static assignment to identify routes and discrete mode choice models to predict the travel behavior of individuals or aggregated groups as they choose a particular mode, route, or destination based on travel characteristics, such as travel time, cost, personal preferences, and socio-demographic factors (Hörl et al., 2019).

The development of agent-based transport simulation considers feedback loops between supply and demand in a more detailed manner, compared to conventional demand forecasting models. These agent-based models account for the dynamic change in demand during the peak period. Using a simulation model to estimate the travel times and other system measures instead of using simple macroscopic equations allows for better estimation of the impact of system performance on travel choices. Researchers have explored the coupling of multi-agent transportation system simulations with activity-based demand generation models and discrete choice models (Auld et al., 2016, Heilig et al., 2018, Reiser et al., 2007). Among the various types of open-sourced multi-agent simulation tools available, the Multi Agent Transport Simulation (MATSim) has been one of the most widely used (Axhausen et al., 2016).

MATSim is an activity-based multi-agent simulation framework implemented in Java. In MATSim, individual travelers (agents) repeatedly optimize their daily activity schedule while in competition for space-time slots with all other agents present within the network. Every agent in the simulation learns, plans, and adapts to the system within an iterative loop. This process is somewhat similar to the route assignment used in other tools but goes beyond route assignment by incorporating other choice dimensions like mode, time, and even destination choice (Balmer et al., 2005b, Grether et al., 2009, Horni et al., 2012).

MATSim is designed to model large-scale scenarios utilizing computationally efficient queue-based approach for modeling network congestion. A MATSim run has a configurable number of iterations, represented by a feedback loop, which is composed of three stages, namely simulation, scoring and re-planning (see Figure 3.1). The *simulation stage* represents the traffic performance estimation of MATSim that provides the travel times in the network. Starting at the beginning of the day, all events coming from each agents' daily plans are processed in order (e.g., leaving the home facility, walking to the bus stop, waiting for the next departure, and so forth). Vehicles traveling on the roads join waiting queues in their paths allowing interferences between the vehicles resulting in congestion. The *scoring stage* considers delays (negative score) and performed activities (positive score). The overall score reveals how well a plan is perceived by the agent. After the scoring, a trip plan undergoes *re-planning* which considers multiple strategies such as rerouting, mode selection, trip departure time, etc. These strategies are chosen based on configurable weights during what is referred to as the *innovation phase*. MATSim allows agents to have multiple plans, which can be configured through a score-based plan removal strategy. The output of this innovation phase is a selected plan and set for further simulation and subsequent scoring in an iterative loop.

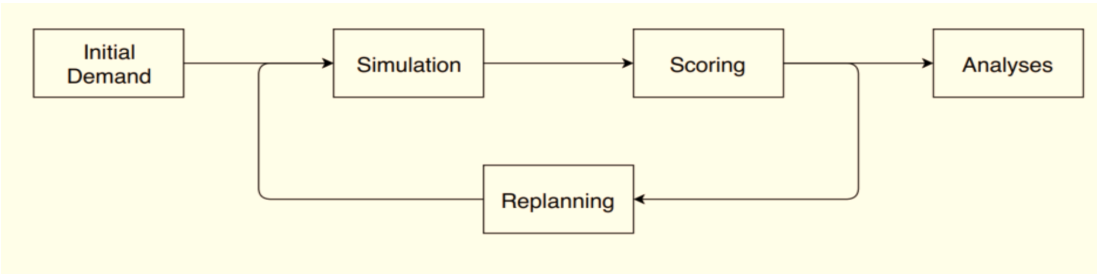


Figure 3. 1: MATSim Framework

Time dependent Origin-Destination (O-D) matrices are typically used in conventional demand forecasting models (Balmer et al., 2005a). O-D data can be used directly as an input to MATSim (Kichkhofer et al., 2016). However, there are limitations with using O-D data as an input for MATSim since activity schedules are not reflected with O-D matrices. Additionally, the use of O-D matrices does not allow detailed consideration of routing adaptation and time adjustments due to time-dependent traffic conditions and toll charges and the impacts of agents' attributes (income, short tour penalties, etc.). Besides, mode choice is also dependent on specific details of individual trips such as walking distance to the public transit stop and unavailability of public transit. These specific details are difficult to account for when performing aggregated analysis using O-D matrices. To tackle this issue, existing activity-based demand models can be directly transformed into agents' plans and used as inputs into MATSim (Charypar et al., 2005).

Unlike conventional travel demand models, MATSim does not use discrete choice models. Instead, the mode choice in MATSim is based on the maximized utility for all agents, which considers a standard scoring function. The default scoring function is the sum of the activity utility in terms of how well is the utility performance and travel disutility which considers factors such as travel time, waiting time, and monetary costs. Each individual agent considers multiple plans, and the selection is based in a manner that maximizes the utility for all agents

This chapter demonstrates the calibration of MATSim and the utilization of the calibrated model to examine the shift in mode choice due to alternative strategies that are designed to motivate such shift.

3.1.1 OBJECTIVE AND SCOPE

The purpose of this work is to demonstrate methods to calibrate a multi-agent simulation model (MATSim) to allow it to estimate the mode split between the passenger car and transit services. The calibration of the multi-agent simulation is based on the regionally approved mode split behavioral model and real-world traffic counts obtained for the network.

3.2 STUDY AREA/ CASE STUDY

The investigation in this study uses the Miami Beach network as a case study. Miami Beach is located in Miami-Dade County in Florida between the Atlantic Ocean and Biscayne Bay. Rapid population growth as well as employment increase in the Miami Beach central business district (CBD) is projected to continue over the next 20 years. Moreover, appealing location attributes such as climate, landscape, entertainment venues, and beaches attract many tourists in the Miami Beach area. Traffic is served by a network of arterials, collectors, and local streets. Miami Beach has two major north-south arterial roadways, namely Collins Avenue and Alton Road. Other major arterials include four east-west roadways within the city, which are the continuation of the four causeways that connect the city to the mainland. The four major causeways are McArthur, John F., Kane Concourse, and Venetian Way. The rest of the major roadways within Miami Beach are collector roads, which form a grid in the South Beach area, with Washington Avenue providing the majority of the north-south connectivity establishing large commercial activity around it. The Metrobus and Trolley comprise the majority of public transportation options in the city of Miami Beach. The Miami Beach network has seven Metrobus routes along with six Trolleys.

3.3 METHODOLOGY

The description of the methodology in this section is organized into four subsections: 1) data preparation, 2) multi-agent simulation, 3) model calibration and 4) scenario development.

3.3.1 DATA PREPARATION

To simulate the scenario of the study area in MATSim, several input files were prepared based on the regional demand model, which is the Southeast Florida Regional Planning Model (SERPM 7.0) and real-world data. The road network of the selected study area was extracted from the SERPM 7.0 model using Open Street Map (OSM) and was integrated into a MATSim layer by using the Java Open Street Map Editor (JOSM) MATSim extension. Links with starting and ending nodes, length, speed, capacity, number of lanes and available modes were integrated into the MATSim as input files. The coded MATSim network consisted of 280 nodes and 698 links.

The use of MATSim requires the generation of full day plans of all agents representing 100% population using the study network. This generation is done in this study based on the Activity Based Model (ABM) of Southeast Florida Regional Planning Model Version 7.0 (SERPM 7.0) This is an activity-based travel demand forecasting disaggregated model, which relies on travel and decision-making behaviors at the household and individual levels. The activity-chain of the decision maker is a sequence of activities with the trips in a whole day (Wang et al., 2001, Fu et al., 2009). The activity-based model consists of information about different household, family, and housing types including detailed household composition, which influences the activity-travel patterns. In SERPM 7.0, the decision-making units for the ABM are the person and the household. The decision-making units are synthesized for each simulation year based on tables of households and persons from the 2005-2009 American Community Survey (ACS) data and forecasted TAZ-level distributions of households and persons by key socio-economic categories (Southeast Regional Planning Model 7.0 Activity-Based Model Users Guide, 2014).

The travel demand and activity patterns of the agents in MATSim for Miami Beach area were collected from SERPM 7.0 from 48 traffic analysis zones (TAZ). A total of 59,474 individuals (agents) with a total of 122,962 trips were modeled. The study then integrated the available facilities, trip purposes, and tip durations in MATSim to develop activity-schedule based on location coordinates and the data collected from the ABM in SERPM 7.0.

The developed Miami Beach MATSim model included three types of public transit modes including the existing regional bus service (Metrobus) and local Trolley service, and a light rail (Metrorail) option. The Metrorail option is specified as a potential future alternative for connecting the City of Miami Beach to Miami-Dade County mainline. The study collected the public transit information such as route profile, stops location, departure time, service frequency, etc. from the Miami-Dade Transit (MDT) SMART Plan (the plan that investigated the improvement of transit services in the region), and the City of Miami Beach website. The route profile describes the stops that the route serves, while the route itself describes the series of links in the network in which the public transit driver has to drive along. MATSim recognize the routes that every transit vehicle has to follow in the network through referenced links.

Table 3.1 summarizes the data source for the input files that were integrated into the MATSim platform as an open scenario. The supply side of the MATSim model is composed of the network and facilities file whereas the demand side is set by preparing the population and activity plans files for the study area based on the SERPM 7.0 model.

Table 3. 1: Data Source for the Input Files

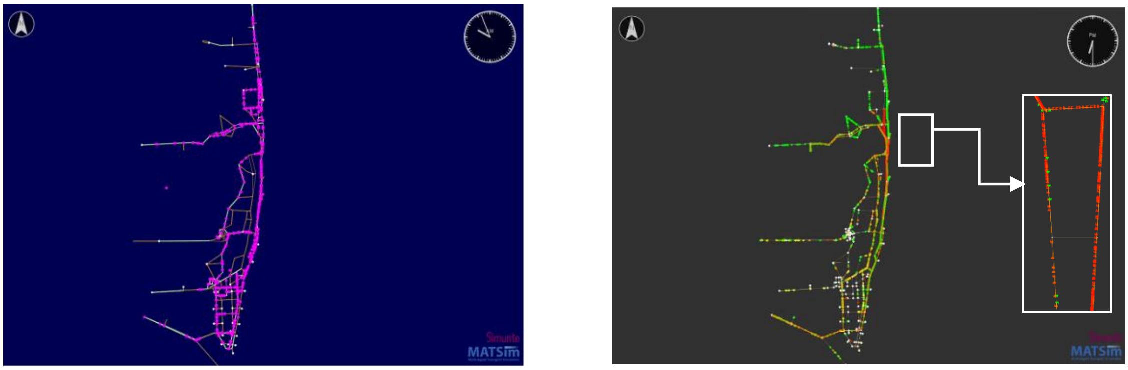
Input	Properties	Source
Network	Links, nodes, location, coordinates, capacity, speed limit	SERPM 7.0
Plans	Trip time, trip purpose, number of agents, available modes, activity start and end time.	SERPM 7.0
Facilities	Facilities related to activity type and schedule	SERPM 7.0
Transit-Metrobus	Details of routes, stops and schedule	Miami-DadeTransit (MDT)
Transit- Trolley	Details of routes, stops and schedule	City of Miami Beach
Transit-Metrorail	Details of routes, stops and schedule	SMART Plan

3.3.2 MULTI AGENT SIMULATION

The parameters to run the MATSim model are set using a *configuration file*. The information in the configuration file controls actions such as the coordinate setting, routing algorithm, and innovative strategy specifications. MATSim uses an algorithm that compares the alternative plans of the agents that are executed in the simulation environment based on the scoring of utility functions. The utility functions are characterized by behavioral parameters and attributes of each alternative plan of the agent. The innovative strategies (re-planning strategies) used in this study are 'Re-Route', 'Best Score' and 'Change Trip Mode'. The re-planning strategies are used to form new plans and alter them based on the conditions. The '*Re-route*' strategy calculates a completely new route for the agent. Consequently, the new plans are added to the inventory of choice-sets of each agent. Poorly scored plans are removed from the agent's choice sets. '*Best Score*' selects the plan with the highest score based on the scores from the previous iteration. '*Change Trip Mode*' randomly picks one of a person's plans and changes the mode of transport. By default, the supported modes are driving a car and using public transport. Only one mode of transport per plan is supported.

In this study, while performing the mode choice, the agents were only allowed to switch between the transport modes private cars (CAR) and Public Transit (PT). The parameter values of the *Fraction of Iteration to Disable Innovation*, *Module Probability for Best Score*, and *Re-routing Proportion* were set as 80%, 70% and 30% respectively. The 80% iteration to disable innovation indicates that for the final 20% of the iterations, the choice set innovation was switched off and all agents only change between plans that exist in their choice set. The values of the Module Probability for Best Score and Re-routing Proportion indicates that 70% of the agents are included in the best scoring and 30% of the agents are rerouted.

Figure 3.2 (a) presents a visualization of the spatial distribution of the transit stops across the entire Miami Beach network. Figure 3.2 (b) illustrates the agents' activities in the entire simulated area for a full day's plan, and the movement of cars and public transit vehicles at 7:30 PM. The red triangles in the network indicate cars in traffic jam, whereas the green triangles show cars in free flow.



(a) **(b)**

Figure 3. 2: (a) Transit Stops and (b) Simulation of the Miami Beach area Network

The *utility function for traveling activity* ($S_{trav,q}$) considers several behavioral parameters and attributes of the alternative plans. Behavioral parameters that are required to be identified and calibrated for use when calculating the utility function are the following:

- *marginal utility of performance* (a pre-factor of car travel time in the mode choice logit model; performing an activity for more time normally increases utility),
- *marginal utility of travelling* (when agents travel by car, they cannot perform an activity; thus, they are marginally and approximately losing utility of performance),
- *marginal utility of money* (conversion of travel cost such as toll or distance cost into utilities; it is the pre-factor of the monetary term in the mode choice logit model),
- *monetary distance rate* (monetary costs increase depending on distance),
- *alternative-specific constant* (the disutility for travelling to activity; for mode choice logit model).

The *travel scoring function* consisting of the marginal utility (and disutility) parameters and monetary distance rate were estimated in this study based on the parameters used in the SERPM 7.0 demand model. These behavioral parameters were fine-tuned to produce traffic flows that are close to those observed based on real-world data. The resulting parameter values are shown in Table 3.2, which are set up during the simulation to match the real-world traffic count data in the network.

Table 3. 2: The Utilized Values of the Behavioral Parameters in MATSim

Parameter	Definition	Default Values in MATSim	Calibrated Value
Marginal utility of money	Rate of increase of utility due to expenditure of money per unit increase of value	1.0 utils/ monetary units	0.03 utils/ $
Marginal utility of performing	Utility gained by activity performance by an agent	6.0 utils/ hour	1.5 utils/ hour
Marginal utility of travelling by car	Utility gained by traveling via mode (car) by an agent	0.0 utils/ hour	0.0 utils/ hour
Marginal utility of travelling by PT	Utility gained by traveling via mode (Pt) by an agent	0.0 utils/ hour	0.0 utils/ hour
Monetary distance rate, CAR	Cost consumed by travelling a distance by car	0.0 money/ m	-0.000125 $/ m
Monetary distance rate, PT	Cost consumed by travelling a distance by Pt	0.0 money/ m	-0.00125 $/ m
Mode Constant, CAR	Default mode-specific constant for selecting car as a mode	0.0	0.01
Mode Constant, PT	Default mode-specific constant for selecting PT as a mode	0.0	-0.883

3.3.3 TRAFFIC COUNT BASED CALIBRATION

MATSim identifies and utilizes full daily plans for the travelers in the network. Initial full day plans for each individual agent within the network can be updated by the MATSim based on real-world traffic counts. There are 50 traffic count stations available in the selected study area. However, 13 count stations do not have data available. Thus, data from the remaining 37 count stations in the Miami Beach network were collected from the Florida Traffic Online – Florida Department of Transportation (FDOT) and used in the study. The count data used from the

Florida Traffic Online website were cross-checked with counts obtained from the demand forecasting SERPM 7.0 model to identify any discrepancies. The study used the MATSim count volumes in hourly time bins as compared to real-world traffic counts to assess the performance of the calibration of the simulation. The Mean Percentage Error (MPE), calculated by Equation 3.1 was used to assess the performance of the traffic flow calibration:

$$\boldsymbol{MPE} = \frac{||\boldsymbol{F_{simulated}} - \boldsymbol{F_{real}}||}{\boldsymbol{F_{real}}} \times \mathbf{100\%} \quad \text{...(3.1)}$$

where,

$F_{simulated}$ = simulated car traffic flow, and

F_{real} = real time car traffic flow.

Figure 3.3 shows the comparison of the volume counts between the real and simulated traffic counts for the six major entry and exit links of the network. It is evident that the daily volume obtained from the MATSim simulation matches very well the real time traffic volume in the specified entry/exit links. Additional comparison of the counts for individual links are presented in the results section, later in this document.

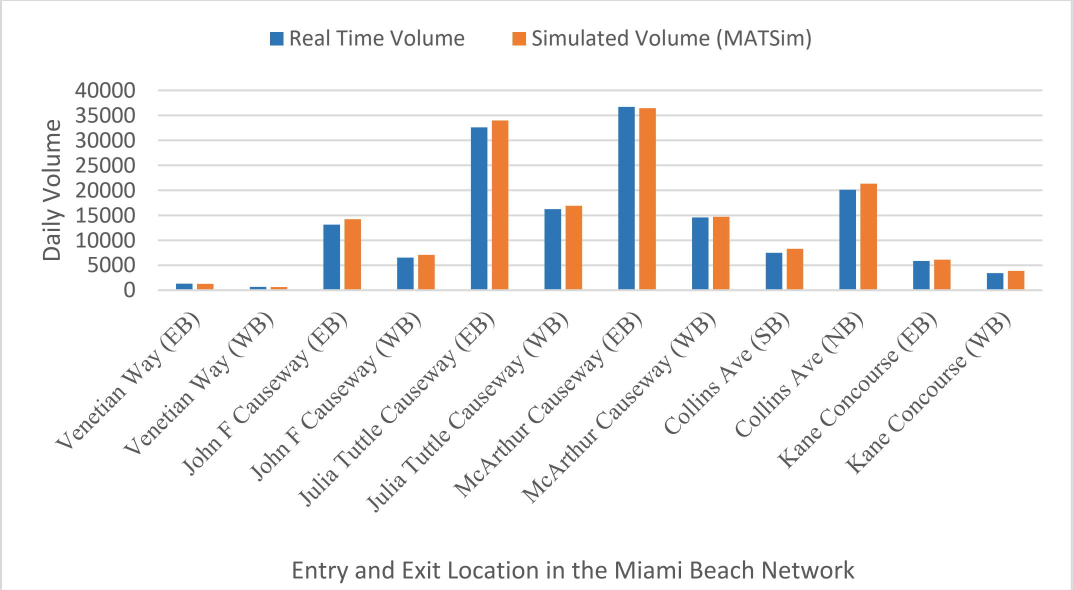


Figure 3. 3: Comparison of the Entry and Exit Location in the Miami Beach Network

3.3.3.1 Mode-Split Calibration using Non-Linear Optimization

The next step was to calibrate the modal split parameters in MATSim. The modal split calibration was treated as an optimization problem that aimed to find the best combination of the probability coefficients that result in the closest match between the modal split equation in the travel demand forecasting model SERPM 7.0 (Equation 3.2) and the equation used in the multi-agent-based simulation model (Equation 3.3).

The study applied the Nelder-Mead non-linear programming (NLP) optimization to solve the optimization problem. The Nelder-Mead method is a heuristic optimization technique similar to genetic algorithm. The optimization problem refers to finding points with an optimal value i.e., maximum or minimum value of an objective function within the searched domain or space. The Sum Squared Error (SSE) is utilized as the objective function for the optimization to minimize the difference in the results obtained using the two equations.

$$U_{i|j,s} = \sum_{k}(\beta_k * Time_k) + \sum_{l}(\beta_l * Cost_l) + \sum_{m}(\beta_m * Location_m) + \sum_{n}(\beta_n * SE_n) + \delta_{i|j} \quad \text{.......(3.2)}$$

Here, travel time variables, travel cost variables, location-specific variables, and socio-economic variables are denoted by the index k, l, m and n, respectively. δ indicates the mode specific constant.

$$S_{trav,q} = C_{mode(q)} + \beta_{trav,mode(q)} \cdot t_{trav,q} + \beta_m \cdot \Delta m_q + (\beta_{d,mode(q)} + \beta_m \cdot \gamma_{d,mode(q)}) \cdot d_{trav,(q)} + \beta_{transfer} \cdot x_{transfer,q} \quad \text{.......(3.3)}$$

$C_{mode(q)}$ and $\gamma_{d,mode(q)}$ indicate mode specific constant and mode specific monetary distance rate, respectively. $\beta_{trav,mode(q)}$, β_m, and $\beta_{d,mode(q)}$ represent direct marginal utility of time spent traveling by mode, marginal utility of money and marginal utility of distance respectively. Travel time and distance travelled between activity locations are denoted by $t_{trav,q}$ and $d_{trav,q}$ respectively. The change in monetary budget of the trip due to fares, tolls etc. is reflected in Δ_{mq}. Public transport transfer penalties and occurrence of transfer are denoted by $\beta_{transfer}$ and $x_{transfer,q}$ respectively.

The heuristic optimization techniques possess a degree of randomness meaning that the algorithm runs with the same problem and the same initial conditions but could yield different results. This is because heuristic methods explore the search space stochastically and identify the most optimal points during the procedure. A simplistic flowchart illustrating the algorithm is shown in Figure 3.4.

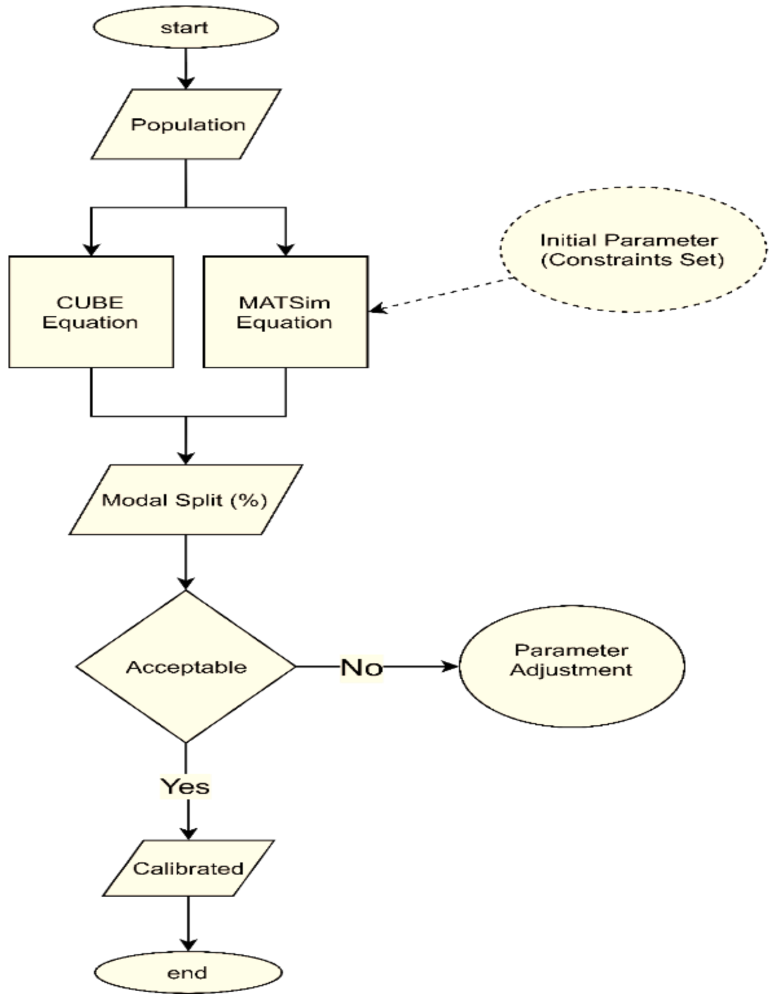


Figure 3. 4: Flowchart of the Non-Linear Optimization of Modal Split based on the SERPM and MATSim Platforms

The comparisons between the two equations are based on three important variables, which are the distance travelled, time spent during the trip, and mode-specific constants. The input parameters are in accordance with the SERPM 7.0 guidelines and MATSim guideline as shown in Table 3.3.

Table 3. 3: Input Parameters for CUBE and MATSim Equation

	Mode	Description	Sub-mode	β_k (Time Coefficient)	k (Average Time)	Mode	$\beta_{trav,mode(q)}$	$T_{trav,q}$ (Time)
	Car	In-Vehicle Time		-0.016	32.6 mins (max)	Car	-6	
	PT	In-Vehicle Time	Urban Rail	-0.0136	32.6 mins (max)	PT	-6	
Time			Bus	-0.0144	32.6 mins (max)			
		First Wait Time		-0.024	1.5 hours (max)	$\beta_{transfer}$	(usually negative)	
		Transfer Wait Time		-0.024	15 min (max)	X_transfer	0/1	
		Walk Access or Egress Time		-0.03	1.9 hours (max)			
	Car		Unit Ratio/ Income Type		(Cost) I ($)	Car	Δmq	βm
		General Cost		-0.038	9.6$ (max = $10)		0 or negative	1
		Auto Operating Cost, Fuel	13.5 c/miles		0.675			
		Auto Operating Cost, Maintenance	6.3 c/miles		0.315			
		Income is incorporated in cost as mentioned in the SERPM Guideline - Average median income in Miami Beach is $43.5k per annum	High Income	-0.0072				
Cost	Car/PT		Medium Income	-0.00118				
			Low Income	-0.00266				
			Very High Income	-0.00032				
	PT					PT	0 or negative	1
		Operating Cost per passenger mile	0.75$/mile		3.75			

	Mode	Description	Sub-mode	β_k (Time Coefficient)	k (Average Time)	Mode	$\beta_{trav,mode(q)}$	$T_{trav,q}$ (Time)
	Car	In-Vehicle Time		-0.016	32.6 mins (max)	Car	-6	
	PT	In-Vehicle Time	Urban Rail	-0.0136	32.6 mins (max)	PT	-6	
Time			Bus	-0.0144	32.6 mins (max)			
		First Wait Time		-0.024	1.5 hours (max)	$\beta_{transfer}$	(usually negative)	
		Transfer Wait Time		-0.024	15 min (max)	X_transfer	0/1	
		Walk Access or Egress Time		-0.03	1.9 hours (max)			
		Total Cost per passenger mile	0.98$/mile		4.9			
							$\beta_{d,car,pt}$ (mode)	$d_{trave,q}$
			Unit		Distance	Car		4.5-5.0 (miles)
Distance	Car				4.5-5.0 (miles)	PT		4.5-5.0 (miles)
	PT				4.5-5.0 (miles)			
		βm (location coefficient)	Location					
Location	Car	0	0					
	PT	0	0					
		Income Average Median income in Miami Beach is $42.5K	High Income	-0.0072				
	Car/PT		Medium Income	-0.00118				
			Low Income	-0.00266				
Socio-Economic Demographic			Very High Income	-0.00032				
		Gender		βn (Socio-economic demographic coefficient				
		Male (52.5%)						
		Female (47.5%)		0.1578				
		Age (For Transit)		βn				

	Mode	Description	Sub-mode	β_k (Time Coefficient)	k (Average Time)	Mode	$\beta_{trav,mode(q)}$	$T_{trav,q}$ (Time)
	Car	In-Vehicle Time		-0.016	32.6 mins (max)	Car	-6	
	PT	In-Vehicle Time	Urban Rail	-0.0136	32.6 mins (max)	PT	-6	
Time			Bus	-0.0144	32.6 mins (max)			
		First Wait Time		-0.024	1.5 hours (max)	$\beta_{transfer}$	(usually negative)	
		Transfer Wait Time		-0.024	15 min (max)	X_transfer	0/1	
		Walk Access or Egress Time		-0.03	1.9 hours (max)			
		16-24 (28.3%)		-0.7947				
		41-55 (41.75%)		-0.423				
		56-64 (18.35%)		-0.4499				
		65 and older (11.6%)		-1.1231				
Mode Specific Constant		δ (Mode-specific constant)						
	Car	1				Car	0	
	PT	-1.708				PT	0	

Table 3.4 shows the parameters of the activity used in the optimization including the mean, standard deviation, and range of the trip length, trip time, transit wait time, and transit ingress and egress times. The parameter values are considered assuming a normal distribution. The optimization uses a synthesized population (100,000 agents) performing a typical home-work activity. Also, constants such as transit fare, mode specific constant, and short trip penalty were set based on the activity-based model.

Table 3. 4: The Parameters of the Home-Work Activity Used in the Optimization

Parameter	Trip Length (Miles)	Trip Time (Minute)	Wait Time (Minutes)	Walk Time (Minutes)
Mean	2.75	15	3	3
Standard Deviation	0.383	1.7	0.34	0.34
Range	1.5-4	10-20	2-4	2-4

3.3.4 Multi-scenario Analysis via MATSim

Three types of scenarios were developed in the MATSim model to examine the ability of the developed method to analyze the change in modal split. The *base scenario* was comprised of the

Miami Beach network replicated from SERPM 7.0 with the inclusion of agents, cars, and public transit as stated above. The *second scenario* was built by adding a Metrorail service to connect the City of Miami Beach to Miami-Dade County mainline utilizing the McArthur Causeway. The *final scenario* considered the implementation of road pricing along McArthur Causeway, a major roadway that is parallel to the Metrorail route. The toll imposed for using the McArthur causeway was set to 50 cents. This scenario was used to observe if the agents would shift to Metrorail due to the implementation of road pricing. A change in modal split in the MATSim output would imply the dynamic nature of MATSim in multiscenario-based analysis in a multi-agent simulation.

3.4 RESULTS

Figure 3.5 shows that the MPE of the simulated traffic flow of all links (both external and internal) relative to real-world traffic flow is significantly lower after adjusting the MATSim model parameters compared to those simulated without adjustment. This demonstrates the importance and value of calibration of model parameters in simulation models in general, and MATSim in particular.

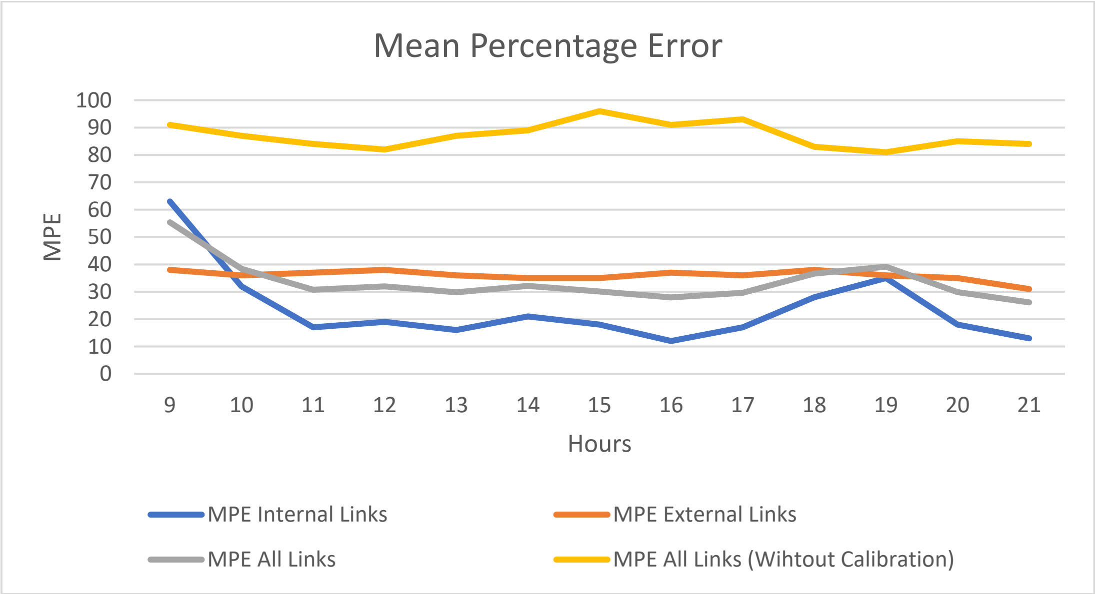


Figure 3. 5 Mean Percentage Error of Link Traffic Flow with and without Calibration

Figures 3.6 through 3.11 display comparisons of hourly traffic counts (real versus simulated) between 9:00 AM and 9:00 PM for three major external and three major internal links in the study network. The simulated hourly traffic counts were generated from the calibrated baseline MATSim model. A good match between the real and simulated count values for the major external and internal links can be observed. For external links, it appears that more traffic is using McArthur Causeway and Julia Turtle Causeway in the simulation compared to the real-world (see Figure 3.6 and 3.8). On the contrary, more vehicles are using John F Causeway in the real-world (see Figure 3.7). Examining Figures 3.6 to 3.8 indicates that further calibration may be needed to shift some of the traffic on John F. Causeway to the other two causeways through route choice and possibly reducing the overall traffic demands entering the network through the three causeways. When examining the internal link results in Figures 3.9 to 3.11, it appears that more traffic is using Washington Avenue in the simulation and, in some of the time intervals, more traffic is using Alton Road and Collins Avenue in the simulation. The results indicate that the model may benefit from further calibration to shift traffic away from Washington Avenue to the other two parallel arterials (Route choice). The reduction of traffic on the causeways mentioned above may also reduce the volume on Washington Avenue producing closer results to real-world conditions.

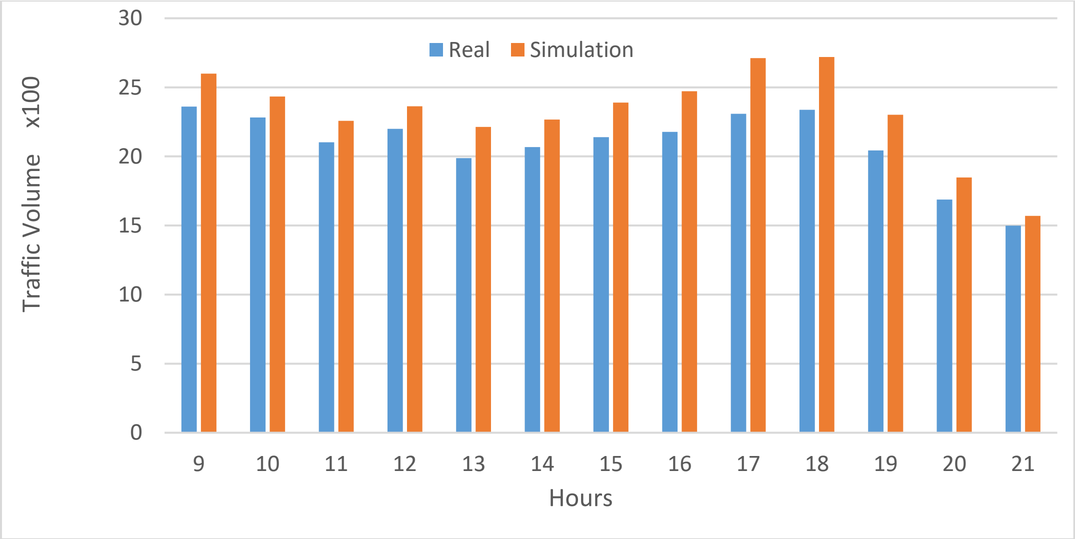


Figure 3. 6 Real versus Simulated Traffic Volume - Link 85 McArthur Causeway (External Link)

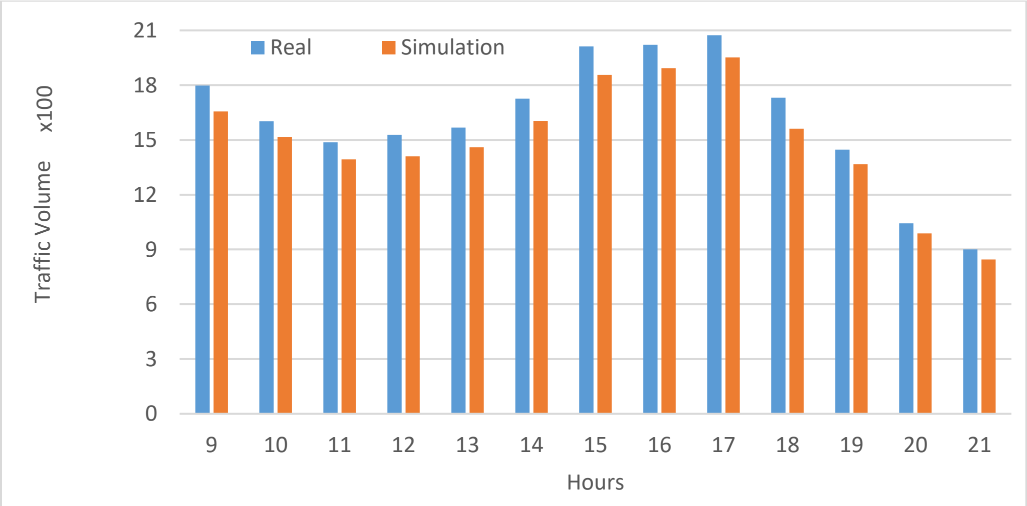


Figure 3. 7 Real versus Simulated Traffic Volume - Link 344 John F Causeway (External Link)

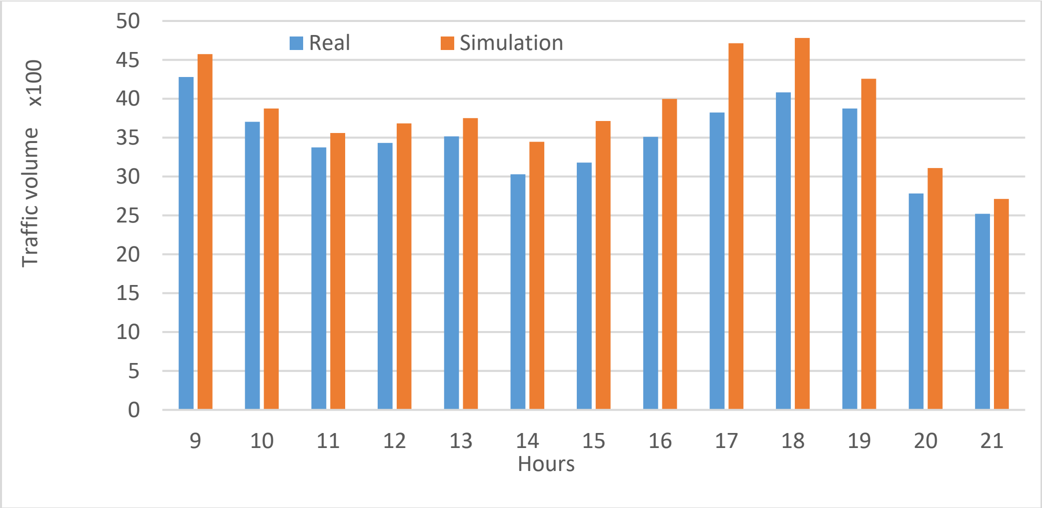


Figure 3. 8 Real versus Simulated Traffic Volume - Link 92 Julia Tuttle Causeway (External Link)

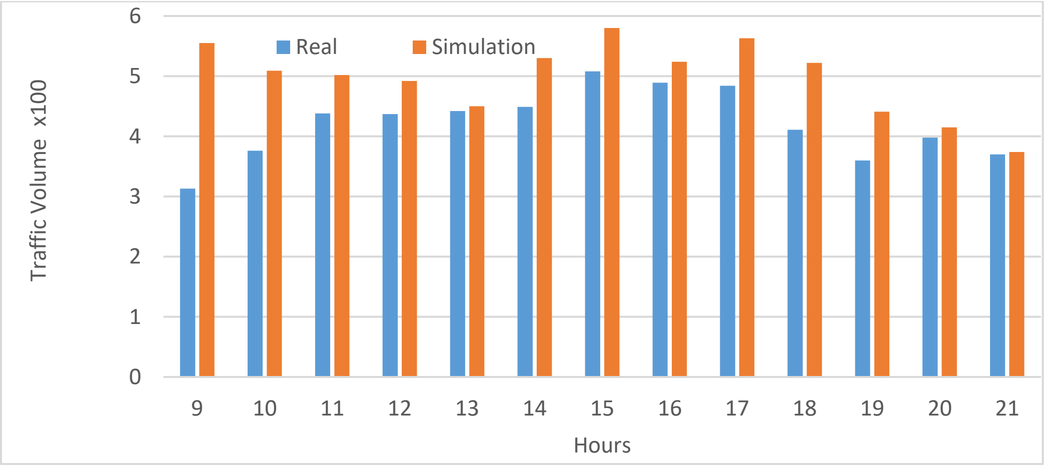


Figure 3. 9 Real versus Simulated Traffic Volume - Link 434 Washington Ave (Internal Link)

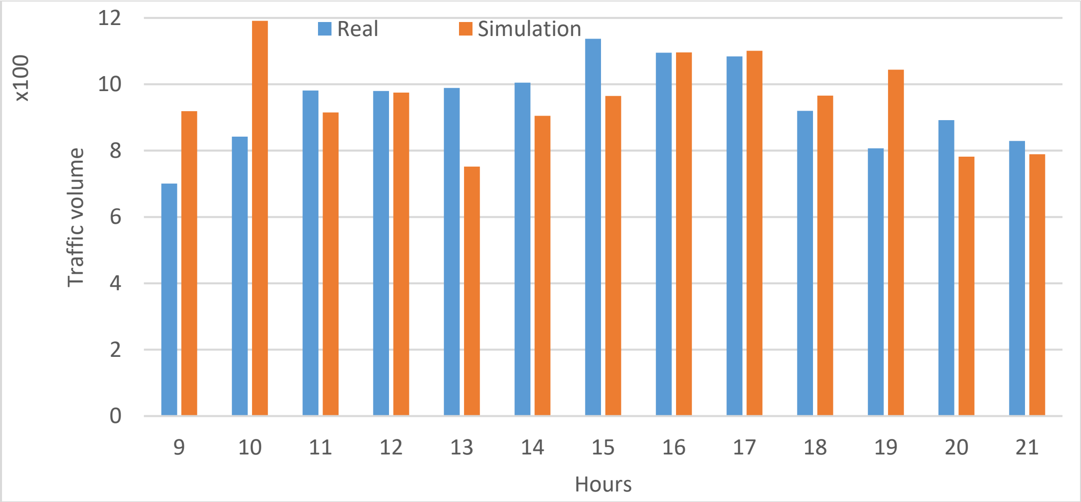


Figure 3. 10 Real versus Simulated Traffic Volume - Link 133 Alton Road (Internal Link)

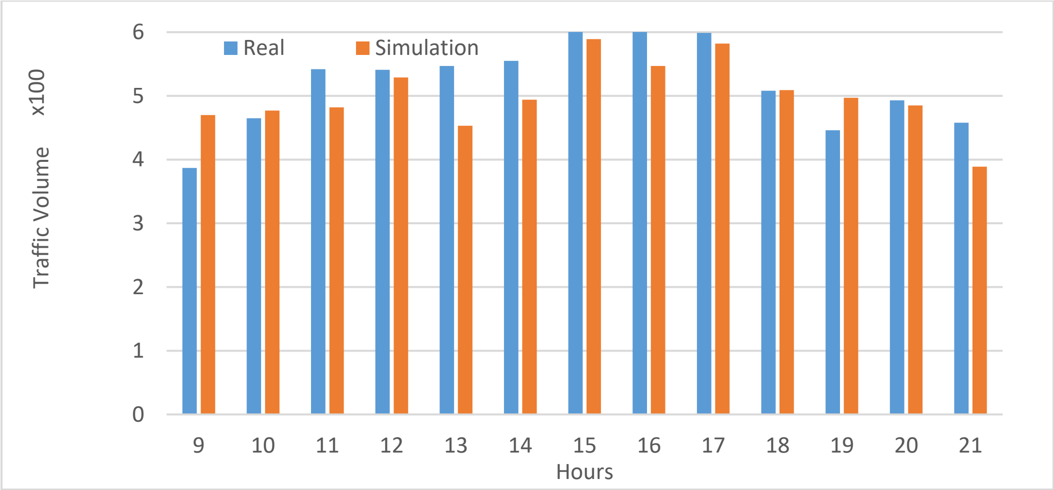


Figure 3. 11 Real versus Simulated Traffic Volume - Link 191 Collins Ave (Internal Link)

For the baseline conditions, the mode split using the regional travel demand model was estimated to be 97.9% CAR and 2.1% PT. In the mode split estimation, each individual agent selects the mode for the activity based on the calculated probability that is related to a utility function in SERPM 7.0 and based on a calculated score in MATSim. As described earlier, this case study utilized a nonlinear optimization process to calibrate the parameters of the mode split equation in MATSim to match the proportion of passenger car to public transit within the network when using the SERPM 7.0 mode split equation. The results from the optimization are values for the mode specific coefficient C_{mode}, cost coefficients $\beta_{d,mode}$, β_m and time coefficients $\beta_{trav,mode}$. The values identified based the optimization were implemented in MATSim. Figure 3.12 shows the mode split with the optimized parameters compared with those obtained using the default parameters in MATSim. With the optimized parameters, MATSim estimates that the transit percentage is 2.07%, which is close to the value estimated by the demand forecasting model and the value estimated based on the survey conducted as part of this study and documented in Chapter 2. Without calibration, MATSim estimated the transit percentage to be 32%, which would be unrealistic given actual travel patterns in the Miami Beach area.

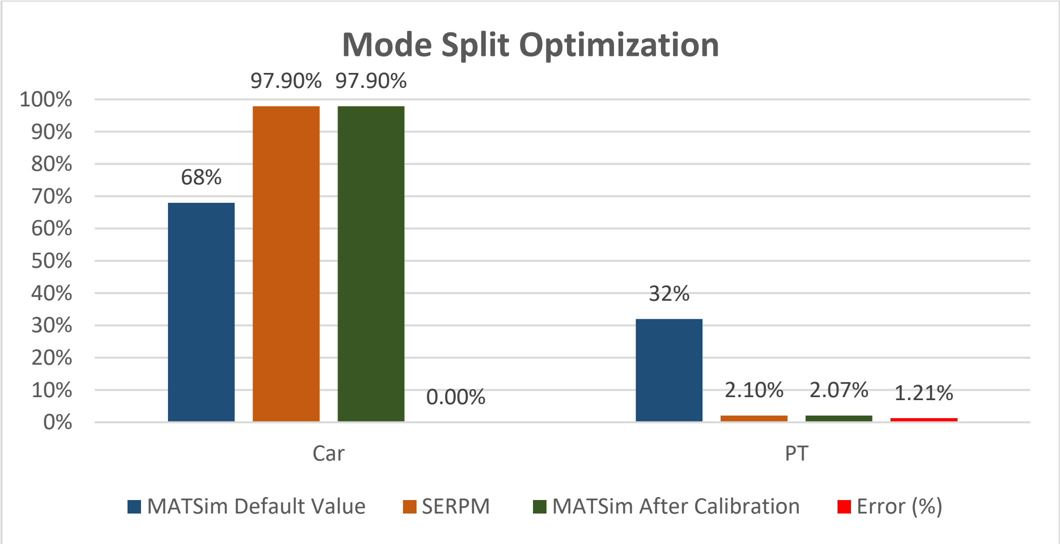


Figure 3. 12: Comparison of the Estimated Mode Split with and without the Use of the Optimization of Model Parameters

Once the Miami Beach MATSim baseline model was calibrated, it was further modified and used to determine the change in the mode split due to the introduction of a light rail system. As described earlier, one scenario considered involved the introduction of a Metrorail (light rail) service parallel to the causeway between the Miami mainland and Miami Beach. In another scenario, in addition to the above-mentioned Metrorail service, a toll was imposed for use of the causeway located nearby. It was hypothesized that these changes will provide a shift from highway traffic to public transit and the shift would be greater in the presence of a road pricing strategy. Table 3.5 shows the results from the MATSim runs, which confirm that the model can predict such shifts. Table 3.5 indicates that there is a shift to transit due to the introduction of the new transit service. Such shift is greater when the new transit option is introduced in combination with a road pricing strategy (4.1% shift towards transit), rather than introducing the new transit service alone (0.5% shift towards transit). Thus, the MATSim results confirmed the study hypotheses.

Table 3. 5 The Impact of the Introduction of Light Rail Service and Highway Toll

Available Mode	MATSim (Default Parameters) Base Scenario	SERPM 7.0	MATSim (Calibrated Parameters)		
			Base Scenario	System w/ Metrorail Addition	System w/ Metrorail Addition & Highway Toll Introduction
Car	68%	97.9%	97.9%	97.4%	93.7%
Public Transit	32%	2.1%	2.1%	2.6%	6.2%

3.5 CONCLUSIONS AND FUTURE WORK

This case study presented a systematic approach for the calibration of the traffic flow and mode choice models of the multi-agent simulation model (MATSim). The study also demonstrated the ability of the MATSim model to estimate the shift in the mode choice due to the introduction of a Metrorail line, a new transit travel option between Miami Beach and Miami mainland. The developed model included full day travel plans of all agents in the network. The plans were generated based on data from the regional demand forecasting model and real-world traffic counts. Additionally, an optimization process was developed to derive the MATSim mode choice parameters based on the results from the established regional model.

The results showed that the utilized calibration process was effective as it produced link traffic flow and mode split values that were much closer to the existing conditions, compared to those generated using MATSim's default parameters. The study also showed the capability of the model to estimate the shift toward public transit due to the introduction of an enhanced transit alternative and the imposition of a toll fee on a nearby highway.

Further fine-tuning of the MATSim model of the Miami Beach network is recommended in future efforts. Such development will provide additional capabilities including:

- Introducing additional modeling scenarios such as the inclusion of Park & Ride, Kiss & Ride, Automated Metro Movers, first-mile-last-mile collector-distributor, ride pooling and hailing, and micro-mobility options, and
- Further improvement and calibration of the traffic flow model and mode split model to allow for modeling the different modes and types of vehicles mentioned in the above list

Future work can shed additional light on the benefits and costs of travel demand modeling with multi-agent simulation models, rather than conventional simulation models. Such knowledge will assist agencies in their decisions to use multi-agent simulation for different applications and to provide justification for these decisions.

3.6 REFERENCES

1. U.S. Department of Transportation, Federal Highway Administration (2010). 2009 National Household Travel Survey (http://nhts.ornl.gov.)
2. Lanken, B., Aarts, H., Van Knippenberg, A., & van Knippenberg, C. (1994). Attitude Versus General Habit: Antecedents of Travel Mode Choice 1. Journal of applied social psychology, 24(4), 285-300.
3. Frank, L., Bradley, M., Kavage, S., Chapman, J., & Lawton, T. K. (2008). Urban form, travel time, and cost relationships with tour complexity and mode choice. Transportation, 35(1), 37-54.
4. Koppelman, F. S. (1989). Multidimensional Model System for Intercity Travel Choice Behavior, Transportation Research Record, 1241, 1-8.

5. Bhat, C.R. (1997), Work Travel Mode Choice and Number of Non-work Commute Stops, Transportation Research Part B, Vol. 31, No. 1, pp. 41-54.
6. Morrison, S. A. & Winston, C. (1985). An Econometric Analysis of the Demand for Intercity Passenger Transportation, In T. E. Keeler (Ed.) Research in Transportation Economics: A Research Annual, Volume 2 (pp. 213-237), JAI Press Inc., Greenwich, Connecticut.
7. Hörl, S., Balać, M., & Axhausen, K. W. (2019). Pairing discrete mode choice models and agent-based transport simulation with MATSim. In 98th Annual Meeting of the Transportation Research Board (TRB 2019).
8. Auld, J., Hope, M., Ley, H., Sokolov, V., Xu, B., & Zhang, K. (2016). POLARIS: Agent-based modeling framework development and implementation for integrated travel demand and network and operations simulations. Transportation Research Part C: Emerging Technologies, 64, 101-116.
9. Heilig, M., N. Mallig, O. Schröder, M. Kagerbauer and P. Vortisch (2018) Implementation of free-floating and station-based carsharing in an agent-based travel demand model, Travel Behavior and Society, 12, 151 – 158.
10. Rieser, M., Nagel, K., Beuck, U., Balmer, M., & Rümenapp, J. (2007). Agent-oriented coupling of activity-based demand generation with multiagent traffic simulation. Transportation Research Record, 2021(1), 10-17.
11. Horni, A., Nagel, K. and Axhausen, K.W. (eds.) (2016). The multi-agent transport simulation MATSim (p. 618). Ubiquity Press.
12. Balmer, M., B. Raney and K. Nagel (2005b) Adjustment of activity timing and duration in an agent-based traffic flow simulation, in: H. Timmermans (ed.) Progress in activity-based analysis, 91–114, Elsevier, Oxford.
13. Grether, D., Y. Chen, M. Rieser and K. Nagel (2009) Effects of a simple mode choice model in a large-scale agent-based transport simulation, in: A. Reggiani and P. Nijkamp (eds.) Complexity and Spatial Networks. In Search of Simplicity, Advances in Spatial Science, 167–186, Springer, doi:10.1007/978-3-642-01554-0.
14. Horni, A. and K. W. Axhausen (2012b) MATSim agent heterogeneity and week scenario, Working Paper, 836, IVT, ETH Zurich, Zurich.
15. Balmer, M., M. Bernard and K. W. Axhausen (2005a) Matching geo-coded graphs, in: 5th Swiss Transport Research Conference, Ascona, March 2005.
16. Kickhofer, B., Hosse, D., Turnera, K., & Tirachinic, A. (2016). Creating an open MATSim scenario from open data: The case of Santiago de Chile. http://www. vsp. tuberline. de/publication: TU Berlin, Transport System Planning and Transport Telematics.
17. Rieser, M., K. Nagel, U. Beuck, M. Balmer and J. Rumenapp (2007b) Truly agent-oriented coupling ¨ of an activity-based demand generation with a multi-agent trac simulation, Transportation Research Record, 2021, 10–17, doi:10.3141/2021-02
18. Charypar, D. and K. Nagel (2005) Generating complete all-day activity plans with genetic algorithms, Transportation, 32 (4) 369–397, doi:10.1007/s11116-004-8287-y.
19. Coordinated Travel – Regional Activity Based Modeling Platform (CT-RAMP). Model Development Report (2014).
20. Wang, J., and Guo, Z. Logistic Regression Models-Methods and Application [M]. Beijing, High Education Publishing Company. 2001.

21. Fu, X., Peng, Z. R., Liang, Q., & Zhang, T. (2009). Aggregated Analysis of Activity-Based Mode Split Models. In International Conference on Transportation Engineering 2009 (pp. 25-30).
22. Southeast Regional Planning Model 7.0 Activity-Based Model Users Guide (2014).
23. Luersen, M. A., & Le Riche, R. (2004). Globalized Nelder–Mead method for engineering optimization. Computers & structures, 82(23-26), 2251-2260

CHAPTER 4: CASE STUDY 3 - POTENTIAL BENEFITS OF INCREASED PUBLIC TRANSIT RIDERSHIP IN MEDIUM SIZED CITIES: A CASE STUDY

4.1 INTRODUCTION

Public transit encompasses a variety of transportation modes and services such as buses, trains, ferries, vanpools, paratransit etc. which are available to the general public (Litman, 2021). Due to its higher occupancy, public transit moves travelers more efficiently than the automobile and can play a crucial role in addressing urban mobility and environmental concerns such as traffic congestion and greenhouse gas emissions. For example, based on data from 2010, 14 percent of global CO_2 emissions by 2010 were solely attributed to the transportation sector (Pachauri et al., 2014) causing 2200 premature deaths and more than $18 billion expenditures in public health in the US (Levy et al., 2010). In 2019, greenhouse gas emissions from transportation accounted for about 29 percent of total U.S. greenhouse gas emissions, making it the largest contributor of U.S. greenhouse gas emissions.

Even though the population of the United States nearly doubled from 1957 to 2017 (172 to 326 million), the number of transit trips over these 60 years remained almost unchanged (10.4 billion to 10.1 billion). The steady increase of private automobile use in the US, in the expense of transit, is attributed to many factors including the development of the US interstate system and the continued expansion of the transportation network infrastructure. Besides the comfort and flexibility of using automobiles, another important reason behind the reduction of transit ridership is urban sprawl. When affordable housing is far away from the job location and is spread in less densely populated areas, transit accessibility becomes limited thus leading to increased automobile use (Blumenberg et al., 2020).

Many communities are rethinking the current transportation and urban planning model and considering the potential benefits from more dense development served by expanded transit options. Using Birmingham, Alabama as a case study, the objective of this work is to document transportation network benefits of shifting automobile trips into transit by expanding public transit options. Despite an estimated Birmingham Metro population of over 1.1M, the public transit options are currently limited to a bus transit system that has faced systemic problems of low ridership and lack of resources and revenues. Reasons behind these issues include the unfavorable image of transit use, lack of resources and support for public transit, and limited-service availability. Therefore, evaluating the potential benefits of increased transit ridership on transportation network operations might help to infer the worthwhile value of investment for public transit in medium-sized cities like Birmingham.

4.2 LITERATURE REVIEW

Many studies examined and documented transportation users' travel preferences and practices. The main stated reasons for selecting the automobile as the preferred mode of transportation are privacy and flexibility, sense of independence, power, control, enjoyment and prestige (Jensen, 1999). Moreover, an automobile journey is fully under the control of the driver, who can drive alone or with chosen persons rather than unknown individuals (Cools et al., 2009). These perceptions along with increased affordability of private automobiles in the US resulted in a continuous increase in the number of automobile trips and contributed to increased automobile ownership and automobile dependency (Gärling et al., 1998). Car ownership creates a strong commitment to use car as well as an attitude to undervalue the alternative transport modes and keep them away from more environmentally friendly public transportation (Tao et al., 2019).

This study explored underlying benefits of using public transportation as they pertain to individual transportation users, transportation authorities and the society. Some of them are summarized and documented in this chapter.

4.2.1 ENVIRONMENTAL BENEFITS

Public transportation can support sustainability initiatives by reducing the frequency of use of private cars and associated environmental impacts. As public transit transports people collectively it is found to produce 45% less CO_2, 95% less CO, and 48% less NO_2 than private vehicles (Kwan & Hashim, 2016). A study after a rail system opening in Taiwan indicates that CO was reduced by 5-15% and another environmental assessment after rail service expansion in Germany indicates the reduction of pollutants such as NO, NO_2 and CO (Beaudoin et al., 2015. It is estimated that if 5% of Americans used public transit instead of private car or if every American used public transit for 5% of their trips during 1970 to 1998, the CO pollution reduction would be more than the CO emitted from all metal processing plants and chemical manufacturing section combined (Shapiro et al., 2002).

4.2.2 ECONOMIC AND SOCIAL BENEFITS

According to a report published by American Public Transportation Association (APTA) in 2009, for one billion dollars of annual investment in public transportation, there would be more than $1.7 billion dollars of added annual GDP (Weisbrod & Reno, 2009). Depending on some factors such as mileage reduction, declining vehicle ownership etc., a shift from automobile to transit provides a variety of cost savings (Litman, 2009) including fuel and oil, insurance costs, parking costs etc.

Available public transit services can be especially beneficial for people with low income who cannot afford automobile ownership and for elderly and disabled persons by offering convenient and affordable service. Thus, it increases social and economic opportunities for physically, socially and economically disadvantaged people along with achieving equity objectives (Litman, 2020).

4.2.3 HEALTH BENEFITS

According to the US Center for Disease Control (CDC) and prevention, at least 30 minutes of daily physical activity such as bicycling, or walking is necessary to stay healthy (Adult Participation in Recommended Levels of Physical Activity --- United States, 2001 and 2003, n.d.). An Atlanta, Georgia survey results show that almost two-thirds of the recommended daily physical activity is achieved by the transit users which is ten times greater than the average walking reported by the non-transit users (Lachapelle & Frank, 2009). It is also worth noting that medical expenses are 32% lower ($1,019 per year) for adults who achieve the recommended physical activity than those who do not ($1,349 per year) (Litman, 2015). According to another study, 21 minutes of walking can help to burn 65.1 to 98.7 calories and 100 kilocalories burn per day might save $12, 500 dollars per person in obesity-associated medical costs (Freeland, 2013). These findings clearly show the value of transit in the wellbeing of transportation users, an issue that is often overlooked by decision makers when they appropriate funding for transportation services and projects.

4.2.4 CONGESTION REDUCTION

The results from a seemingly unrelated regression equation (SURE) model showed that 50 percent increase of city bus routes in highly populated areas of Taiwan reduce car usage by 1.4% which corresponds to 300,000 vehicles (Jou & Chen, 2014). Another study on Bay Area Rapid Transit (BART) system indicates that in absence of BART services during the morning peak, driving times increased more than four times in multiple corridors (BART, 2016). Similar studies (Adler & Ommeren, 2016); (Anderson, 2014) on investigating the effects of transit absence indicated increased traffic volumes, and longer delays. On the contrary, a study on Salt Lake City's University TRAX light-rail system in 2014 found that typical vehicle traffic reduced by nearly 50% with the expansion of the light rail system (Ewing et al., 2014). Some other studies (Liu & Cirillo, 2014); (Mulalic et al., 2016) also provide links between transit availability and transportation mode choices and highlight the potential positive impacts from introduction or expansion of transit services in a region for individuals, the transportation network operations, and the community. However, given local differences, it is important to conduct local studies to gain an understanding of potential impacts of transit ridership increase on local congestion and quantify such impacts.

4.3 METHODOLOGY

4.3.1 SIMULATION PLATFORM

Based on findings from earlier work (Sisiopiku and Salman, 2019), the Multi-Agent Transport Simulation platform (MATSim) agent-based transportation simulation was used to simulate the impact of changes in transit ridership in Birmingham, AL. 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. It starts with an initial population demand (a.k.a. plans) in the studied area. Each agent in the population maintains a pool of up to 5 day-plans. As also described in Chapter 3, in each iteration, the MATSim's "mobsim" simulation executor first runs the selected plans of the agents in the synthetic road network environment. 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). Afterwards, the re-planning 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.

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 for the collection of simulation statistics for individual links. MATSim provides overview summaries of traffic counts and other performance measures for the whole network as well as for individual links. Also, a google maps-based visualization is available, displaying simulation output results for each station in a pop-up window.

4.3.2 CASE STUDY DESIGN

The Birmingham MATSim simulation model is an extension of a pilot model of Birmingham developed for STRIDE Project B (Sisiopiku et al., 2019; Guo et al., 2019; Guo et al., 2019b). The model was refined and expanded to include public transit. More specifically, a base model considering private cars only (base) and three scenarios including public transit were developed using MATSim. The Population was generated by population synthesis (Ramadan and Sisiopiku, 2019; Mulalic et al., 2016) using travel diary data from a survey in Birmingham (Sarjana et al., 2020) and open source data (Guo et al., 2019). To speed up computational performance, and similarly to MATSIM-based studies reported in the literature from Sweden, Austria, South Korea, Berlin, Germany, Paris, France, in Zurich, Switzerland (Bischoff, et al., 2019; MATSim. *Scenario Gallery*. n.d. [cited 2020 10.19]), 10% of the total population was used for the simulation. Criteria considered for the development of each scenario are listed in Table 4.1. For the first scenario, the probability of choosing public transit is set as equal or less than 0.1 because with this probability, 1.1% of total agents (travelers) choose public transit. This percentage represents public transit ridership in Birmingham at the time of study (BJTCA, 2018; Census profile: Birmingham-Hoover, AL Metro Area, 2019). To determine the impact of increased ridership, the probability was then set as less than or equal to 0.5 and 0.9 respectively for scenario 2 and scenario 3, respectively. Table 4.1 also shows the expected market share of the various modes considered in the study (namely car, transit, and walk) under the 3 scenarios tested (namely probability of ≤0.1, ≤0.5, ≤0.9).

Table 4. 1: Scenario design criteria

	Scenario 1	Scenario 2	Scenario 3
Probability of Choosing Public Transit (PT)	≤0.1	≤0.5	≤0.9
Car Percentage	97.5%	87%	76.4%
Public Transit Percentage	1.1%	5.7%	10.1%
Walk Percentage	1.4%	7.3%	13.5%

4.3.3. MODEL CALIBRATION AND VALIDATION

Simulation model validation is an important step toward increasing the confidence in use of any simulation model. An approach commonly used to validate simulation models, including MATSim, is by comparing traffic volumes produced by the simulation model with actual traffic counts (Bischoff, et al., 2019). In this study, traffic volumes from the MATSim model (scenario 1) were compared to traffic counts collected by the Alabama Department of Transportation (ALDOT) for two time periods (7 AM to 8 AM and 5 PM to 6 PM). A total of 90 links along I-459 N, I-459 S, AL 25 S, I-65 N, AL 3 N, AL3 S, AL 5 E, AL 38 E, and the junction of I-65 and I-59 located in the Jefferson and Shelby counties were used for validation.

Figure 4.1 and Figure 4.2 show results from the Birmingham MATSim model validation, where the X-axis represents the field volume data and Y-axis represents the simulated volumes. The three diagonal lines in the graphs represent the simulated versus real volume ratio of 2:1, 1:1 and 1:2 which are named as 2 counts, 1 count and 0.5 count respectively. Counts falling between 2 and 0.5 counts are considered acceptable (Van der Merwe, 2011). It can be observed that most of the data points are within these boundaries, thus the model validation is deemed acceptable.

4.3.4 LINK ID SELECTION

MATSim generates output according to the link ID, which refers to the identity of the roadway sections in the MATSim platform. This study analyzed the performance of road network as a result of increased transit ridership by examining links near the bus stops. The procedure used for selecting 93 links from the 1761 bus stop links within the study network is shown in Figure 4.3 and the selected links in the study area are shown in Figure 4.4 with blue dots.

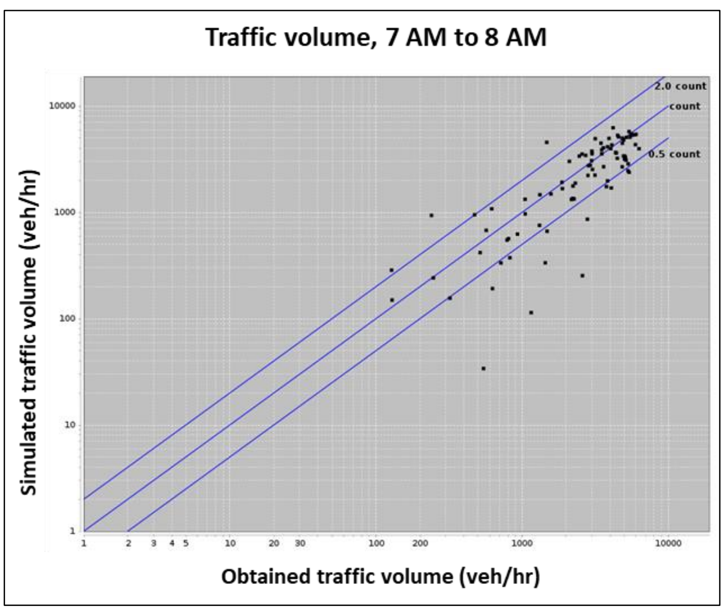


Figure 4. 1 Comparison between simulated and field traffic volumes for validation links for 7 AM to 8 AM

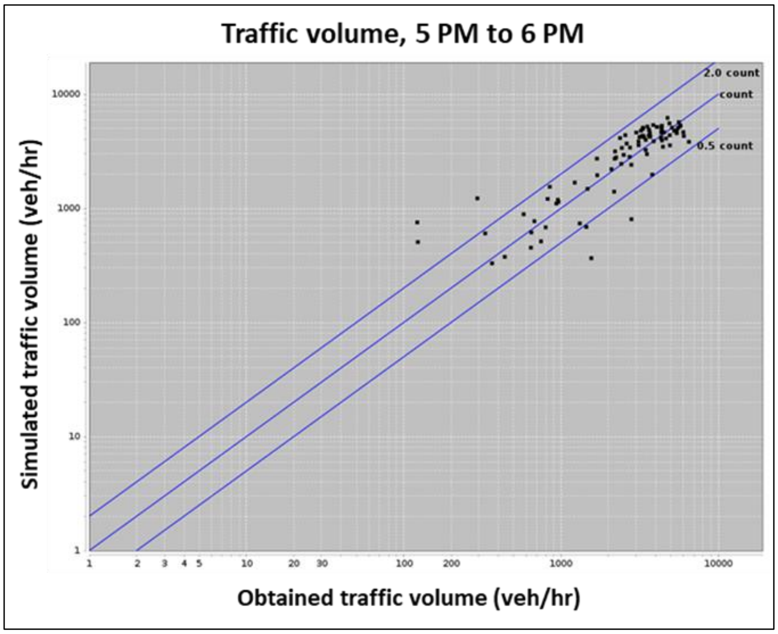


Figure 4. 2 Comparison between simulated and obtained traffic volumes of validation links for 5 PM to 6 PM

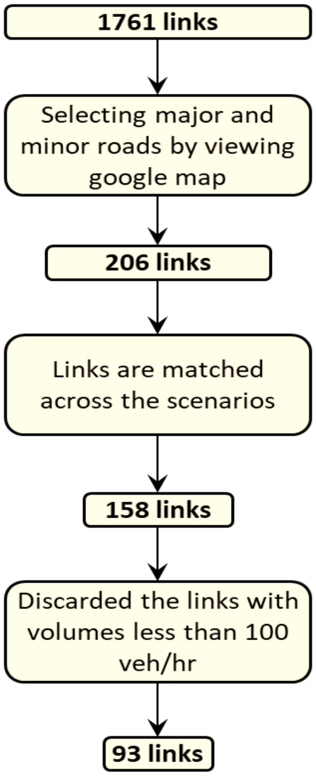


Figure 4. 3 Flow chart of the link selection

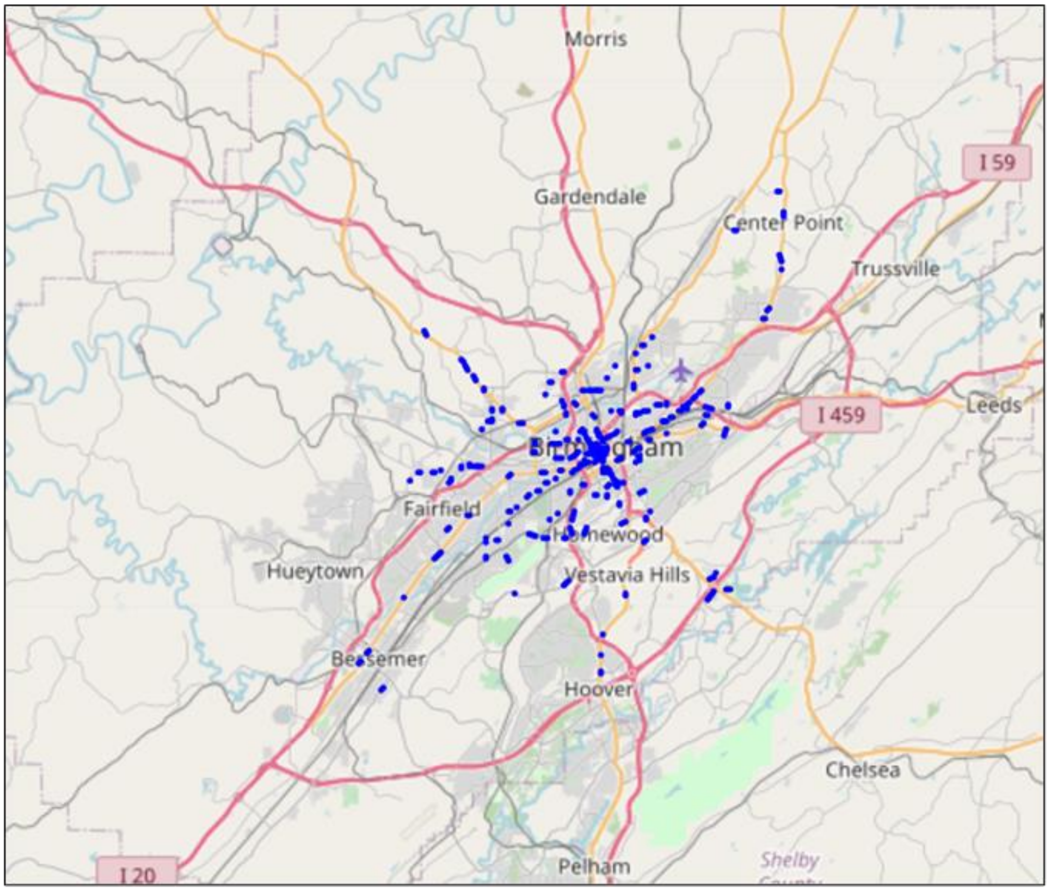


Figure 4. 4 Location of the selected links (blue dot)

Based on the data available at the live traffic website (TomTom. *Birmingham traffic*. 2020), traffic congestion in the Birmingham region peaks from 5 to 6 PM for a typical weekday. During this time slot, an addition 50% of travel time is needed on average to complete a trip, compared to the travel time under free flow conditions. As far as congestion severity time period is concerned, the 5 PM to 6 PM period is followed by 4 PM to 5 PM, 3 PM to 4 PM, 7AM to 8 AM and 8 AM to 9 AM, during which travel times are 43%, 36%, 33% and 27% higher respectively, as compared to travel time under free flow conditions. The study selected the same five time periods for further analysis.

To accommodate the big range of volumes (120 veh/hr to 2,520 veh/hr), links were classified in five different groups based on the procedure described in Figure 4.5. Resulting link groupings from this procedure are summarized in Table 4.2. It should be noted that the simulated traffic volumes are 10% of the total. For example, group 3 includes simulated volumes of 51-90 veh/hr which refer to links with actual traffic volumes of 510-900 veh/hr as seen in Table 4.2. Also, the volumes are total volumes per link (directional volumes, all lanes combined).

To determine the performance of the road network in the event of increased transit ridership, traffic volume, speed, and average travel time were analyzed for the selected 5 time slots. The results are summarized next.

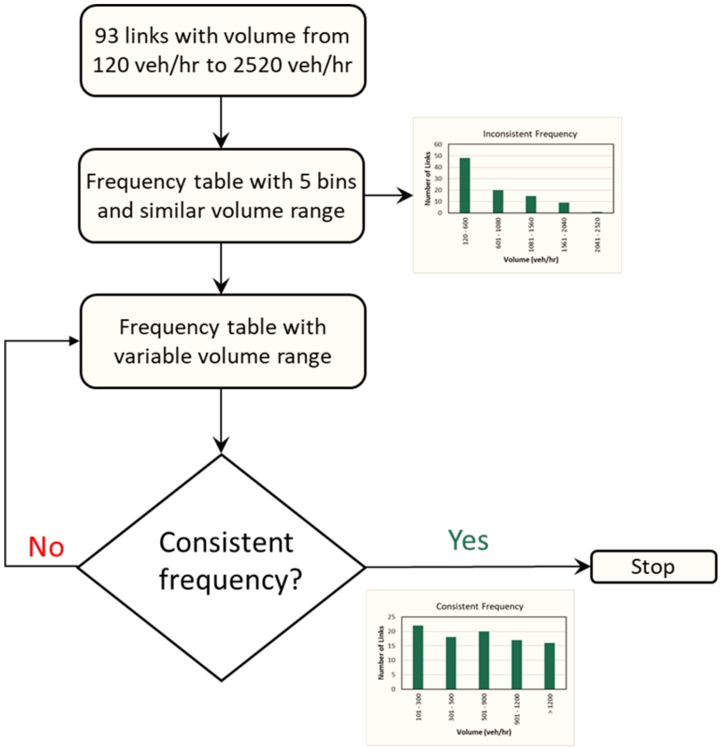


Figure 4. 5 Group identification process

Table 4. 2: Grouping of the study links

Groups	Traffic Volume (veh/hr)	Number of Links
Group 1	101 - 300	22
Group 2	301 - 500	18
Group 3	501 - 900	20
Group 4	901 – 1200	17
Group 5	>1200	16

4.4 RESULTS AND DISCUSSION

4.4.1 TRAFFIC VOLUME ANALYSIS

Figure 4.6 shows the mean traffic volume for three of the scenarios, 5 time slots and 5 groups. The purpose of showing traffic volume data for the study scenarios based on the groups is to showcase how increased transit ridership affects network operations under different volumes levels. The black line, blue line, red line, green line and purple line in Figure 4.6 stands for the traffic volumes between 7 AM to 8 AM, 8 AM to 9 AM, 3 PM to 4 PM, 4 PM to 5 PM and 5 PM to 6 PM respectively. To facilitate comparisons, the line color scheme used in the analysis represents the same time slots for all the groups.

4.4.1.1 Traffic Volume Change for Group 1

Mean traffic volumes are shown for group 1, which consists of the study links with volumes of

101-300 veh/hr. The probability of travelers choosing public transit increases from scenario 1 (public transit share of trips = 1.1%) to scenario 2 (public transit share of trips = 5.7%) and from scenario 2 to scenario 3 (public transit share of trips = 10.1%). With increased transit ridership, a decrease in automobile trips is seen, especially in the PM time periods. It should be noted, that the effects on traffic volume between scenarios 1 and 2 are negligible under low traffic demand conditions during AM time periods (black and blue lines). It can be further observed that the mean traffic volume is reduced by 100 veh/hr from scenario 1 to scenario 3 from 4-6 PM (green and purple lines). A volume reduction (40 veh/hr) occurs also from scenario 2 to scenario 3 from 3-6PM. Overall, the traffic volume for this group of roadway links is reduced due to the increase in public transit probability.

4.4.1.2 Traffic Volume Change for Group 2

Group 2 represents links with volume from 301-500 veh/hr. Group 2 shows traffic volume reduction throughout the 5 study time periods considered for increasing transit probability both from scenario 1 to scenario 2 and from scenario 1 to scenario 3. The greatest impacts are observed during the 3 PM to 4 PM time period (red line) between scenarios 1 and 3 (nearly 170 veh/hr), noticeable with the steep downward slope.

4.4.1.3 Traffic Volume Change for Group 3

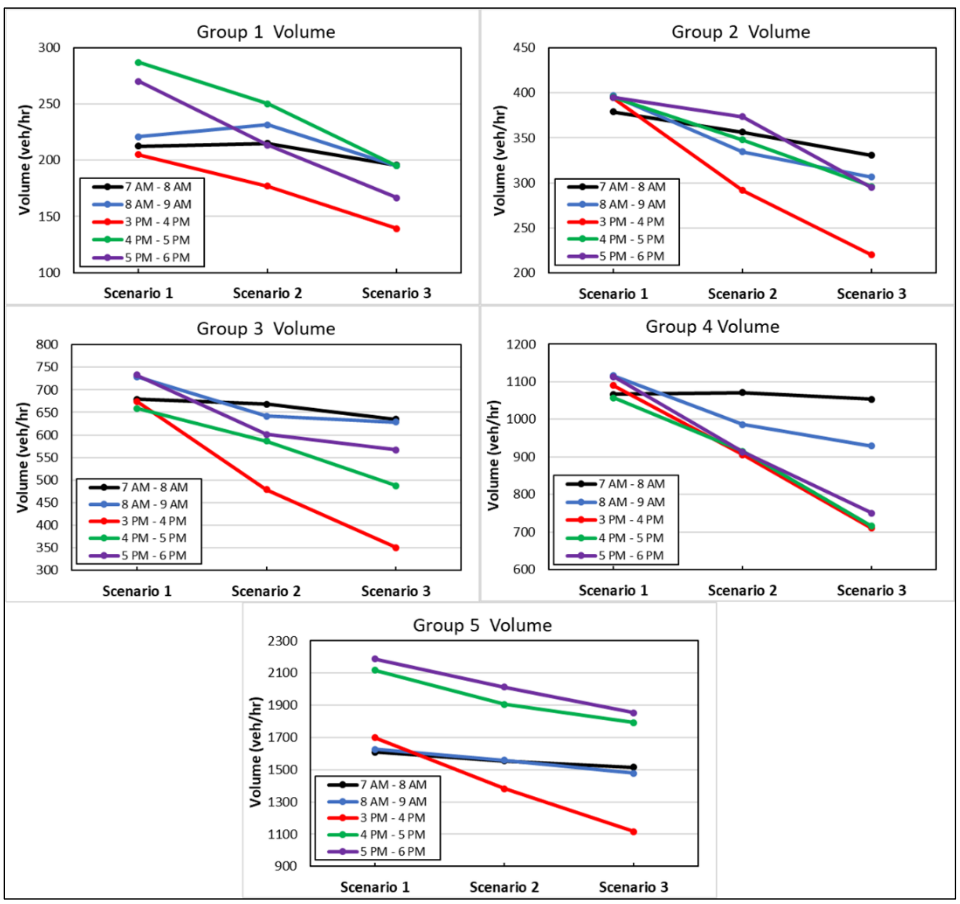


Figure 4. 6 Traffic volume variation with increased transit ridership in 5 groups

From scenario 1 to scenario 3, all the time slots for group 3 (501-900 veh/hr) show noticeable traffic volume reductions as the probability of transit use increases. The highest reduction (320 veh/hr) takes place during the 3 PM to 4 PM time period (red line). The volume reduction is also noticeable from scenario 1 to scenario 2, from 3 PM to 4 PM (red line) and 5 PM to 6 PM (purple line) which show mean volume reductions of 300 veh/hr and 130 veh/hr respectively.

4.4.1.4 Traffic Volume Change for Group 4

Comparison of simulation outputs for Group 4 (901-1200 veh/hr) links shows little to no change in the traffic volumes during the 7 AM to 8 AM time slot from scenario 1 to scenario 2 to scenario 3 (black line). During all other time periods, traffic volume drops as transit ridership increases. Once again, the highest such impact is observed during the 3 PM to 4 PM time slot (red line) where the mean traffic volume reduction between scenario 1 and 3 is 380 veh/hr.

4.4.1.5. Traffic Volume Changes for Group 5

In the group 5 links which carry volumes in excess of 1200 veh/hr, mean traffic volumes are reduced for all 5 study timeframes and for both changes in transit use considered (i.e., scenario 2 and 3) as compared to the scenario 1. The most significant impact is observed during the 3 PM to 4 PM time period (red line) where mode shift toward transit (from scenario 1 to scenario 3) results in reduction of average traffic volume on group 5 links from 1700 veh/hr to 1100veh/hr (or 580 veh/hr).

To further quantify the impacts on traffic volume as a result of changes in transit ridership, percentage flow reductions are calculated from no transit availability (base model) to scenarios 1, 2, and 3 that assume a future increase in transit ridership from 1.1% to 5.7% to 10.1% respectively using Equation (4.1), Equation (4.2), and Equation (4.3) respectively. The results are summarized in Table 4.3.

$$\frac{Base\ model\ volume - Scenario\ 1\ volume}{Base\ model\ volume} \times 100 \ \ldots\ldots\ldots\ldots\ldots \quad (4.1)$$

$$\frac{Scenario\ 1\ volume - Scenario\ 2\ volume}{Scenario\ 1\ volume} \times 100 \ \ldots\ldots\ldots\ldots\ldots \quad (4.2)$$

$$\frac{Scenario\ 1\ volume - Scenario\ 3\ volume}{Scenario\ 1\ volume} \times 100 \ \ldots\ldots\ldots\ldots\ldots \quad (4.3)$$

Table 4. 3: Traffic volume reduction percentages in different scenarios

Time Periods	Base model to scenario 1	Scenario 1 to scenario 2	Scenario 1 to scenario 3
7 AM - 8 AM	-4.2%	2.0 %	5.6%
8 AM - 9 AM	3.6%	7.3%	12.5%
3 PM - 4 PM	3.3%	20.0%	36.6%
4 PM - 5 PM	0.7%	10.5 %	18.1%
5 PM - 6 PM	-3.4%	9.1%	16.8%

According to the results shown in Table 4.3, current transit ridership (scenario 1) had no impact in reducing traffic volume in three time periods (8 AM to 9 AM, 3 PM to 4 PM, 4 PM to 5 PM), when compared to the no transit option (base). However, increasing the transit ridership in scenario 2, resulted in noticeable traffic volume reductions. Further increase of the public transit ridership in scenario 3 reduces the traffic volumes even further, with the highest reduction percentage of 36.6% occurring between 3 PM to 4 PM.

4.4.2 TRAFFIC SPEED ANALYSIS

As the probability of choosing public transit is 0.1 in scenario 1 and then increased to 0.5 and 0.9 in scenarios 2 and 3 respectively, an improvement in traffic performance is expected in terms of speed increases associated with higher transit ridership. It is postulated that increased transit ridership will result in increases in the average vehicular speeds. Figure 4.7 shows the simulated speeds for the three scenarios, 5 time slots and 5 groups studied.

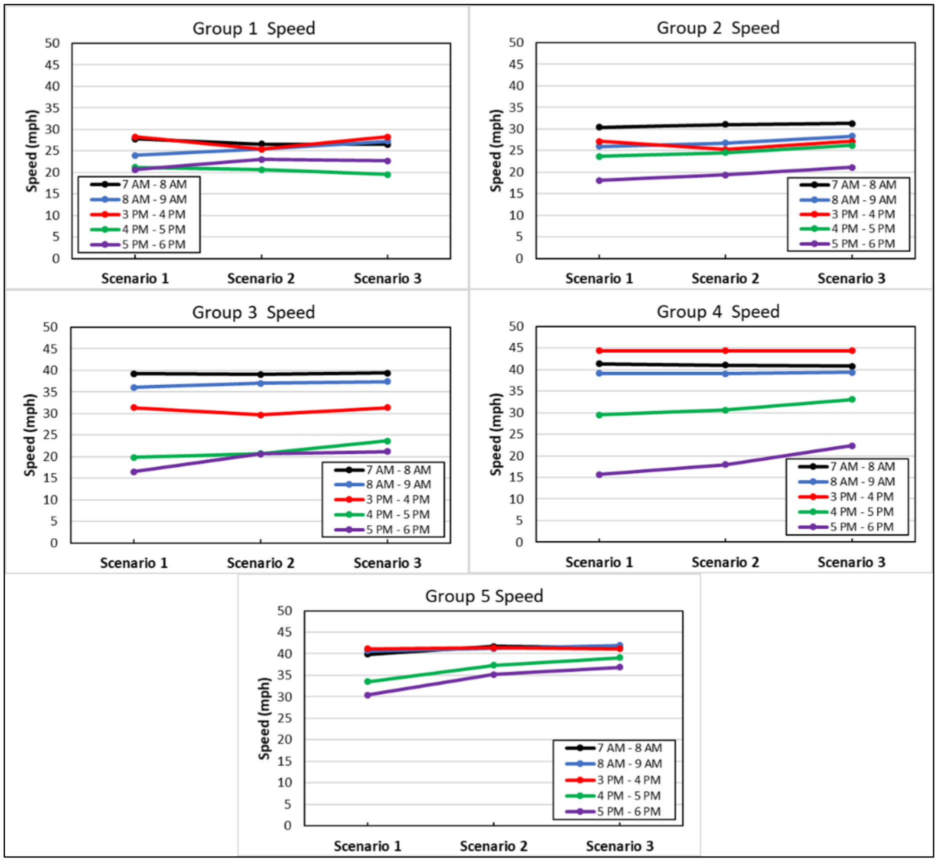


Figure 4. 7 Traffic speed variation with increased transit ridership in 5 groups

4.4.2.1 Traffic Speed Change for Group 1

Free flow speed refers to the average speed in the absence of congestion or adverse conditions in a roadway. The MATSim results show that for free flow traffic conditions (group 1) the impact of mode shifts from automobile to transit on speed is negligible. As shown in Figure 4.7, speed differences are small in group 1 for all time periods considered when comparing results from scenario 1 to scenario 2 to scenario 3.

4.4.2.2 Traffic Speed Change for Group 2

Group 2 represents near free flow conditions. For this traffic group, the effects on speed from ridership shifts toward transit are still small. Figure 4.7 shows that speed increased overall by 2 to 3 mph except during the 3 PM to 4 PM (red line) time period for group 2.

4.4.2.3 Traffic Speed Change for Group 3

The study results confirm that for group 3 links, speeds increased by 3 to 4 mph during the afternoon peak times (green line- 4 PM to 5 PM and purple line- 5 PM to 6 PM) both in scenario 2 and scenario 3. Speeds for the other three time periods remain almost constant or slightly decrease till the execution of scenario 3 for this group.

4.4.2.4 Traffic Speed Change for Group 4

Links in group 4 show no speed changes in response to shifts in ridership (red, blue and black lines). As transit ridership increases (from scenario 2 to 3) and during afternoon peak times (green line- 4 PM to 5 PM and purple line- 5 PM to 6 PM) the results show overall speed increases of nearly 4 mph and 7 mph respectively.

4.4.2.5. Traffic Speed Change for Group 5

For study links with traffic conditions described by group 5, the highest speed increase happens during the afternoon peak (the purple line- 5 PM to 6 PM) with a mean speed increase of 5 mph in scenario 2 and 7 mph in scenario 3, as compared to scenario 1. The second highest increase is visible from 4 PM to 5 PM (green line) with average speed increase of 3 mph in scenario 2 and 5 mph in scenario 3, as compared to scenario 1. The speeds of remaining three time slots studied are almost similar throughout the three scenarios.

Though the volume reduction is noticed to be higher for 3 PM to 4 PM, speed increase for this time period is almost zero for group 1 to group 4. Thus, free flow speeds of the 93 links were compared with the operating speed for scenario 1 in 3 PM to 4 PM. The findings from the observation showed that, most of the links, have near free flow condition (Table 4.4). Therefore, speed was not affected as a result of the volume reduction in scenario 2 and scenario 3.

Table 4. 4: Traffic speed changes for different scenarios

Link ID	Operating speed in scenario 1 (mph)	Free Flow speed (mph)	Difference between Free flow and operating speed (mph)
107920507_14	36	37	1
259336961_0	36	37	1
323899401_8	49	50	1
7782325_7_r	37	37	0
259970324_1_r	26	28	2
7742120_1_r	36	37	1
592215806_4_r	37	37	0
259311994_2	36	37	1
7740932_2	27	28	1
394283610_3	19	19	0
7782325_5	37	37	0
7740932_0	28	28	0

A similar comparison setup was followed to document the percent speed increase resulting from the assumed increase in transit ridership in Birmingham as expressed by increased transit use probability in scenarios 2 and 3. The results are summarized in Table 4.5 and are calculated using Equation (4.4), Equation (4.5), and Equation (4.6), below.

$$\frac{Scenario\ 1\ speed - Base\ model\ speed}{Base\ model\ speed} * 100 \quad (4.4)$$

$$\frac{Scenario\ 2\ speed - Scenario\ 1\ speed}{Scenario\ 1\ speed} * 100 \quad (4.5)$$

$$\frac{Scenario\ 3\ speed - Scenario\ 1\ speed}{Scenario\ 1\ speed} * 100 \quad (4.6)$$

Table 4. 5: Traffic speed changes for different scenarios

Time Periods	Base model to scenario 1	Scenario 1 to scenario 2	Scenario 1 to scenario 3
7 AM - 8 AM	0.4%	0.2%	0.2%
8 AM - 9 AM	0.1%	2.1%	4.8%
3 PM - 4 PM	8.8%	-2.5%	0.0 %
4 PM - 5 PM	-2.4%	8.8%	15.2%
5 PM - 6 PM	-2.3%	16.0%	22.0%

According to Table 4.5, increase in transit ridership contributes to automobile speed increase during peak traffic time periods. Iincreases in speeds were found between 4 PM to 5 PM and 5 PM to 6 PM while speeds in remaining three time periods had very little or no increase in speed.

4.4.3 TRAVEL TIME ANALYSIS

Travel time for a roadway link refers to the time needed to drive from start point to the end point of that link. The expectation is that modal shifts from automobile to transit may result in reduction of link travel times, thus resulting in an improvement of traffic network performance. To verify this hypothesis, an analysis of travel time data was performed for the study links and for the 5 study time periods. Average travel times for the three scenarios and 5 groups are shown in Figure 4.8.

4.4.3.1 Travel Time Change for Group 1

The impacts on average travel times in response to modal shifts towards transit are small (1 to 2 sec) for all time slots considered in group 1.

4.4.3.2 Travel Time Change for Group 2

Under group 2 conditions, there is little to no change observed to average travel time as transit ridership increases from scenario 1 to scenario 2 and from scenario 1 to scenario 3. An average travel time decrease by around 3 seconds is seen from 5 PM to 6 PM (purple line) from scenario 1 to 2 and around 6.5 seconds from scenario 1 to scenario 3.

4.4.3.3 Travel Time Change for Group 3

Under group 3 traffic conditions, some more noticable reductions in travel times are realized during the afternoon peak time periods (green line- 4 PM to 5 PM and purple line- 5 PM to 6 PM) as transit ridership increases in scenario 2 and scenario 3. The average travel time is decreased by almost 10 seconds from 4 PM to 5 PM in scenario 3 and almost 8 seconds 5 PM to 6 PM.

4.4.3.4 Travel Time Change for Group 4

There is no visible impact on link travel times from changes in transit ridership during the 7 AM to 8 AM (black line), 8 AM to 9 AM (blue line) and 3 PM to 4 PM (red line) time periods in group 4. The opposite is true for the afternoon peak times, as shown by the green line and purple line in Figure 4.8.

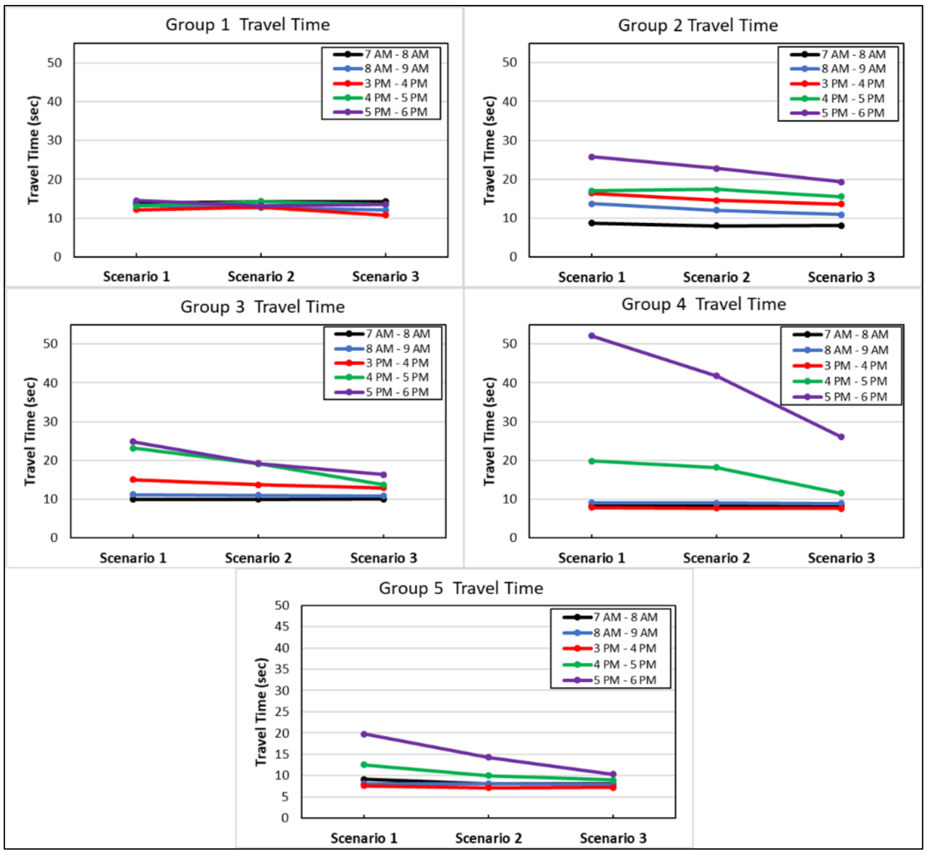


Figure 4. 8 Travel Time variation with increased transit ridership in 5 groups

Based on the simulation results, the average travel time during the 5 PM to 6 PM time period (purple line) decreased by 9 seconds from scenario 1 to scenario 2 and a total of 26 seconds from scenario 1 to scenario 3. The decrease in travel times during the 4 PM to 5 PM time period (green line) are around 1 second and 8 seconds between scenario 1 and scenarios 2 and 3 respectively.

4.4.3.5 Travel Time Change for Group 5

From the simulation results for group 5 conditions, it can be seen that during 7 AM to 8 AM, 8 AM to 9 AM and 3 PM to 4 PM (black, blue and red lines respectively) there is very little or no change in average link travel times for the two scenarios considered as compared with scenario 1. However, during the 4 PM to 5 PM time period (green line) as well as the 5 PM to 6 PM time period (purple line), link travel times decreased by 3 seconds and 9 seconds respectively, when transit ridership changed from scenario 1 to scenario 3 conditions.

Using a similar comparison setup as the one used for determining impacts on traffic volumes and speed, the percent change of average travel time is calculated using Equation (4.7), Equation (4.8), and Equation (4.9). The results are reported in Table 4.6.

$$\frac{Base\ model\ travel\ time - Scenario\ 1\ travel\ time}{Base\ model\ travel\ time} * 100 \ldots\ldots\ldots\ldots\ldots\ldots \quad (4.7)$$

$$\frac{Scenario\ 1\ travel\ time - Scenario\ 2\ travel\ time}{Scenario\ 1\ travel\ time} * 100 \ldots\ldots\ldots\ldots\ldots\ldots \quad (4.8)$$

$$\frac{Scenario\ 1\ travel\ time - Scenario\ 3\ travel\ time}{Scenario\ 1\ travel\ time} * 100 \ldots\ldots\ldots\ldots\ldots\ldots \quad (4.9)$$

Table 4. 6: Travel Time percentage change in different scenarios

Time Periods	Base model to scenario 1	Scenario 1 to scenario 2	Scenario 1 to scenario 3
7 AM - 8 AM	0.8%	2.3%	1.9%
8 AM - 9 AM	-1.0%	4.5%	8.5%
3 PM - 4 PM	10.3%	6.8%	11.4%
4 PM - 5 PM	-2.4%	13.8%	29.1%
5 PM - 6 PM	-5.2%	23.1%	40.7%

The results indicate that transit ridership during the study time periods has no effect in reducing average travel time (except for the time between 3 PM to 4 PM), however, travel time is reduced while transit ridership is increased in scenario 2 and scenario 3. The reduction is higher for the scenario with more transit probability (i.e., scenario 3) and for groups 4 and 5 where the network carries heavier traffic loads.

4.4.4 SIGNIFICANCE TEST ANALYSIS

To understand whether the traffic flow reduction associated with changes in transit ridership is statistically significant, several significance tests were performed. First, a significance test was performed to test difference within the three scenarios. If the traffic flow reduction was significant within the scenarios, then tests between the scenarios were also performed. The tests were performed using the 95% confidence level, hence, a p-value less than 0.05 indicates that the difference is statistically significant between/within groups. Table 4.7 shows the test scores (p-value) of the significance tests performed herein.

The results from the statistical tests between the base model and scenario 1 imply that the current level of public transit use does not have any significant impact on traffic volumes. This is evident from the high p-values documented in Table 4.7 resulting from the comparison of traffic volumes between the base model and with scenario 1. Furthermore, results show that traffic flow reduction is statistically significant when comparing results from scenario 1 to scenario 2 as well as scenario 1 to scenario 3. The only exception is for group 1 which shows a p-

Table 4. 7: Statistical significance test scores (p-value) for traffic volume changes

Groups	Base model with scenario 1	Within three scenarios	Scenario 1 with scenario 2	Scenario 1 with scenario 3
Group 1	0.7801	**0.0169**	0.2240	**0.0175**
Group 2	0.6078	**0.0006**	**0.0222**	**0.0055**
Group 3	0.5706	**0.0285**	**0.0323**	**0.0327**
Group 4	0.6728	**0.0050**	**0.012**	**0.021**
Group 5	0.9968	0.2827		

value of 0.224>0.05, indicating that there is no evidence to support that there is a statistical difference in traffic volumes from scenario 1 to scenario 2 during free flow conditions (group 1).

Table 4.8 and Table 4.9 show the p-values of the significance tests done for speed and travel time changes respectively. The results show that there is not enough evidence to suggest that there are statistically significant differences in speed or travel time associated with the increase in transit ridership as per the study scenarios.

Table 4. 8: Statistical significance test scores (p-values) for speed changes

Groups	Base model with scenario 1	Within three scenarios
Group 1	0.8958	0.9517
Group 2	0.6599	0.7837
Group 3	0.8967	0.9394
Group 4	0.9413	0.9536
Group 5	0.9322	0.2999

Table 4. 9: Statistical significance test scores (p-values) for travel time changes

Groups	Base model with scenario 1	Within three scenarios
Group 1	0.9738	0.5976
Group 2	0.7644	0.7221
Group 3	0.7142	0.4394
Group 4	0.8730	0.7504
Group 5	0.6853	0.4117

4.5 CONCLUSIONS AND RECOMMENDATIONS

This case study used a comprehensive activity-based simulation model of the Birmingham area to simulate traffic operations under various transit ridership scenarios ranging from 0% (base) to 1.1% (scenario 1-current) to 5.7% (scenario 2-future) to 10.1% (scenario 3-future). The analysis considered links with various levels of traffic demand (Groups 1 to 5) and 5 hourly time slots. The main findings from this study are summarized below.

- Current transit ridership has no significant effect on traffic volume reduction for the roadway sections with low traffic demand. As transit ridership increases, traffic volume reductions are reported and, with a few exceptions, the traffic flow reductions are statistically significant when comparing results from scenario 1 to scenario 2 as well as scenario 1 to scenario 3.

- Increase the transit ridership results in speed increases for 4 of the 5 time periods considered in this study, however, those changes are not statistically significant.
- Similarly, with the increase of transit ridership, travel time is decreasing, but not significantly.

Based on the study findings we conclude that higher levels of modal shifts from private cars to transit modes might be necessary in order to materialize significant differences in speed and travel time. Still, some network performance improvement was documented in this study as a result of increased transit ridership as demonstrated by the percent reduction of volume and travel times, and percent increase of speed. This implies that benefits of increasing transit ridership in medium-size cities like Birmingham can contribute to improving the performance of the road network.

Initiatives that can increase transit ridership include expanding the number of bus routes and/or frequency of service, strategically positioning new bus stops near the residential areas, providing transit use incentives, and improving quality of transit service. Additionally, initiating a feeder service to provide first and last mile service connections for the distant passengers shows a good potential to increase transit ridership (Luk & Olszewski, 2003).

The findings from this study highlight the potential benefits of increased bus ridership on transportation network operations in medium-size US cities such as Birmingham. The study results may encourage the local transit authority, transportation planners and decision makers to think of ways to improve and/or expand public transit services in order to attract new public transit users. The Birmingham MATSim model developed in this study can be used to test future transit scenarios and provide scientific evidence of the expected impacts on transportation network operational performance from transit service changes in the region.

4.6 REFERENCES

1. Adler, M. W., & van Ommeren, J. (2016). *Does public transit reduce car travel externalities? Quasi-natural experiments' evidence from transit strikes.* Journal of Urban Economics, 2016. 92(C), 106-119. https://econpapers.repec.org/article/eeejuecon/v_3a92_3ay_3a2016_3ai_3ac_3ap_3a106-119.htm
2. Anderson, M. L. (2014). Subways, Strikes, and Slowdowns: The Impacts of Public Transit on Traffic Congestion. American Economic Review. 104(9), 2763-96.
3. BART, *BART’s Role in the Region*. 2016.
4. Beaudoin, J., Farzin, Y. H., & Lin Lawell, C.-Y. C. L. (2015). *Public transit investment and sustainable transportation: A review of studies of transit's impact on traffic congestion and air quality.* Research in Transportation Economics, 52, 15-22.

5. Bischoff, J., et al., (2019). *Impacts of vehicle fleet electrification in Sweden–a simulation-based assessment of long-distance trips*. in *2019 6th International Conference on Models and Technologies for Intelligent Transportation Systems (MT-ITS)*. IEEE.
6. BJTCA, (2018). *Transit Development Plan-Technical Memorandum 1*.
7. Blumenberg, E., Garrett, M., King, H., Paul, J., Ruvolo, M., Schouten, A., Taylor, B. D., & Wasserman, J. L. (2020). What's Behind Recent Transit Ridership Trends in the Bay Area? Volume I: Overview and Analysis of Underlying Factors. *Escholarship.org*. https://doi.org/10.17610/T6PC7Q
8. Cools, M., Moons, E., Janssens, B., & Wets, G. (2009). Shifting towards environment-friendly modes: profiling travelers using Q-methodology. *Transportation*, *36*(4), 437–453. https://ideas.repec.org/a/kap/transp/v36y2009i4p437-453.html
9. Ewing, R., Tian, G., & Spain, A. (2014). Effect of Light-Rail Transit on Traffic in a Travel Corridor. TREC Final Reports. https://doi.org/10.15760/trec.56
10. Freeland, A. L., Banerjee, S. N., Dannenberg, A. L., & Wendel, A. M. (2013). *Walking associated with public transit: moving toward increased physical activity in the United States.* American journal of public health, 2013. 103(3), 536-542.
11. Gärling, T., Gillholm, R. & Gärling, A. (1998). Reintroducing attitude theory in travel behavior research: The validity of an interactive interview procedure to predict car use. *Transportation* 25, 129–146. https://doi.org/10.1023/A:1005004311776
12. Guo, G., Khalil, J.M., Yan, D., and Sisiopiku V. (2019). "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, Proceedings of 2019 IEEE International Conference on Big Data, pp. 4512-4519, Los Angeles, CA.
13. Guo, G., Khalil, J.M., Yan, D., and Sisiopiku V. (2019). "Realistic Transport Simulation with Open Data", BigData 2019: 6066-6068, Los Angeles, CA.
14. Jensen, M. (1999). Passion and heart in transport -- a sociological analysis on transport behaviour. Transport Policy, 6(1), 19–33.
15. Jou, R.-C. and Chen, T.-Y. (2014). *Factors affecting public transportation, car, and motorcycle usage.* Transportation Research Part A: Policy and Practice. 61, 186-198.
16. Kwan, S., & Hashim, J. (2016). A review on co-benefits of mass public transportation in climate change mitigation. *World Transit Research*. https://www.worldtransitresearch.info/research/5859/
17. Levy, J. I., Buonocore, J. J., & von Stackelberg, K. (2010). Evaluation of the public health impacts of traffic congestion: a health risk assessment. *Environmental Health*, *9*(1). https://doi.org/10.1186/1476-069x-9-65
18. Lachapelle, U., & Frank, L. D. (2009). *Transit and health: mode of transport, employer-sponsored public transit pass programs, and physical activity.* Journal of public health policy, 30(S1), S73-S94.
19. Litman, T. (2015). *Evaluating Public Transportation Health Benefits*. Victoria Transport Policy Institute For The American Public Transportation Association.

20. Litman, T. (2009).*Transportation cost and benefit analysis.* Victoria Transport Policy Institute, 31.
21. Litman, T. (2021). *Evaluating public transit benefits and costs*. 2020: Victoria Transport Policy Institute Victoria, BC, Canada.
22. Liu, Y., & Cirillo, C. (2014). Measuring transit service impacts on vehicle ownership and use. Public Transport, 7(2), 203–222. https://doi.org/10.1007/s12469-014-0098-8
23. Luk, J., & Olszewski, P., (2003). *Integrated public transport in Singapore and Hong Kong.* Road & Transport Research, 12(4), 41.
24. MATSim. *Scenario Gallery*. n.d. [cited 2020 10.19]; Available from: https://matsim.org/gallery/berlin.
25. Mulalic, I., Pilegaard, N., & Rouwendal, J. (2016, January 4). Does Improving Public Transport Decrease Car Ownership? Evidence from the Copenhagen Metropolitan Area. Papers.ssrn.com. https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2710547
26. Pachauri, R. K., Allen, M. R., Barros, V. R., Broome, J., Cramer, W., Christ, R., Church, J. A., Clarke, L., Dahe, Q., Dasgupta, P., Dubash, N. K., Edenhofer, O., Elgizouli, I., Field, C. B., Forster, P., Friedlingstein, P., Fuglestvedt, J., Gomez-Echeverri, L., Hallegatte, S., & Hegerl, G. (2014). Climate Change 2014: Synthesis Report. Contribution of Working Groups I, II and III to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change. In R. K. Pachauri & L. Meyer (Eds.), *epic.awi.de*. IPCC. https://epic.awi.de/id/eprint/37530/
27. Ramadan, OE, and **Sisiopiku, V.P.** (2019). "A Critical Review on Population Synthesis for Activity- and Agent-Based Transportation Models", In *Transportation*, Stefano De Luca and Roberta Di Pace Editors, ISBN 978-953-51-6877-5, DOI: 10.5772/intechopen.86307.
28. Shapiro, R., Hassett, K., & Arnold, F. (2002). *Conserving energy and preserving the environment: The role of public transportation.*
29. 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), December 2019.
30. Sisiopiku, V.P., and Salman, F. (2019). "Simulation Options for Modeling Shared Mobility", Proceedings of the 2019 AlaSim International Conference and Exhibition, Huntsville, AL.
31. Tao, S., He, S.Y., and Thøgersen, J. (2019). *The role of car ownership in attitudes towards public transport: A comparative study of Guangzhou and Brisbane.* Transportation Research Part F: Traffic Psychology and Behaviour, 60, 685-699.
32. TomTom. *Birmingham traffic*. 2020 [cited 2020; Available from: https://www.tomtom.com/en_gb/traffic-index/birmingham-alabama-traffic/.
33. U.S. Census Bureau. *American Community Survey 1-year estimates*. 2019 [cited 2020; Available from: https://censusreporter.org/profiles/31000US13820-birmingham-hoover-al-metro-area/.

34. Van der Merwe, J., (2011). *Agent-based transport demand modelling for the South African commuter environment*.University of Pretoria.
35. Weisbrod, G., & Reno, A. (2009). *ECONOMIC IMPACT OF PUBLIC TRANSPORTATION INVESTMENT*. https://onlinepubs.trb.org/onlinepubs/tcrp/docs/TCRPJ-11Task7-FR.pdf
36. *Adult Participation in Recommended Levels of Physical Activity --- United States, 2001 and 2003*. (n.d.). Www.cdc.gov. Retrieved January 5, 2022, from https://www.cdc.gov/mmwr/preview/mmwrhtml/mm5447a3.htm#:~:text=The%20findings%20in%20this%20report

CHAPTER 5: CASE STUDY 4 - REALISTIC URBAN TRAFFIC SIMULATION WITH RIDE-HAILING SERVICES

5.1 INTRODUCTION

The advent of shared-economy and smartphones have made on-demand transportation services possible. This brought new options to travelers but also increased the complexity to the study of urban mobility. Companies that offer these services are called Transportation Network Companies (TNCs) due to their internet-based nature. Although ride-hailing is the most known service TNCs provide, little is known about how such services interfere in traffic conditions while replacing other transportation modes. Recently, Li et al. (2022) proposed two hypotheses: (1) Uber entry reduces traffic congestion in sprawling urban areas, and (2) Uber entry increases traffic congestion in compact urban areas. Their argument is that, compared with compact areas, people in sprawling areas are faced with a more serious (first or) last mile problem because of the absence of large traffic generating centers and large separation in land uses in sprawling areas; so on-demand ride-hailing services have a greater potential to be integrated with existing bus routes to address the (first or) last mile problem, expanding the service areas of public transits. In contrast, the average trip length tends to be shorter in compact areas than sprawling areas, so people are more likely to substitute public transit trips with point-to-point on-demand ride-hailing service trips because of convenience and less waiting time. Tirachini (2020) also indicates that ride-hailing is expected to reduce parking pressure in busy areas, but results on the degree of complementarity and substitution between ride-hailing and public transport is mixed. Currently, most studies on the impact of ride-hailing services were conducted in big cities such as Shenzhen (Beojone and Geroliminis 2021), NYC (Quian et al. 2020) and San Francisco (Erhardt et al. 2019) where data are easier to collect, and the findings are that ride-hailing services intensify congestion. While these findings align with the second hypothesis of Li et al. (2022), most cities are moderate-sized and few studies have been conducted for them, let alone verifying the first hypothesis of Li et al. (2022).

There are two main reasons for the lack of more studies to address the actual impact of TNCs on traffic operations. This first is the lack of available TNC travel data and the second is the lack of commercially available simulation platforms to model TNCs operations in conjunction with other travel modes. Even though Uber and other TNC providers operate in many US markets for over a decade, up to this point, detailed TNC trip-level data have been closely guarded by privately held firms, thus limiting the opportunity for quantitative travel behavior analysis (Grahn et al. 2021). In fact, efforts to reach out to some TNC companies about accessing anonymized ride data for research purpose in our study region of Birmingham, Alabama were not fruitful as TNC providers replied that they could not share data due to legal concerns. Similar experiences have been reported for other cities such as San Francisco (Cooper et al. 2018). Some TNC companies such as Didi also experienced controversy when sharing ride data . Furthermore, and despite the fact that there is a wide variety of traffic simulation software that can evaluate operational performance of transportation networks, such tools have not yet incorporated modules to allow for integration of ride-hailing services.

Building on our earlier work to develop a prototype agent-based model for the city of Birmingham (Sisiopiku et al., 2019) and expand the model to include transit and shared mobility options (Khalil et al. 2021, Khalil et al. 2021b, Sultana et al. 2021), this case study uses novel methods to incorporate Uber trips into the Birmingham MATSim simulation model. The aim of the study is to demonstrate the feasibility of modeling ride-hailing services using an agent-based simulation platform and evaluate the impact of such services on traffic operations in the Birmingham region. To accomplish this aim, we proposed a general data collection method and matched it with an effective analytic pipeline so that the whole data collection and analytic pipeline can be generalized and used in other medium-sized cities where TNCs operate.

5.2 METHODOLOGY

In this section, we describe our approach to building a digital twin of Birmingham's transportation, which is a transportation simulation model that incorporates transportation modes such as public transit and ride-hailing services, in addition to private vehicles that constitute most of Birmingham's traffic. With this digital twin, transportation engineers can analyze the impact of ridehailing services under different scenarios, such as "if the number of Uber drivers doubles" which could happen in the near future. The digital twin may also enable new opportunities, such as being used as an environment to develop reinforcement learning models that learn good policies to regulate ride-hailing service deployment.

Our digital twin is created in three steps: (1) to fit a spatiotemporal distribution of Uber rides using the collected data via network kernel density estimation (KDE), (2) to generate travel day-plans for the Birmingham population including those containing Uber rides, and (3) to execute these day-plans on the Birmingham Road network using an agent-based simulation software MATSim[1]. The digital twin can be configured with different parameters such as the number of Uber drivers, and simulation output can be used for various downstream traffic analytic tasks. The distribution combined with population statistics from census data are then used to generate realistic Uber rides for agent-based simulation (see Section 5.2.3).

The following subsections discuss the Uber data collection and extraction process; summarizes the MATSim simulation model features, and describes the efforts undertaken by the study team to modify the Birmingham MATSim model in order to allow for simulation of automobile, transit, and Uber trips in an integrated manner.

[1] matsim.org

5.2.1 DATA COLLECTION

Due to the lack of available TNC data, transportation researchers have explored various methods to collect their own data about Uber/Lyft rides for data-driven analysis. A recent study by Beojone and Geroliminis (2021) obtained GPS coordinates of 20,000 taxis collected every 30 seconds for 20 hours in the city of Shenzhen, China. However, this method does not hold promise in the US as US Uber/Lyft drivers drive private vehicles rather than taxis (as in China). Qian et al. (Qian et al. 2020) built a crawler, which simulates the ride requesting behavior on the mobile app of Uber, to fetch the trajectories near a number of data collection stations placed with proper spacing and collection frequency in NYC. This crawler is not released and is fragile to changes of Uber API. Erhardt et al. (2019) studied the impact of TNC rides in San Francisco using data scraped by Cooper et al. (2018) from Uber and Lyft for San Francisco. However, such datasets are not available for other regions, including the Southeast. Circella et al. (2018) collected ride-hailing data with an online survey, but their study focuses on adoption statistics such as age groups and frequency of use, with little spatiotemporal information not to mention ride trajectories. Grahn et al. (2021) analyzed the interaction between TNCs and bus services in Pittsburgh, PA, by using the data of Uber surge multipliers (a pricing strategy to balance supply and demand) to approximate TNC usage (i.e., demand over supply ratio) for ten predefined points of interest throughout the city. The information, however, is insufficient to fit a distribution of Uber rides from various origins to various destinations, as is needed for a agent-based transportation simulation such as the one in the Birmingham MATSim model.

To address the TNC data issue in our study, we proposed a novel approach to collecting ride-hailing trip data with the following desirabilities: (i) the data collection procedure is general and easy to carry out in any city, without the need to survey a large number of transportation users, (ii) the data extraction procedure is easy to carry out by a layman without any domain expertise in order to facilitate crowdsourced data extraction, and (iii) the extracted data contain detailed spatiotemporal ride information, sufficient to fit a spatiotemporal distribution of ride-hailing trips.

In our study, we gathered details about Uber trips that took place in Birmingham in 2019 and 2021, before and after the surge of the COVID pandemic. The Uber trip records were obtained from Uber/Uber Eats drivers who drove in Jefferson/Shelby, AL counties during the study period. Qualified study participants received compensation for enrolling and participating in the study, according to the number of trip records that they provided. Incentive rewards were also offered for referrals of other drivers.

Our data gathering approach followed two stages: (1) *Uber driver survey*, which collected app screenshots of Uber rides from recruited Uber drivers, and (2) *ride information extraction* from app screenshots which was performed by crowdsourcing. Both stages were designed so that those involved would not need special skills to carry out their assigned responsibilities, making it easy to recruit participants and complete the necessary tasks.

Specifically, in Stage (1) we recruited Uber drivers who operated the Birmingham metropolitan region and consented to participate in the study and share their detailed trip records. Since each driver served a diverse group of travelers on demand, we only needed to recruit a small number of drivers, as compared to if we surveyed travelers directly. This made the recruitment process far more manageable and helped to ensure that the Uber trip database was large in size and diverse, with respect to trip characteristics including O-D, time of the day, and purpose. Our study approach made the task simple: After obtaining the necessary IRB approvals and consent, Uber drivers who agreed to participate met a member of our research team, who worked together with each driver to take screenshot of each ride summary in the Uber app and to send the screenshots to the research team for processing. We found this process to be more efficient, compared with letting drivers upload their ride information online on their own, which could be error-prone and deter many potential participants who are not familiar with technology. Our efforts led to the recruitment of eight Uber drivers and collection of 4,229 screenshots of Uber trips.

Stage (2) dealt with the actual data extraction. This process faced two challenges in its design. Since a screenshot is just an array of pixels: (i) crowdworkers need to extract text information manually such as "Duration" and "Time Requested" from a large number of screenshots; (ii) there is a thumbnail image of the ride trajectory overlayed on a background map in each screenshot, but a need exists to extract the exact GPS coordinates of the origin, destination, and key locations on the trajectory. We addressed challenge (i) by developing an Optical Character Recognition (OCR) tool to extract the text fields directly into our survey table, and by using crowdworkers to examine the extracted text against the screenshot in order to identify and fix any errors caused by OCR. We addressed challenge (ii) by providing detailed instructions to crowdworkers on how to use Georeferencer [2] to perform image-to-map alignment. After the screenshot image was map-aligned, crowdworkers could extract the coordinates of trajectory points directly from the map. Section 5.1.2 presents more details about our data processing approach.

[2] https://www.georeferencer.com/

5.2.2 TRAJECTORY EXTRACTION

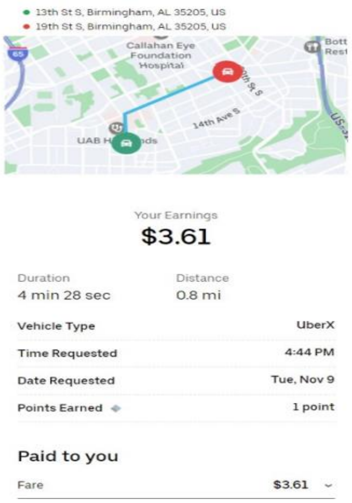


Figure 5. 13 A Screenshot of Ride Summary from the Uber App

Figure 5.1 shows a sample screenshot, which contains text information about the Uber trip origin, destination, duration, and time requested. We developed an OCR tool to help crowdworkers to automatically extract those text fields into a table and requested them to examine the extracted text against the screenshot to fix any OCR errors.

One challenge in extracting the origin and destination coordinates is that a screenshot image only provides origin and destination street names, rather than the specific locations along the streets (e.g., intersections or exact address). This is a privacy protection mechanism used by Uber to hide the exact passenger destination addresses (e.g., home addresses) from drivers. For the need of our study, we needed to recover the location coordinates with reasonable accuracy. This is possible since there is also a thumbnail image of ride trajectory included in the ride summary screenshot, which is overlayed on top of a map (see Figure 5.1). Specifically, we asked crowdworkers to use Georeferencer to first align the map of the thumbnail image to the actual map from Georeferencer using some critical points that can be easily identified, such as the three points illustrated in Figure 5.2. Such critical points can be identified by looking for blocks, parks and highways of unique shapes, once the crowdworkers zoom the map of Georeferencer into the target region covering the origin and destination streets. Figure 5.3 shows the Georeferencer interface after the screenshot image is map-aligned using the critical points, where the transparency of the screenshot image is set to 50% to allow the map beneath it to be seen.

Figure 5. 14 Critical Points for Map Alignment

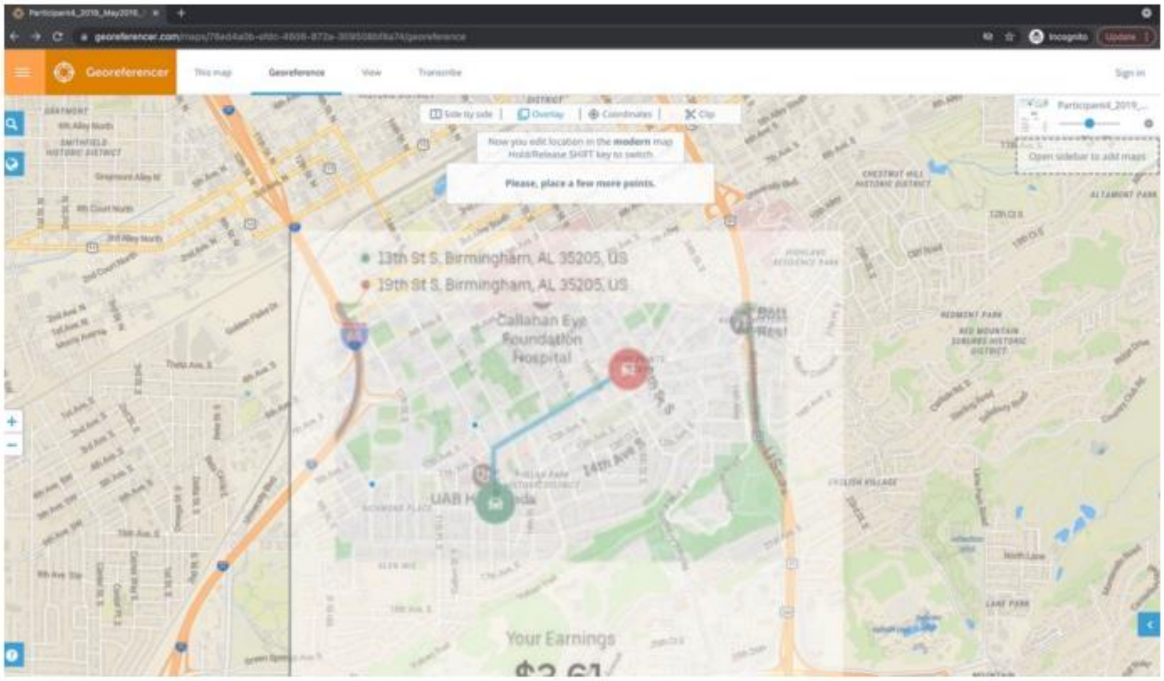


Figure 5. 15 A Map Aligned Screenshot

It is worth noting that, after alignment, we can further double check if the blue highway segments of the thumbnail image sit right on top of their orange counterparts in Georeferencer, as well as checking the alignment of other parts such as “green” areas. If the result was not visually perfect, we asked crowdworkers to re-select the critical points to repeat the map-alignment process. We also require the crowdworkers to attach a computer screenshot of their alignment like the one in Figure 5.3 as a proof of proper alignment.

Once the thumbnail image is properly map-aligned, a crowdworker may then find the origin and destination locations that are in the green and red circles, respectively, but also on the streets indicated at the top of the app screenshot image. Figure 5.4 illustrates this process, where we select the origin (resp. destination) location to be a point in the green (resp. red) circle that is also on the 13th (resp. 19th) Street South. In rare cases, where a circle does not cover the street indicated in the text, we choose the location on the street that is closest to the circle. As shown in Figure 5.5, after we click the origin (Point #3) and destination (Point #4), we can also click critical points of the blue trajectory, i.e., those trajectory locations that change directions such as Point #5. For long ride trajectories, there could be quite a few turning points and we required a crowdworker to click all of them to extract their coordinates.

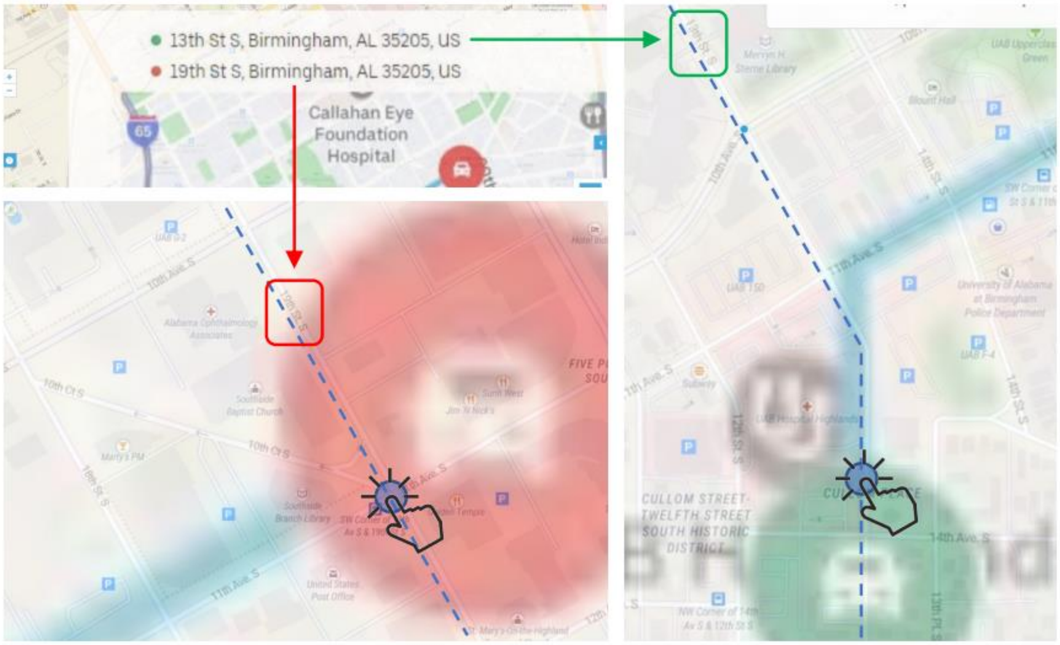


Figure 5. 16 Locating Origin and Destination

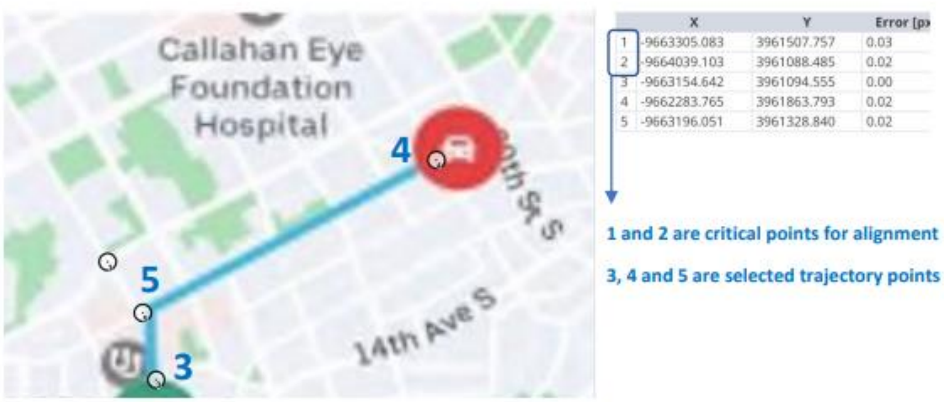


Figure 5. 17 Critical Point Extraction for a Trajectory

Note from Figure 5.5 that Georeferencer uses the coordinate system of Pseudo Mercator projection (EPSG:3857). Thus, we convert the extracted coordinates to GPS coordinates (EPSG:4326) using Python's "pyproj" library, to facilitate map matching later that recovers the entire trajectory from the sequence of critical points using OpenStreetMap's road network (which is based on the GPS coordinate system).

While one can use any crowdsourcing platform such as Amazon Mechanical Turk to conduct the extraction tasks, we assigned the tasks to the students taking CS 685/785: "Foundations of Data Science" at UAB's Department of Computer Science as a voluntary bonus assignment. The class size was around 200 and 155 students opted in. We divided the screenshot images evenly among the students so that each image was covered by 3 students. Some collected images were problematic such as missing the destination, or corresponding to canceled rides, so we removed them and got only 3,922 valid screenshot images out of the 4,229 collected, among which most were UberX rides (2,632), followed by 1,061 delivery rides and 119 order-and-pay rides from Uber Eats. There were also a small number of other ride types such as UberXL and Uber Pet. We provided detailed project instructions to the 155 students and taught them how to effectively use Georeferencer with a live in-class demo that was also recorded for later watching. We also allowed the students to do one resubmission to fix the issues we identified after examining their first submissions, and each final screenshot extraction was given a quality rating such as good, usable, or unusable. Most screenshot images had at least one good/usable extraction, but a small number of images had no usable extraction for which our research team then worked on their extraction by ourselves.

After obtaining all the extracted ride trajectories, we mapped them to the underlying Birmingham area road network obtained from OpenStreetMap, using the FMM[3] library for trajectory map matching. Figure 5.6 shows a downtown-to-airport raw trajectory extracted as a sequence of points plotted with a red polyline, and the map-matched trajectory plotted with a green line. We can see that the map-matched trajectory is along the roads and highways, which is properly recovered from the sparse set of raw trajectory points.

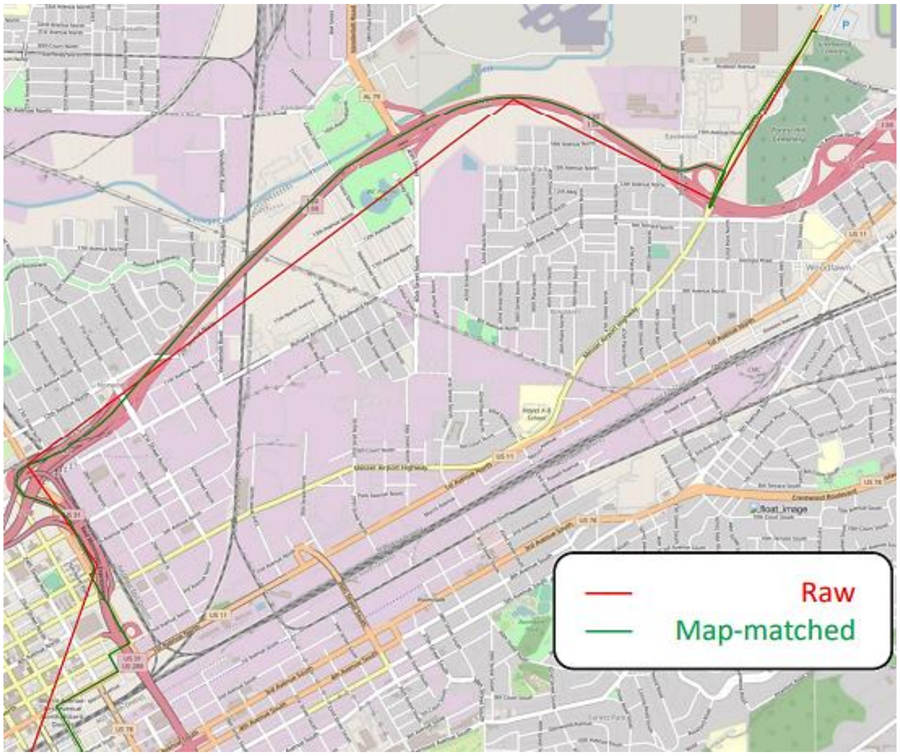


Figure 5. 18 A Raw Trajectory and its Map-Matched Version

[3] https://github.com/cyang-kth/fmm

We next present some visualization results to summarize our extracted trajectories. We uniformly sampled 100,000 points from the trajectories after map matching and plotted a heat map using Python's Folium library. Figure 5.7 shows the heat map for 2 different zoom levels, where we can observe that the traffic hotspots of Uber rides are along the key highway corridors clustered towards the downtown of Birmingham. We also grouped rides by origin-destination (OD) zipcode areas (defined by TigerWeb) and created an outbound and an inbound trip D3.js interactive visualization.

Figure 5.8 shows the outbound trip visualization, where a count-threshold bar is used to filter out edges (origin-zipcode, destination-zipcode, count) with low counts. For visualization purposes it is important to note that the darker blue an area is, the more Uber ride origins it contains. The large figure displayed in Figure 5.8 shows the outbound edges from zipcode area 35233 (UAB) highlighted in red. The small figure insert in Figure 5.8 shows the impact from raising the count-threshold to the highest level, where we can see that the UAB campus and the Birmingham airport area are two biggest Uber ride hotspots.

5.2.3 NETWORK KERNEL DENSITY ESTIMATION (KDE)

While Kernel density estimation (KDE) can be generalized to 2D or 3D Euclidean space, our application of transportation simulation requires that origin and destination locations are on a road network, as an arbitrary location on the 2D map may be invalid (e.g., resulting in an origin or destination appearing in the middle of a lake). Xie and Yan (2008) proposed a network KDE approach to estimate the density of traffic accidents strictly over a road network, based on a set of historical traffic accident events with locations. This work was later extended by Romano and Jiang (2017) to incorporate the temporal domain.

Figure 5.9 shows a road network fragment with two observed sample locations; assume that each sample contributes to the density of road network locations within radius r = 20 m from it (in terms of shortest network path distance), then the red segments around each sample in Figure 5.9 indicates its impacted locations.

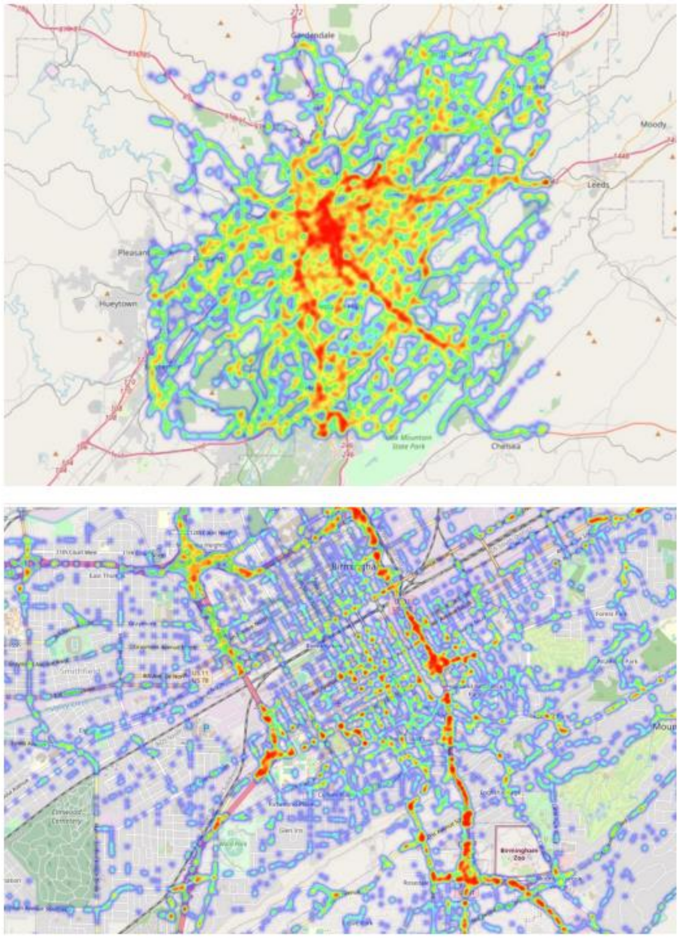


Figure 5. 19 Heat Maps of Trajectory Points for 2 Zoom Levels

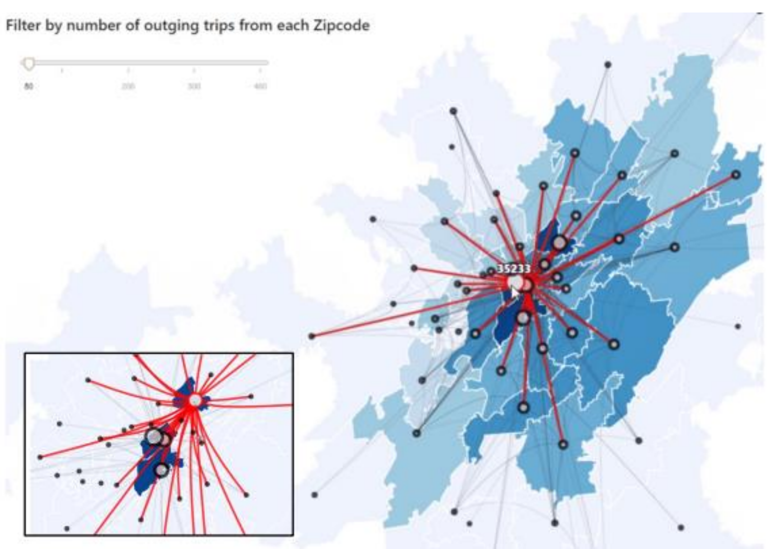


Figure 5. 20 Critical Point Extraction for a Trajectory

Both Xie and Yan (2008) and Romano and Jiang (2017) discretize a road network into small units called lixels (linear pixels) to improve the computation efficiency, an approach which we also followed in our work. Figure 5.10 illustrates the lixel network created from the road network in Figure 5.9, where a basic lixel unit is a line segment of length of 5 m. A road segment is decomposed into lixels of length of 5 m, except for the last one whose length could be less than 5 m and called a residual.

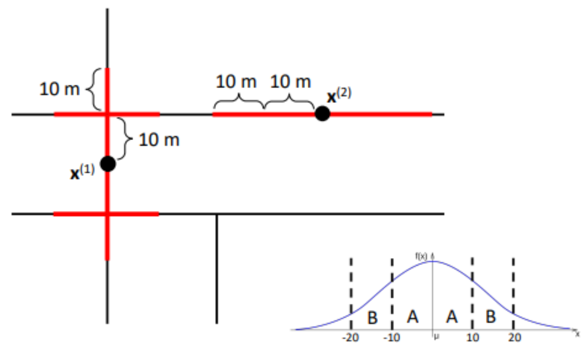


Figure 5. 21 Road Network Weight Function

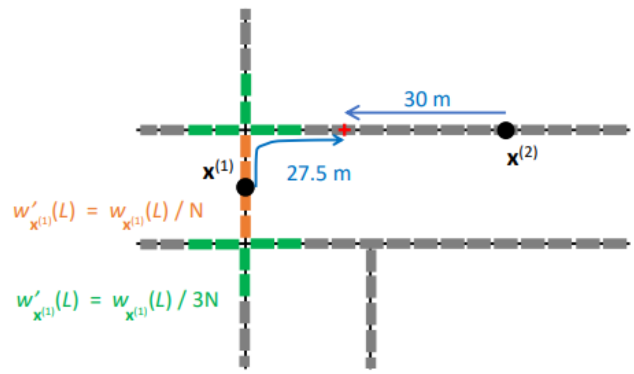


Figure 5. 22 The Lixel Network

Figure 5.11 (left side) shows the network KDE results fit from the Uber pickup locations when we use Gaussian kernels with σ = 100 m, 200 m, and 500 m plotted using QGIS. There, let us denote the KDE density times 104 by ρ, then blue (resp. orange, red) means $\rho \in \{0.0, 0.1, 0.2\}$ (resp. {0.3, 0.4, 0.5, 0.6}, {0.7, 0.8, 0.9}) with gradient color change (e.g., the color of 0.25 is the mean of red and orange in Red Green Blue- RGB values). Figure 5.11 shows lixel distributions fit by our spatial KDE for Birmingham, where the lower left corner is the UAB campus and the upper right corner is the Birmingham airport. We can see that while σ = 100 m is very concentrated at the pickup locations collected by our survey, σ = 200 m already spreads out the orange density portions to cover most of the urban road network in the Birmingham study region.

Figure 5.11 (right side) shows arixel distributions fit by our spatiotemproal KDE for Birmingham, where we use the “Box” plot of the Qgis2threejs plug-in to show the arixel densities around UAB Green which is located at the center of the UAB campus, for five 1-hour long periods [7:00, 8:00], [8:00, 9:00], [9:00, 10:00], [10:00, 11:00], and [11:00, 12:00]. We can see that σ = 200 m still spreads out the density well, even though the temporal dimension makes space sparser. As a result, we set σ = 200 m as default.

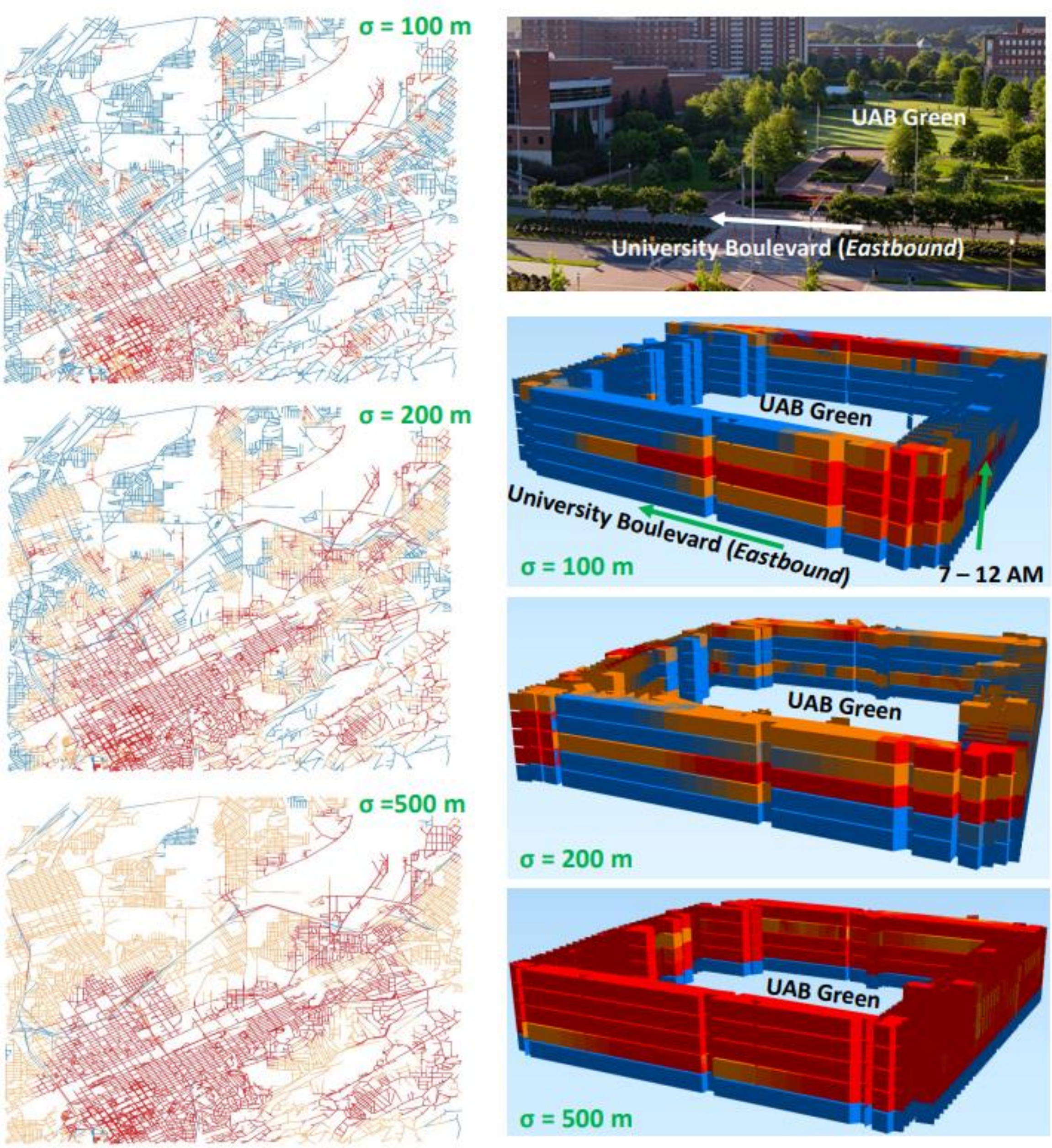


Figure 5. 23 KDE Distributions for Lixels and Arixels

5.2.4 MATSim REVIEW

As stated earlier, the simulation model of choice for this case study was the Multi-Agent Transport Simulation (MATSim). The comparison of features, capabilities, and limitations of many transportation simulation options performed by Sisiopiku and Salman (2019) identified MATSim as the most promising and well-established available platform for modeling shared mobility and ridesourcing services (such as Uber and Lyft).

As discussed also in Chapters 3 and 4, MATSim takes the travel day plans of a population and executes them on the underlying road network (obtained from OpenStreetMap) to generate simulated traffic. Each simulation job of MATSim executes in iterations where each iteration executes the selected travel plans of all agents over an underlying road network. Figure 5.12 shows the execution flow of a MATSim job. It starts with an initial population demand (aka. 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, the MATSim's "mobsim" simulation executor first runs the selected plans of the agents in the synthetic road network environment. Next, a score is assigned to each executed plan based on the agent's experience. Then, the replanning step selects a candidate plan for each agent based on the plan scores in the agent's plan pool, and may modify (or, mutate) this plan for execution in the next iteration.

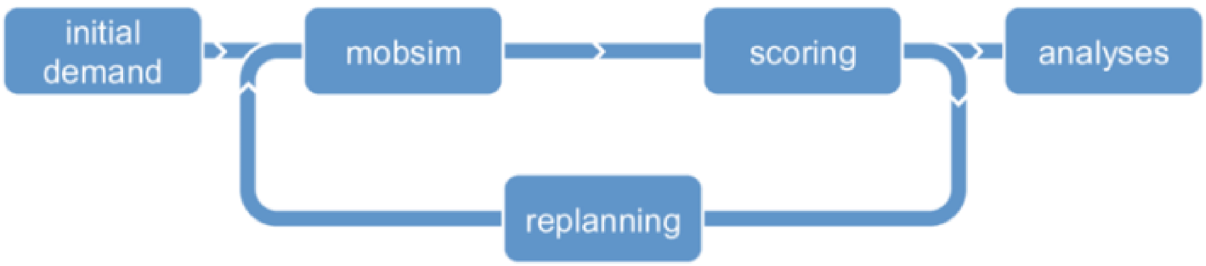


Figure 5. 24 The Co-Evolutionary Algorithm of MATSim

A key challenge in utilizing the MATSim platform for traffic simulation modeling is how to generate a realistic synthetic population along with their travel day-plans, a required input to MATSim. In our prior work (Guo et al. 2019; Guo et al. 2019b; and Sisiopiku et al. 2019), 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 only considers private vehicles, which accounts for the vast majority of Birmingham's traffic. An extension of this work was presented in Chapter 4, where transit trips were incorporated into the Birmingham MATSim simulation model. In the case study described herein, we explored how we can use those private-vehicle day-plans as the background traffic and add ride-hailing and public transportation trips into our simulation, in order to create an enhanced Birmingham MATSim model that allows modeling of Uber, public transit, and private vehicle trips in an integrated manner.

5.3 AGENT-BASED TRANSPORT SIMULATION OF THE BIRMINGHAM NETWORK

This section explains how we conducted agent-based transportation simulation of the Birmingham network by integrating the knowledge obtained from our Uber driver survey. We use a multi-agent simulation software called MATSim (www.matsim.org) to build a digital twin of Birmingham's transportation. In our digital twin, we reused the realistic day-plans of travel mode “private vehicles” that we created for the population in Birmingham in our prior work (Sisiopiku et al. 2019, Guo et al. 2019; Guo et al. 2019b) as the background traffic and generated new day-plans for transportation users who do not use private vehicles (e.g., carless travelers including UAB student and those living in poverty, visitors, etc.). These transportation users have the mode options of (1) taking a bus, (2) requesting an Uber ride, and (3) walking. Each agent selects a mode based on the plan scores. For example, an Uber ride could be preferred for long-distance traveling since the travel time is much shorter than taking a bus or walking, which contributes to a better plan score in MATSim. In contrast, walking could be preferred for a short-distance travel since the time to wait for the Uber driver to pick up is saved. Note, however, that the outputted final plans after the co-evolutionary algorithm of MATSim (see Figure 5.12) may have a different mode from the initial ones, since agents may switch between alternative travel modes based on previously executed plan scores.

MATSim supports public transit mode but requires the transit schedule as an input. In Birmingham, MAX buses are currently the only public transit option, and their schedules are available[4] as GTFS files that can be inputted to MATSim using the GTFS2MATSim[5] tool. More details on incorporation of public transit trips into the Birmingham MATSim are available in Chapter 4. We next explain how we configured MATSim and generated initial travel day-plans for Uber trips.

5.3.1 GENERATING INITIAL UBER RIDES

Let us assume that the total number of Uber rides in a day is N. Our goal is to generate N Uber ride plans of the form (p_s, t_s, p_e), where p_s (resp. p_e) is the passenger pickup (resp. drop-off) location, and ts is the Uber request time. To ensure that the generated Uber ride plans match the spatiotemporal ride distributions from our Uber driver survey, we fit a spatiotemporal distribution of (p_s, t_s) using the spatiotemporal KDE approach described in Section 5.2.3. We denote this distribution as P_s, which is basically a set of arixels each with a probability mass. Sampling P_s samples an arixel respecting the probability mass function defined by the discrete P_s arixel distribution. We also generate a spatiotemporal distribution P_e using spatiotemporal KDE, from the set of drop-off locations of the Uber rides in our survey, where we estimate the time of drop-off by adding “Duration” to “Time Requested”.

From now on, let us denote P_s (resp. P_e) as a weighted arixel set {(p_s, t_s), w_s} (resp. {(p_e, t_e), w_e}) for ease of presentation. An easiy way to create a ride (p_s, t_s, p_e) is to sample (p_s, t_s) from P_s and sample (p_e, t_e) from P_e. However, this simplistic approach has some issues:

- Arixels are associated with 1-hour time intervals rather than a specific time, which we address by uniformly sampling a time in the interval (e.g., 8:20 am for [8 am, 9 am)).

[4] https://maxtransit.org/rider-tools/gtfs-information
[5] https://github.com/matsim-org/GTFS2MATSim

- It may generate a ride out of two arixels where t_s is 9 am but t_e is 8 am or 5 pm, which is unrealistic. We want to ensure that $t_e > t_s$ and the trip duration ($t_e - t_s$) is not too long (e.g., within 1 hour) as is typical for an Uber ride in Birmingham.
- We only surveyed a limited number of Uber drivers each with preference on the service regions (e.g., near their home), so some areas in Birmingham may not be sufficiently covered.

To address the issue of survey data sparsity, we smooth the origin-destination (OD) matrix of Uber rides from our survey data with the background population OD matrix. Recall that we grouped Uber rides by origin-destination (OD) zipcode areas and let us denote this OD matrix by A_{uber} where $A_{ub}[i][j]$ keeps the number of rides from zipcode area i to zipcode area j. We also obtain another OD matrix of the background Birmingham population, denoted by A_{pop}. Here, $A_{pop}[i]\ [j] = n_i \times n_j$, where n_i denotes the population size in zipcode area i and is obtained from the American Community Survey (ACS)[6] of the US Census Bureau, using Variable B01003 (Total Population) obtained at the ZIP Code Tabulation Area (ZCTA) level. ZCTA is US Census Bureau's terminology for zipcode area, so we will use it in our discussion hereafter.

We generate each of the n_{ij} Uber rides as follows:

- Sample (p_s, t_s) from $P_s[i]$;
- Uniformly sample a time τ_s from the 1-hour interval t_s;
- Remove all arixels (p_e, t_e) of $P_e[j]$ with $t_e < t_s$ or $t_e > t_s + 1$, and denote the pruned distribution by $P_e{}^*[j]$; and
- Sample (p_e, t_e) from $P_e{}^*[j]$.

In this way, we obtain an Uber ride plan (p_s, τ_s, p_e) where p_s is in ZCTA i, p_e is in ZCTA j, and the sampled destination arixel has a drop-off time that is neither before t_s, nor beyond 1 hour after t_s. One issue remains: it is possible that in Step 3, we cannot find any valid arixel for $P_e{}^*[j]$ from $P_e[j]$, in which case we can repeat these steps again to resample. However, if the probability of this sampling failure is high, the procedure can get stuck in resampling in hope to generate a valid Uber ride plan. Fortunately, this failure rate is low in our case, and we set the maximum number of resampling trials to 2 after which we simply give up generating this plan. With N = 3200 total initial Uber ride plans to generate, our sampling method generates 3134 plans without resampling, 38 plans with 1 resampling trial, 13 plans with 2 resampling trials; only 15 plans are not successfully generated due to failing 2 resampling trials, which is a very small fraction that can be safely ignored.

[6] https://www.census.gov/programs-surveys/acs

5.3.2 BIRMINGHAM MATSim CONFIGURATION FOR PUBLIC TRANSIT AND UBER DRIVERS

MATSim supports public transit mode but requires the transit schedule as an input. In Birmingham, buses are the only public transit, and their schedules are available[7] as GTFS files that can be inputted to MATSim using the GTFS2MATSim[8] tool. In Birmingham, buses are not a transportation mode of choice because service is limited, inconvenient and infrequent. Most current transit users are captive riders, typically low-income residents who cannot afford private vehicles. To model this population, we replace those background day-plans that originated from poor ZCTAs (i.e., ZCTAs where at least 1,000 people living in property, obtained using ACS Variable C17002 "Ratio of Income to Poverty Level in the Past 12 Months") with bus day-plans, and when choosing POIs (e.g., grocery stores, banks; POIs are obtained from OpenStreetMap) for a bus-mode day-plan, we only consider those POIs within 1 km from a bus station, so that the POIs are reachable by walking.

MATSim uses its dynamic vehicle routing package (DVRP) to solve dynamic vehicle routing problem. To provide tdemand-responsive transport service functionality, two extensions were developed, namely Dynamic Transport Services (DRT) and Taxi. The difference is that DRT allows rides to be shared, and the occupied vehicles can be diverted from the current destination to pick up new passengers on the way. However, ride sharing is not supported by Uber service in the Birmingham metropolitan area, so we decide to use MATSim's Taxi (org.matsim.contrib.taxi) extension. This extension requires us to specify the number of drivers and their starting locations, which we generate by sampling from P_S.

5.4 EXPERIMENTS

This section shows how our digital twin can be used to study the impact of Uber services in Birmingham. While no TNC data are available, we estimate the number of daily Uber rides in Birmingham to be 3,500 - 4,500 trips daily and we generate 3,200 initial Uber ride plans using the approach described in Section 5.3.2. Transportation researchers may adjust this number as needed in our digital twin, if a more accurate estimate becomes available.

We tried different number of Uber drivers to configure MATSim's Taxi extension, including 50, 100, 200, 400 and 800 drivers. Note that when there are only 50 Uber drivers, in order to complete the 3,200 daily Uber requests, each driver needs to serve as many as 64 rides. This is unrealistic and as a result, a lot of the Uber ride plans cannot complete at the end of the day. Figure 5.13 shows the number of Uber ride plans that are en-route (requested but not completed), departed, and arrived in each hour. We set Uber drivers to stop working after the 24th hour, so next-day requests are not served and the green curve for en-route plan number becomes flat. We can see that when there are only 50 or 100 drivers, most Uber ride requests cannot be completed. When the number of drivers is 200, most requests are satisfied at the end

[7] https://maxtransit.org/rider-tools/gtfs-information
[8] https://github.com/matsim-org/GTFS2MATSim

of the day but there are still a non-negligible number of unserved requests. Finally, when there are 400 drivers, the requests can be cleared at the end of the day. Given these observations, we considered 200, 400, and 800 Uber drivers for the remaining analysis.

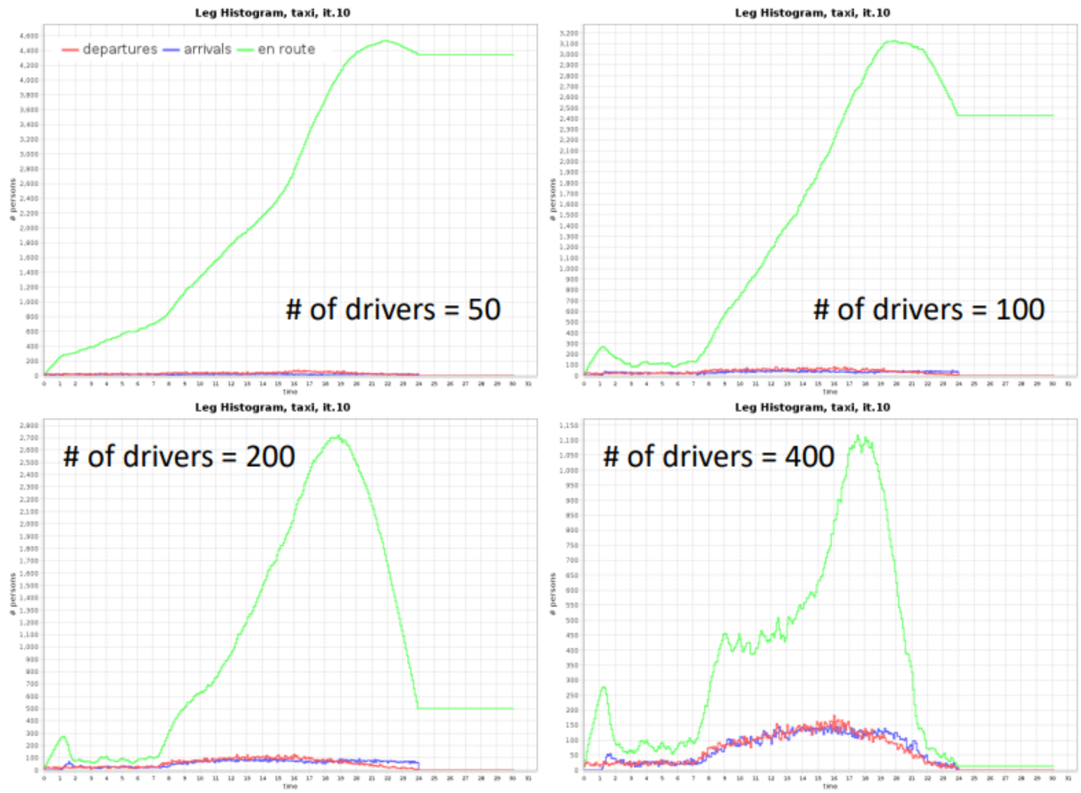


Figure 5. 25 Number of Rides at Different Status by Hours

Table 5.1 shows the statistics of our executed plans outputted by MATSim using a population size of 69,826, where 56,252 are private-vehicle trips, and the remaining are Uber rides, bus rides and walking trips that may switch between each other based on the scores of plan execution. Initially, we only generate 3200 Uber ride plans, but more Uber rides may appear in the simulation output, which are switched from the initial bus ride plans.

Table 5. 10: Statistics of Executed Plans-Trips by Mode

# of Drivers	Uber Trips	Bus Trips	Walk Trips	Car Trips
200	7,997 (11.5%)	3,339 (4.8%)	2,238 (3.2%)	56,252 (80.6%)
400	10,353 (14.8%)	2,418 (3.5%)	803 (1.2%)	56,252 (80.6%)
800	10,691 (15.3%)	2,212 (3.2%)	617 (1.0%)	56,252 (80.6%)

From Table 5.1, we can observe a significant increment in the number of Uber rides when the number of Uber drivers increases from 200 to 400. The modal shift from bus and walking modes toward Uber is due to better plan scores resulted from a shorter travel time. This matches with intuition since when there not enough Uber drivers in a city, the waiting time for pickup is too long to persuade people to use ride-hailing services. In contrast, we only observe a marginal increment in the number of Uber rides when the driver number increases from 400 to 800, indicating that the demand for Uber service almost saturates and more Uber drivers would not significantly increase the total revenue they earn collectively.

Figure 5.14 reports on the Uber driver status (i.e., empty drive, occupied drive; pickup; drop off; idle status) on an hour-by-hour basis. We can see that with 200 Uber drivers, most drivers are occupied (gray) especially between 8:00AM and 9:00PM, while most drivers are idle (green) staying at their last drop-off locations outside those times. With 400 drivers, Uber services are saturated in the afternoon, while many drivers are not occupied in the morning. A similar trend is observed in the case of 800 Uber drivers, with a peak occurring from 4:00 PM to 7:00 PM. Figure 5.14 also shows, that the Birmingham Uber market can optimally support a maximum of around 530 Uber drivers, even at its peak time that between 5:00 and 6:00PM.

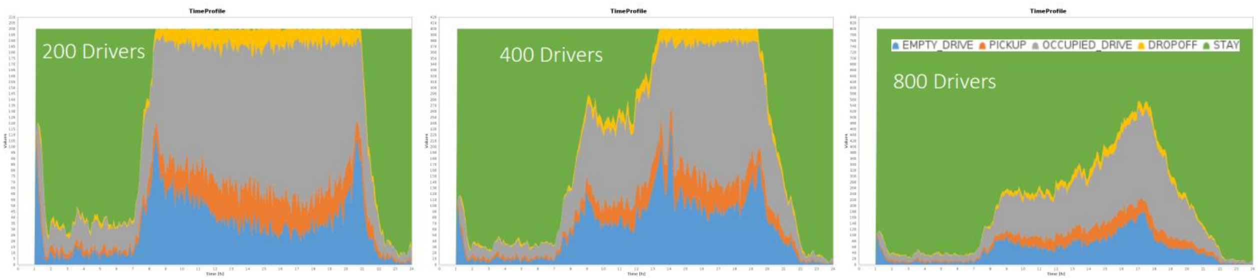


Figure 5. 26 Uber Driver Status Statistics

Results from the Birmingham MATSim model were also used to evaluate the operational performance of transportation facilities. As an example, Figures 5.15 and 5.16 report the hourly average speed and hourly vehicle volume, respectively, for U.S. Highway 31 near UAB in downtown Birmingham. The results show some gains in speed when Uber services are available, especially during peak hours (e.g., 10:00AM, 4:00PM–6:00PM). Examination of Figure 5.16 confirms that most hourly volumes also increased in the presence of Uber services, especially during the morning peak time (7:00AM–9:00AM). The results on other highways and major roadways follow similar trends, indicating that Uber operations result in small improvements in network performance. These findings indicate that ride-hailing services actually help reduce the congestion in Birmingham (a moderate-sized city) during peak hours, which is a finding different from those reported by prior works that evaluated TNC operational impacts in big cities which suggested increase in congestion due to the availability of TNC services.

5.5 DISCUSSION AND CONCLUSION

In this case study, we presented a digital twin of transportation in the Birmingham metropolitan area, created out of an Uber driver survey and other open data sources. The goal was to create a realistic transportation simulation model for transportation engineering researchers to explore the impact of different services and conditions to inform better decision making in transportation planning. This is different from prior traffic generators such as MNTG (Mokbel et al. 2013) that create valid traffic on road networks from user-specified parameters with the aim to benchmark spatial query processing algorithms.

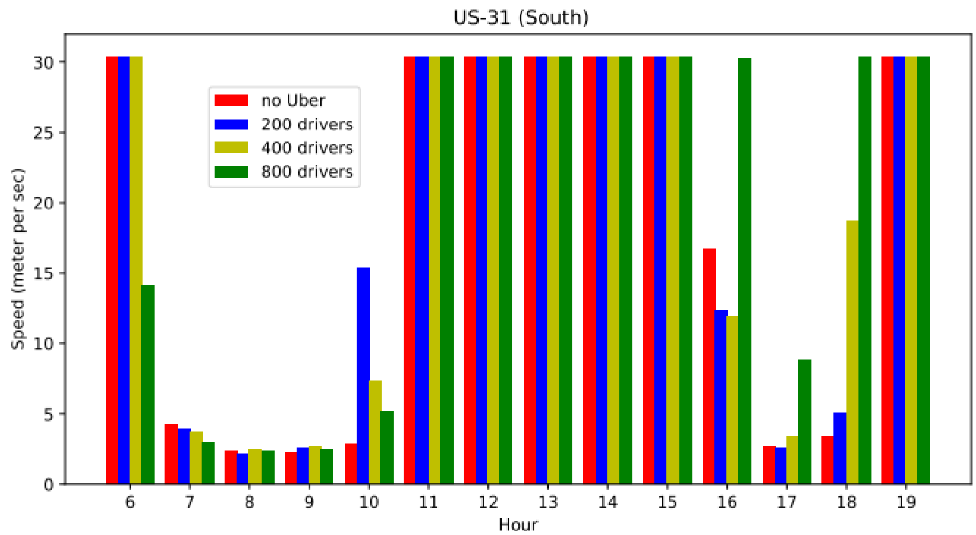


Figure 5. 27 Hourly Average Speed of U.S. Highway 31

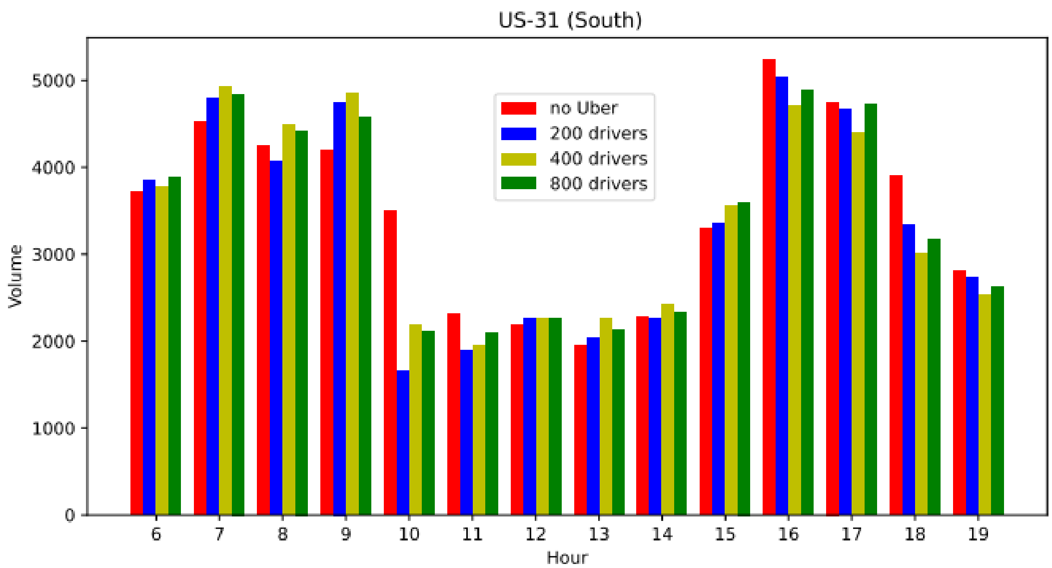
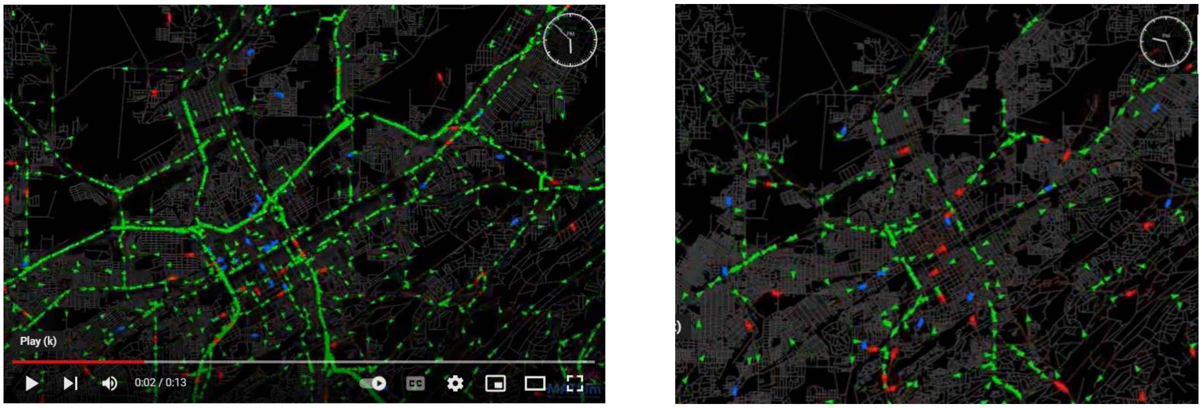


Figure 5. 28 Hourly Vehicle Volume of U.S. Highway 31

On the methodological level, our work introduced a new approach for collecting detailed Uber trip data along with improvements on the previous version of the Birmingham MATSim model, which allowed us to model automobile, transit, and Uber trips all at once. Figure 5.17

shows a screenshot from the Birmingham MATSim simulation visualization[9] showing generated trips in the Birmingham downtown area, color coded by mode. It should be noted that the work presented here shows modal shifts from transit and walk trips toward Uber, when there are enough Uber drivers operating to keep waiting times low and can help determine optimal number of Uber drivers on an hour-by-hour basis. It also provides information on the operational performance of the transportation network when TNC services are present and can be used as a planning or policy tool.



Red: Uber **Blue:** Public Transit **Green:** Private Cars; Birmingham Downtown; 5:50PM (L) and 9:25PM (R)

Figure 5. 29 Traffic Simulation of Birmingham Downtown

Our main contributions are summarized as follows:

- We proposed a general data collection approach that is easy to carry out in any city, where only a small number of TNC drivers need to be surveyed to collect their ride summary screenshots in TNC apps.
- We proposed an effective crowdsourcing approach to extract spatiotemporal ride information from app screenshots of TNC drivers, without additional skills from crowdworkers.
- We used the network KDE method to effectively fit the TNC ride distribution from the extracted data.
- We designed algorithms to generate realistic travel day-plans for the entire Birmingham population, using the learned TNC ride distribution and other open data sources such as US Census Bureau, which enables effective agent-based simulation on MATSim.

In future work, we plan to explore the impact of shared mobility on transportation network operations by considering (a) shifts from automobile trips toward TNC in the presence of UberPOOL (i.e., where riders share the same Uber vehicle with other riders), and (b) shifts

[9] https://www.youtube.com/watch?v=NGCIqhzRFo4

from automobile trips toward transit in an integrated TNC and public transit scenario where Uber service can be used to address the first mile/last mile problem.

5.6 REFERENCES

1. Mokbel, M.F., Alarabi, L., Bao, J., Eldawy, A., Magdy, A., Sarwat, M., Waytas, E. and Yackel, S. 2013. MNTG: An Extensible Web-Based Traffic Generator. In SSTD (Lecture Notes in Computer Science), Vol. 8098. Springer, 38–55.
2. Qian, X., Lei, T., Xue, J., Lei, Z. and Ukkusuri, S.V. 2020. Impact of transportation network companies on urban congestion: Evidence from large-scale trajectory data. Sustainable Cities and Society 55 (2020), 102053.
3. Romano, B., and Jiang, Z. 2017. Visualizing Traffic Accident Hotspots Based on Spatial-Temporal Network Kernel Density Estimation. In SIGSPATIAL. ACM, 98:1–98:4.
4. Sisiopiku, V.P., and Salman, F. (2019). "Simulation Options for Modeling Shared Mobility", Proceedings of the 2019 AlaSim International Conference and Exhibition, Huntsville, AL.
5. 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), December 2019.
6. Sultana, T., Sisiopiku, V.P., Khalil, J. and Yan, D. 2021. Potential Benefits of Increased Public Transit Ridership in Medium Sized Cities: A Case Study. Journal of Transportation Technologies 12, 1 (2021), 59–79.
7. Tirachini, A. 2020. Ride-hailing, travel behaviour and sustainable mobility: an international review. Transportation 47, 4 (2020), 2011–2047.
8. Xie, Z. and Yan, J. 2008. Kernel Density Estimation of traffic accidents in a network space. Comput. Environ. Urban Syst. 32, 5 (2008), 396–406.

6.0 SUMMARY FINDINGS AND STUDY CONTRIBUTIONS & LIMITATIONS

This project conducted a comprehensive study to examine the influence of transit- and TNC operations on individual travelers' mode choices and assess the impact of such choices on transportation network performance. Four case studies have been completed in this project:

In *Case Study 1* we analyzed 790 questionnaire surveys of transportation system users across the Birmingham and Miami Beach regions with the purpose of establishing links between travel behaviors and TNC use in the Southeastern US. The survey responses confirmed that TNC service coverage and user characteristics influence user perceptions and adoption of such services. In the Birmingham region, TNC use was strongly correlated with vehicle availability and waiting time. In the Miami Beach region, variables identified as significant predictors of TNC use included vehicle ownership, vehicle availability, availability of Uber/Lyft, age range (18-29), high income (>$75K), and residency. Overall, the case study findings established valuable links between travel behaviors and TNC use. These can inform transportation agencies about the needs and opportunities for TNC services in the local markets and can guide future efforts to create targeted marketing plans in support of mode switching toward shared modes, including TNCs.

In *Case Study 2* we developed a base model of the Miami Beach network using the multi-agent simulation model (MATSim). The model was calibrated extensively using a regionally approved mode split behavioral model and real-world traffic counts. We used the calibrated Miami Beach MATSim model to estimate modal shifts between the passenger car and transit services due to the introduction of (a) an enhanced transit alternative and (b) road pricing, i.e., a toll fee for using a nearby highway facility. The case study findings confirmed that the modal shift towards transit is greater when the new transit option is introduced in combination with a road pricing strategy (4.1% shift towards transit), rather than alone (0.5% shift towards transit). Another contribution of this case study was to demonstrate a method for calibration of a MATSim model using a regionally approved mode split behavioral model and real-world traffic counts.

In *Case Study 3* we examined potential impacts from expanding public transit options in the Birmingham region on network performance. Performance measures considered include traffic volumes, speeds, and travel times. We used a comprehensive activity-based simulation model of Birmingham developed in MATSim to simulate traffic operations under various transit ridership scenarios ranging from 0% (base) to 1.1% (scenario 1-current) to 5.7% (scenario 2-future) to 10.1% (scenario 3-future). The findings identified ridership scenarios, demand levels, and time periods when such effects on traffic operations were significant. Overall, this case study highlighted the potential benefits of increased bus ridership on transportation network operations in medium-size US cities such as Birmingham.

Case Study 4 demonstrated the feasibility of modeling TNC services using the MATSim agent-based simulation platform and evaluated the impact of such services on traffic operations in the Birmingham region. The findings of the study helped to determine the optimal number of Uber vehicles in the network on an hour-by-hour basis, given the varying demand for TNC service. Moreover, examination of speed and volume data confirmed that the availability of Uber services did not cause any additional congestion in Birmingham, compared to the base case scenario (no TNC service). This case study showcased the development of a realistic multi modal transportation simulation model in MATSim that transportation engineers can used to explore the impact of different services on network operations. Such information can benefit transportation agencies and TNC industry partners alike and facilitate better planning and decision making in the future.

In conclusion, the study findings expand our understanding on TNC use, the influence of TNCs on urban mobility, and opportunities for TNC integration with other modes in mid-size cities. Furthermore, the methods proposed and tested in this study can be used in other medium-sized cities to understand travelers' mode choices in the presence of public transit and TNC services, and access the impacts of such preferences on their travel patterns and, in turn, the performance of the transportation network.

Spectrum bias and self-selection bias may have impacted the findings of our study, as is often the case with studies that collect information about perceptions, preferences, and choices of transportation users. Therefore, we caution that study findings are generalizable only to groups with a similar demographic profile as the one described in our research. Also, the scope of this study was limited to Birmingham, AL and Miami Beach, FL. While the methods proposed and tested in this project may be transferable to other locations, the actual findings may be influenced by local conditions in the study test beds. Thus, future studies at additional locations where TNC services are available are recommended to help researchers understand similarities and differences in TNC use and impacts on traffic congestion due to local conditions in the Southeast and across the nation.

7.0 APPENDICES

7.1 Appendix A – Acronyms

ABM	Activity Based Model
ACS	American Community Survey
APIs	Application Programming Interfaces
APTA	American Public Transportation Association
CBD	Central Business District
DRT	Dynamic Transport Services
DVRP	Dynamic Vehicle Routing Package
KDE	Kernel Density Estimation
Lasso	Least Absolute Shrinkage and Selection Operator
MATSim	Multi Agent Transport Simulation
MDT	Miami-Dade Transit
MPE	Mean Percentage Error
NLP	Non-Linear Programming
OD	Origin-Destination
OCR	Optical Character Recognition
PT	Public Transit
SERPM	Southeast Florida Regional Planning Model
STRIDE	Southeastern Transportation Research, Innovation, Development and Education Center
SSE	Sum Squared Error
TAZ	Traffic Analysis Zones
TNCs	Transportation Network Companies
US	United States
VMT	Vehicle Miles Traveled
ZCTA	ZIP Code Tabulation Area

7.2 Appendix B – Associated websites, data, etc., produced

Project data has been uploaded to Zenodo.

7.2 Appendix C –Survey Instrument – Birmingham, AL

BIRMINGHAM SURVEY

1/24/2019 Qualtrics Survey Software

Informed Consent

Welcome to the UAB travel diary survey!

Dr. Virginia Sisiopiku (UAB) invites you to be part of a research project that studies technology influence on travel demand and behavior. Your feedback is very important, as it will help UAB researchers to understand and model travel behavior in the Birmingham region.

If you agree to participate, you will be asked to complete a survey about your travel preferences and practices as you travel on a typical weekday in and around Birmingham. The survey should take approximately 10 minutes to complete and your participation is voluntary. Please be assured that your responses will be kept completely confidential and exempt from public disclosure by law. Please note that this survey will be best displayed on a laptop or desktop computer. While you can complete the survey on a mobile device, some features may be less compatible for use on a mobile device.

Your kind assistance in providing input through the completion of this survey is greatly appreciated. If you have questions about the survey or research study, you can contact Dr. Sisiopiku, UAB, Civil, Construction, and Environmental Engineering, Birmingham, AL 35294, or via email at vps@uab.edu.

If you have questions about your rights as a research participant, or concerns or complaints about the research, you may contact the UAB Office of the IRB (OIRB) at 205-934-3789 or toll free at 1-855-860-3789. Regular hours for the OIRB are 8:00 a.m. to 5:00 p.m. CT, Monday through Friday.

By clicking the consent button below, you acknowledge that your participation in the study is voluntary, you are 18 years of age, and that you are aware that you may choose to terminate your participation in the study at any time and for any reason.

- ○ I consent, begin the study
- ○ I do not consent, I do not wish to participate

ZIP Code Validation

Home ZIP Code

Travel Preferences

I have used the following in the Birmingham region at least once in the past year:
Check all that apply

- ☐ Transportation Network Companies (Uber, Lyft, etc.)

https://uab.co1.qualtrics.com/WRQualtricsControlPanel/Ajax.php?action=GetSurveyPrintPreview 1/8

- Public Transit
- Organized ride sharing program
- Bicycle
- None of the above

Last trip with Transportation Network Companies (Uber, Lyft, etc.)

- Within the past 7 days
- Within the past 30 days
- Within the past two months
- Within the past 6 months
- Within the past year

Reason(s) for using Transportation Network Companies (Uber, Lyft, etc.)
Check all that apply

- Convenience
- Cheaper than other alternatives
- Destination has little or no parking availability
- Parking at destination is expensive
- Safety/to avoid driving under the influence
- Car is not available
- Transit is not accessible
- Transit is not reliable
- Other modes are not available
- Other reason (fill in)

Trip purpose(s) for using Transportation Network Companies (Uber, Lyft, etc.)
Check all that apply

- Commute to school/work
- Run an errand (e.g. shopping, medical/dental appointment, etc.)
- Special events where parking is an issue
- Nightlife (or any other activity impairing driving)
- Shopping
- Other (fill in)

Reason(s) for not using Transportation Network Companies (Uber, Lyft, etc.)
Check all that apply

- ☐ Not convenient
- ☐ Expensive
- ☐ Not available / Area not serviced
- ☐ Safety concerns
- ☐ Other (fill in)

Travel Diary Consent

We care about the quality of our survey data and hope to receive the most accurate measures of the trips of your day. It is important to us that you thoughtfully consider and record each trip of your day over a 24-hour period.

Do you commit to providing your thoughtful and honest answers to recording all the trips of your day over a 24-hour period?

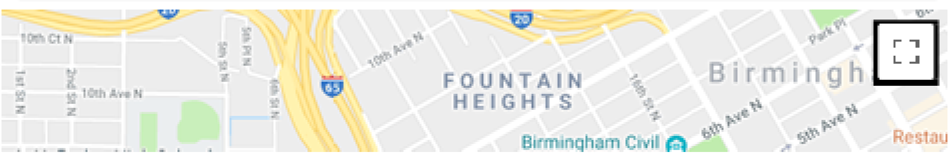
- ○ I will provide my best answers
- ○ I will not provide my best answers
- ○ I cannot promise either way

Initial Location

Please tell us about your trips during a typical weekday
Considering your trips yesterday or on a typical weekday, indicate every place you visited from the beginning of the day and for a 24-hour period. For the purpose of this survey, the day starts at 12:00 AM (midnight). Please also list walk trips that are 10 minutes or longer.

a. Please provide address (or closest intersection) to your initial location at 12:00AM (midnight)

Enter a location



Location Type

Home

Block 9

Please tell us about your trips during a typical weekday.
Considering your trips yesterday or on a typical weekday, indicate every place you visited from the beginning of the day and for a 24-hour period. For the purpose of this survey, the day starts at 12:00 AM (midnight). Please also list walk trips that are 10 minutes or longer.

b. ${lm://Field/1} Trip/Place Visited (address or closest intersection)

Enter a location

Trip Time

	hh	mm	AM/PM
Trip Start Time	00	00	AM

	hh	mm	AM/PM
Trip End Time	00	00	AM

Trip Purpose

Home

Mode

- Car
- Carpool/Vanpool
- Car rental
- Taxi
- Uber/Lyft
- Transit
- Bike
- Walk

Please share your experience with Transportation Network Companies (Uber, Lyft, etc.)
For each location you normally Uber/ Lyft or similar rides, indicate the typical wait time and car availability. Car availibility means the number of Uber/ Lyft cars you typically see swarming your location when using the mobile app.

Wait Time in minutes	Company	Uber/ Lyft Car Availability
0 - 5 minutes	Uber	1-2

Is this your last trip of the day (before you go to bed) ?

- Yes, this was my last trip

- No, I took another trip

Block 8

I would like to see more of the following where I live.
Check all that apply

- Public Transit (bus, light rail)
- Transportation Network Companies services (Uber/ Lyft, etc)
- Bicycle lanes
- Sidewalks
- Parking lots

General Information

Gender at birth

- Male
- Female

Age

18 to 24 years ▾

Current employment status

Full time ▾

Occupation

Management, business, science, and arts occupations ▾

Industry

Agriculture, forestry, fishing and hunting, and mining ▾

Annual Household Income

▾

Highest Degree

High school diploma ▾

Auto Ownership

- ○ I own a car
- ○ I have regular access to a vehicle that someone else in my household owns
- ○ I do not own or have regular access to a car

Please provide home address or closest intersection

Enter a location

7.2 Appendix D –Survey Instrument – Miami Beach, FL

MIAMI BEACH SURVEY

iQ Score: Fair

Published

Default Question Block

Block Options

Q1

The review involves filling an online survey that will take approximately 30 minutes. Your responses will be confidential and we do not collect identifying information such as your name, email address, or IP address. The survey questions will be about technology influence on travelers' attitudes, preferences, and choices and their potential impact on transportation services in the Miami Beach Network. More specifically, the study investigated the influence of Transportation Network Companies (TNCs) such as Uber and Lyft on travelers' behavior in the Miami Beach Network, Florida.

All data is stored in a password protected electronic format. To help protect your confidentiality, the surveys will not contain information that will personally identify you. The results of this study will be used for scholarly purposes only and may be shared with Qualtrics University representatives.

Q50

Are you a Miami Beach resident?

- Yes
- No

Q52

Have you been to Miami Beach within the past four months?

- Yes
- No

Q2

Do you have access to the following transportation mode ?

- Car
- E-hailing Service (Uber/Lyft etc)
- Public Transit Services
- Motorbike

Q3

How many vehicles do you own?

- 0
- 1
- 1+

Q4 **I have used the following at least once in the past year**

- Lyft
- Uber
- Organized Ride Sharing Program
- Public Transit Service
- × None of the above

Q5 **How many times you have used the E-hailing service (e.g. Uber, Lyft etc.) in the past 30 days?**

- 0
- 1-10
- 11-20
- 21-30
- 31-40
- 41-50
- More than 50

Q6 **How many times you have used the E-hailing service (e.g. Uber, Lyft etc.) in the month of January and February of 2020?**

- 0
- 1-5
- 6-10
- 10-20
- More than 20

Q7 **How many times you have used the Public Transit in the past 30 days?**

- 0
- 1-10
- 11-20
- 21-30
- 31-40
- 41-50
- More than 50

Q8 **How many times you have used the Public Transit in the month of January and February?**

- 0
- 1-5
- 6-10
- 10-20
- More than 20

Q9

If you do not use e-hailing service then which mode of transport do you prefer to use as an alternative?

- Miami Dade Bus
- City of Miami Beach Trolley
- Taxi
- Motorbike
- Bicycle
- Walk
- e-Scooter
- × None of the above

Q12

I would like to see future expansion of the following service in my area.

- Miami Dade Bus
- City of Miami Beach Trolley
- TNS (Uber, Lyft, etc.)
- Bicycle Lane
- Sidewalk
- Parking Lots

Q11

Are you willing to use the e-hailing services (e.g. Uber, Lyft etc.) for the first or last mile to reach home or to the nearby stop-station?

- Yes
- No
- Maybe

Q10

Are you willing to use the e-hailing services (e.g. Uber, Lyft etc.) for the first or last mile if incentives (discounts) are provided?

- Yes
- No
- Maybe

Q13

Are you willing to use the e-hailing services (e.g. Uber, Lyft etc.) for commuting if rewards point system is introduced by your office?

- Yes
- No
- Maybe

Q14

Are you willing to use the e-hailing services (e.g. Uber, Lyft etc.) for commuting if the transit service is made free to use the service?

- Yes
- No
- Maybe

6/10/2020 Qualtrics Survey Software

Q15

If a Metrorail is available from Downtown to South Beach, are you willing to take the service?

- Yes
- No
- Maybe

Condition: No Is Selected. Skip To: End of Block.

Q16

Which mode of transportation would you use to get to the Metrorail station?

- Walk/Bicycle
- TNC (Uber, Lyft, etc.)
- Automated People Mover
- City of Miami Beach Trolley
- Car (Park & Ride Facility)
- e-Scooter

Add Block

Weekday Block Options

Q17

Considering your typical weekday (being Monday to Friday), indicate every place you visited from the beginning of the day and for a 24-hour period. For the purpose of this survey, the day starts at 12:00 AM.

a. Initial Location at 12:00

https://fiu.ca1.qualtrics.com/Q/EditSection/Blocks?SurveyID=SV_6Xv4Zez5j3IyoVD 4/17

6/10/2020 Qualtrics Survey Software

0 18 AM Location

https://fiu.ca1.qualtrics.com/Q/E ...SV_6Xv4Zez5j31yoVO 5/17

Q19

Location Type

- ○ Home
- ○ Work
- ○ School
- ○ Eat / Get take-out
- ○ Nightlife / Bar
- ○ Shopping – Grocery
- ○ Shopping – Retail
- ○ Services (e.g. Bank, post office)
- ○ Pick-up passenger
- ○ Drop-off passenger

Q56

Departure Time

- ○ 12AM-3AM
- ○ 3AM-6AM
- ○ 6AM-9AM
- ○ 9AM-12pM
- ○ 12PM-3PM
- ○ 3PM-6PM
- ○ 6PM-9PM
- ○ 9PM-12AM

Add Block

Weekday Destination Block Options

Q21

b. Trips/ Places Visited

022

Location

6/10/2020 Qualtrics Survey Software

Q23

Trip Purpose

- Home
- Work
- School
- Eat / Get take-out
- Nightlife / Bar
- Shopping – Grocery
- Shopping – Retail
- Service (e.g. Bank)
- Pick-up passenger
- Drop-off passenger

Q25

Mode

- Car
- Carpool/Vanpool
- Car Rental
- Taxi
- Uber or Lyft
- Transit
- Bike
- Walk

Q57

Arrival Time

- 12AM-3AM
- 3AM-6AM
- 6AM-9AM
- 9AM-12pM
- 12PM-3PM
- 3PM-6PM
- 6PM-9PM
- 9PM-12AM

Q26

Have you visited another location

- Yes
- No

Condition: No Is Selected. Skip To: End of Block.

Add Block

https://fiu.ca1.qualtrics.com/Q/EditSection/Blocks?SurveyID=SV_6Xv4Zez5j3lyoVD 8/17

Weekend Block Options

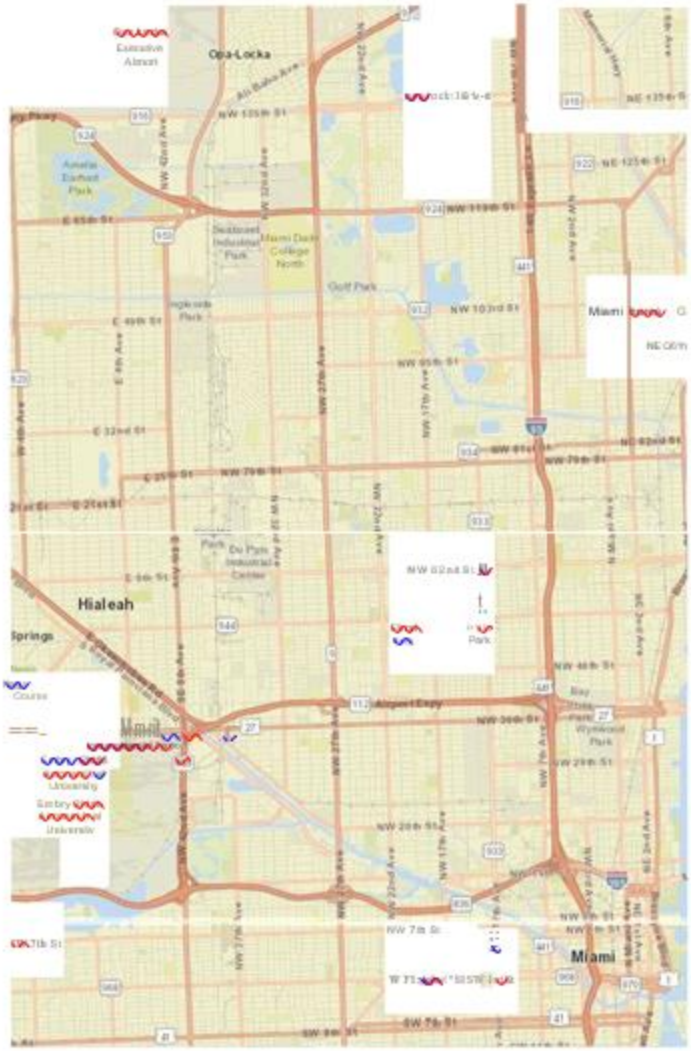
Q27

Considering your typical weekend (being Saturday or Sunday), indicate every place you visited from the beginning of the day and for a 24-hour period. For the purpose of this survey, the day starts at 12:00 AM.

a. Initial Location at 12:00

6/10/2020 Edit Survey | Qualtrics Survey Software

028 AM Location



https://fiu.ca1.qualtrics.com/Q/EditSection/Blocks/Ajax/GetSurveyPrintPreview?ContextSurveyID=SV_6Xv4Zez5j31yoVO 10/17

6/10/2020 Qualtrics Survey Software

Q29

Location Type

- Home
- Work
- School
- Eat/Gate take-out
- Nightlife/Bar
- Shopping – Grocery
- Shopping – Retail
- Service (e.g. Bank)
- Pick-up passenger
- Drop-off passenger

Q59

Departure Time

- 12AM-3AM
- 3AM-6AM
- 6AM-9AM
- 9AM-12pM
- 12PM-3PM
- 3PM-6PM
- 6PM-9PM
- 9PM-12AM

Add Block

Weekend Destination Block Options

Q31

b. Trips/ Places Visited

https://fiu.ca1.qualtrics.com/Q/EditSection/Blocks?SurveyID=SV_6Xv4Zez5j3lyoVD 11/17

6/10/2020 Edit Survey | Qualtrics Survey Software

032

Location

https://fiu.ca1.qualtrics.com/Q/E...SV_6Xv4Zez5j31yoVO 12/17

Q33

Trip Purpose

- Home
- Work
- School
- Eat / Get take-out
- Nightlife / Bar
- Shopping – Grocery
- Shopping – Retail
- Service (e.g. Bank)
- Pick-up passenger
- Drop-off passenger

Q61

Arrival Time

- 12AM-3AM
- 3AM-6AM
- 6AM-9AM
- 9AM-12pM
- 12PM-3PM
- 3PM-6PM
- 6PM-9PM
- 9PM-12AM

Q35

Mode

- Car
- Carpool/Vanpool
- Car Rental
- Taxi
- Uber or Lyft
- Transit
- Bike
- Walk

Q36

Have you visited another location?

- Yes
- No

Condition: No Is Selected. Skip To: End of Block.

Add Block

E-hailing Experience

Block Options

Q37

Please share your experience with e-Hailing services (Uber, Lyft, etc.)

For each location you normally request Uber/Lyft or similar rides, indicate the typical wait time and car availability. Car availability means the number of Uber/Lyft cars you typically see swarming your location when using the mobile app.

6/10/2020 Qualtrics Survey Software

Q47 **Location**

Q38 **Availability of E-hailing Service?**

- ○ Yes
- ○ Maybe
- ○ No

https://fiu.ca1.qualtrics.com/Q/EditSection/Blocks?SurveyID=SV_6Xv4Zez5j3IyoVD 15/17

6/10/2020 Qualtrics Survey Software

Q48

Car Availability in your Area

- 0
- 0-5
- More than 5

Q40

Waiting Time

- Less than 5 minutes
- 5-15 minutes
- 15-30 minutes
- More than 30 minutes

Q39

Company Name

- Uber
- Lyft
- Other

Q49

I Have More Experience

- Yes
- No

Add Block

Demographic Question Block Options

Q42

What is your Age?

- <18
- 18-29
- 30-49
- >50

Q44

What is your gender?

- Male
- Female
- Others

https://fiu.ca1.qualtrics.com/Q/EditSection/Blocks?SurveyID=SV_6Xv4Zez5j3lyoVD 16/17

6/10/2020 Qualtrics Survey Software

Q43

What is your education level?

- High School or less
- Some college credit, no degree
- College Graduate
- Graduate School

Q45

How much is your annual income?

- < $30,000
- $30,000-$74.99,000
- >75,000

Q46

State your race.

- White
- Black or African American
- American Indian or Alaska Native
- Asian
- Native Hawaiian or Pacific Islander
- Other

Add Block

End of Survey Survey Termination Options...

https://fiu.ca1.qualtrics.com/Q/EditSection/Blocks?SurveyID=SV_6Xv4Zez5j3IyoVD 17/17

7.2 Appendix E –Significant Variables – Birmingham, AL and Miami Beach, FL

Significant Variables	
Birmingham Region	
1	Car Availability
2	Waiting Time
Miami Beach Region	
1	Miami Beach Resident
2	Access to E-hailing Service (Uber/Lyft e.t.c)
3	Owner of 1 vehicle
4	Have used none of the transportation modes in the past year
5	Have used public transit service at least once in the past year
6	Have used uber at least once in the past year
7	E-hailing service user in the past 30 days
8	Used public transit more than 50 times in the past 30 days
9	Used public transit from 1 to 5 times in the month of January and February
10	No preferred alternative mode of transport to E-hailing service
11	Would like to see expansion of Miami Dade bus service
12	Willing to use the e-hailing services (e.g. Uber, Lyft etc.) for commuting if the transit service is made free to use the service
13	Would use Car (Park & Ride Facility) to get to the Metrorail station
14	Would use TNC (Uber, Lyft, etc.) to get to the Metrorail station
15	Location Type: Eat/Get take out
16	Location Type: Services (e.g. Bank, post office)
17	Location Type: Worl
18	Trip Purpose: Eat / Get take-out
19	Trip Purpose: Home
20	Trip Purpose: Shopping-Grocery
21	Mode: Car
22	Mode: Uber or Lyft
23	Location 2 Type: Work
24	Trip 2 Purpose: Pick Up passenger
25	Trip 2 Purpose: Service (e.g. Bank)
26	E-hailing Service Available
27	0 to 5 Cars are avaibale in the area
28	Education Level: Graduate School
29	Education Level: High School or Less
30	Education Level: Some college credit, no degree
31	Annual Income >$75,000
32	Race: Other

7.3 Appendix C – Summary of Accomplishments

Date	Type of Accomplishment	Detailed Description
08/22/2022	Conference Presentation and Publication	Khalil, J., Yan, D., Yuan, L., Adhikari, S., Jafarzadehfadaki, M., Sisiopiku, V. and Jiang, Z. “Realistic Urban Traffic Simulation with Ride-Hailing Services: A Revisit to Network Kernel Density Estimation”. Paper accepted for presentation and publication to SIGSPATIAL 2022 proceedings. Paper selected as *finalist for best paper award.*
03/09/2022	Student Accomplishment or Award	Mostafa Jafarzadehfadaki received the 2021-2022 Outstanding International Ph.D. Student Award for the UAB School of Engineering.
2/24/2022	Conference Presentation	Morshed S.A., Hadi, M., and Sisiopiku, V.P. “A Novel Multi-Agent Based Simulation Study on the Extension of Metrorail in Miami Beach Region”. 7th Annual Conference for the Southeastern Region, Boca Raton, FL, March 2022.
12/07/2021	Publication	Sultana, T., Sisiopiku, V.P., Khalil, J., and Yan, D. (2022). “Potential Benefits of Increased Public Transit Ridership in Medium Sized Cities: A Case Study”, Journal of Transportation Technologies, Vol. 12, No. 1, January, 2022. Available at https://www.scirp.org/journal/paperinformation.aspx?paperid=114470
04/22/2021	Publication	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), April 2021. Available at https://www.scirp.org/journal/paperinformation.aspx?paperid=108245
01/29/2021	Student Accomplishment or Award	People’s Choice Award, STRIDE Student’s Poster Showcase and Competition (TRBAM 2021). Poster Title: Comparative Mode Split Analysis Between Multi-Agent Simulation and Activity-Based Demand Model: Miami Beach Open Scenario. Presenter: Syed Ahnaf Morshed (FIU).
12/28/20	Poster Submitted to STRIDE Student Poster Competition	Morshed, S.A. “Comparative Mode Split Analysis Between Multi-Agent Simulation and Activity-Based Demand Model”

12/28/20	Poster Submitted to STRIDE Student Poster Competition	Sultana, T. "Justifying Public Transit Investment in Medium-sized Cities by Evaluating their Utility using Agent-based Simulation"
8/1/20	Paper Submitted	Sisiopiku, V.P., Morshed, S.A., Sarjana, S., and Hadi M. (2020). "Transportation Users' Attitudes and Choices of Ride-Hailing Services in Two Cities with Different Attributes Transportation Research Board". Submitted for consideration for presentation at the 100th Annual Meeting of the Transportation Research Board.
6/17/20	Journal Paper	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, Available at https://www.scirp.org/pdf/jtts_2020061810033242.pdf
5/13/20	Webinar Presentation	STRIDE Webinar presentation by project PI Dr. Virginia Sisiopiku. "Technology Influence on Travel Demand and Behaviors".
2/24/20	Conference Presentation	Accepted for presentation at the 7th Annual UTC Conference for the Southeastern Region, Boca Raton, FL, April 2020. Title: "Planning for Mobility as a Service (MaaS) in the Birmingham Region by Integrating TNCs and Public Transit".
01/13/20	Student Accomplishment or Award	Ms. Sahila Sarjana received the 1st place award at the STRIDE Student Poster Competition that took place at the 99th Annual Meeting of the TRB (Washington, DC; 01/13/20). The poster title was "Examining the Impact of Transportation Network Companies (TNCs') Presence on Travel Preferences and Patterns of Birmingham Transportation Users".
12/10/19	Conference Paper	Paper presentation at 2019 IEEE International Conference on Big Data. Citation follows: Guo, G., Khalil, J.M., Yan, D., and Sisiopiku V. (2019). "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, Los Angeles, CA.
11/22/19	Student Accomplishment or Award	Ms. Sarjana's abstract was accepted for STRIDE Student Poster Competition at 2020 TRB. The poster titled "Examining the Impact of Transportation Network Companies (TNCs') Presence on Travel Preferences and Patterns of Birmingham Transportation Users" by Sarjana S., Sisiopiku V., and Ramadan, O. was prepared and submitted on 12/05/19.

10/1/19	Student Accomplishment or Award	Ms. Sahila Sarjana was named the recipient of the UAB National Alumni Society Mobola Kukoyi Scholarship. She is a Graduate Research Assistant at the TREND Lab working on transportation users' perceptions and attitudes toward Transportation Network Services (TNC) such as Uber and Lyft. The award was presented on 10/01/19.
8/19/19	Conference Presentation	Dr. Sisiopiku is scheduled to present a paper titled "Simulation Options for Modeling Shared Mobility" at the 2019 AlaSim M&S Conference, Huntsville, AL (planned). The paper was accepted for presentation on 8/19/19.
4/1/2019	Conference Presentation	Dr. Sisiopiku presented a paper titled "Mobility Patterns and Mode Preferences of Birmingham Travelers" at the 2019 SDITE/MCDITE Annual Meeting in Arlington, VA.
03/29/19	Student Accomplishment or Award	Ms. Sahila Sarjana won first place in the Southern District ITE (SDITE) graduate student technical paper contest. Her paper was titled: "Identifying Micro-Level Determinants that Influence the Transportation Network Companies (TNCs) Growth through Analysis of Transportation Users' Preferences and Attitudes".
2/22/2019	Other	Dr. Sisiopiku offered a seminar to UAB Civil Engineering Undergraduate and Graduate students titled "Traffic Congestion: Needs, Opportunities, and UAB Research Contributions" highlighting ongoing transportation research at the TRENDLab, including Project I2.
2/20/2019	Conference Presentation	Abstract titled "Mobility Patterns and Mode Preferences of Birmingham Travelers" was selected for presentation at the 2019 SDITE/MCDITE Annual Meeting in Arlington, VA
10/12/18	Faculty Accomplishment or Award	Dr. Sisiopiku received a 2018 Certificate of Meritorious Achievement award from the Southern District Institute of Transportation Engineers
9/20/18	Student Accomplishment or Award	Dr. Ossama Ramadan received the 2018 UAB Most Esteemed Postdoc Award