

The Pennsylvania State University

The Graduate School

**MODELING AND EVALUATION OF MICROMOBILITY
IN A MULTI-MODAL TRANSPORTATION SYSTEM**

A Dissertation in

Civil and Environment Engineering

by

Muyang Lu

© 2022 Muyang Lu

Submitted in Partial Fulfillment
of the Requirements
for the Degree of

Doctor of Philosophy

December 2022

The dissertation of MUYANG LU was reviewed and approved by the following:

Ilgin S. Guler

Associate Professor of Civil and Environment Engineering

Dissertation Adviser

Chair of Committee

Yiqi Zhang

Assistant Professor of Industrial and Manufacturing Engineering

Vikash V. Gayah

Associate Professor of Civil and Environment Engineering

Xianbiao Hu

Assistant Professor of Civil and Environment Engineering

Patrick Fox

Head of the Department of the Graduate Program

ABSTRACT

Micromobility, including e-scooters and bikes, has recently become a popular travel mode and rapidly proliferated across big cities. Appropriately modeling and evaluating the performance of micromobility in a multi-modal transportation system is critical to the design and operation of a transportation system. This research focuses on analyzing the performance of micromobility in three aspects, e.g., impact on travel demand, impact on traffic safety and impact on traffic operations.

Firstly, the spatiotemporal trip characteristics of shared micromobility when operating in the same urban area with public transit and motorized vehicles are analyzed using data collected from the City of Austin. A decomposition model of the variations in daily trip volumes and a GWR model to reveal the spatiotemporal usage patterns of each mode are developed. The results reveal that temporal patterns of shared bike usage are more consistent with commuting compared to the temporal patterns of e-scooters. From a spatial perspective, the results suggest that if e-scooter ridership complements public transit ridership in a census tract, shared bike ridership will show a competing effect with public transit, and vice versa. Overall, e-scooters and shared bikes with transit could reduce car trips and congestion if implemented carefully to improve environmental sustainability.

Next, the impact of non-motorized modes on traffic safety is considered by incorporating use of non-motorized modes to macroscopic traffic safety models. A macro-level crash prediction model for the Manhattan area of New York City is developed that considers bike share trip information, roadway and demographic variables, subway flows, taxi movements, and person-trips to various points of interest (POI) as measures of travel exposure. The models are developed using negative binomial regression and various functional forms are considered. The spatial variation of the correlations between crash frequency and exposure variables are also explored with geographically weighted negative binomial regression. The results suggest that the number of shared bike trips and

POI visits are positively associated with an increase in pedestrian and cyclist crash frequencies; however, these features are less descriptive of motorist crash frequency. Considering the spatial interactions, the correlation between cyclists' crash frequencies and exposure factors is generally homogeneous across the Manhattan area and is not significantly impacted by the local land use. On the other hand, pedestrian and motorist crash frequencies are more sensitive to the geographic features of an area.

Finally, to measure the influence of bicycle trips on traffic operations, a mixed car-bike traffic flow model is proposed for an undivided two-way roadway based on the Lax-Hopf equation. Slow-moving bicycles are treated as a moving bottleneck in the proposed model. The impact of bike flow on traffic delay with different car and bike flows, different opposing traffic flows, and different level of uncertainties are evaluated based on the proposed model. The results show that the speed and flow of cars and bikes can significantly influence the expected traffic delay, and the specific interactions of shockwaves of queues play a large role in the expected total delay in a homogeneous traffic environment. By comparing the total delay to a scenario with a dedicated bike lane, the domains of application of shared vs. dedicated bike lanes can be identified. As the speed reduction resulting from implementing a dedicated bike lane reduces, or the opposing traffic flow reduces, the threshold of bike flow that can achieve lower total traffic delay by implementing a dedicated bike lane increase. The proposed method can help improve traffic in environments where cars and bikes coexist.

TABLE OF CONTENTS

LIST OF FIGURES	vii
LIST OF TABLES	x
ACKNOWLEDGMENTS	xi
1 INTRODUCTION	1
1.1 LITERATURE REVIEW.....	3
1.1.1 <i>Shared e-scooter travel and usage patterns compared to shared bikes</i>	3
1.1.2 <i>Interaction between micromobility and other traffic modes</i>	4
1.1.3 <i>Traffic crash frequency models</i>	5
1.1.4 <i>Impact factors of crash frequency</i>	6
1.1.5 <i>Impact of shared bikes on traffic safety</i>	7
1.1.6 <i>Traffic flow modeling</i>	8
1.1.7 <i>Car-bike mixed traffic flow modeling</i>	9
1.1.8 <i>Implement of dedicated bike lane</i>	9
1.2 RESEARCH GAPS.....	10
1.3 RESEARCH OBJECTIVES.....	10
2 SHARED MICROMOBILITY AND PUBLIC TRANSIT	12
2.1 DATA DESCRIPTION.....	12
2.1.1 <i>Shared micromobility dataset</i>	13
2.1.2 <i>Public transit dataset</i>	15
2.1.3 <i>Supplemental datasets</i>	16
2.2 METHODOLOGY.....	18
2.2.1 <i>Temporal variation analysis of each traffic mode</i>	19

2.2.2	<i>Spatial correlation analysis among different modes</i>	20
2.3	RESULTS AND DISCUSSIONS	22
2.3.1	<i>Individual characteristics of shared bike, e-scooter, and public transit</i>	22
2.3.2	<i>Spatial correlation among shared bike, e-scooter, and public transit</i>	28
2.3.3	<i>Discussion</i>	35
2.4	SUMMARY	38
3	SHARED BIKE AND TRAFFIC SAFETY	40
3.1	Data Description	40
3.1.1	<i>Crash dataset</i>	41
3.1.2	<i>Shared bike trip records</i>	42
3.1.3	<i>POIs visit volumes</i>	42
3.1.4	<i>Traffic surrogate datasets</i>	42
3.1.5	<i>Demographic information</i>	43
3.1.6	<i>Geographic route profile</i>	44
3.2	Methodology	46
3.2.1	<i>Negative binomial model</i>	47
3.2.2	<i>Geographically weighted negative binomial model</i>	48
3.3	Results and discussions	49
3.3.1	<i>Cyclists model</i>	50
3.3.2	<i>Pedestrian model</i>	55
3.3.3	<i>Motorist model</i>	57
3.3.4	<i>GWNBR model comparison</i>	60
3.3.5	<i>Model discussion</i>	62
3.4	SUMMARY	64
4	SHARED BIKE AND TRAFFIC EFFICIENCY	66
4.1	METHODOLOGY	66

4.1.1	<i>Mixed car and bikes flow modeling</i>	67
4.1.2	<i>Impact of bike lanes on traffic efficiency</i>	76
4.2	RESULTS AND DISCUSSION	77
4.2.1	<i>The total delay in mixed traffic flow</i>	78
4.2.2	<i>Impact of variations in mixed traffic flow on delay</i>	86
4.2.3	<i>Domains of application of dedicated bike lanes</i>	91
4.3	SUMMARY	98
5	CONCLUSIONS	100
5.1	ENGINEERING SIGNIFICANCE.....	101
5.2	LIMITATIONS	102
	REFERENCES	103

LIST OF FIGURES

Figure 2-1 Spatial distribution of different traffic modes in the City of Austin	15
Figure 2-2 Temporal distribution of different traffic modes across one week. Data are normalized based on total weekly trip volume and aggregated over analysis window	24
Figure 2-3 Euclidean distance between POI visits temporal patterns and traffic modes usage patterns. .	24
Figure 2-4 Normalized trip volume vs. time-of-day decomposition models. (The dashed lines show the results of the fitting curve for each component, and the solid curves are the fitting curve for the entire day. The shaded areas represent the real data.)	26
Figure 2-5 Percentage of each component for different traffic modes	27
Figure 2-6 R2 distribution of Transit-GWR model	32
Figure 2-7 Local coefficients of variables in the Transit-GWR model	35
Figure 3-1 The relationships between cyclists crash prediction and POIs data in model 1 and model 2.	51
Figure 3-2 CURE plots for cyclists involved crash models.....	54
Figure 3-3 CURE plots for pedestrians involved crash models.....	57
Figure 3-4 CURE plots for motorists involved crash models.....	60
Figure 3-5 Coefficients estimation of $\log(\text{Essential POI visits})$ from GWNBR model for three types of crash frequency model	61
Figure 3-6 Coefficients estimation of $\text{Log}(\text{Shared bike trips})$ from GWNBR model for two types of crash frequency model	62
Figure 4-1 (a) Illustrative time space diagram of bike and cars in the analysis and opposing direction, (b) illustration of cars on the roadway at time instance one and two.	68
Figure 4-2 Analytical calculation of single moving bottleneck.....	71
Figure 4-3 Illustration of Lax-Hopf conditions	72
Figure 4-4 Trajectories of bikes with and without bike lanes.....	77
Figure 4-5 Maximum opposing traffic flow for different car speeds	79
Figure 4-6 Comparison of shockwaves (and traffic density) for a moving bottleneck.....	80

Figure 4-7 Total delay (veh * sec) caused by a single bike.....	82
Figure 4-8 Total delay (veh * sec) caused by a bike flow of 100 bike/h	83
Figure 4-9 Different interaction patterns between bottlenecks	84
Figure 4-10 Total delay from the Lax-Hopf framework with different bike flows and car flows (veh * sec).....	86
Figure 4-11 Total delay of mixed traffic flow with different bike headway variation (veh * sec)	88
Figure 4-12 Total delay of mixed traffic flow with different bike speed variation (veh * sec)	89
Figure 4-13 Total delay of mixed traffic flow with different opposing car spacing variation (veh * sec)	90
Figure 4-14 Delays of the basic experiment in unit of person * sec.....	93
Figure 4-15 Change in total delay due to the presence of a dedicated bike lane for different car speeds reductions (person * sec).....	94
Figure 4-16 Change in total delay due to the presence of a dedicated bike lane for different opposing traffic flow (person * sec).....	95
Figure 4-17 Change in total delay due to the presence of a dedicated bike lane for different opposing spacing variation (person * sec).....	96

LIST OF TABLES

Table 2-1 Samples of shared micromobility dataset.....	13
Table 2-2 Samples of Capital Metro APC ridership dataset.....	16
Table 2-3 POI footage dataset summary.....	17
Table 2-4 Selected features for e-scooter, bike or transit usage prediction.....	29
Table 2-5 Spatial regression results of Transit-GWR model.....	33
Table 3-1 Statistical summary of the census tract level data used in the study.....	44
Table 3-2 Coefficient estimations of cyclists involved crash model.....	52
Table 3-3 Coefficient estimations of pedestrian involved crash model.....	56
Table 3-4 Coefficient estimations of motorists involved crash model.....	58
Table 4-1 Reduction in free-flow speed (km/h).....	76
Table 4-2 Comparison of Lax-Hopf approach and analytical equations for different time steps.	80

ACKNOWLEDGMENTS

I would never be able to pursue a PhD degree without the tremendous support from my parents, especially during such a tough pandemic time! Studying overseas for three and half contentious years is a huge challenge both for me and my mom and dad. I still clearly remember that the first time I share my thought with my parents about pursuing a higher degree in another country, I was astonished by the calmness and limitless support. The same support also comes when I miss home, feel stressed in the work, and feel lost about the future. Those constant supports give me great power and courage to move on and pursue my dreams and push me beyond my limitations to be a better man. Therefore, I would like to sincerely thank my mom and dad for their understanding, support, and dedication during my PhD study. I wish them to be safe, healthy, and happy every day of their life. I would like to thank my grandmother and grandfather for their encouragement and concern and my sister Ping and brother-in-law Lushun for their company and help. I also would like to thank my cutest nephew Allen who came to this world in the second year of my PhD study for the joy and lough he brought to me and my whole family. I'm proud to have such a wonderful family and my heart is full of appreciation and satisfaction in the whole journey of my life.

I cannot express enough how lucky and appreciative I am to have Dr. Ilgin S. Guler as my advisor, mentor, and friend over the past three and half years. Three years ago, she gave me the precious chance to pursue a PhD degree at Penn State which opened a new world to me and changed the course of my life from then on. She is full of love and patient with me during my study. About 160 times of one-on-one regular meetings and about 63 times of paper draft reviews and revisions ensured every achievement I acquired in my academic research. She leads me out of my confusion, lightens my path, and corrects my mistakes with her wisdom, experience, and sharp judgment in my academic research journey. She inspires, supports, and encourages my every decision and nourishes my

unmatured ideas to make them come true. She also set a great example to me about a work-life balanced lifestyle and thought me how to work hard while taking good care of our family. Dr. Vikash V. Gayah is also an important mentor, teacher, and friend in my PhD study. He passes the profound knowledge to us in a clear and simple way in the classes and provides creative and constructional ideas for my academic research. I would like to sincerely thank Dr. Guler and Dr. Gayah for their support and help throughout my entire PhD study.

I would like to thank my doctoral committee members Dr. Yiqi Zhang, Dr. Xianbiao Hu, Dr. Gayah, and Dr. Guler for the consultants, comments, and suggestions in the writing of this dissertation. This research cannot be complete without their tremendous contributions. I would like to thank my colleagues and coauthors of this study Dr. Elizabeth J. Traut, Dr. Hao Liu, and Dr. Kan Wu for their contributions to this dissertation. I would like to thank my respectful teachers in my course study Dr. Aleksandra Radlińska, Dr. Eric Donnell, Dr. Martin T. Pietrucha, Dr. Kostas Papakonstantinou, and others for bringing knowledge to me and for their lovely lectures.

I am flattered to have a number of great friends Binyang, Mingming, Lixiang, Pengsen, Jiachang, Ruofei, Rui, Zhanzhao, Fan, Abby, Qicheng, Tian, Xiaoyan, Xusheng, Asif, Alex, Lijuan, Xiaohua, Chris, Jay, Guan hao, Dongqin, Agni, Abhishek and so many others who helped me in need and support me in the darkness. We talk about our deepest minds, share our dreams, find the best version of ourselves, and overcome our weaknesses. The time we spend together defines who we are and let us know who we want to be. Finally, I would like to thank myself for carrying on all the way along and never giving up in face of any difficulties. I believe the characters and experiences I built up in these three and half years will lighten my way in my future and provide me strength to be the person I want to be.

My PhD study was contentiously supported by funds available through the Center for Integrated

Asset Management for Multi-modal Transportation Infrastructure Systems (CIAMTIS): Region 3
University Transportation Center. The content of this dissertation reflects solely the views of the
authors who are responsible for the facts and the accuracy of the data presented herein.

1 INTRODUCTION

Micromobility, such as bicycles, e-bikes, e-scooters, and e-skateboards, are a range of small, lightweight vehicles that are typically operated at speeds below 25 km/h by individual users. While these modes have been used frequently, recently they have become more popular as shared mobility options. In 2008, Washington, D.C. launched the first modern bike-share system in the U.S. named smartBike DC, a 10-station, 120-shared bike pilot program. Since 2010, over 343 million trips have been taken in the U.S. by micromobility (NACTO, 2020). E-scooters have seen a large increase in popularity since 2018 and the e-scooter market is expected to grow to be more than \$30 Billion by 2025 (Renub Research, 2020). As micromobility becomes a mainstream transportation mode, its impact on the transportation system could become even more significant and diverse. Thus, an appropriate policy that seamlessly accommodates these modes into the existing traffic environment is important to a city.

Shared bike, shared e-scooter and public transit are the major non-automobile transportation modes used in urban areas for commuting and leisure trips. Despite the apparent similarity of shared e-scooters and shared bikes, these two micromobility modes have significant differences related to the users' travel patterns and behaviors. Therefore, they are usually unevenly distributed across different areas and serve different roles when jointly participating in a multi-modal transportation network, especially when complementing or competing with public transit. Understanding the temporal and spatial characteristics of multiple micromobility modes that coexist with public transit can help improve policy making in cities, however, this knowledge is lacking from the existing literature.

Despite the opportunities that shared micromobility modes bring to mobility, the interaction of multiple modes in the same urban space creates unique safety challenges. Therefore, a successful shared micromobility adoption strategy requires to achieve the balance between benefits and safety. Hence, it is important to model the impact of these modes on overall safety. perhaps due to data

availability, existing traffic safety models tend to use only exposure attributes related to vehicular traffic and rarely include exposure attributes related to non-motorized travel. Those that do consider non-motorized travel only include at most one explanatory variable.

Another major impact from micromobility on the existing traffic system is its impacts on traffic efficiency. Despite the improvement in infrastructure and many newly installed dedicated bike lanes, there are still a lot of undivided two-way roads where bikes share the same road with motorized vehicles, and more interactions between bikes and cars occur. Cyclists, including the riders of traditional bikes and electrical bikes, usually move slower than cars and are extremely vulnerable. For a two-lane roadway, cars need to borrow the opposing lane to pass a slow-moving bike while maintaining a safe distance. When there are no gaps in the opposing lane, cars cannot pass the bike and the bike becomes a moving bottleneck. In this situation, the bike can lead to car delays. The interactions between cars and bikes on two-directional roadway, where bikes would have the most impact on car traffic, have not been studied in the literature. Further, there are no analytical tools to evaluate the need for a dedicated bike lane considering total system delay.

The goal of this dissertation is to systematically evaluate micromobility in a multi-modal transportation system from three aspects, e.g., impact on travel demand, impact on traffic safety and impact on traffic operations. Specifically, the impact of shared micromobility on travel demand will be analyzed based on the spatial interactions of shared bike, shared e-scooter and public transit and to explore their different roles in a multi-modal transportation system. The impact on traffic safety will be evaluated by incorporating several exposure metrics that capture non-motorized and public transportation use, such as roadway and demographic variables, as well as bike share trip information, subway flows, taxi movements, and person-trips to various points of interest (POI) as measures of travel exposure, at the census-tract level. Finally, impact on traffic operation will be explored by proposing a multimodal approach to understanding the impacts of bikes on traffic flow, using a macroscopic car-bike mixed traffic flow model based on the Lax-Hopf equations. The impact of

slow-moving bikes on traffic delay will be quantified thereby.

The remainder of the dissertation is organized as follows. First a comprehensive literature review will be conducted, and the research objectives will be identified in the rest of this chapter. Then, the interaction between shared micromobility and public transit, the impact of shared bike on traffic safety, and impact of bikes on traffic efficiency are studied in the following three consecutive chapters. Finally, the major conclusions from this study are summarized in the last chapter.

1.1 LITERATURE REVIEW

The literature review aims at providing background information on multimodal trip characteristics, multimodal traffic safety and multimodal traffic analysis. To this end, the multimodal trip characteristics are discussed in two subsections: 1) comparing shared e-scooter to shared bikes, and 2) interaction between micromobility and other traffic modes. The traffic safety literature is presented in three subsections that first describe general information on traffic crash frequency models and variables considered in crash frequency modeling. Next, specific literature on multi-modal traffic safety models is discussed. Finally, the multimodal traffic analysis literature is discussed in three subsections that include information on general traffic flow modeling, followed by car-bike mixed traffic flow modeling and implementation of dedicated bike lanes.

1.1.1 Shared e-scooter travel and usage patterns compared to shared bikes.

Shared e-scooters have many characteristics in common with shared bikes such as use in urban areas, flexibility, convenience, and can be docked or dockless. However, e-scooters have some distinctive features, including requiring fewer riding skills, less bike-friendly clothing requirements, and limited driving range. Studies show that shared e-scooter users typically skew young, and affluent (Laa and Leth, 2020), and achieve greater gender parity compared to docked bike sharing services (Clewlow, 2019). However, the low deck and safer performance of e-scooters make this mode more user-friendly for older users as well compared to shared bikes (Gitelman et al., 2017). A survey-based study found that, compared to motorized bicycles, e-scooters statistically have a shorter travel

distance (Jordehi et al., 2013); another study conducted in Indianapolis found that more than half of trips taken by e-scooters were less than 10 minutes long and travelled less than one mile (Mathew et al., 2019). E-scooter trips were also found to have lower speeds compared to e-bikes (Almanaa et al., 2021).

Due to the different characteristics, the trip purposes of shared e-scooters and shared bikes are different: shared e-scooters are more likely to be used for leisure trips while shared bikes tend to be used for commuting (Gitelman et al., 2017; Hardt and Bogenberger, 2019; McKenzie, 2019; Zou et al., 2020). Different dock types also lead to different travel purpose, where docked micromobility modes are more preferred for commuting. Studies further concluded that the usage patterns of shared e-scooters are more similar to usage patterns of non-member bike-share users compared to member bike-share users (Reck et al., 2021; Younes et al., 2020). Further, weather was observed to be less of a disutility for shared e-scooter users than for bike-share users, regardless of membership type. The temporal usage patterns of shared e-scooters were found to vary based on location, e.g., in Austin the largest shared e-scooter volumes were observed during afternoon and weekends, while in Minneapolis the largest volumes were experienced during the evening (Bai and Jiao, 2020).

1.1.2 Interaction between micromobility and other traffic modes.

Micromobility can benefit the general traffic network, e.g., by providing an alternative transport mode to motorized vehicles, providing last-mile connectivity with public transit, or complementing other modes of transport (Médard de Chardon, 2019). Understanding how micromobility complements and competes with other traffic modes is crucial to the understanding of a multi-modal transportation systems.

Typically, shared bike systems, shared e-scooters, and public transit are modes that are used for similar purposes. Various studies have tried to understand the interactions between these modes. Survey-based studies found that shared e-scooters, similar to shared bikes, mostly replace walking

and public transport for short distance travel (Hardt and Bogenberger, 2019; Laa and Leth, 2020; Sanders et al., 2020). Studies also concluded that bike share systems can complement public transit, e.g., by serving as the first-and-last mile solution (Ma et al., 2015; Pan et al., 2010), or substituting for public transit by replacing short distance trips (Campbell et al., 2016; Campbell and Brakewood, 2017; Guidon et al., 2019). These effects can vary both spatially and temporally (Kong et al., 2020; Ma et al., 2019). From the temporal perspective, it was found that the correlation between dockless shared bikes and public transit varies from weekdays, when there is a positive correlation, to weekends, when there is a negative correlation (Kong et al., 2020; Ma et al., 2019). Another study found that bikes and public transit substitute each other for short-term trips, e.g., individual trips, but complement in the long-term, e.g., considering multiday or multi-season results (Singleton and Clifton, 2014). From the spatial perspective, a survey conducted in Washington D.C. and Minneapolis found that the direction of mode shift between shared bike and public transit depends on demographic attributes such as age, gender, commute distance, and location of residences (Martin and Shaheen, 2014). It was also found that the introduction of a new transit system can influence shared bike usage (Gu et al., 2019). There are only few studies that analyzed the interaction between public transit and e-scooter usage. One recent study found that transit accessibility increased shared e-scooter usage (Bai and Jiao, 2020). Further, a few different works have considered the interaction between shared bikes and public transportation to understand how the presence of the two modes together could change their ridership (Campbell et al., 2016; Campbell and Brakewood, 2017; Guidon et al., 2019; Ma et al., 2019; Pan et al., 2010).

1.1.3 Traffic crash frequency models

Research that analyzes the impact of influential factors on traffic crash frequency can be categorized into two groups: micro-level and macro-level. Micro-level studies focus on relationships between crash frequency on individual roadway elements (i.e., roadway segment or intersections) and features associated with that element. The explanatory variables included in such micro-level models include

geometric characteristics (Anastasopoulos et al., 2012; Gooch et al., 2018, 2016; Jung et al., 2014; Zhang et al., 2020), *traffic volume* (Zhang et al., 2020; Jung et al., 2014; Anastasopoulos et al., 2012; Gayah and Donnell, 2021; Roque and Cardoso, 2014; Bao et al., 2019; Caliendo et al., 2013), *infrastructure condition* (Caliendo et al., 2013), and *weather* (Bao et al., 2019; Cheng et al., 2017; Jung et al., 2014). The outcome of micro-level crash prediction models provides the predicted crash frequency on any individual roadway element. On the other hand, macro-level studies seek to relate the average crash frequency within a given geographic area—such as a city, a borough, or a census tract—with features associated with the region. These models consider demographics, land use, socioeconomics, and transportation infrastructure information as input and predict crash frequency within the region (Wang et al., 2019). The present study focuses on macro-level crash prediction and how it might be influenced by multi-modal travel demand patterns and emerging mobility services.

1.1.4 Impact factors of crash frequency

Many researchers have studied the impacts of different factors on crash frequency at the macroscopic scale. Such macroscopic models usually utilize Poisson or negative binomial (NB) models to account for crash data that can only take integer values (Islam et al., 2022; Mahmud and Gayah, 2021; Sengupta et al., 2021). However, such models do not account for the spatial interactions between the variables. Further, existing macroscopic safety models mostly rely on traditional vehicular-based explanatory variables (Bao et al., 2019; Jia et al., 2018; Reynolds et al., 2009). Only a few studies have considered area-wide measures of exposure that may include both vehicular and non-vehicular traffic. For example, one study evaluated the impact of person-trips to specific points-of-interests (POIs) on traffic crash frequency and found that densely visited areas, such as residential areas, banks, and hospitals, are associated with increased crash frequency (Jia et al., 2018). Another study considered the impact of transportation infrastructure on bicycling injuries and crashes and concluded that the presence of sidewalks, multi-use trails, and major roads were associated with increased bicyclist crash risk, while the presence of a bicycle facility decreased the crash risk (Reynolds et al.,

2009). Besides those non-vehicular traffic factors mentioned above, vehicular factors such as traffic flow are tightly related to crash frequency as well, and often used as explanatory variables in macroscopic safety models (La et al., 2013).

1.1.5 Impact of shared bikes on traffic safety

Shared bikes have recently become more popular and can help improve mobility in a city; however, the impacts on safety are unknown. The major known impacts of shared bike trips include mode substitution from private cars, public transit, and walking (Buck et al., 2013; Shaheen et al., 2013; Song et al., 2020; Zhang and Zhou, 2019), congestion mitigation (Buck et al., 2013; Shaheen et al., 2013; Wang and Zhou, 2017) and environmental benefits (Shaheen et al., 2013; Zhang and Mi, 2018). Unfortunately, cyclists are often some of the most vulnerable users on the roadway network (David, 2016; Guler and Grembek, 2016). In general, the safety of bicyclists has been extensively studied in the literature, generally in the context of safety in numbers. The main goal of these studies is to examine if bicyclists (or pedestrians) become safer as more of them are present on the roadway (Jacobsen, 2003; McAndrews, 2012; Pucher and Dijkstra, 2003). These studies have mixed conclusions on whether an increase in the number of bicyclists can reduce crash risk – Jacobsen and Pucher concluded that a motorist is less likely to collide with a person walking and bicycling if more people walk or bicycle. Policies that increase the numbers of people walking and bicycling appear to be an effective route to improving the safety of people walking and bicycling. However, McAndrews found that whereas motor vehicle occupants account for the highest share of fatalities in the region, pedestrians and bicyclists have the highest risk relative to their mobility. A recent study that has considered shared bikes specifically has shown that shared bikes might have a lower crash risk compared to traditional cycling, due to the better structural design of the bikes, more noticeable colors, reflectors, and lights, and lower speed limited by the gear ratio (Zheng and Li, 2020). However, the impacts of shared bike trips while controlling for other factors such as POI visits, taxi trips, etc. on bicycle, pedestrian, or motorcyclist crashes remains unknown.

1.1.6 Traffic flow modeling

The focus of this work is to analyze mixed car-bike flow on two-lane roadways where cars need to cross over to the opposing traffic lane to pass a bike. The mixed car-bike traffic flow is modeled at both the microscopic and macroscopic levels in this work. Typically, microscopic level traffic work considers individual entities such as the social force model for pedestrians (Qu et al., 2017), cellular automata model (Chen and Wang, 2016) or car following models (Li and Sun, 2012) for cars, and agent-based models for traffic demand (Agarwal et al., 2015). The macroscopic approach, on the other hand, considers parameters of homogenous traffic aggregated over time or space, such as traffic flow, density, or average speed. The macroscopic approach is computationally more efficient to determine expected traffic delay based on the traffic condition in a given time-space domain. The fundamental macroscopic traffic model, the Lighthill and Whitham and Richards (LWR) model, is a first-order kinematic wave model that describes traffic states between traffic flow and density, where the speed of traffic is determined through the fundamental diagram (Lighthill and Whitham, 1955; Richards, 1956). Many solutions to the LWR model have been proposed over the past decades, such as finite difference method (Godunov and Bohachevsky, 1959), the cell transmission model (Daganzo, 1994), the Lax-hopf Equation (Daganzo, 2005), etc. The Lax-Hopf equations is widely used for traffic flow modeling (Costeseque, 2018; Vishnoi and Claudel, 2022). Recently, a fast simulation framework to solve the dynamic Lax-Hopf function (Hopf, 1969; Lax, 1957) has been proposed (Claudel and Bayen, 2010a, 2010b). In this solution, the cumulative count at each time-space combination is determined as the minimum value resulting from a series of conditions. The internal conditions, which represents a series of moving bottlenecks, were then also incorporated in the Lax-Hopf framework ((Simoni and Claudel, 2020). This algorithm was made computationally efficient by discarding some of the operations associated with internal conditions. More recently, the Lax-Hopf equations has been utilized to incorporate slow-moving bus as internal conditions in a bus-car mixed traffic flow (Wu et al., 2020).

1.1.7 Car-bike mixed traffic flow modeling

To address the impact of bikes on traffic flow in a macroscopic framework, studies focused on modeling the speed change of vehicles in mixed traffic flow. One study developed two-dimensional speed functions to model the speed changes of cars when overtaking a bike (Wierbos et al., 2018). In that work, it is assumed that cars and bikes share the same road, that bikes are allowed to pass a car queue freely and cars can pass a bike at reduced speed. Another work considered the impact of bikes on traffic speed and concluded that bike lane-sharing behavior significantly influenced the equilibrium speed-density relationships in the LWR model (Chunchu and Kuzhiyamkunnath, 2014). Another line of work considered the modeling of non-lane-based heterogeneous traffic flow, where several classes of vehicles with different desired speeds share the roadway (Metkari et al., 2013). In this non-lane-based scenario, the roadway geometry, prevailing traffic conditions, and static and dynamic properties of vehicles in the traffic stream dominate the traffic (Khan and Maini, 1999). Overall, all existing macroscopic approaches to model mixed traffic flow are based on wide lanes, where cars and bikes can be in the same lane laterally side-by-side, which enables easy overtaking.

1.1.8 Implementation of dedicated bike lane

Creating a dedicated bike lane is a common approach to diminish the car-bike interactions. Several objectives have been considered when determining where to implement bike lanes on a network. These include a utility function that considers the trade-off between coverage of bike trips and continuity of bike lanes (Liu et al., 2021), budget, ease of construction, expected bike lane utilization (He et al., 2020), overall traffic cost, distance and time (Barwaldt et al., 2014). These studies did not consider the negative impact of bike lanes on general traffic. To the author's knowledge only one paper considered both the benefits to the cyclists and potential dis-benefit to the normal traffic flow of bike lanes (Mesbah and Thompson, 2011). However, this study was based on the user behavior choice model where utility functions were used to measure the cost of using different paths but didn't

quantify the impact of bike lanes on the traffic efficiency.

1.2 RESEARCH GAPS

To summarize, the literature typically has focused on the travel and usage patterns of shared bikes or shared e-scooters considering temporal or spatial variations. Most of these works have considered systems where either shared bikes or e-scooters operate individually (not together) and have identified the differences in travel behavior between shared bikes and shared e-scooters. To the author's knowledge, there are no studies that analyze the spatial interaction among shared bike, shared e-scooter, and public transit within the same city. Here, interactions may include competition, where one mode competes or replaces another, or complement, where travelers must rely on two or more than two modes to complete their trips. Given the unique characteristics of shared e-scooters and shared bikes, it is necessary to consider the differences between them when complementing or competing with public transit to better distribute them across the city. For the macroscopic traffic safety models that predict crash frequency over regions of a transportation network, perhaps due to data availability, these models tend to use only exposure attributes related to vehicular traffic and rarely include exposure attributes related to non-motorized travel. Those that do consider non-motorized travel only include at most one explanatory variable. With respect to mixed traffic flow modeling in the presence of slow-moving bikes, the interactions between two directional car-bike mixed traffic flows have not been studied in the literature, especially when cars need to borrow space from the opposing direction to pass a bike. Additionally, the domains of application of bike lanes from a systems perspective, e.g., considering total delay as the sum of bike and car delay, has also not been developed.

1.3 RESEARCH OBJECTIVES

Based on the research gaps, this research will extend the current literature by conducting systematic research surrounding three aspects, e.g., traffic demand, traffic safety, and traffic operation. Three

main goals will be achieved by doing three tasks in this research.

1. Interactions among shared bike, shared e-scooter and public transit

This task is to explore the interaction amongst three non-automobile transportation modes, e.g., shared bike, shared e-scooter, and public transit, that operate within the same urban area. This goal is accomplished by using data from Austin, TX. First the individual spatiotemporal characteristics of shared bike, shared e-scooter and public transit are summarized, and then the spatial interaction among these three modes is modeled considering demographic factors and general trip purposes.

2. Impact of Shared bike on Traffic Safety

The second task to incorporate the impact of shared bike trips on safety performance from a macroscopic (i.e., regional) perspective while taking into account spatial interactions. Traditional explanatory factors—such as traffic network characteristics, roadway network information, and demographics—are also incorporated, along with a novel exposure metric—person trips to various points of interest—to develop a macro-level safety performance function utilizing a geographically weighted negative binomial model for the Manhattan area of New York City. The conclusions will be helpful to understand the relative impacts of these factors on crash risk and contribute to the understanding of how shared mobility services might influence safety performance.

3. Car-bike mixed traffic flow modeling

The last task of this study is to propose a macroscopic car-bike mixed traffic flow model to evaluate the impact of bikes on traffic operations. Specifically, the objectives include 1) to develop a mixed car-bike traffic flow model based on the Lax-Hopf equations considering traffic flow in two directions, 2) to quantifying the overall delay caused by bike flow and evaluate the impact of key traffic flow parameters on the total delay, and 3) to evaluate the traffic efficiency with and without dedicated bike lanes to recommend the application domain of dedicated bike lanes.

2 SHARED MICROMOBILITY AND PUBLIC TRANSIT

Shared bike, shared e-scooter, and public transit make up most public transportation modes in big cities. Their combination can provide a convenient, efficient, and flexible multi-modal transportation service. Despite the obvious similarity among them, differences exist in the roles that they play in a multi-modal transportation system. A case study in the City of Austin, where shared bike, shared e-scooter and public transit coexist, is used to explore the unique characteristics and how they spatially complement or compete with each other.

2.1 DATA DESCRIPTION

In this research, the open access trip datasets from the City of Austin, Texas, in the United States is used. The dataset includes both shared bikes and shared e-scooters operating within a network also served by public transit, and thus provides a unique opportunity to observe the interactions between multiple modes. It should be noted that different cities may have different preference of shared micromobility service based on their own demographics, geography, and policy regulation. Some cities are dominated by shared bike, such as Houston, New York, and Miami, while others are proliferated with shared e-scooter, such as San Antonio. Comparison of shared bike and shared e-scooter between different cities is not meaningful since they are highly influenced by the local characteristics. Only few cities have a relatively balanced development of shared bike and shared e-scooter at the same time, and the City of Austin is an example of them, and thus becomes a good target city for this research.

All datasets are collected in a same time range, e.g., from September 1, 2018, to September 1, 2019, to make them comparable and to rule out the impacts of the COVID-19 pandemic. The origins and destinations of trips from all traffic modes, e.g., shared e-scooter, shared bike, and public transit, are aggregated at the census tract level based on the availability of data.

2.1.1 Shared micromobility dataset

The *Shared Micromobility Vehicle Trips* dataset contains reports of shared dockless electric powered scooter (e-scooter) or shared dockless bike trips provided to the City of Austin Transportation Department as a part of the Shared Small Vehicle Mobility Systems operating rules (Economic Development, 2018). There is only one type of two-wheeled standing shared e-scooter that operates in the City of Austin, and all of the e-scooter data is available in the *Shared Micromobility Vehicle Trips* dataset. On the other hand, there are two types of shared bike services, e.g., the shared dockless bike, which are operated by companies such as Bird, Lime, and the shared docked bike is operated by B-Cycle (now renamed as MetroBike) (“AustinTexas.gov,” 2022). The *Shared Micromobility Vehicle Trips* only includes data from the first of these two services. In 2019 only 12 out of 158 census tracts, mainly in and around downtown, had shared docked bikes, while 97 census tracts had dockless shared bikes. Therefore, this study only takes the shared dockless bike trips into consideration to represent the shared bike trips. Further, the shared bike system in the City of Austin was mostly comprised of regular human-powered bikes during the analysis window in this study, and only a small portion of those bikes were upgraded to electricity-assisted after 2020.

This dataset contains the vehicle type, start time and end time, trip duration and distance, as well as the start and end census tract of each trip, starting from April 2018. A sample of this dataset is shown as Table 2-1.

Table 2-1 Samples of shared micromobility dataset

ID	Vehicle Type	Trip Duration (sec)	Trip Distance (m)	Start Time	End Time	Census Tract Start	Census Tract End
1	scooter	570	826	4/29/19 17:15	4/29/19 17:15	48453001100	48453001100

2	scooter	438	1815	4/29/19 17:30	4/29/19 17:45	48453000601	48453000500
3	scooter	391	1519	4/29/19 17:30	4/29/19 17:30	48453000401	48453000307
4	bicycle	153	0	4/29/19 17:00	4/29/19 17:00	48453000604	48453000604
5	scooter	570	1843	4/29/19 17:30	4/29/19 17:45	48453001100	48453001100
6	scooter	668	2829	4/29/19 17:15	4/29/19 17:30	48453000603	48453001200
7	bicycle	390	1283	4/29/19 17:30	4/29/19 17:45	48453001100	48453002304
8	bicycle	432	1698	4/29/19 17:30	4/29/19 17:30	48453000601	48453000603
9	scooter	196	796	4/29/19 17:15	4/29/19 17:30	48453000603	48453000603
10	scooter	622	3010	4/29/19 17:15	4/29/19 17:30	48453001505	48453001818

The dataset is cleaned to remove rows containing empty values, distances less than or equal to zero or larger than 20 km, durations less than or equal to zero or larger than 2 hours, and incorrect formats, which eliminated 0.37% of the data. After this data cleaning, the dataset spans September 1, 2018, to September 1, 2019, and contains a total of 7.03 million trips, of which 6.68 million are e-scooter trips and 0.35 million are shared bike trips. Even though the shared e-scooter and shared bike trip columns are not well-balanced, the sample sizes are sufficient to reflect the spatial pattern of each of them. The spatial distribution of the origins of the micromobility trips in the City of Austin is shown in Figure 2-1a. From this figure it can be observed that micromobility trips are mostly concentrated within the center of the city.

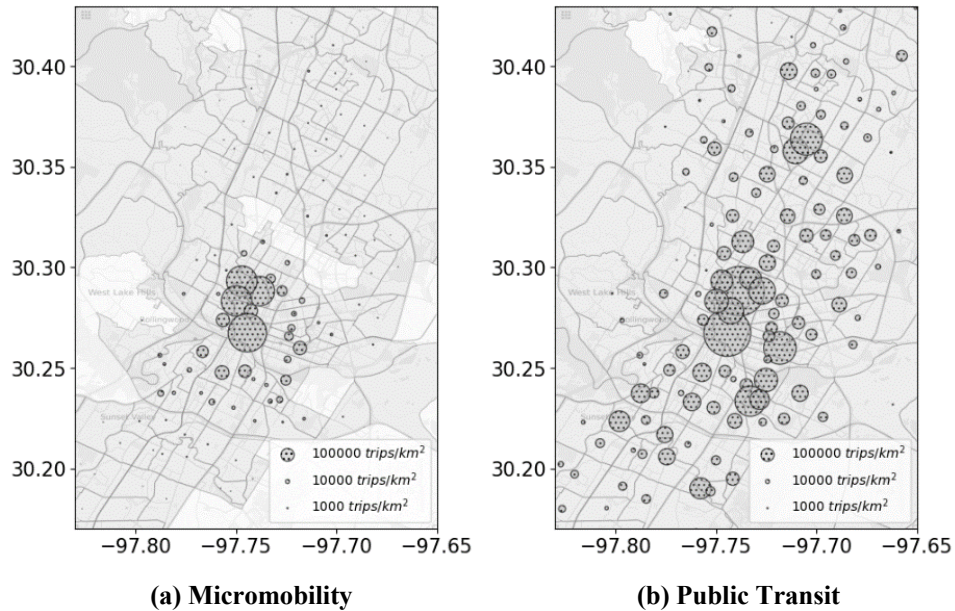


Figure 2-1 Spatial distribution of different traffic modes in the City of Austin

2.1.2 Public transit dataset

Capital Metro is the City of Austin's regional public transit provider with services including bus, shuttles, and freight rail. The *Capital Metro Automatic Passenger Counter (APC) Ridership* dataset is obtained directly from the Automatic Vehicle Location System (AVL) installed on buses (Capital Metro, 2020). Each row contains the specific bus door open and close time at each stop accompanied by the geographic coordinates and the number of passenger that board and alight through all doors at this stop. Samples of the APC dataset are shown in Table 2-2, with key attributes collected from three times alighting and boarding process at the stop. The raw APC data contains abnormal records; therefore, a data cleaning and pre-processing were adopted first. The APC data is available from January 1, 2016, to December 31, 2019. However, due to service changes or missing data, some of the bus routes do not have complete data during this period. When bus routes with more than 5 months of missing data are removed, 57 bus routes remain, and these bus routes account for the majority (96%) of the total passenger volumes across the City of Austin. Therefore, the analysis is conducted using the number of alighting passengers on these 57 bus routes aggregated at the census tract level

during the analysis window. The spatial distribution of public transit ridership is shown in Figure 2-1b.

Table 2-2 Samples of Capital Metro APC ridership dataset

Attributes	Sample 1	Sample 2	Sample 3
Door open date time	8/17/2019 23:57:32	8/18/2019 0:05:31	8/18/2019 9:29:12
Door close date time	8/18/2019 0:00:01	8/18/2019 0:05:43	8/18/2019 9:29:24
Dwell time (sec)	149	12	12
Number of max loads	7	3	4
Number of alighting	0	1	1
Number of boarding	3	0	0
Vehicle latitude	30.3941	30.2679	30.2869
Vehicle longitude	-97.7484	-97.7428	-97.7335
Route ID	3	1	7
Vehicle ID	2373	2055	2625

2.1.3 Supplemental datasets

To further understand factors that lead to spatial variation in the relationship between shared micromobility modes and public transit, supplemental datasets are utilized. The goal is to understand

how different trip purposes might impact usage of different modes. Hence, the *footage data of point of interests (POIs)* dataset available from *SafeGraph* is utilized. This dataset contains the hourly visit volumes of more than 6 million POIs across the US and is collected through various methods mainly based on mobile phone devices (SafeGraph Inc., 2020). This different POIs could help describe both trip purpose and land use within a census tract. After removing the abnormal data points, empty rows, and outliers, the POIs located within the City of Austin are extracted and the visit volumes in the analysis window are aggregated within each census tract. To simplify the model and clearly conclude which type of traffic demands associate with higher competition among shared micromobility modes and public transit, all the 33,590 POIs within the City of Austin are grouped into 9 categories considering similar purposes as shown in Table 2-3.

It is noteworthy that even though this study does not consider the transfer between different traffic modes, the category of *transportation and motor vehicle POIs* consists of the visit volumes of transportation hubs, terminals, airports, and car rental companies, which can be used to estimate the impact of other traffic modes. Summary statistics on each group of POIs are provided in Table 2-3.

Table 2-3 POI footage dataset summary

Groups	Total Visits	No. of OPIs	Average Visits	Max Visits	Min Visits	Visits Std.
Health Care	9,563,082	6,827	7,001	727,249	0	33,171
Manufacturing and Wholesale	630,527	357	469	145,354	0	4,230
Miscellaneous and Grocery Stores	26,611,822	5,856	20,208	1,387,663	0	89,466

Schools	7,721,565	2,812	5,628	396,259	0	23,528
Amusement and Recreation	41,036,456	3,191	30,965	2,114,891	0	106,300
Transportation and Motor Vehicle	19,017,290	1,362	14,646	12,821,792	0	282,680
Food and Drink	70,845,702	10,178	53,613	3,134,511	0	213,116
House Maintenance	4,288,505	565	3,208	1,047,156	0	36,096
Public Services	1,953,962	2,445	1,452	117,530	0	6,815

Further, publicly available census data from the U.S. Census Bureau is utilized to represent the demographic characteristics of each census tract. For the purposes of this analysis three categories of information are collected: individual information (e.g., gender, education, birthplace, occupation), household information (e.g., household size, household type, number of workers per household, vehicle ownership) and transportation related information (e.g., time of departure from home for commute, commute mode of transportation, commute travel duration). There are a total of 69 demographic variables available in this dataset. The demographic characteristics used in the models are provided in the *RESULTS AND DISCUSSIONS* section.

2.2 METHODOLOGY

The methods used in this study are briefly described in this section. First, a decomposition model to identify the difference in temporal travel patterns amongst public transit, shared bike and shared e-scooters is discussed. Next, geographically weighted regression to explore the spatial relationship is presented.

2.2.1 Temporal variation analysis of each traffic mode

A decomposition model is designed to identify the unique characteristic of each mode's travel behavior pattern. The rationale is that temporal variations in the data that is observed is the outcome of multiple travel behaviors overlapping together. As such, to understand the underlying travel behaviors, the aggregated observations are decomposed to multiple parts.

The number of trips on a given day (Monday-Sunday) is firstly grouped into bins of 1-hour and averaged for the same day for all the available data. In other words, the final data represents the average number of trips observed by a given mode, on a given day (Monday-Sunday) for a given 1-hr time period. Then, the total weekly trip volumes of each mode are normalized to be between 0 and 1 to make them comparable with each other. The temporal variation is calculated as **Equation 2-1**.

$$C_w(t) = \frac{1}{C_T} \sum_{i=1}^N C_{i,w}(t, t + 1) \quad 2-1$$

where t is the time of day in units of hours; $w = 1, 2, 3, \dots, 7$ represents Monday to Sunday; $C_w(t)$ is the temporal variation at time t , for day w ; C_T is the total trips volumes of the analysis year; N is the number of weeks in the analysis year; $C_{i,w}(t, t + 1)$ is the trip volumes of day w in i^{th} week from time t to $t + 1$.

Next, to quantitatively compare the temporal variation of each traffic mode, the daily trip curve is decomposed into multiple curves (represented by distributions), each presenting a major unique trip behavior pattern in a day. The decomposed results offer insights into the underlying patterns. A decomposition model is implemented for each dataset (i.e., shared bike trips, e-scooter trips, and public transit trips) as in **Equation 2-2**:

$$C(t) = \sum_{i=1}^N \delta_i F_i(t) \quad 2-2$$

where $C(t)$ is the the number of trips by a given mode that occur at time, t (in hours); N is the number of components indexed by i (for example, if we decompose the daily aggregated curve into morning peak, afternoon peak and off peak periods, we will have $N = 3$); $F_i(t)$ is the probability density

function of component i ; and δ_i is the coefficient, or a scaler, of component i . The components represent broad categories of types of trips that happen, e.g., morning commute, afternoon commute, and general day trips. Here, even though the data uses discrete time steps, t is used as a continuous variable such that continuous functions of $F_i(t)$ can be used.

According to the characteristics of the trip distribution, $F_i(t)$ can take different forms, such as normal, lognormal, or Weibull distribution etc. Based on testing experiments, the normal distribution performed very well on the three traffic volume datasets and achieved comparable relative error compared with the asymmetric lognormal distribution. On the other hand, normal distribution has the simplest structure and can improve the model efficiency significantly compared to the other distributions. Therefore, $F_i(t)$ is assumed to follow normal distribution for the morning peak, afternoon peak, and the general day trips as **Equation 2-3** in this study

$$F_i(t) = \frac{1}{\sqrt{2\pi}\sigma_i} e^{-\frac{(t-\mu_i)^2}{2\sigma_i^2}} \quad 2-3$$

where μ_i and σ_i are the mean and standard deviation of component i , respectively, in units of hours.

2.2.2 Spatial correlation analysis among different modes

A common method to analyze the spatial correlation among different traffic modes is to use a regression model. For example, when analyzing the impact of shared bike and e-scooter trips on public transit ridership, public transit ridership could be selected as the dependent variable and the trip volumes of shared bike and e-scooter, along with other impact factors selected as explanatory variables. If a positive (negative) coefficient is estimated for the shared bike trip volumes this would imply that higher shared bike usage stimulates more (less) public transit ridership, indicating a complementary (competing) relationship between shared bike and public transit.

A global regression model can capture the general relationship between the explanatory variables and

the response variable by assuming these relationships are spatially independent. However, for large cities like Austin, the topography, population density, and land use type vary across different parts of the city. Latent factors such as bus service accessibility, topography, and operation policy that vary in a spatial dimension can make these assumptions invalid. The interactions between traffic modes could be influenced by the local characteristics. Thus, a more flexible model is necessary to model these variations across different geographic regions. Instead of estimating a constant coefficient for each explanatory variable in a global regression model, the geographically weighted regression (GWR) assigns a set of coefficients of explanatory variables for each spatial unit (e.g., census tract in this study). GWR is well-known for its capability to model heterogeneity (e.g., non-stationarity) across different areas and is widely used for traffic demand prediction. The sub-model for each spatial unit can be expressed as in **Equation 2-4**.

$$y_j = \sum_{k=0}^N \beta_{jk} x_{jk} + \varepsilon_j \quad 2-4$$

where y_j is the dependent variable (e.g., number of e-scooter trips) in a census tract, j ; N is the number of explanatory variables used in the model; x_{jk} is the k^{th} explanatory variable in census tract j ; β_{jk} is the coefficient of the k^{th} explanatory variable in census tract j ; and the ε_j is the residual in census tract j .

An ordinary least squares model (OLS) is also developed for comparison. Traditionally, the coefficients of an OLS are estimated using **Equation 2-5**. However, the coefficient estimation in GWR is also impacted by the weighted neighbors as shown in **Equation 2-6**.

$$\hat{\beta} = (X^T X)^{-1} X^T y \quad 2-5$$

$$\hat{\beta}_j = (X^T W_j X)^{-1} X^T W_j y \quad 2-6$$

where X is the input matrix; y is the output matrix; W_j is a matrix of weights specific to census tract j such that observations nearer to j are given greater weight than observations further away.

The weights of these neighbors, W_j , follows a distance-decay function to emphasize the impacts of

near census tracts and to ignore the ones further than a certain distance. Hence, the number of neighbors that have influence, referred to as the bandwidth, also needs to be selected. A small bandwidth may result in an unstable fit in a small regional area, while a large bandwidth may introduce bias, and neglect the regional variation (Munira and Sener, 2020). The bandwidth can be defined based in two ways, either on a distance threshold or a specific number of neighbors. The corrected Akaike Information Criterion (AICc) is generally used as a criterion to evaluate the model performance and, therefore, can be used to select the optimum bandwidth to achieve the best model performance. The AICc is calculated as **Equation 2-7**.

$$AICc = 2k - 2\ln(L) + \frac{2k^2 + 2k}{n - k - 1} \quad 2-7$$

Where, k is the number of estimated parameters in the model; L is the maximum value of the likelihood function; n is the sample size of the dataset.

The model that achieves the lowest AICc is selected as the best model and the corresponding bandwidth is selected as the optimal bandwidth (Akaike, 1974). The mathematical solution of the GWR can be found in (Brunsdon et al., 1998). For this work, an open-source package named as *pysal/mgwr* in the Python platform is used to estimate the parameters (Oshan et al., 2018). The outcome of GWR consists of a set of coefficients for each census tract for each explanatory variable and the corresponding R^2 value which measures the goodness-of-fit.

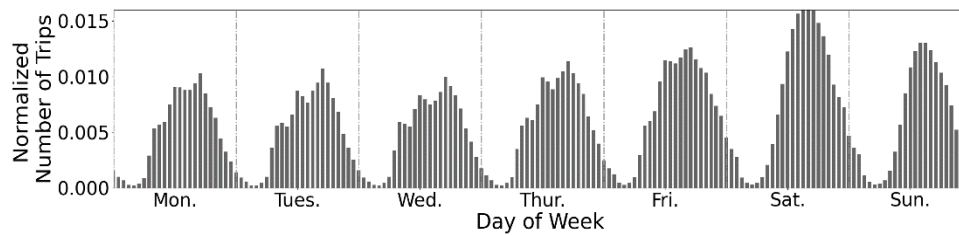
2.3 RESULTS AND DISCUSSIONS

2.3.1 Individual characteristics of shared bike, e-scooter, and public transit

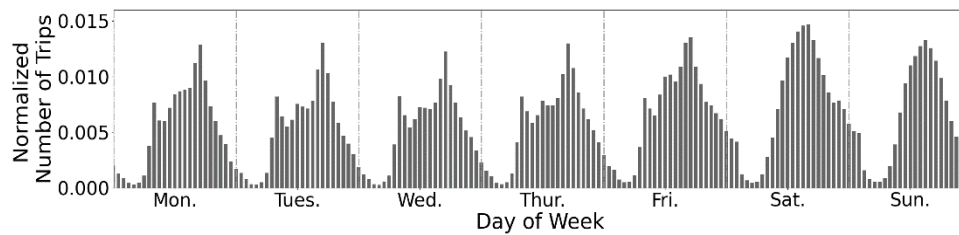
First, some basic statistics are calculated to understand the general trends in data. The average distance, duration, and speed of an e-scooter trip is found to be 1.51 km, 10.65 mins, and 9.60 km/h, respectively. Compared to the e-scooter trips, shared bike trips have a longer average distance and duration of 2.62 km and 14.62 mins, respectively. The average speed of shared bikes is also higher

than e-scooters at 12.14 km/h. While the specific average speed of travel by public transportation in Austin is not provided, in the United States the average speed of travel by public transportation is 22.69 km/h. Also, the average travel distance on Capital Metro was 5.5 km according to the APTA factbook (American Public Transportation Association, 2020). Hence, the general statistics suggest that micromobility is used on average for shorter trips than public transit.

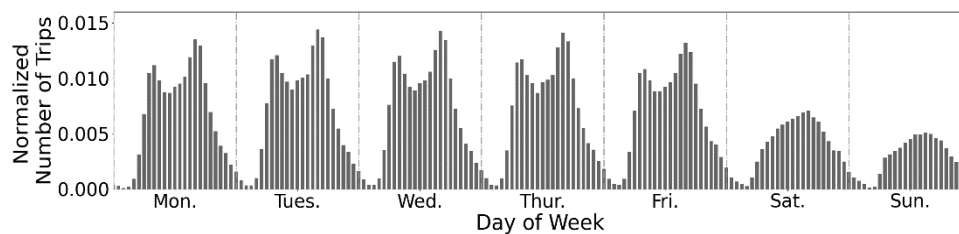
Details in the temporal variations of trips are explored by aggregating trips taken by each mode every hour considering the day of the week. Then, the average hourly trips in the analysis window for each day is determined to show the temporal distribution of each mode over a week. The temporal distributions of e-scooter, shared bike, public transit, and motorized vehicle trips across one week are shown in Figure 2-2.



(a) E-Scooter



(b) Shared Bike



(c) Public Transit

Figure 2-2 Temporal distribution of different traffic modes across one week. Data are normalized based on total weekly trip volume and aggregated over analysis window

From Figure 2-2, the daily trip volume on weekend days is higher than during weekdays for micromobility modes but lower for the public transit mode. The temporal distribution of different traffic modes indicate that they are preferred to be used for different travel purpose. To further explore the possible travel purpose of those three traffic modes, they are compared to the different type of POI footage temporal distribution shown in Table 2. The average temporal distribution of the 9 groups on each day of a week in a bin of 1 hour is extracted and normalized to compare with the traffic modes usage pattern. The similarity between the POI visit pattern and traffic modes usage pattern is measured by the Euclidean distance between each pair, and the distance is shown in Figure 2-3.

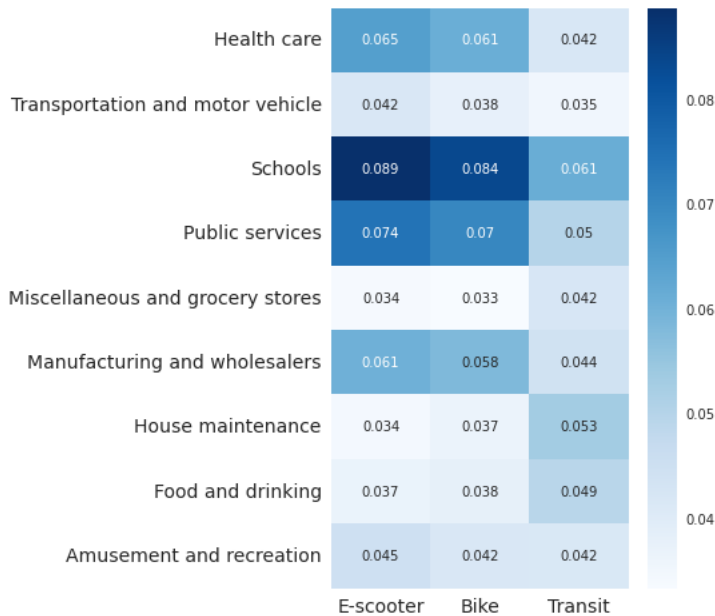


Figure 2-3 Euclidean distance between POI visits temporal patterns and traffic modes usage patterns.

From Figure 2-3, it can be seen that public transit has shorter distance from the school, public service,

and manufacturing and wholesalers POIs, which indicates that the public transit tends to be used as a commute tool to work. On the other hand, the temporal patterns of shared e-scooter and shared bike are similar to the footage associated with food and drinking, grocery stores and house maintenance, which is more like for leisure purpose.”

Next, the decomposition model for all seven days of the week and all three traffic modes is implemented considering three normally distributed components (e.g., 21 different decompositions are estimated). The decomposition models are estimated based on a smaller bin, e.g., 15 mins, to ensure the accuracy of the fitting curves. Given the good fit of a three-component decomposition model more components are not explored. The three-components are assumed to mimic morning peak, afternoon peak and a general day travel. The results suggest that trends for weekdays are similar to each other, and trends for weekend days are also similar. Hence, to demonstrate the results of the decomposition model for each traffic mode, Tuesday and Saturday are chosen as a typical weekday and weekend day, respectively. To accommodate the trips generated after 23:59 each day, the daily trip distributions are plotted from 6:00 to 5:59 (+1 day) for the shared e-scooter and shared bike, and from 4:00 to 3:59 (+1 day) for public transit ridership. The fitting curves are shown in Figure 2-4. In this figure, the dotted lines represent the three different components, while the solid line represents their summation. The real data is shown as a gray area. In general, it can be seen that the morning peak appears as the left-most distribution, and the afternoon peak appears as the right-most distribution. In Figure 2-4e and Figure 2-4f the morning the morning peak is not visible because it is very small. Note that it is possible to have more than three components in Equation. 2. For example, Figure 4e show that there is another component of shared bikes around 2:00 on Saturday. However, since this late evening component (around 2:00) only happen in shared bike trips during weekend and not observed in other traffic modes on other days, for simplicity, it was not considered in this study.

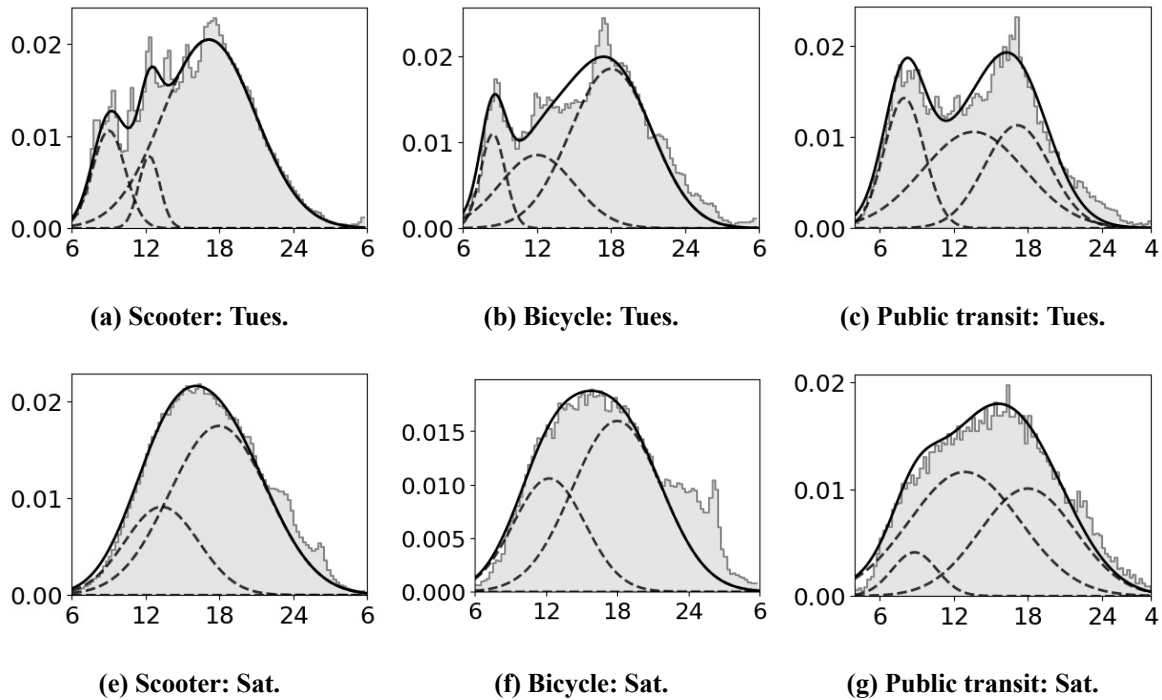
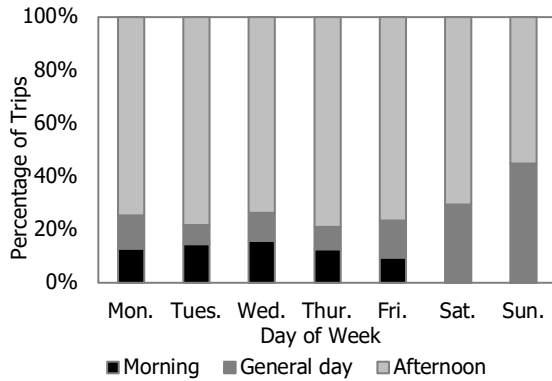
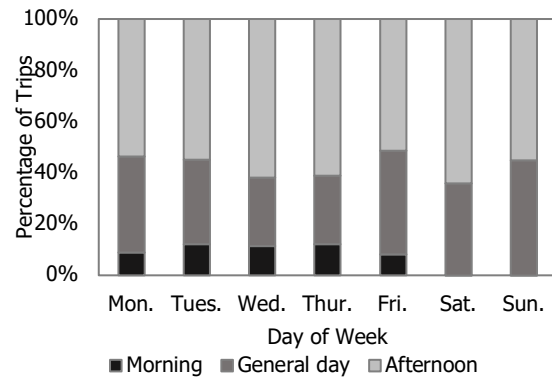


Figure 2-4 Normalized trip volume vs. time-of-day decomposition models. (The dashed lines show the results of the fitting curve for each component, and the solid curves are the fitting curve for the entire day. The shaded areas represent the real data.)

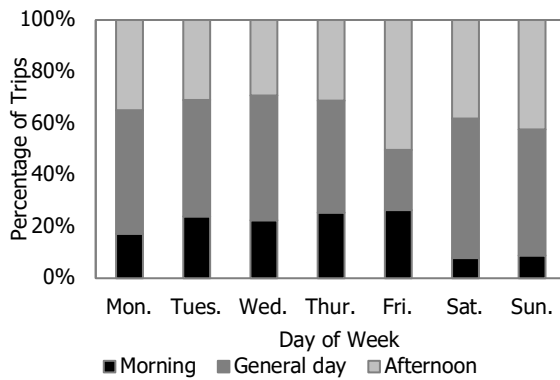
Using the decomposition results, the percentages of trips that account for the morning peak, afternoon peak and general day trips are calculated and shown in Figure 2-5. The results suggest that public transit exhibit a strong morning peak during the weekdays, which is not the case for weekends. 58% of the public transit trips can be attributed to the morning peak and afternoon peak components. The ratio of the morning peak to afternoon peak is also the highest, about 0.65, which indicates a temporal symmetric distribution associating with commuting trips (Bordagaray et al., 2016).



a). E-scooter



b). Bicycle



c). Public Transit

Figure 2-5 Percentage of each component for different traffic modes

From the e-scooter trips decomposition results in Figure 2-5, it can be found that 61% of all e-scooter trips can be attributed to the afternoon peak component, and this percentage increases to approximately 70% when considering only Friday and Saturdays. Additionally, the percentage of trips attributed to the morning peak for e-scooters is very low compared to other traffic modes, and the ratio of the morning peak to afternoon peak (0.09) shows that e-scooter usage is asymmetric during a day. These results indicate that e-scooters are more likely to be used as a leisure travel mode during the evening, especially on weekends, rather than a commute tool. Shared bike trips also tend to be attributed to the afternoon peak component, however the ratio of morning to afternoon peak trips is slightly higher at 0.19. Also, recalling that the shared bike mode has a longer trip duration and

travel distance, it can be concluded that the shared bike mode is more frequently used for commuting than e-scooters. Further, combining with Figure 2-2 it can be observed that for both e-scooter and shared bike trips, there is a slightly higher tail followed by a shoulder in the evening (around 11 pm) on Fridays and Saturdays. This indicates that micromobility is likely chosen as a leisure travel mode for weekend late evenings. These conclusions are consistent with findings in the literature (McKenzie, 2019) and further provide a quantitative comparison between three traffic modes which indicate that they have specific usage and travel patterns.

To summarize, the data indicates that public transit is mainly used for commuting while micromobility modes are mostly used for leisure activities in the evening and on weekends. Specifically, the shared bike is more likely to be used for commuting compared to e-scooter. This quantitative analysis of each traffic mode can help to distribute different traffic modes and build a harmonious traffic network for the different demands.

2.3.2 Spatial correlation among shared bike, e-scooter, and public transit

2.3.2.1 Feature selection

All data are aggregated at the census tract level within the entire analysis year to perform the spatial regression. After combining all available demographic characteristics with the nine POI categories, a total of 78 potential variables that might influence micromobility ridership exist in the spatial domain. Including all of the potential variables into the model is not reasonable since the true impact of individual variables cannot be shown. The commonly used feature selection method in statistical models is to manually select variables based on their significance level, e.g., p-value. However, when the size of potential variables is large, this could be inefficient since it requires re-estimating the model many times, and the interaction among different variables could be missed in this process. Therefore, a Boruta feature selection, which is commonly used in machine learning methods, is adopted to pre-select the relevant variables. The Boruta method has two advantages: 1) the

effectiveness of each variable is determined considering how they improve the model compared to a randomized shadow of themselves, hence variables do not compete with each other, and 2) repetition – the results become robust through iterations. The importance of each variable is determined by considering the average results of all iterations (Mazzanti, 2020). The selection is based on the demographic characteristics, and general POIs information for a given census tract and the relevant micromobility or public transit ridership. Based on the importance level of variables from the Boruta feature selection, 13 variables are that can impact e-scooter ridership, 12 variables that can impact shared bike ridership, and 9 variables that can impact public transit ridership are selected as shown in Table 2-4.

Table 2-4 Selected features for e-scooter, bike or transit usage prediction

Variable name	Description	Selected for model		
		E-scooter	Bike	Transit
Travel time - 5 to 9 mins	# of people whose commute time is 5 to 9 mins	✓	✓	✓
No vehicle	# of households that do not own a car	✓	✓	✓
POIs - amusement and recreation	# of visits to amusement and recreation POIs	✓	✓	✓
POIs - food and drinking	# of visits to food and drinking POIs	✓	✓	✓
POIs - miscellaneous and grocery stores	# of visits to miscellaneous and grocery stores POIs	✓	✓	✓
Public transit	Public transit ridership	✓	✓	✗
Transportation - motorbike, bike or other	# of people who commute by motorcycle, bicycle or other.	✓	✓	✗

Household type - nonfamily	# of nonfamily type households	✓	✓	✗
Household - 1	# of single-person households	✓	✓	✗
Vehicle - 1	# of households with one vehicle	✓	✓	✗
Leaving home - 830 859	# of people departing home between 8:30 and 8:59 am	✓	✓	✗
POIs - transportation and motor vehicle	# of visits to transportation and motor vehicle POIs	✓	✓	✗
Transportation - walked	# of people who walk to work	✓	✗	✗
Scooter trips	#of shared e-scooter trips	✗	✗	✓
Bicycle trips	# of shared bike trips	✗	✗	✓
POIs – public services	# of visits to public service POIs	✗	✗	✓
Transportation – public transit or taxicab	# of people who commute by public transit or taxicab	✗	✗	✓

The feature selection results suggest that a few of the variables impact the usage of all three modes considered, including a short travel time to work, no vehicle ownership, and POIs related to amusement and recreation, food and drinking, and miscellaneous and grocery stores. Most of these POIs are associated with leisure travel, which is consistent with the literature.

As expected, it is found that public transit ridership is also an important feature in determining the shared e-scooter or shared bike usage. Other variables that impact both shared e-scooter and shared bike usage, but that don't impact public transit ridership are number of people who commute by taxicab, motorcycle or bicycle, nonfamily type or single-person households, households with a single vehicle, departure time from home and number of visits to transportation and motor vehicle POIs. Most of these variables make sense as they can describe a need for additional modes of transportation. Overall, the factors that impact e-scooter or bike ridership are very similar but only differ in one

variable, which is whether people walk to work in a given census tract or not. This is reasonable since the average travel distance by shared bikes is longer than by e-scooters and in general shared bikes are not expected to replace walking trips like e-scooters do.

There are four features that are unique to transit usage, which are scooter and bicycle trips, public service POIs and number of people who commute by public transportation. In general, the features relevant for transit are consistent with the literature (38, 39).

2.3.2.2 Geographically weighted regression results

To explore the spatial correlations among shared bikes, e-scooters, and public transit, three GWR models are developed by selecting the shared bike trip volumes, e-scooter trip volumes and public transit ridership as response variable, respectively. The results suggest that the shared bike model and the shared e-scooter model are dominated by each other, and the other variables become insignificant due to the high correlation between shared bike and shared e-scooter usage. Hence, here only the results of the Transit-GWR model, which considers public transit ridership as the dependent variable and the other two as independent variables, is considered. This transit model, which takes shared bike and e-scooter trip volumes along with other impact factors as explanatory variables, demonstrates a relationship between public transit usage and the other two micromobility modes. These relationships help to depict a clear spatial correlation pattern among these three public traffic modes. The detailed specification and the results of the model are introduced here.

The optimal bandwidth for the Transit-GWR model is determined as 42 using the AICc criterion. For comparison, a global linear regression model, is also estimated using ordinary least squares (OLS) method. In terms of fit, a higher log-likelihood ratio was achieved for the Transit-GWR model (-2213.78) as compared to the OLS model (-2303.61) implying better fit. The Transit-GWR also achieves smaller AICc values, (4556.97, compared to 4627.22 for OLS) and higher average R^2 (0.93, compared to 0.78 for OLS), which suggests that the Transit-GWR model outperforms the OLS model

significantly and achieves higher prediction accuracy. The R^2 for each census tract in the GWR model is shown in Figure 2-6. From Figure 2-6, it can be seen that the GWR model achieved high reliability in the center and north areas of the City of Austin and most of the census tracts have R^2 larger than 0.8.

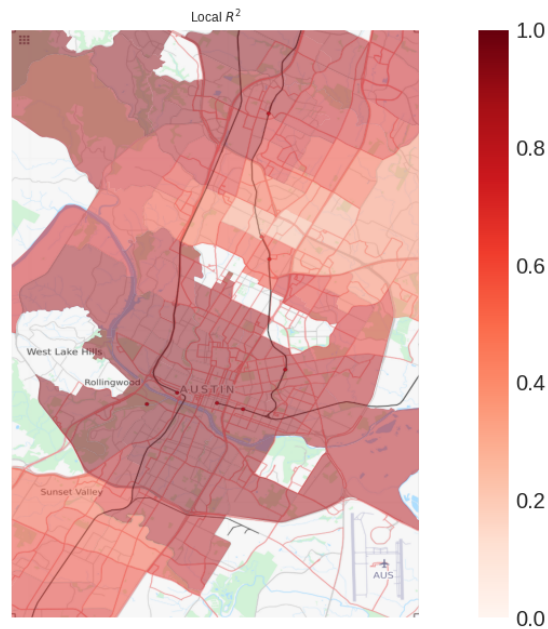


Figure 2-6 R^2 distribution of Transit-GWR model

The coefficient estimations of the OLS results and statistical summary of the Transit-GWR results are shown in Table 2-5. The OLS estimation results show that shared bike is positively correlated with public transit with 95% confidence, but e-scooter trips are found to be insignificant to public transit ridership in the global model. Different from the global model, the Transit-GWR model provides more detailed insight. For the relationship between shared bike usage and public transit ridership, instead of a general positive correlation across the whole city, the Transit-GWR model suggests that the shared bike usage is significantly correlated with public transit ridership in 87 census tracts, in which 53 of these have a positive correlation while the remainder indicate a negative

correlation. This suggests that the relationship between shared bike usage and public transit ridership is not homogenous as the global model indicates. Similar results are observed for the relationship between e-scooter usage and public transit ridership as well; the Transit-GWR model suggests that e-scooter usage is significantly correlated with public transit in 86 census tracts, and 57 of them are positively correlated with each other. Those spatial variations of correlation between micromobility modes and public transit capture the impacts of regional factors, some of which are present in the model. Most of the POI groups are not evenly distributed across the city, such as public service and amusement and recreation, and the preferred transportation mode as well as the expected travel time to the workplace are significantly different across census tracts. These differences are captured in the GWR model by assigning different coefficients to different census tracts as shown in Table 2-5. Besides those differences, the global model and GWR model also get some consistent results with respect to several impact factors. For example, the OLS results suggest that the visit volumes of food and drinking POI should be significantly and positively correlated with public transit, which is confirmed by the GWR results for 79 census tracts.

Table 2-5 Spatial regression results of Transit-GWR model

Variables	OLS Results		GWR Results			
	Coef.	P> z	Mean	# Signif Coef	# (+) Coef.	# (-) Coef.
Constant	-36000	0.70	39600	22	20	2
Scooter trips	-1.0	0.44	40.8	86	57	29
Travel time - 5 to 9 mins	-288.4	0.41	-42.9	31	10	21
Bicycle trips	69.1	0.01	8701.5	87	53	34
POIs - public services	2.3	0.33	-0.7	37	6	31

POIs - miscellaneous and grocery stores	0.4	0.67	1.9	28	28	0
POIs - amusement and recreation	1.1	0.31	0.3	55	39	16
POIs - food and drinking	3.2	0	3.1	79	79	0
Vehicle - no	314.9	0.51	192.9	55	33	22
Transportation – public transit or taxicab	836.9	0.10	121.2	41	22	19

Figure 2-7 shows the spatial distribution of the significant coefficients of the GWR model. The most pronounced conclusion is that the correlation between public transit and shared micromobility modes are different from east to west of the city, where the shared bike and e-scooter have approximately reversed correlations with public transit. For most census tracts where the e-scooter ridership is positively correlated with public transit ridership, shared bike usage is negatively correlated. This appears to imply that only one of these micromobility forms, either shared bike or e-scooter, will tend to serve as a complementary mode to public transit based on specific local characteristics. E-scooter appears to be more likely to be positively correlated with public transit in the west of the City of Austin and negatively correlated in the east areas. It was noteworthy that the shared bike system in the City of Austin was dominated by the human-powered bikes during the analysis window in this study, and only a small portion of those bikes are upgraded to electricity-powered after 2020. Therefore, a possible explanation of this spatial pattern is that the west of the City of Austin is more mountainous and hence e-scooters are more attractive in this region due to the electric power. Other characteristics of the city also vary between the east and west sides and may contribute to this pattern, though. The demographic variables of 5 to 9 mins of commuting time, and households without vehicle ownership show negative correlation with public transit in the city center and positive correlation in the north and south areas. This is consistent with expectation that in the center of the city, more alternative modes present, such as shared micromobility which are more convenient and efficient than public transit in a congesting traffic network. Hence, travelers might be more like to

use the shared e-scooter or shared bike in the city center instead of public transit.



Figure 2-7 Local coefficients of variables in the Transit-GWR model

2.3.3 Discussion

This study analyzes the usage patterns of shared bike, e-scooter, and public transit, in a market that they co-exist, from both individual and interacting perspectives. The unique datasets and spatial regression model used in this study allows us to offer insights to the existing literature about the shared micromobility usage in urban areas.

First, the individual characteristics of shared bikes, e-scooter and public transit usage are highlighted through a temporal pattern analysis. The decomposed components of a whole day's trips distribution suggest that public transit usage patterns mimic commuting patterns, while shared micromobility modes exhibit trip distributions that are not typical for commuting. Public transit is more likely to be used during weekdays and has a symmetric distribution on morning and afternoon peaks, while the micromobility modes are more preferred during the weekend, especially in the afternoon and late evening. When comparing the temporal usage pattern of shared bikes to e-scooters, shared bikes demonstrate usage patterns more similar to commuting patterns than the e-scooter, evidenced by the more pronounced morning and afternoon peaks.

The spatial interactions among shared bike, e-scooter and public transit are then explored by developing a set of GWR models for public transit ridership considering the e-scooter, shared bike, POIs visits and demographic information. Instead of incorporating static public transit variables such as stop density or route density used in the literature, this chapter used the trips volume of the public transit within each census tract to reflect the dynamic demand of public transit. The POI trip information is also included into the model with the actual visit volumes instead of POI density or other static land used indicator. Those unique datasets improved the reliability of the model results.

This study filled existing research gaps by providing a detailed comparison of temporal and spatial characteristics among shared bike, e-scooter, and public transit. Different from the existing literature that analyzed their travel patterns and impacts of e-scooters and shared bikes separately (McKenzie, 2019; Reck et al., 2021; Younes et al., 2020), this study analyzed the interactions among the two

types of micromobility modes and public transit within the same area. Some literature concluded that the correlation between shared bike and public transit may be different between short distance trips and long distance trips (Guidon et al., 2019; Kong et al., 2020; Levy et al., 2019), this study shows that this division exists between shared bike and e-scooter as well, where e-scooters complement public transit in the west of the City of Austin and shared bikes promote public transit usage in the east areas. This division might be caused by the demographic information, geographical characteristics, or the electrical assistance of e-scooters. These conclusions extended the understanding of micromobility especially with respect to differing interactions with public transit.

2.3.3.1 Practical significance

The conclusions from this research can provide a solid reference to multi-mode transportation practitioners when developing and operating shared micromobility systems. It suggested that the development of one type of shared micromobility service requires to fully considering the existing traffic network to achieve the maximum benefit. The land usage, demographic information as well as the geographical information could significantly influence users' preference. Both as a type of micromobility services, shared e-scooter and shared bike share some similarity. However, they are still slightly differentiated in respect of user group, travel time and trip purpose due to the power system, convenience, flexibility and ability to ride a long distance. Based on the case of the City of Austin, it's recommended that the shared e-scooter should be distributed around recreation center, famous sites and grocery stores, while shared bike will be more benefit to distributed around the public transit stops, transportation hubs, and commercial business district to provide a connection between different modes and service for the last mile commuting. In the mountainous area, it's preferred to adopt some electric powered modes, such as e-scooters and e-bikes.

The decomposition model and geographically weighted regression developed in this study is based on the specific datasets of Austin, however, the methodology can be implemented to any other city

that are planning to adopt new shared micromobility programs or optimize the existing distribution of those micromobility facilities.

2.4 SUMMARY

This chapter analyzed spatiotemporal travel patterns of two types of micromobility modes and their interaction with public transit considering the impact of POIs visit volumes, and demographic information. Specifically, the research gap of the different roles of e-scooters and shared bikes when coexisting with public transit is addressed. To do so, shared dockless e-scooter and shared dockless bike data in the City of Austin, Texas, are used. Since both modes co-exist along with public transit in this city, users can select different traffic modes freely and thus are able to reflect the choices of two types of micromobility modes in a background of public transit. The results of the individual characteristics analysis indicated that e-scooters are more likely to be used for leisure trips than public transit and shared bikes based on the temporal distribution of usage and the ratio of the morning peak to the afternoon peak trips. Similarly, shared bikes are also used for leisure trip but appears to have more usage during the peak hours, too, compared to e-scooters. The Transit-GWR model showed that, while e-scooter and shared bike usage is highly positive correlated with each other, e-scooters tend to complement public transit, particularly in the west areas with more mountainous terrain, by providing a connector, collector, and distributor for public transit trips. However, shared bikes are more likely to complement public transit and promote it in the east of the city, where the population is less affluent and public transit density is larger. These differences suggest that shared bike and shared e-scooter have different roles in relation to public transit considering the built environment and the demographics.

This chapter suggests that implementing a single mode strategically could help improve access to public transit, but when two of the micromobility modes co-existing, residents could have a

preference for one over the other based on the specific regional environment. In general, only one of these modes will complement public transportation system in a given area. Since shared bikes and e-scooters attract users with different travel purposes, the implementation of these modes to complement public transportation can be chosen based on land use, demographic, and terrain characteristics in different areas. This work can inform local policies that regulate micromobility with goals of reducing congestion and increasing environmental sustainability.

A limitation of this work is the actual trip purpose of the different traffic modes users was unknown. A trip purposes survey in the future could help to examine the relationship between micromobility usage and geographical information to derive more detailed reasons for the different correlation distributions.

3 SHARED BIKE AND TRAFFIC SAFETY

Macroscopic traffic safety models that predict crash frequency over regions of a transportation network are becoming increasingly common. However, perhaps due to data availability, these models tend to use only exposure attributes related to vehicular traffic and rarely include exposure attributes related to non-motorized travel. Those that do consider non-motorized travel only include at most one explanatory variable. This chapter seeks to extend these existing studies by incorporating several exposure metrics that capture non-motorized and public transportation use at the census-tract level. A macro-level crash prediction model for the Manhattan area of New York City is developed that considers roadway and demographic variables, as well as bike share trip information, subway flows, taxi movements, and person-trips to various points of interest (POI) as measures of travel exposure. The models are developed using negative binomial regression and various functional forms are considered. The spatial variations between crash frequency and the selected exposure metrics are further explored by developing a set of geographically weighted negative binomial regression (GWBNR) models for three different types of crash frequency.

3.1 Data Description

Several unique datasets were obtained and combined for use in this study. Data on individual crashes in New York City were obtained from the New York City Open Data Portal to be used as the dependent variable in the model (Police Department (NYPD), 2022). Shared bike trip data were obtained from the Citi Bike system to evaluate the impact of mass shared bike trips on the traffic safety (Citi Bike, 2022). In addition, Point of Interests (POIs) visits data was obtained from *SafeGraph* to be used as a surrogate for general trip demand. These data were further supplemented with taxi trip records and subway stations' entrance and exit volumes from the NYC Open Data Portal which served as a general surrogate for traffic volumes. Finally, the data were combined with demographic and geographic information since these features have known relationships to

macroscopic safety performance in the literature.

The analysis period used for this study was a period with a time overlap with all the datasets described above. It was further shortened to not include data during the period of the COVID-19 pandemic, as both shared bike usage and POIs visits changed dramatically during this period. As a result, the analysis period considered here was from January 1, 2018, to January 1, 2020, for a total of two years of data. The shared bike system is mainly distributed within the borough of Manhattan, therefore, to make the datasets compatible, all the data are collected from the Manhattan area in New York City and aggregated at the census tract level, which consists of 278 census tracts in total. In order to develop the geographically weighted model in a spatial dimension, the two-year data are summed within each census tract, which makes up 278 observations in total. The description of each dataset and the data cleaning and processing are discussed below.

3.1.1 Crash dataset

The crash dataset contains details of all police-reported crashes in NYC from January 2013 to September 2020. The police report (form MV104-AN) is a requirement for fatal or injury crashes, along with crashes with more than \$1,000 worth of property damage. A total of 1.71 million events are recorded in this dataset and each record contains the specific crash time and geolocation of it. The number of persons, pedestrians, cyclists, and motorists that were injured and killed is recorded as well, which are extracted and used in the present study.

To develop a robust and reliable aggregated crash prediction model, this study used all fatal and injury crashes that occurred within Manhattan. The fatal and injury crashes were further categorized based on mode of the travelers that experienced the fatality or injury. Three types of categories were considered based on the type of mode involved in the crash: 1) cyclist; 2) pedestrian; and 3) motorist. The statistics summary of the crash data is shown in Table 3-1.

3.1.2 Shared bike trip records

Citi Bike is the dominant shared bike service provider in New York City. A publicly accessible dataset provided by Citi Bike was used, which contains the trip duration, start and end time, date, and station names of each trip. In this study, each trip is aggregated at the census tract level based on the geolocation of its start station; however, similar results are obtained when aggregating based on the end station. To be compatible with other datasets, the data starting from January 1st, 2018, to January 1st, 2020, are extracted in this study. The statistical summary of the shared bike dataset is shown in Table 3-1.

3.1.3 POIs visit volumes

The POI footage pattern dataset was obtained from *SafeGraph*, which contains the hourly visit volumes of more than 6 million points of interest across the US (SafeGraph, 2020). The data is collected via various methods, the most common of which is from mobile phone devices. The geographic location of each POI is provided, along with the city and region name. Each POI is also categorized based on the type of location. For this study, all the POIs within the Manhattan area in New York City during the time frame of the study were extracted. There are about 150 categories for a total of 26,780 POIs within Manhattan. To simplify this dataset, these 150 categories are further grouped into 9 categories based on their similarity as shown in Table 3-1.

3.1.4 Traffic surrogate datasets

Ideally, actual traffic volumes on roadways within each census tract would be used to predict the crash frequency. Unfortunately, such a dataset was not available for all roadways in the Manhattan area. Instead, two traffic surrogate datasets were used in this study: taxi trips and the number of people entering and exiting subway stations. These data were used to reflect the influence of traffic networks on crash frequency.

The taxi trips dataset includes the trip records of yellow taxis, green taxis, and For-Hire vehicles (such as Uber, Lyft, Via, and Juno) (NYC Taxi and Limousine Commission, 2022). Each record consists of the pickup and drop-off date, time, location, trip distance, fare, and driver-reported passenger counts. Since the pickup and drop-off location of each taxi trip comes with the Taxi Zone, which is larger than the census tract, the trip volumes for the taxi zones are assigned to the census tract proportional to the area of the census tracts.

The subway data consists of entrance and exit counts to each subway station aggregated every four hours. However, the subway stations are only available in a small fraction of the census tracts, hence most of the census tracts have 0 value. Thus, the subway data is transformed to a dummy variable in this study to indicate a census tract with or without subway stations. The summary of the traffic surrogate datasets is shown in Table 3-1.

3.1.5 Demographic information

The demographic data of each census tract is collected from the open census data available through *SafeGraph*, which includes data from the American Community Survey 5-year estimates. The data is comprised of three parts: individual information (such as age, gender), household information (such as family income, age of household), and transportation-related information (such as travel time from home to work, traffic mode to work). The information is typically provided in a categorical format; for example, age is divided into 3 groups, young (under 17 years), middle (18-49), and old (50 years and older). Hence the data summarizes the number of people belonging to each group for each census tract, or the number of households that belong to different categories. The statistical summary of the demographic characteristics used in this study is shown in Table 3-1. To focus on the different impacts of demographic characteristics, the percentages of people in each demographic group instead of the absolute values of the population are used in this study, and the total population within a census tract is also used as an independent variable in the models.

3.1.6 Geographic route profile

The roadway density is also collected as a part of this dataset as it could also be a surrogate for traffic related variables and significantly impact crash frequency. The total route length is calculated from the centerline profile which is a single line representation of New York City streets containing address ranges and other information such as traffic directions, road types, segment types (Police Department (NYPD), 2022). The road type include street, highway, bridge, tunnel, boardwalk, path/trail, step street and so on. All type of road is considered in this study. The roadways are distributed to census tracts by combining the route centerline profiles with the shapefile of census tracts. The roadway density is determined by summing the route lengths within a census tract and dividing by the area of the corresponding census tract, see **Equation 3-1**.

$$roadway\ density = \frac{\sum_{all\ route} Route\ length}{Area\ of\ census\ tract} \quad 3-1$$

Besides the roadway density, the total route length, the number of intersections within each census tract are also extracted as independent variables to be considered.

Table 3-1 Statistical summary of the census tract level data used in the study

Dataset	Variable (Unit)	Average	Max.	Min.	Std.
Crash dataset	Number of pedestrian-involved crashes (crashes)	16	57	0	12
	Number of cyclist-involved crashes (crashes)	10	52	0	8
	Number of motorist-involved crashes (crashes)	31	195	0	25
	Total number of fatal and injury crashes (crashes)	57	209	0	36

Shared bike	Number of shared bike trips (trip)	99452	1103973	0	131370
POI Data	Health Care (person trips)	12043	727249	0	75712
	Manufacturing and Wholesalers (person trips)	2221	145354	0	19695
	Miscellaneous and Grocery Stores (person trips)	31329	1387663	0	210962
	Schools (person trips)	10806	396259	0	61963
	Amusement and Recreation (person trips)	36943	2114891	0	208690
	Transportation and Motor Vehicle (person trips)	13133	12821792	0	193160
	Food and Drinking (person trips)	88505	3134511	0	521280
	House Maintenance	3208	1047156	0	36096
	Public Services (person trips)	5563	117530	0	39892
Traffic surrogates	Taxi trips (trips)	2709805	14445574	134453	2277586
	Subway Dummy	0.35	1.00	0.00	0.48
	Population	11515	32079	1	6221
	Percentage of people <= 17 years old	0.14	0.32	0.00	0.07

	Percentage of people > 50 years old	0.32	0.79	0.00	0.11
	Percentage of males (person)	0.47	0.88	0.00	0.08
	Percentage of households with family income less than \$ 19,999 per year (household)	0.20	0.63	0.00	0.13
Geographic route profile	Total route length (miles)	8.20	74.18	1.25	7.21
	Roadway density (mi/mi ²)	104.62	490.96	31.74	43.65
	Number of intersections (count)	174	974	16	100

3.2 Methodology

Crash count data take the form of non-negative integers and are commonly modeled by linear regression, Poisson regression, and negative binomial regression (Abdulhafedh, 2016; Ye et al., 2018). Compared to linear regression, Poisson and negative binomial regression models perform better by directly addressing the counts nature of crashes (i.e., crash frequencies take non-negative integer values) and are thus more frequently used in the crash frequency modeling (Ye et al., 2018). However, the Poisson model assumes that the mean of the count data equals its variance. This assumption limits its performance on the crash frequency modeling since crash data is typically subject to overdispersion, i.e., the variance of the data is larger than the mean. The negative binomial distribution, on the other hand, relaxes this constraint and allows for overdispersion, and therefore achieves more flexibility and better performance (Hilbe, 2011). Hence, the negative binomial (NB) regression is adopted in this study to model the different types of crash frequencies, i.e., pedestrian crashes, bicyclist crashes, and motorist crashes, considering various covariates. The Hoerl functional

form is also adopted to improve the performance of the models.

The global NB model assumes the relationship between response variables and explanatory variables is constant across the research areas. However, when the research area is big or the land use characteristics are significantly different from place to place, this assumption might be questionable and therefore limit the performance of the global model. Therefore, a geographically weighted regional model, e.g., the GWNBR model, is developed to capture the spatial variations.

The basic theory of negative binomial regression and Hoerl functional form, and the GWNBR model are briefly described below.

3.2.1 Negative binomial model

Assuming that Y is the response variable, which is the crash count data in this study, \mathbf{x} is a vector of covariates, and $\mu(\mathbf{x})$ is a non-negative function of \mathbf{x} , the probability mass function for a negative binomial distribution can be expressed as:

$$Pr(Y = y | \mathbf{x}) = \frac{\Gamma(y + \alpha^{-1})}{y! \Gamma(\alpha^{-1})} \left(\frac{\alpha \mu(\mathbf{x})}{1 + \alpha \mu(\mathbf{x})} \right)^y \left(\frac{1}{1 + \alpha \mu(\mathbf{x})} \right)^{\alpha^{-1}} \quad 3-2$$

The basic model structure for the NB model is an exponential function, which can be expressed as in **Equation 3-3**.

$$\mu(\mathbf{x}) = \exp(\boldsymbol{\beta} \mathbf{x}) = \exp(\beta_0 + \beta_1 x_1 + \beta_2 x_2 + \dots + \beta_n x_n) \quad 3-3$$

where, x_1, x_2, \dots, x_n are the elements of covariates vector \mathbf{x} , $\beta_0, \beta_1, \dots, \beta_n$ are the elements of coefficient vector $\boldsymbol{\beta}$, n is the number of covariates.

As a result, the mean and variance of NB distribution can be expressed as follows:

$$E(\mathbf{x}) = \mu(\mathbf{x}) \quad 3-4$$

$$Var(\mathbf{x}) = \mu(\mathbf{x}) + \alpha \mu(\mathbf{x})^2 \quad 3-5$$

From **Equation 3-5**, it can be found that the relationship between the mean and variance of the NB

model depends on the dispersion parameter, α .

In the standard NB model, exposure variables are typically entered in a log form (i.e., $\log(x_1)$), while the remaining explanatory variables are incorporated into the model by an exponential function as shown in **Equation 3-3**. However, Hauer (2015) argued that the increase of one covariate may not always lead to an exponential increase of response variables (Hauer, 2015). Therefore, a more flexible function, such as the sigmoid function or Hoerl function should be used. The Hoerl function can take convex or concave shape relationships with inflection points and therefore is considered in this study (Wang et al., 2020). For a given exposure variable x_1 , the Hoerl function can be written as:

$$\mu(x) = \exp(\beta_0 + \beta_1 x_1 + \beta_{n+1} \log(x_1) + \beta_2 x_2 + \dots + \beta_n x_n) \quad 3-6$$

Equation 3-6 suggests that the Hoerl function can be implemented by adding the log form of covariates into the dataset as additional variables.

3.2.2 Geographically weighted negative binomial model

Instead of estimating one set of parameters, β , for all the spatial units, e.g., census tracts in this study, the GWNBR model assigns a set of parameters β_i for each census tract i specifically. The basic structure of the GWNBR model can be expressed as in **3-7**.

$$\mu_i(x_i) = \exp(\beta_i x_i) \quad 3-7$$

This regional model $\mu_i(x_i)$ is estimated considering that each census tract has a fixed number of neighbors around it. The number of neighbors is determined by the bandwidth, b , which can be defined based on a threshold distance or based on a number of neighbors. In this study, the bandwidth is defined as a specific distance from the centroid of each census tract, and all centroids of census tracts within this distance are selected as neighbors to this census tract. Based on this bandwidth, the geographical weights of neighbors around census tract i is defined as the adaptive bi-squared kernel as **Equation 3-8**.

$$w_{i,j} = \begin{cases} \left(1 - \left(\frac{d_{ij}}{b}\right)^2\right)^2 & \text{if } j \text{ is the neighbors of } i \\ 0 & \text{otherwise} \end{cases} \quad 3-8$$

where, d_{ij} is the distance between the centroids of census tracts i and j .

The optimal bandwidth is determined before estimating the GWBNR model by considering the b that provides the minimum corrected AIC criterion (AICc). The AICc is calculated as **Equation 3-9**.

$$AIC_C = -2L(\boldsymbol{\beta}) + 2k + \frac{2k(k+1)}{m-k-1} \quad 3-9$$

where, $L(\boldsymbol{\beta})$ is the likelihood function of the GWBNR model, k is the number of the effective number of parameters in the model, and m is the number of observations.

The model is estimated using the *SAS/IML Macro* code provided by da Silva and Rodrigues (da Silva and Rodrigues, 2014).

3.3 Results and discussions

This section provides the results of the models developed to describe cyclist, pedestrian, and motorist crash frequency. Each model uses as a dependent variable the number of people from this mode injured or killed within a given census tract. Three different types of global models and one GWBNR model are considered for each crash category considered. *Model 1* is the basic NB model that considers all relevant explanatory variables in the traditional NB form shown in **Equation 3-3**. It is specified as a starting point for model optimization and considers explanatory variables selected based on the significance level and the performance of the model. Note that POI information considered in this first model represents all POI person-visits within each census tract. *Model 2* extends the first model by using the Hoerl functional form for exposure variables to improve the model fit to the observed data, as discussed in the methodology section. *Model 3* uses the second model and extends it to consider only a subset of POI categories that can be used to improve model fit. Incorporating all POI categories into the dataset might lead to arbitrary results since some of the

categories may not directly impact crash risk, such as person-trips to manufacturing factories and household maintenance companies. Therefore, to further improve the model, only POI categories that represent daily needs are considered in Model 3. These categories include Health Care, Miscellaneous and Grocery Stores, Schools, Transportation, and Motor Vehicle, Food and Drinking, Public Services, and are combined to form the *Essential POI visits*. A GWNBR model is then developed based on the selected variable in Model 3 and the results are compared to the global model. The spatial influence of the explanatory variables on crash frequency is also analyzed based on the parameter estimations in GWNBR model.

3.3.1 Cyclists model

Coefficients for all models estimated to predict fatal and injury counts of cyclists in Manhattan, along with various metrics to assess model performance (i.e., log-likelihood; corrected Akaike information criterion, AICc; R^2 ; root-mean square error in the unit of number of fatal or injured during 2-years analysis window within each census tract, RMSE; and mean absolute error in the unit of number of fatal or injured during 2-years analysis window within each census tract, MAE), are shown in Table 3-2.

3.3.1.1 Model 1 – Best traditional NB model

The results of Model 1 suggest that the cyclist crash frequency is expected to increase with total number of person visits to POIs within a census tract, the number of shared bike trips within the census tract, the number of intersections within the census tract and the percentage of low-income families. Cyclist crash frequency is expected to decrease with the density of roadways within the census tract, and fraction of both older and younger population.

3.3.1.2 Model 2 – Hoerl function NB model

Different types of Hoerl functions are tested and found that transforming the *Shared bike trips* and

Number of intersections variables into logarithm form and adding the logarithm form of *Total POI visits* can improve the model fit. To account for 0 values in the logarithm transformation of the *Shared bike trips*, a value of 1 is added to each of these variables, which is very small compared to their general magnitude.

Comparing model fit between models 1 and 2 reveals that the Hoerl function improved the model performance. In addition, the significance level of some model coefficients increased compared to model 1. To better illustrate the benefit of using the Hoerl function, Figure 3-1 shows the relationship between POI visits and the prediction of crash frequencies for model 1 and 2 for a sample census tract. From this diagram, it can be seen that while model 1 has a monotonically increasing function between predicted crash frequency and POIs visits, model 2 has a non-monotonic prediction, i.e., the predicted crash frequency increases with POI for low volumes of POI visits (lower than 10,000) and decreases with POI for large POI volumes. This indicates that model 1 and model 2 have a similar performance for low POI values, but model 2 is more realistic when the POI values are high.

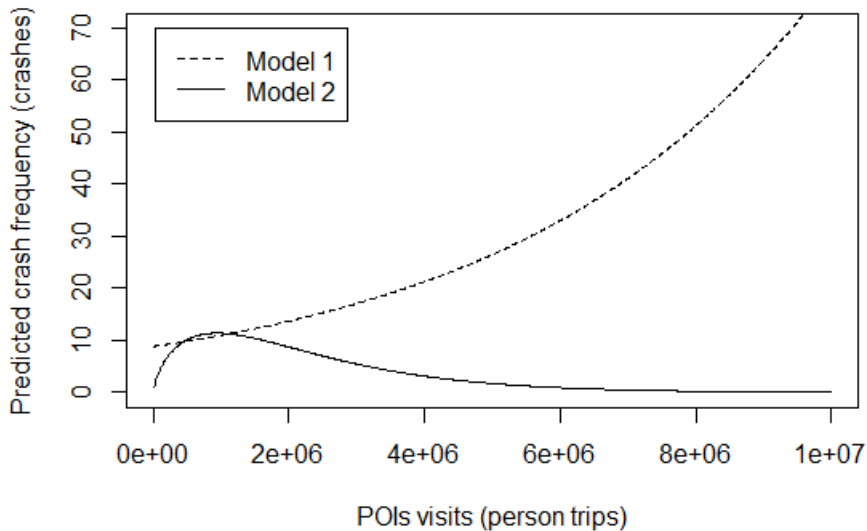


Figure 3-1 The relationships between cyclists crash prediction and POIs data in model 1 and model 2

3.3.1.3 Model 3 – Hoerl function NB model with Essential POIs

The comparison between the results of model 2 and model 3 suggests that *Essential POI visits* increased the general fit of the model by improving the performance metrics. This is consistent with the expectation that the Essential POIs are a better surrogate for general demand and hence can predict cyclists' crashes better.

3.3.1.4 Model 4 – GWNBR model

First, the bandwidth of the GWNBR model is selected based on AICc criteria. The maximum distance between the centroids of two census tracts within the Manhattan area is 21.90 km. The AICc criteria suggested that the optimal bandwidth is equal to the maximum distance, which implies that there is little spatial variation and the impact of exposure matrix to the cyclist involved crash is homogeneous within the Manhattan area. Therefore, the GWNBR model is essentially reduced to the global NB model (Model 3), and no significant improvement is observed compared to the global model.

Table 3-2 Coefficient estimations of cyclists involved crash model

Variable	Model 1		Model 2		Model 3		GWNBR Coef.		
	Coef.	P(> z)	Coef.	P(> z)	Coef.	P(> z)	Mean	Min	Max
(Intercept)	1.66	0.00	-8.74	0.00	-8.72	0.00	-8.81	-9.58	-7.74
Total POI visits	1.7e-7	0.00	0.00	0.00	-	-	-	-	-
Log (Total POI visits)	-	-	0.69	0.00	-	-	-	-	-
Essential POI visits	-	-	-	-	-4.6e-7	0.00	-4.6e-7	-5.3e-7	-3.2e-7
Log (Essential POI visits)	-	-	-	-	0.74	0.00	0.75	0.64	0.82

Shared bike trips	1.2e-6	0.00	-	-	-	-	-	-	-
Log (Shared bike trips)			0.04	0.00	0.04	0.00	0.04	0.03	0.04
Roadway density	-2.2e-3	0.03	-1.6e-3	0.06	-1.4e-3	0.10	-1.4e-3	-1.5e-3	-1.2e-3
Number of intersections	2e-03	0.00	-	-	-	-	-	-	-
Log (Number of intersections)	-	-	0.37	0.00	0.33	0.00	0.33	0.30	0.38
Male percentage	1.15	0.03	0.77	0.08	-	-	-	-	-
Subway station dummy variable	0.17	0.06	0.15	0.05	0.13	0.07	0.13	0.12	0.14
Fraction of people < 18 years old	-1.65	0.02	-1.44	0.02	-1.54	0.02	-1.45	-1.75	-0.64
Fraction of people > 49 years old	-0.81	0.04	-1.30	0.00	-1.12	0.00	-1.17	-1.44	-1.01
Percentage of households with family income less than \$20,000 per year	-	-	0.95	0.00	1.03	0.00	0.98	0.64	1.17
Log-Likelihood	-869.16		-815.64		-808.30		-805.59		
AICc	1758.32		1655.28		1638.61		1638.39		
R ²	0.32		0.52		0.55		0.56		
RMSE	9.53		5.52		5.28		5.25		
MAE	5.11		3.86		3.75		3.73		
Bandwidth	-		-		-		21.90 km		

The results shown in Table 3-2 suggest that the log-likelihood and R^2 increased significantly and AICc also decreased from model 1 to model 2 to model 3. The error matrix, RMSE, and MAE also decreases, which indicates that model 3 has better predictive power than models 1 and 2. The GWNBR model achieved comparative results with model 3. To further investigate model fit, Cumulative Residual (CURE) plots of the three global models for cyclists involved crash prediction are shown in Figure 3-2. CURE plots provide the cumulative residuals as a function of a given variable. Hauer (2015) explains that the CURE plots can be compared to the confidence intervals obtained assuming a random walk process to assess the overall model fit to the data (Hauer, 2015). Models with a good fit to the data should have no significant trend and instead oscillate around zero and generally stay within the confidence intervals. The CURE plots in Figure 3-2 demonstrate that model 3 provides the best fit of the observed data since the observed cumulative residuals have less systematic trends and more points fall within the 95% confidence intervals.

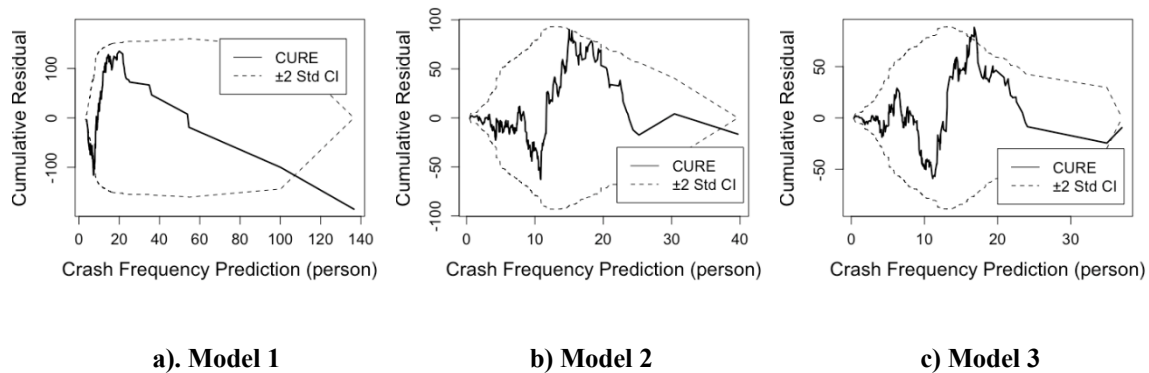


Figure 3-2 CURE plots for cyclists involved crash models

The results of all those models suggest that the number of shared bike trips is significantly positively correlated to the cyclist crash frequency. And more specifically, they are positively correlated for every census tract as suggested by the GWNBR model. This is consistent with our expectation that

the introduction of shared bikes can increase the risk of cyclists involved in crashes. This could simply be an indication of the exposure variable being larger, i.e., more bike trips imply a larger number of bicycles at risk. However, it could also be due to other reasons such as more disruptions to vehicular traffic that lead to risky maneuvers or the presence of inexperienced bicycle users on the roads. Therefore, it is critical to consider the safety impact when introducing the shared bike system to a specific area. Further considering the results of model 3, the log form of *Essential POI visits* and subway station dummy variable has positive coefficients, which indicates that these POI categories and subway stations increased the cyclists' crash risk around them. Surprisingly, the roadway density has a negative coefficient, which indicates that the areas that have dense road networks have a lower frequency of cyclist crashes. In general, roadway density is larger in the center of the city and is more likely to be equipped with bike lanes or lower speed limits and travel speeds, which could decrease the crash risk. However, the positive coefficient of the number of intersections confirms that these facilities pose a significant crash risk to bicycle users. For the age groups, the middle age group, 18-49 years old, is selected as the baseline, and the fraction of younger (<18 years old) and older (>49 years old) people have negative coefficients. This indicated that the middle-aged population is associated with the highest hazard for cyclist crashes. Areas with lower family income have a higher risk compared to wealthy areas, which is selected as a baseline in the model. The large bandwidth of the GWNBR model indicated a homogeneous relationship between impact factors and cyclist crashes. One potential reason is that shared bike users might be comprised mostly of tourists within the Manhattan area and thereby their travel pattern is independent to the land use within the city.

3.3.2 Pedestrian model

Table 3-3 provides the best models obtained for pedestrian involved crashes within the individual census tracts. From the evaluation matrix in Table 3-3 and CURE plots in Figure 3-3, Model 3 outperforms Model 1 and Model 2 for the prediction of pedestrian involved crashes. The optimal bandwidth for the GWNBR model is found to be 7.49 km, which indicated that there is spatial

variation in the pedestrian involved crashes. When the spatial variation is addressed by the GWNBR model, the predictive accuracy improved significantly increasing the R^2 from 0.50 to 0.59.

The selection of significant variables reveals that the number of shared bike trips is significantly correlated to the pedestrian crashes as well, but not as significant as to the cyclist crashes. The coefficient of $\log(\text{shared bike trips})$ is positive, while some census tracts have a negative estimation in the GWNBR model, which means that the shared bike generally poses hazards to the pedestrians. The log form of *Essential POI visits* also has a positive coefficient, which indicates that the distribution of POIs is tightly correlated with pedestrian crashes. Different from the cyclist crash model, the young group population is not significantly different from the baseline, the middle age group, while the old group population is negatively correlated with pedestrian crashes. This implies that the young and middle age group have similar risk levels of pedestrian involved crashes.

Table 3-3 Coefficient estimations of pedestrian involved crash model

Variable	Model 1		Model 2		Model 3		GWNBR Coef.		
	Coef.	P(> z)	Coef.	P(> z)	Coef.	P(> z)	Mean	Min	Max
(Intercept)	1.88	0.00	-5.19	0.00	-5.17	0.00	-5.92	-8.24	-2.05
Total POI visits	0.00	0.00	-	-	-	-	-	-	-
Log (Total POI visits)	-	-	0.49	0.00	-	-	-	-	-
Log (Essential POI visits)	-	-	-	-	0.51	0.00	0.57	0.11	0.76
Shared bike trips	7.1e-7	0.03	-	-	-	-	-	-	-
Log (Shared bike trips)	-	-	0.01	0.13	0.01	0.12	0.02	-0.12	0.03
Number of intersections	1.6e-3	0.00	1.6e-3	0.00	1.6e-3	0.00	1.3e-3	1.9e-5	4.0e-3
Population	2.5e-5	0.00	-	-	-	-	-	-	-

Log (Population)	-	-	0.11	0.00	0.10	0.00	0.08	-0.10	0.31
Fraction of people >49 years old	-0.55	0.12	-0.79	0.01	-0.69	0.02	-0.59	-4.23	1.11
Percentage of households with family income less than \$20,000 per year	-	-	0.55	0.04	0.52	0.04	0.48	-0.30	2.63
Log-Likelihood	-994.34		-962.53		-952.68		-922.88		
AICc	2002.67		1941.06		1921.36		1892.00		
R ²	0.34		0.46		0.50		0.59		
RMSE	12.77		8.55		8.26		7.63		
MAE	8.01		6.35		6.16		5.73		
Bandwidth	-		-		-		7.49 km		

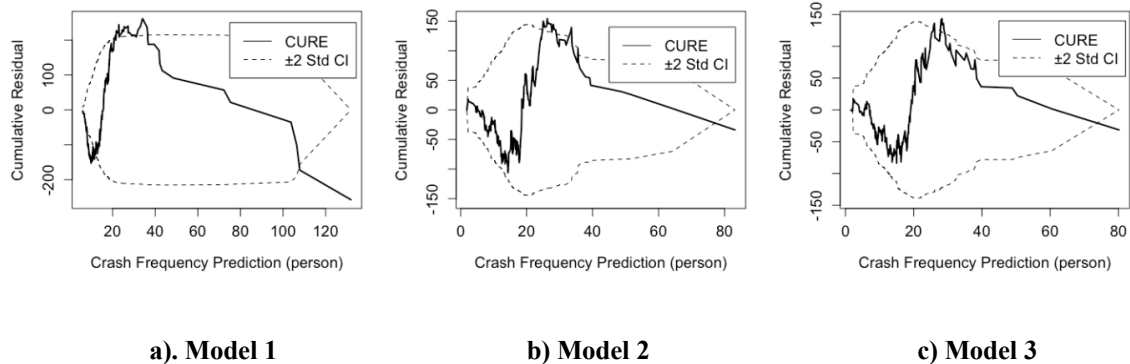


Figure 3-3 CURE plots for pedestrians involved crash models

3.3.3 Motorist model

The model estimation results for the motorist crash models are presented in Table 3-4, and the CURE plots for the three global models are shown in Figure 3-4. The model fit results suggest that Model 3 has the best predictive power and model fit. Interestingly, the motorist involved GWNBR model has

an optimal bandwidth that is very small, only 3.08 km, which indicates a significant spatial variation in the relationship between exposure matrix and crash frequency. Accounting for the spatial variation in the GWNBR model significantly improves the model accuracy, increasing the R^2 from 0.35 to 0.58. This may indicate that the regional characteristics, such as land use, traffic network profile, or transportation infrastructure characteristics, impact the motorists involved crash frequency significantly across different areas. This is consistent with expectation since motorist involved crashes are more tightly related to the route profile and the surface transportation network, rather than characteristics related to other modes.

The shared bike trip volumes are not significant in these models, which implies that the impact of shared bike is mostly on the cyclist and pedestrian involved crashes. However, the number of taxi trips, which is considered a surrogate for general traffic, is shown to be significant and positively related to motorist crashes. The motorist crash frequency is positively related to the route length and number of intersections. This indicates that the quality of the route network is critical to motorized vehicle safety. More detailed correlation between route profile, such as curvature, pavement condition, AADT, and traffic regulation, and the motorized vehicle crash frequency can be further analyzed by developing micro-level crash prediction models.

Table 3-4 Coefficient estimations of motorists involved crash model

Variable	Model 1		Model 2		Model 3		GWNBR Coef.		
	Coef.	P(> z)	Coef.	P(> z)	Coef.	P(> z)	Mean	Min	Max
(Intercept)	1.96	0.00	-3.01	0.00	-2.80	0.00	-4.06	-19.32	4.21
Total POI visits	0.00	0.44	-	-	-	-	-	-	-

Log (Total POI visits)	-	-	0.16	0.02	-	-	-	-	-
Log (Essential POI visits)	-	-	-	-	0.16	0.01	0.12	-0.32	0.61
Total route length	0.03	0.00	-	-	-	-	-	-	-
Log (Total route length)	-	-	0.45	0.00	0.48	0.00	0.48	-0.34	1.35
Number of intersections	1.7e-3	0.00	-	-	-	-	-	-	-
Log (Number of intersections)	-	-	0.24	0.00	0.22	0.01	0.19	-0.60	0.70
Total taxi trips	8.4e-8	0.01	-	-	-	-	-	-	-
Log (Total taxi trips)	-	-	0.12	0.09	0.11	0.10	0.26	-0.37	1.03
Fraction of people <18 years old	1.70	0.03	1.62	0.02	1.55	0.03	0.86	-5.89	9.36
Percentage of household with family income less than \$20,000 per year	1.59	0.00	1.54	0.00	1.50	0.00	1.00	-2.19	5.34
Log-Likelihood	-1199.45		-1181.63		-1181.23		-1108.78		
Degree of freedom	8.00		8.00		8.00		55.34		
AICc	2414.91		2379.25		2378.46		2355.63		
R ²	0.26		0.34		0.35		0.58		
RMSE	81.61		21.18		21.28		17.20		
MAE	19.86		14.26		14.25		11.48		

Bandwidth	-	-	-	3.08 km
-----------	---	---	---	---------

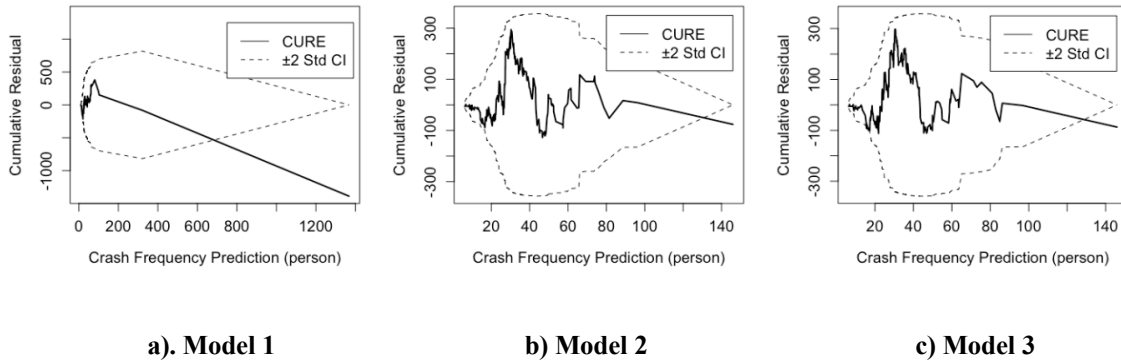
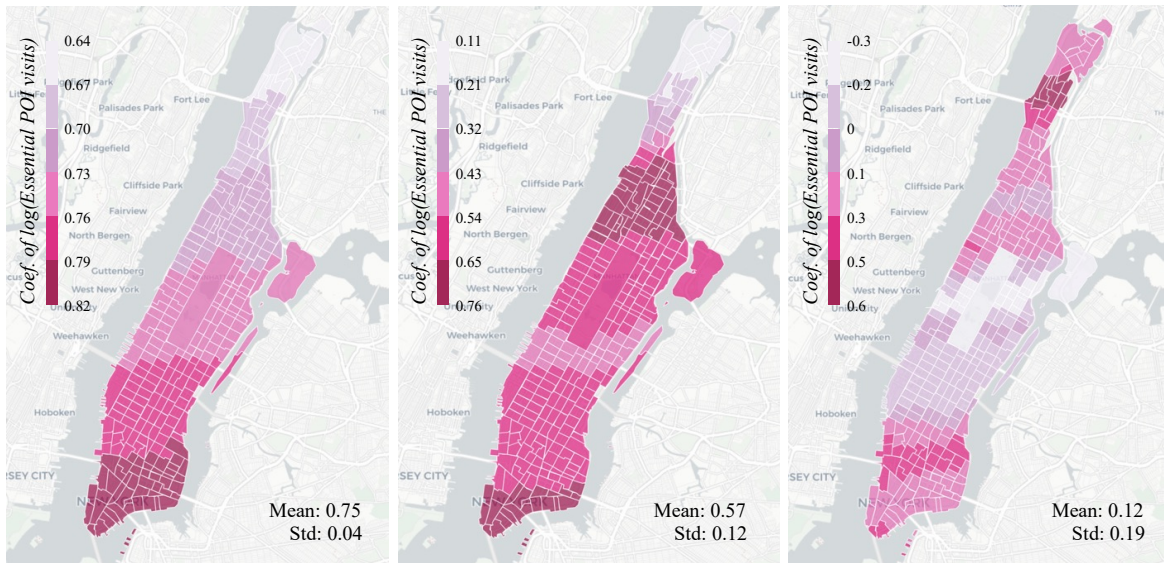


Figure 3-4 CURE plots for motorists involved crash models

3.3.4 GWNBR model comparison

The GWNBR model reveals the different relationship between exposure matrix and crash frequency. The coefficients estimated for different explanatory variables for each census tract reflect this relationship at a local level. Take the *Log (Essential POI visits)*, which is one of the mutual variables for three types of crash frequency predictive model, as an example. The coefficients of this variable, *Log (Essential POI visits)*, for each census tract from three crash type models are plotted in Figure 3-5. From these figures, it can be seen that the coefficients in the cyclists involved crash model do not vary much spatially, and only have a small range from 0.64 to 0.82. To the contrary, the motorist involved crash and pedestrian involved crash models demonstrated a more uneven distribution of the coefficient of this variable. In both, the uptown and downtown have higher coefficients than the Middletown. In the motorist involved model, the coefficients of *log (Essential POI visits)* becomes negative near Central Park, implying that the less essential POI visits there are, the likelihood of motorist crashes increases. This makes sense considering that the essential POI visits are tightly associated with people’s daily needs while Central Park serves as more of a recreation

area; thus, the essential POI visits may not be a major contributor to the motorist crashes compared to the uptown and downtown.



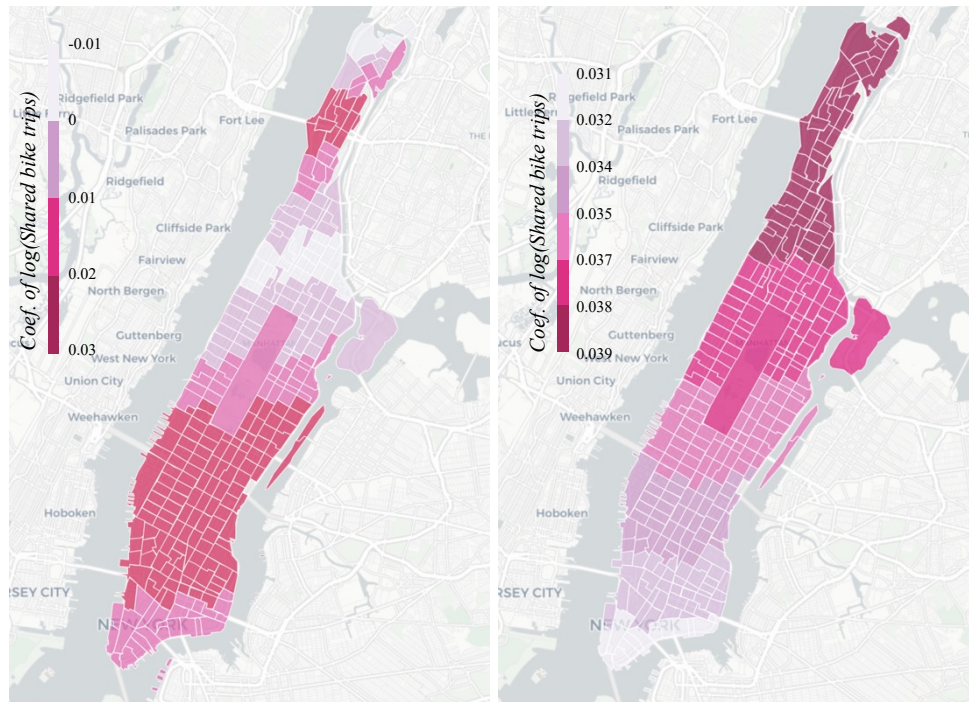
a). Cyclist crash model

b). Pedestrian crash model

c). Motorist crash model

Figure 3-5 Coefficients estimation of $\log(\text{Essential POI visits})$ from GWNBR model for three types of crash frequency model

The shared bike trips are only observed to be significant in the pedestrian and cyclist involved model, and the spatial variation of this coefficient obtained from the GWNBR model are shown in Figure 3-6 for these two models.



a). Pedestrian involved crash model

b). Cyclist involved crash model

Figure 3-6 Coefficients estimation of *Log (Shared bike trips)* from GWNBR model for two types of crash frequency model

The coefficient distributions imply that the shared bike trips have a more complicated impact on pedestrian involved crash than the cyclist involved crashes. While the cyclist involved model shows that the coefficient of the shared bike trips varies only slightly from 0.031 to 0.039, the pedestrian model has a larger range of variation and some of census tracts even have negative coefficients. These results might suggest that the usage of shared bikes in some of the northern parts of Manhattan can reduce pedestrian crash by replacing the more dangerous motorized vehicle for the daily trips. The consistency of the coefficients in the cyclist involved model suggest that the positive correlation between shared bike usage and cyclist involved crash is evidenced more uniformly across Manhattan.

3.3.5 Model discussion

The optimal NB model specifications for three types of crash frequency indicate that different types

of crashes are sensitive to different impact factors. Shared bike trips have a significant positive impact on cyclist and pedestrian crashes, but not on motorist crashes. Therefore, the shared bike system should be carefully distributed considering the potential risk, especially for areas where young people are concentrated. The associated hazard can be minimized by posing appropriate regulations or providing necessary bike facilities, such as bike lanes and more accessible helmet distributions.

The traffic network characteristics have different impacts on the crash frequency. The dense route network is usually coupled with bike facilities and low-speed limits; therefore, the crash frequency of cyclists is negatively correlated with roadway density. Intersections, on the other hand, are always associated with a high risk of all types of crashes. This is reasonable as intersections are locations where all modes (pedestrians, cyclists, motorists) must interact and also when vehicles moving in different directions cross paths.

Among the three types of crash frequency prediction models, the cyclist crash model achieved the highest accuracy, with R^2 equal to 0.55, while the pedestrian crash model and motorist crash model only have R^2 of 0.50, and 0.35, respectively. However, when accounting for the spatial variation of the explanatory variables using the GWNBR model, all three models achieved R^2 values between 0.55-0.60, with the motorist crash model showing the largest increase in R^2 followed by the pedestrian crash model. This indicates that motorist crashes and pedestrian crashes are more likely to be influenced by the spatial dependencies of the explanatory variables.

The conclusion in this study highlights the crash risk brought by the shared bike trips. As the shared micromobility brought convenience and flexibility to the society, the traffic operators should be fully aware of the safety concerns associates with them and make efficiently corresponding polices to moderate the safety risk. Fox example, as the shared bike trips increase, the dedicated bike lane should be implemented to separate the motorized traffic flow from the cyclists and more detailed traffic regulation and operation strategy should be designed. Even the motorists involved crash

frequency is not found significantly related to the shared bike trips, it accounts for the highest share of fatalities. This study indicated the impact factor of motorists involved crash is highly spatially variant and specifically traffic regulation that reduced the crash risk should be made in a regional level considering the local land use, demographic informal and traffic network characteristics. The similar conclusions can be extrapolated to other cities as well.

3.4 SUMMARY

This chapter collected several crash-related datasets to understand how various measures of travel exposure – shared bike use, person-trip information, subway flows – may impact crash frequency for bicyclists, pedestrians, and motorists. Negative binomial regression models were used to predict fatal + injury crash frequencies of three different modes. Multiple specifications were considered and GWNBR models were further developed to explore the spatial variation. Comparison of the models indicated that the Hoerl function can improve the performance of the crash prediction model, and the POIs that are related to necessary trips are more tightly related to the crash frequency of all modes compared to using all POI trips. Spatial variation in the relationship was not observed in crashes involving cyclists; however, such variation was observed in crashes involving pedestrians and motorists. The model results indicate that an increase in shared bikes can increase the crash risk of cyclists and pedestrians but does not influence motorist crash risk. Other variables that are related to crash risk include the density of the roadway network, which reduces cyclist crash risk; the number of intersections, which increases crash risk for all modes; and younger population and lower family income, which increases crash risk.

Despite the large dataset assembled for this work, not all factors that contribute to crashes were captured. Mainly, the motorist volumes were missing from the models which likely contributed to the lower predictive accuracy of the models. The model performance should be able further improved with novel datasets once available in the future.

4 SHARED BIKE AND TRAFFIC EFFICIENCY

Bikes often share the road with cars on low-volume two-way roads and cars need to move across the centerline to use the opposing lane to pass bikes. When the opposing lane is occupied, the car cannot pass the bike safely and the bike creates a moving bottleneck for cars. Delay occurring due to this moving bottleneck largely depends on the speed and flow of cars and bikes, and the specific interactions of shockwaves of queues. To quantify the impact of bikes on normal traffic flow in this situation, a numerical solution to the Lax-Hopf equation is used to develop a macroscopic car bike mixed traffic flow model. Numerical experiments to understand the impacts of bike headway and headway variation, bike speed and speed variation, traffic volumes, and headway variations on the total traffic delay are conducted. The variations of bike speed, bike headway, and opposing traffic spacing are shown to diminish the impact of the interaction of queues on total delay. Finally, dedicated bike lanes are compared to mixed bike-car lanes to determine the application domains of dedicated bike lanes considering total, i.e., bike and car, delay.

4.1 METHODOLOGY

In this section, the methodology of mixed car and bike flow modeling is presented. The methodology applies to an isolated two-way undivided roadway with one lane in each direction. Cars traveling in the same direction as the bikes (called analysis direction traffic from here on) must move across the centerline of the roadway and borrow space from the opposing lane to pass a slow-moving bike. If there is no safe gap in the opposing direction, cars cannot pass a bike and have to queue until the opposing lane is free. Some assumptions are made for the methodology to be implemented:

- 1) Car flow is homogeneous if not specifically noted.
- 2) Traffic flows in both directions, and the free-flow speed in each direction is the same. Bicycle flow only exists in the analysis direction. However, the model can be modified to address

two-directional bike flow.

- 3) Bicycles keep a constant speed and can pass each other freely without crossing into the opposing lane direction.
- 4) Cars can change their speed instantaneously without accelerating or decelerating, and traffic flow can be modeled with a triangular fundamental diagram.
- 5) The cars in the analysis direction always yield to cars from the opposing direction when passing a bike.

4.1.1 Mixed car and bikes flow modeling

First, the conditions under which cars can pass bikes are analytically derived. This derivation includes the determination of the maximum opposing flow, the duration that bikes block cars, and the maximum capacity of cars that can pass the bike. Next, the delay and queue length caused by a single bike is analytically derived assuming homogeneous traffic flow. Finally, the Lax-Hopf equations-based numerical traffic flow model is described to evaluate scenarios in which multiple bikes create several moving bottlenecks that interact with each other.

4.1.1.1 Conditions that allow for cars to pass bikes

Figure 4-1 depicts a scenario in which a bike is travelling in the analysis direction, two cars also travelling in the analysis direction, and two cars travelling in the opposing direction. Note that in this figure all cars are shown with their lengths, and also a gap in front of them to depict a safe distance between vehicles. The time in between encountering consecutive opposing direction cars for a given bike, t_t , can be determined as shown in **Equation 4-1**.

$$t_t = h_{opp} \frac{v_c}{v_b + v_c} \quad 4-1$$

where h_{opp} is the headway of cars in the opposing direction (hr), v_c is the speed of the cars (km/h),

and v_b is the speed of the bikes (km/h).

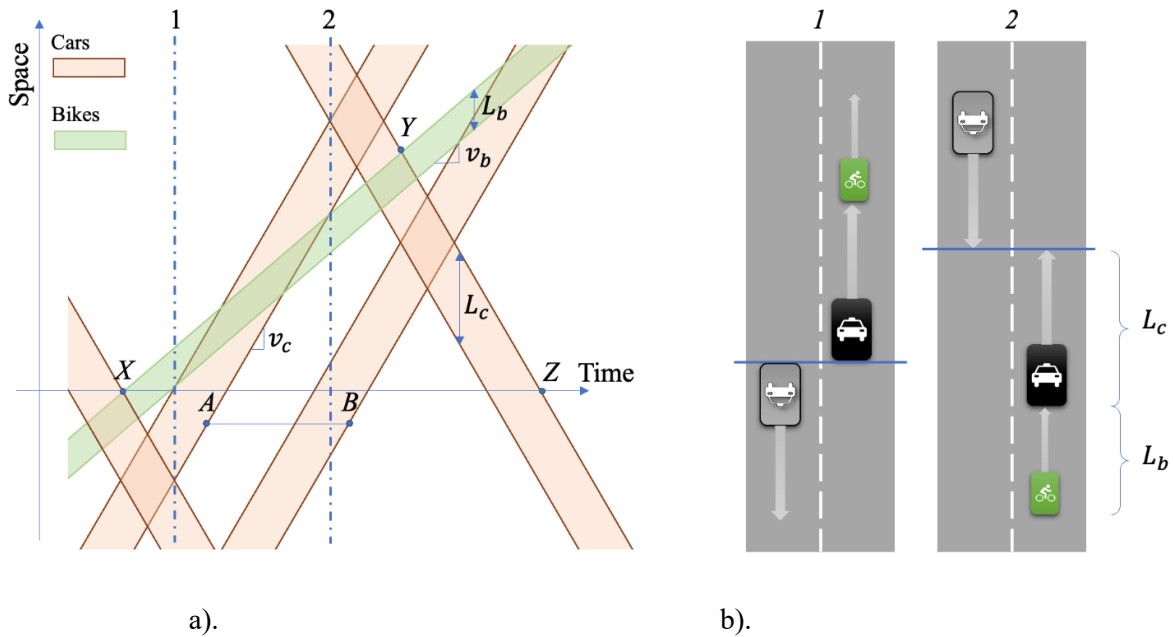


Figure 4-1 (a) Illustrative time space diagram of bike and cars in the analysis and opposing direction, (b) illustration of cars on the roadway at time instance one and two.

Cars travelling in the analysis direction will not have large enough gaps to pass the bike for a small t_t . Hence, there is a minimum t_t (or maximum flow) of cars travelling in the opposing direction to allow cars travelling in the analysis direction a large enough gap to pass the bike. To this end, Figure 4-1a depicts the case where only a single car (referred to as the target car) travelling behind a bike is able to pass the bike between two cars travelling in the opposing direction. Note that there are two key features in this figure that allow this maneuver to happen. The first key feature is that at the time shown as 1, the backs of the cars travelling in the two directions are at the same point (see Figure 4-1b). This is the earliest time point in which the target car can start using the opposing travel lane. The second key feature is that at the time point shown as 2, the fronts of the cars (including the safety margin) travelling in the two directions are at the same point. This is the latest time point in which the target car needs to be back in its own lane. Hence, the car needs to be able to complete its passing

maneuver between points 1 and 2. Therefore, the time in between the two opposing cars in Figure 4-1 is the minimum required time to allow one car to pass. Considering Figure 4-1, denote the time in between the bike encountering two opposing cars that will only allow one car to pass as t_{tmin} . This can be seen as the time in between points X and Y in Figure 4-1. This time can be calculated as:

$$t_{tmin} = 2 \left(\frac{L_b + L_c}{v_c - v_b} + \frac{L_c}{v_c} \right) \frac{v_c}{v_b + v_c} \quad 4-2$$

where L_c is the length of a car plus a safety margin, L_b is the length of a bike plus a safety margin. Then, the time during which no cars can pass the bike can be determined as the difference between t_{tmin} and the time it takes a car to pass the bike, t_p , as in Equation 4-3.

$$t_p = \frac{L_c}{v_c - v_b} \quad 4-3$$

Note that in this equation the length of the bike is not considered. This is only to aid in the implementation of the formulation in the Lax-Hopf framework. The Lax-Hopf framework assumes that the bike (i.e., moving bottleneck) does not have a length, and hence, the time it would take the car to pass the bike in the Lax-Hopf formulation is calculated rather than the actual time it takes the car to pass the bike. This simplification makes the Lax-Hopf model consistent with the analytical formulation. Finally, the duration in which a car cannot pass a bike, named the blocking time, t_b , is determined as:

$$t_b = t_{tmin} - t_p = \frac{2v_c}{v_b + v_c} \left(\frac{L_b + L_c}{v_c - v_b} + \frac{L_c}{v_c} \right) - \frac{L_c}{v_c - v_b} \quad 4-4$$

This blocking time is used as an input to the Lax-Hopf formulation to determine when the moving bottleneck has a capacity of zero.

The above equations can also be used to determine the maximum opposing flow that will allow for cars to pass the bike, Q_{opp} , and the lower bound of the maximum capacity of the roadway for cars in the presence of a bike, Q_{max} , which is achieved when the opposing flow is Q_{opp} . Considering Figure 4-1 again, it can be seen that if the headway between opposing direction cars is ε smaller than what

is shown in the figure (where $\varepsilon \rightarrow 0$), then no cars can pass the bike in the gaps between opposing cars. Hence, the maximum opposing flow, Q_{opp} , can be determined considering the headway of the opposing cars as shown in this figure (i.e., time between points X and Z), and calculated in Equation 4-5.

$$Q_{opp} = \frac{1}{t_{min}(1+\frac{v_b}{v_c})} \quad 4-5$$

Similarly, for a given Q_{opp} , the time in between two consecutive cars that can pass a given bike is as shown in Figure 4-1, see time between points A and B . This time can be used to determine a lower bound on the maximum capacity of the analysis direction in the presence of a bike, Q_{max} . This is considered a lower bound since as Q_{opp} decreases, more cars can pass the bike between each opposing car. At the extreme end where $Q_{opp} \rightarrow 0$, the upper bound on analysis direction capacity is observed, i.e., $Q_{max} \rightarrow Q_C$, where Q_C is the unrestricted capacity of the roadway. This lower bound on the maximum capacity of the analysis direction in the presence of a bike, Q_{max} , can be determined as in Equation 4-6.

$$Q_{max} = \frac{1}{2\left(\frac{L_b+L_c}{v_c-v_b} + \frac{L_c}{v_c}\right)\frac{v_c-v_b}{v_b+v_c}} = \frac{v_c(v_c+v_b)}{2(L_b v_c + 2L_c v_c - L_c v_b)} \quad 4-6$$

4.1.1.2 Traffic flow theory equations for a single bike

The impacts of a single bike can be determined utilizing traffic flow theory. Here, assume that traffic flow follows a triangular fundamental diagram with a free flow speed v_c , congested backward wave speed $-w$ ($w > 0$), maximum flow Q_C , and jam density K_J . Assume that the traffic demand in the analysis direction is Q_A , see Figure 4-2a for the corresponding fundamental diagram. It is assumed that the bike creates a moving bottleneck with a passing rate of zero when it is active for a duration of t_b as derived above. The corresponding time-space diagram of the moving bottleneck caused by the bike, and the flow/density states that arise as a result including the shockwaves are shown in Figure 4-2b.

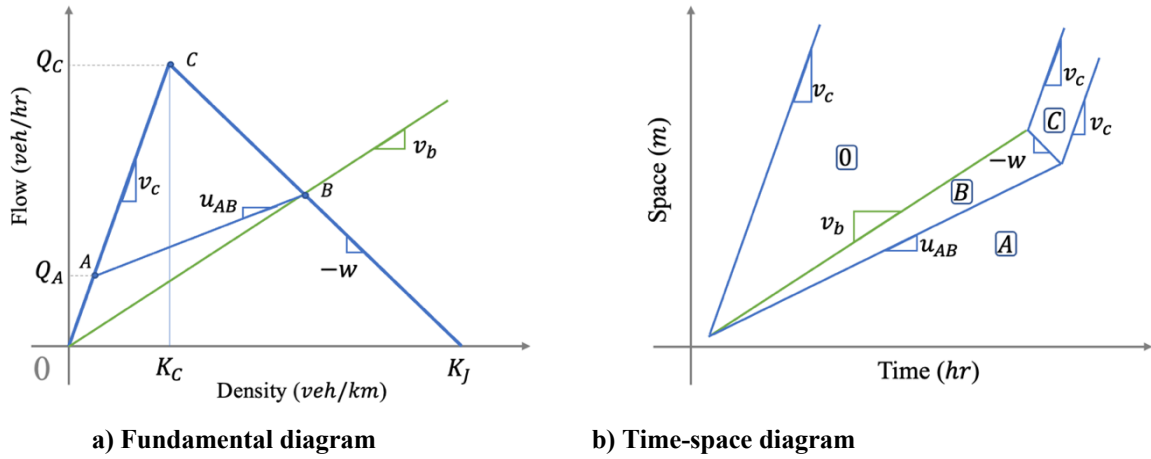


Figure 4-2 Analytical calculation of single moving bottleneck

From Figure 4-2a, the speed of the shockwave of the queue that grows behind the bike, u_{AB} , can be calculated as:

$$u_{AB} = \frac{\frac{wv_bK_J - Q_A}{v_c + w} - \frac{Q_A}{v_c}}{\frac{wK_J - Q_A}{v_b + w} - \frac{Q_A}{v_c}} = \frac{v_c(wK_Jv_b - Q_Av_b - Q_Aw)}{wK_Jv_c - Q_Av_b - Q_Aw} \quad 4-7$$

The moving bottleneck, of length t_b , then will create a short queue behind the bike. The total delay, caused by this single bottleneck can then be calculated using the time-space diagram shown in Figure 4-2b as (derivation not shown for brevity):

$$D_{total} = \frac{Q_A t_b^2 (v_c - u_{AB})(v_b + w)(v_c - v_b)}{2v_c^2(w + u_{AB})} \quad 4-8$$

Note that it is very difficult to extend the analytical approach to multiple bikes due to the potential of interaction between congested states that arise from multiple bikes. Hence, next the Lax-Hopf approach is proposed.

4.1.1.3 Lax-Hopf approach for multiple bikes

A numerical solution to the Lax-Hopf equations to solve the Lighthill-Whitham-Richards (LWR) problem has been proposed for when traffic flow follows a triangular fundamental diagram (Wu et

al., 2020; Hopf, 1969; Lax, 1957). This solution method is computationally efficient since it can be used to estimate the traffic states in the downstream of the road segment during the analysis window, T , without evaluating the entire time-space domain. It is also reliable since its solution converges to the analytical calculation when the simulation step is small enough. Further, a moving bottleneck can easily be incorporated into the Lax-Hopf framework. Therefore, the Lax-Hopf equations are adopted in this study to model the car-bike mixed-flow while treating the slow-moving bikes as moving bottlenecks.

The Lax-Hopf equation considers the traffic states that may arise, including a cumulative count of vehicles, density, and flow, at every time-space point resulting from an initial condition, upstream conditions, downstream conditions and internal conditions, see Figure 4-3. Initial conditions are the traffic states on the road segment at time zero, upstream and downstream boundary conditions are the traffic states at the upstream point and downstream point of the segment during the analysis time window, T , respectively, and internal conditions represent the moving bottleneck.

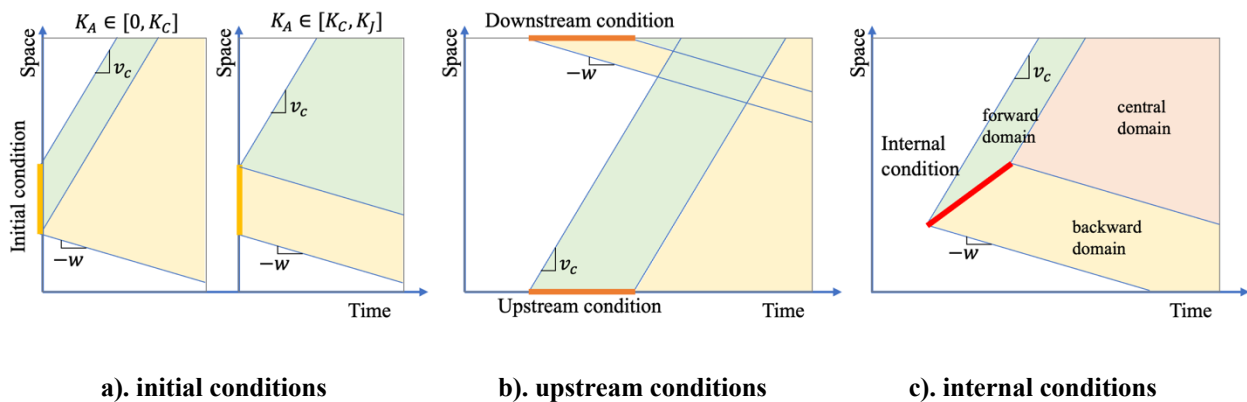


Figure 4-3 Illustration of Lax-Hopf conditions

The implementation of the Lax-Hopf equation to model the car-bike mixed traffic flow is described below and more details about the Lax-Hopf equations can be found in the literature (Wu et al., 2020).

Step 1. Initialize the time-space diagram

The Lax-Hopf numerical evaluation is based on a matrix representing the cumulative count of vehicles at every time, t , and location, x , $N(t, x)$. Here it is assumed that the time/space domain is discretized with a time step, Δt , and a space step, Δx . To initialize the N , the initial traffic density along the road segment at time 0 is assumed to be K_A , the upstream traffic flow during time T is assumed to be Q_A , The downstream is constrained by the capacity, Q_c , and $N(0,0)$ is assumed to be 0. Then, the initial conditions, upstream conditions, and downstream conditions can be derived.

Step 2. Initialized time-space location of the internal conditions

Assume that there is a homogeneous opposing flow of Q_{opp} traveling at a speed of v_c . The time in between a bike encountering two consecutive opposing direction cars, t_t , can be calculated using **Equation 4-1**. In between two opposing cars, it is assumed that there is a time period, t_b , during which a car cannot pass a bike in the presence of a car in the opposing lane as calculated in **Equation 4-4**. Hence, there is a blocking time of t_b during each t_t that bike blocks following cars from passing it. Therefore, each bike generates a series of intermittent moving bottlenecks from (T_s, X_s) to (T_e, X_e) along the analysis direction:

$$T_s = t_0 + n * t_t \quad 4-9$$

$$X_s = n * t_t * v_b \quad 4-10$$

$$T_e = t_0 + n * t_t + t_b \quad 4-11$$

$$X_e = n * t_t * v_b + t_b * v_b \quad 4-12$$

where, t_0 is the time when a bike enters the analysis segment, n is the index of each moving bottleneck generated by this bike, $n = 1, 2, 3, \dots$ until $T_s \geq T$ or $X_s \geq X$. Note that this can be generated independently of the traffic conditions in the analysis direction because the opposing flow and the bike have priority, i.e., the traffic operations in the analysis direction do not impact the opposing direction or the bike trajectory.

Step 3. Calculate the cumulative vehicle count at the beginning of the internal conditions.

The cumulative vehicle counts at each starting point of internal conditions are first estimated. The estimation of cumulative vehicle counts of interval conditions are similar to the methods for the downstream points but follow a temporal order from the left to the right on the timeline. Only the cumulative vehicle count at the starting point of an interval condition needs to be estimated since there is no passing through the bike and the cumulative vehicle count at the ending point remains the same.

For each point, (t, x) , the cumulative count of a vehicle equals the minimum considering all conditions (i.e., initial, upstream, downstream, and internal). Denote the location of the beginning of the condition i as (t_i, x_i) , the cumulative vehicle count of the beginning of condition i is N_i , the density at point (t, x) as K_A , the flow at point (t, x) as Q_A .

For initial condition, i , if $k \in [0, K_c]$, the resulted cumulative count of the vehicle at point (t, x) is

$C_{ini,i}(t, x)$

$$C_{ini,i}(t, x) = \begin{cases} N_i + K_A(x - (x_i + v_c t)) & x \in [x_i + v_c t, x_{i+1} + v_c t] \\ N_i + K_c(x - (x_i + v_c t)) & x \in [x_i - wt, x_i + v_c t] \end{cases} \quad 4-13$$

If $K_A \in [K_c, K_j]$, the resulted cumulative count of the vehicle at point (t, x) is:

$$C_{ini,i}(t, x) = \begin{cases} N_i + K_j wt - K_A(x - (x_i - wt)) & x \in [x_i - wt, x_{i+1} - wt] \\ N_{i+1} - K_c(x - (x_i + v_c t)) & x \in [x_{i+1} - wt, x_{i+1} + v_c t] \end{cases} \quad 4-14$$

For upstream condition, i , the resulted cumulative count of the vehicle at point (t, x) is $C_{up,i}(t, x)$.

$$C_{up,i} = \begin{cases} N_i + Q_A \left(t - t_i - \frac{x}{v_c} \right) & t \in \left[t_i + \frac{x}{v_c}, t_{i+1} + \frac{x}{v_c} \right] \\ N_{i+1} + Q_c \left(t - t_{i+1} - \frac{x}{v_c} \right) & t \in \left[t_{i+1} + \frac{x}{v_c}, T \right] \end{cases} \quad 4-15$$

For downstream condition, i , the resulted cumulative count of the vehicle at point (t, x) is

$C_{down,i}(t, x)$.

$$C_{down,i} = \begin{cases} N_i + K_j(X - x) + Q_A \left(t - t_i - \frac{X-x}{w} \right) & t \in \left[t_i + \frac{X-x}{w}, t_{i+1} + \frac{X-x}{w} \right] \\ N_{i+1} + K_j(X - x) + Q_C \left(t - t_{i+1} - \frac{X-x}{w} \right) & t \in \left[t_{i+1} + \frac{X-x}{v_c}, T \right] \end{cases} \quad 4-16$$

For internal conditions, i , the resulted cumulative count of the vehicle at point (t, x) is $C_{int,i}(t, x)$.

$$C_{int,i} = N_i + q_r(t - t' + t_i) + K_C v' t' \quad 4-17$$

where, q_r is the passing rate of the internal condition, which is assumed to be 0 in this study, t', v' are given by the following equations for the three domains (as shown in Figure 4-3).

a). when (t, x) within the forward domain,

$$t' = \frac{x - (x_i + v_b(t - t_i))}{v_c - v_b}, v' = 0 \quad 4-18$$

b). when (t, x) within the backward domain,

$$t' = \frac{(x_i + v_b(t - t_i)) - x}{v_b + w}, v' = v_c + w \quad 4-19$$

c). when (t, x) within the central domain,

$$t' = t - t_{i+1}, v' = v_c - \frac{(x - x_{i+1})}{t'} \quad 4-20$$

The final cumulative count of the vehicle at (t, x) is the minimum value of all conditions:

$$N(t, x) = \min \left(\min_i(C_{ini,i}), \min_i(C_{iup,i}), \min_i(C_{down,i}), \min_i(C_{int,i}) \right) \quad 4-21$$

All internal conditions are calculated by repeating **Step 3** for every moving bottleneck of every bike.

Step 4. Calculate the cumulative vehicle count at the downstream boundary.

Once the internal conditions are all calculated, the cumulative vehicle count downstream can be updated considering all conditions as described in **Step 3**.

The cumulative vehicle count downstream and upstream can be then used to determine the total delay that cars encounter on a given roadway segment within an analysis window.

4.1.2 Impact of bike lanes on traffic efficiency

A dedicated bike lane can be implemented by taking away space from cars such that the car and bikes do not interact with each other on the roadway. On a two-lane roadway this is typically done by narrowing the car travel lanes. This is expected to have two impacts: 1) the free-flow speed of cars will be reduced due to the narrowing of the lane; and 2) bikes would not be able to pass each other due to the lack of space. The car and bike delays associated with these two phenomena are discussed next.

4.1.2.1 Car delay with the bike lane

A dedicated bike lane is often implemented by taking space away from cars and dedicating it to bikes, which narrows the width of the motorized vehicle lane. The narrowed lane causes cars to travel at lower speeds. The Highway Capacity Manual provides values for the reduction in free-flow speed on two-lane roadways as a function of lane width and shoulder width (National Academies of Sciences, 2022). The adjustment assumes the standard lane width to be 3.66 m (12 ft) and the standard shoulder width to be 1.82 m (6 feet). The adjustment for lane width and shoulder width from *Highway Capacity Manual*, converted to the metric system for consistency, is shown in Table 4-1 (National Academies of Sciences, 2022).

Table 4-1 Reduction in free-flow speed (km/h)

Lane width (m)	Shoulder width (m)		
	[0.61, 1.22)	[1.22, 1.82)	[1.82, ~)
[3.05, 3.35)	5.95	3.86	1.77
[3.35, 3.66)	4.83	2.74	0.64
[3.66, ~)	4.18	2.09	0.00

To get a generalizable conclusion, different speed reduction levels are tested in this study and domains of application of bike lanes are determined.

4.1.2.2 Bike delay with the bike lane

When the bikes are not allowed to pass each other, bikes with different desired speeds can lead to delay. Figure 4-4 shows the trajectory of bicycles with dedicated bike lane in solid lines, note how some need to slow down when they catch up with other bicycles. The virtual trajectories of the same bicycles if they could pass one another is shown as a grey dotted line. Then the delay from bikes is calculated as the difference between trajectories with and without bike lane at the downstream end of the roadway.

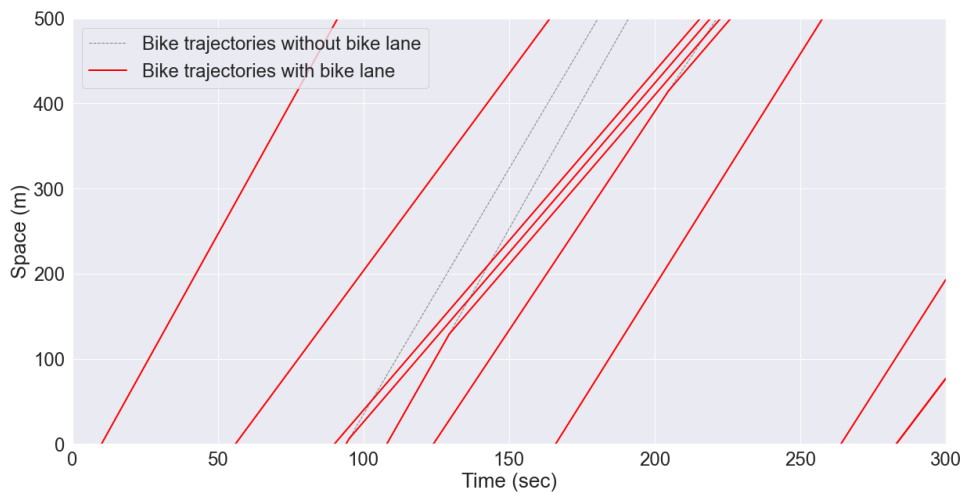


Figure 4-4 Trajectories of bikes with and without bike lanes

4.2 RESULTS AND DISCUSSION

In this section, the car-bike mixed traffic flow model is implemented for different scenarios. The model is first evaluated compared to the analytical calculation results to determine its reliability. A

series of experiments are then conducted to analyze the impact of several parameters on the total mixed traffic delay, followed by the analysis of impacts of variations in real traffic flow. Finally, based on the total delay calculation, the application domains of dedicated bike lane are derived for different scenarios.

4.2.1 The total delay in mixed traffic flow

In this section, first, the maximum opposing traffic flow is analyzed. Then, a simple case study considering the moving bottleneck created by a single bike is conducted to illustrate the accuracy of the Lax-Hopf framework. The impact of car and bike speeds and flows on traffic delay is explored based on a single bike and multiple bikes in two consecutive sub-sections.

Unless specified otherwise, the following parameters are used for the evaluation of the proposed models: car flow in the analysis direction, $Q_A = 250 \text{ veh/h}$; car flow in the opposing direction, 150 veh/h ; car free-flow speed (both directions), $v_c = 45 \text{ km/h}$; bike speed, $v_b = 20 \text{ km/h}$, roadway capacity, $Q_c = 1500 \text{ veh/h}$; and backward wave speed, $w = -v_c/4$. The minimum spacing in the traffic flow is set as the average length of a vehicle plus a clearance space of 2 secs driving distance when driving at their current speed, e.g., $L_c = d_c + 2 * v_c/3.6$, $L_b = d_b + 2 * v_b/3.6$, in units of meter, where d_c is the average length of a passenger car which is 5 meters in this study, and d_b is the average length of a bike, which is 2 meters in this study. For this set of parameters, t_b , is calculated using **Equation 4-3** as 7.6 sec.

4.2.1.1 Maximum opposing traffic flow for different car and bike speeds

The maximum opposing traffic flow is the critical opposing flow beyond which the cars in the analysis direction would not have large enough gaps to pass a bike. The results of maximum opposing traffic flow for different combinations of car speeds and bike speeds, obtained from **Equation 4-5**, are plotted in Figure 4-5. For a given bike speed, a higher car speed allows a higher opposing traffic

flow due to a shorter passing time. Similarly, as the bike speed decreases, the passing time becomes even shorter, and the maximum opposing traffic flow becomes larger. However, the maximum opposing flow does not increase linearly with car speed. Generally, the opposing traffic flow needs to be lower than 300 *veh/h* (assuming homogenous flow) to allow a car-bike mixed traffic flow. Otherwise, cars will experience significant delays since there will not be long enough gaps to pass a slow-moving bike. Scenarios as such would warrant the implementation of a dedicated bike lane to reduce traffic delays. Based on these results, the baseline opposing traffic flow for all following scenarios is set as 150 *veh/h*, which is lower than the maximum opposing flow.

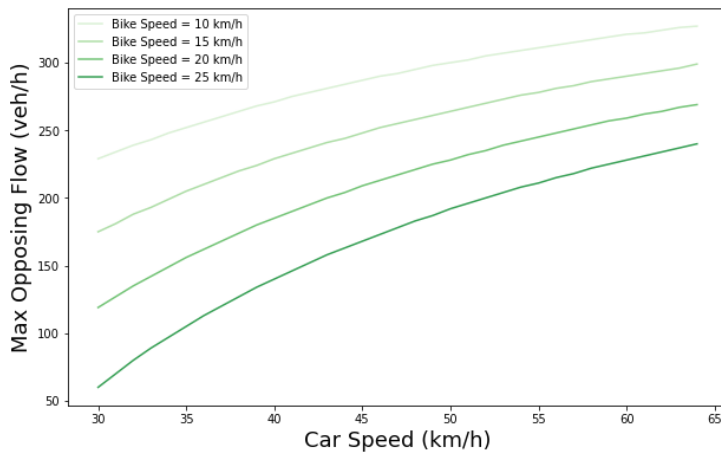


Figure 4-5 Maximum opposing traffic flow for different car speeds

4.2.1.2 Evaluation of the Lax-Hopf equations

To compare the analytical equations to the Lax-Hopf framework, the total delay of cars resulting from the moving bottleneck of a single bike is calculated using both approaches. A time-space domain of 30 *sec* by 100 *meters* with a moving bottleneck starting from $T_1 = 10$ *s*, $X_1 = 20$ *m* is considered.

Figure 4-6a shows the analytical shockwaves predicted for a single bike in between two cars travelling in the opposing direction, assuming there is enough time to clear the queue behind the bike before the next opposing direction vehicle is encountered. Figure 4-6b shows the same scenario predicted by the Lax-Hopf framework, which displays the traffic density of each time-space point. The figures visually appear to align the Lax-Hopf results with the analytical results very well.

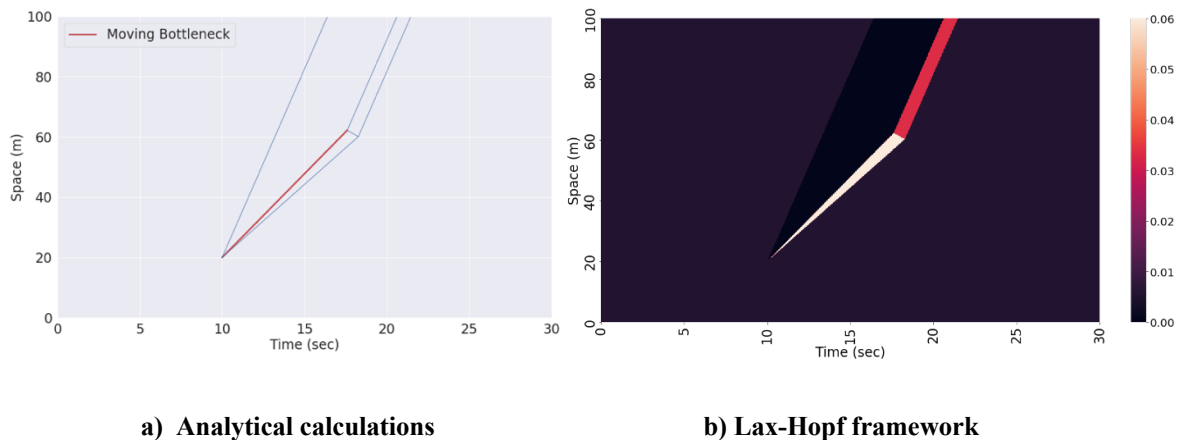


Figure 4-6 Comparison of shockwaves (and traffic density) for a moving bottleneck

For the same case study, Table 4-2 the results of the delay obtained from the analytical approach and the Lax-Hopf framework. Different time steps and space steps of the numerical evaluation are considered since this impacts the accuracy of the Lax-Hopf model. A small step will result in a high accuracy but could be computationally expensive. The results suggest that when the evaluation time step is less than 0.05 *sec* and the space step is less than 0.05 *meter*, the errors between the Lax-Hopf approach and the analytical results are less than 1%, and the corresponding computational time is also acceptable. Therefore, an evaluation time step of 0.05 *sec* and space step of 0.05 *meter* are selected for the remaining analyses.

Table 4-2 Comparison of Lax-Hopf approach and analytical equations for different time steps.

Time step (<i>sec</i>)	Space step (<i>meter</i>)	Total Delay from Lax-Hopf Equation (<i>veh * sec</i>)	Total Delay from Analytical Calculation (<i>veh * sec</i>)	Error ratio	Simulation time (<i>sec</i>)
1	1	1.1077	1.1542	4.03%	0.0107
0.5	0.5	1.2076	1.1542	4.62%	0.0110
0.1	0.1	1.1358	1.1542	1.61%	0.0189
0.05	0.05	1.1451	1.1542	0.81%	0.0349
0.01	0.01	1.1524	1.1542	0.19%	0.3733
0.005	0.005	1.1533	1.1542	0.08%	1.2875

4.2.1.3 Impact of different car and bike speeds

Next, the impact of free-flow speed of cars and bike speeds are explored to analyze their impact on the total delay. A time-space domain of 300 *sec* by 1000 *meters* is adopted here. The bike speed is varied from 15 *km/h* to 25 *km/h* and the car free-flow speed is varied from 30 *km/h* to 60 *km/h*. Note that here it is assumed that all bikes travel at the same speed, and within the analysis window each bike generates several moving bottlenecks with a passing rate of zero.

First, only a single bike is considered to avoid spillover of queues from one bike to another. The total delay predicted from the analytical calculations, as shown in **Equation 4-8**, is compared to the outputs of the Lax-Hopf numerical evaluations, see Figure 4-7. For the analytical approach, it is assumed that each blocking time of a bike creates an independent queue, and hence the delays from each blocking instance are summed to get the overall delay. However, this approach ignores the interactions that may arise between different queues.

Figure 4-7 shows that the analytical calculations and the Lax-Hopf model result in mostly consistent delay values. For a specific car speed, as the bike speed increases generally the total delay reduces

since the speed difference between the car and the bike reduces. However, when the car speeds are low and the bike speeds are high (i.e., the difference between the bike and car speeds become too small), the total delay significantly increases. This is observed since the blocking time of cars (which depends on the inverse of the difference between the car and bike speed, see **Equation 4-4**) becomes too large. Hence, there is not enough time between opposing cars to allow any cars to pass the bike. In this scenario, the car needs to queue behind the bike for the entire analysis window leading to large delays. On the other hand, for a specific bike speed, as the car speed increases, the total delay is expected to increase as well. However, the impact of the car speed is not as significant as the bike speed.

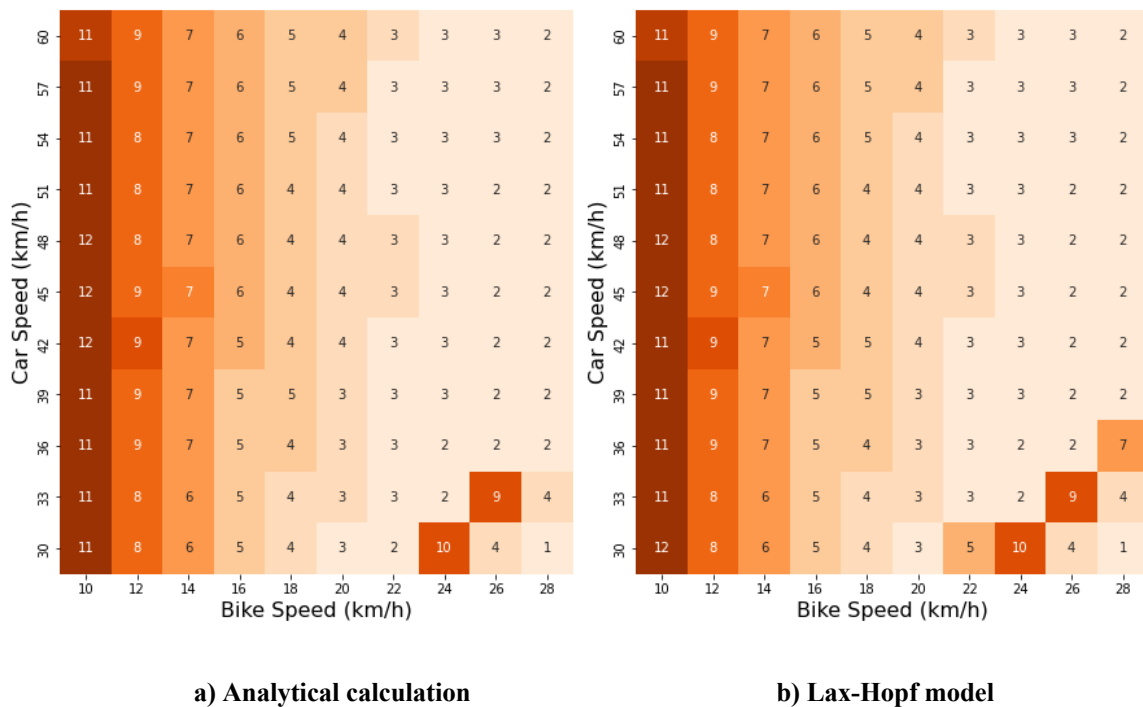


Figure 4-7 Total delay ($veh * sec$) caused by a single bike

Next, a relatively high bike flow of 100 $bike/h$ is selected to explore how queue spillovers might change the results. This corresponds to about 8 bikes in a 300 $secs$ analysis window. The total delay

from the analytical calculation and the Lax-Hopf framework is shown in Figure 4-8.

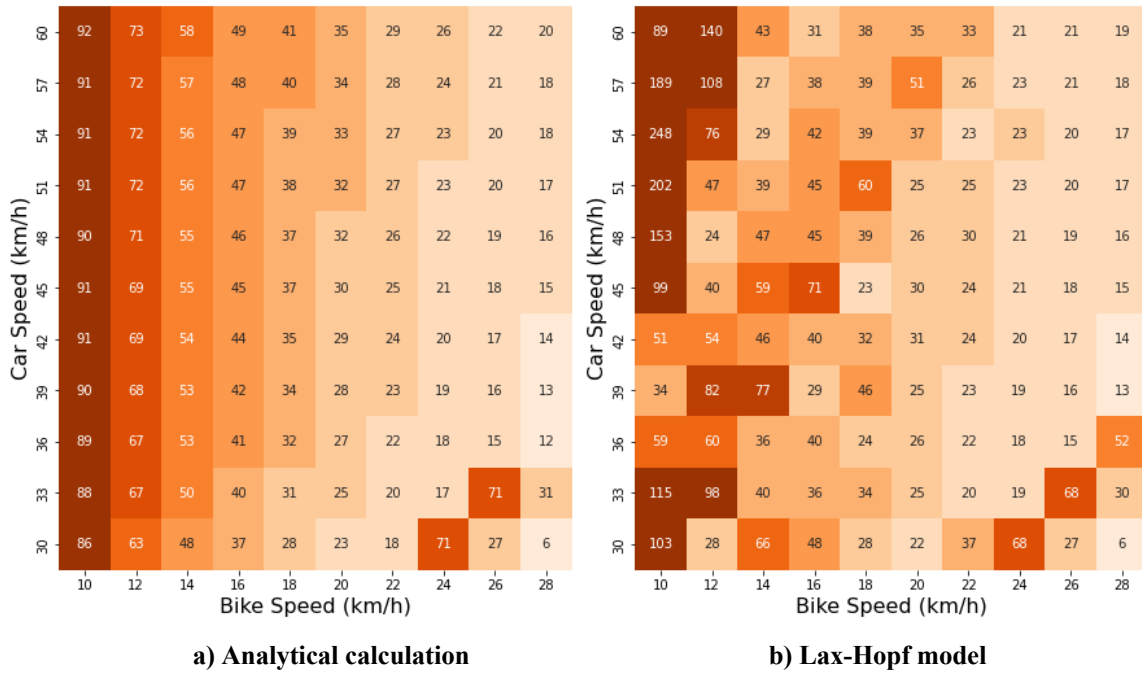
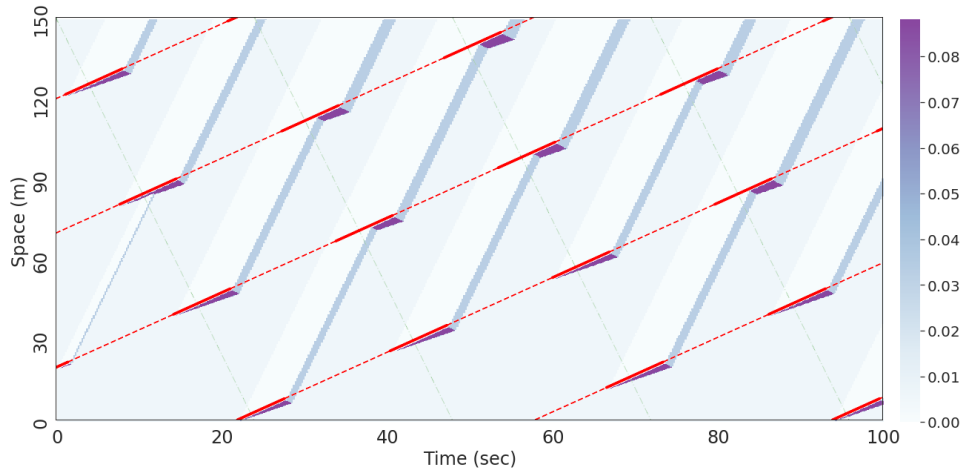


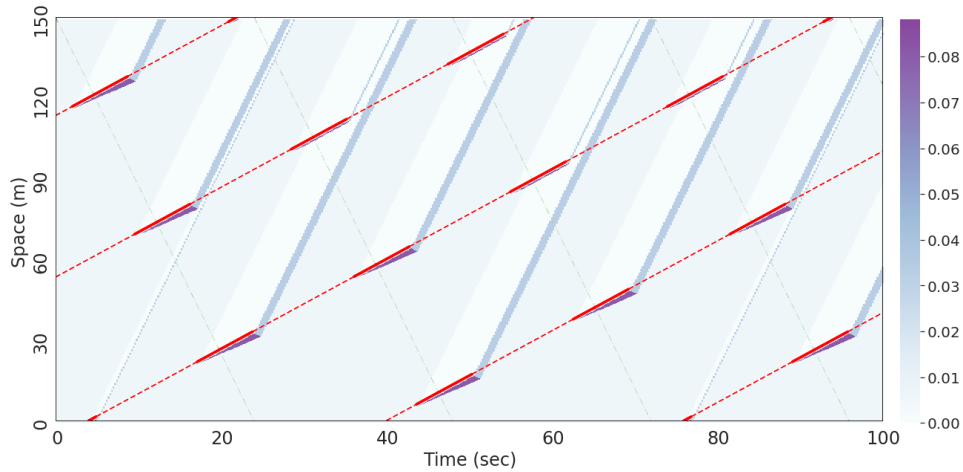
Figure 4-8 Total delay ($veh * sec$) caused by a bike flow of 100 bike/h

From Figure 4-8, the analytical calculation results demonstrate a similar pattern to what was observed for a single bike in Figure 4-7 since the total delay is only a summation of the expected delay from each individual bottleneck. However, the Lax-Hopf results display a more complicated pattern when the car and bike speeds vary. The interaction among different bottlenecks can either increase or decrease the total delay.

To understand this better, the time-space diagrams for when the car speed is 45 km/h , and the bike speed is 10 km/h and 12 km/h , are shown in Figure 4-9.



a). Time-space diagram of traffic density when the bike speed is 10 km/h



b). Time-space diagram of traffic density when the bike speed is 12 km/h

Figure 4-9 Different interaction patterns between bottlenecks

From Figure 4-9, it can be seen that when the bike speed is 10 km/h (Figure 4-9a), the traffic discharged from the upstream bottleneck will queue behind another bike downstream again and therefore has a higher total delay (98.69 veh * sec) than the analytical delay (90.72 veh * sec). On the other hand, when the bike speed increases to 12 km/h, the interaction among different bottlenecks possesses another pattern that reduces the total delay – the bottlenecks in the upstream block most of the traffic from arriving at the downstream bottlenecks, therefore, the downstream bottlenecks

become inactive and the total delay ($40.34 \text{ veh} * \text{sec}$) is lower compared to the analytical results ($69.49 \text{ veh} * \text{sec}$). Hence, it can be seen that the analytical equations are unreliable when multiple bikes are considered.

It is important to note that these results assume completely regular headways of bikes and cars traveling in the opposing direction. Hence, the impacts of variability in bike and car headways are explored in the next section.

4.2.1.4 Impact of different car and bike flows

The bike flow and car flow have a significant impact on the total traffic delay. Generally, both higher bike flow and car flow are expected to result in higher traffic delays. However, the interaction between different bikes can make the situation complex. To understand this, the bike flow is varied from 50 bike/h to 150 bike/h , and the car flow is varied from 100 veh/h to 600 veh/h . The total delay from the Lax-Hopf evaluation with different bike flow and car flow for different bike speeds is plotted in Figure 4-10. Note that the analytical equations are no longer used since previously it was found that analytical equations were not reliable when considering multiple bikes simultaneously.

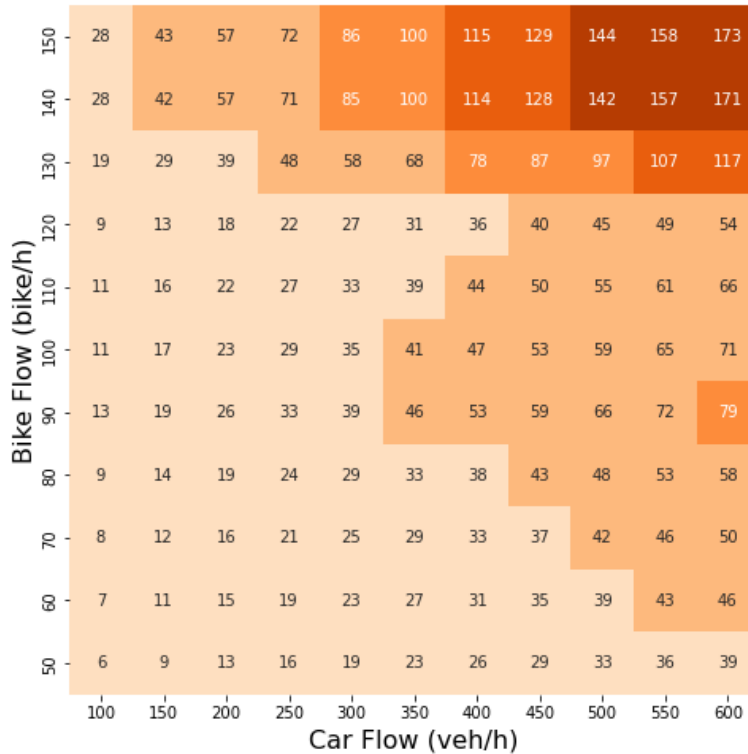


Figure 4-10 Total delay from the Lax-Hopf framework with different bike flows and car flows ($veh * sec$)

From Figure 4-10, it can be seen that as car flow increases, the total delay will increase as expected. As the bike flow increases, on the other hand, the total delay has an irregular pattern but generally increases as well. This irregular pattern is due to the different interactions among bottlenecks as shown in Figure 4-9.

From the above analysis it can be determined that the specific interaction pattern of bikes and opposing cars impacts the total delay significantly. In reality, variations in traffic flow might break those patterns and change the traffic delay. In the next section, the impact of variation in different parameters on delay is explored.

4.2.2 Impact of variations in mixed traffic flow on delay

In the previous section, traffic flow is assumed to be homogenous, and all cars and bikes are uniformly distributed on the roadway. Some variations that break the homogeneity, such as bike speed variation, bike headway variation, and opposing traffic spacing variation, may impact the interaction of the shockwaves caused by different bikes. In this section, the impacts of non-homogeneous bike headway, varying bike speed, and non-homogeneous opposing car spacing on delay are analyzed.

4.2.2.1 Sensitivity to the bike headway variation

Here it is assumed that the bike headway follows a normal distribution, with a mean equal to 36 secs (which corresponds to a bike flow of 100 *bike/h*), and the standard deviation is varied from 0 to 10 sec. For each headway variation and bike speed combination, 100 headway samples are generated and the corresponding average total delay for these 100 bikes is calculated as shown in Figure 4-11.

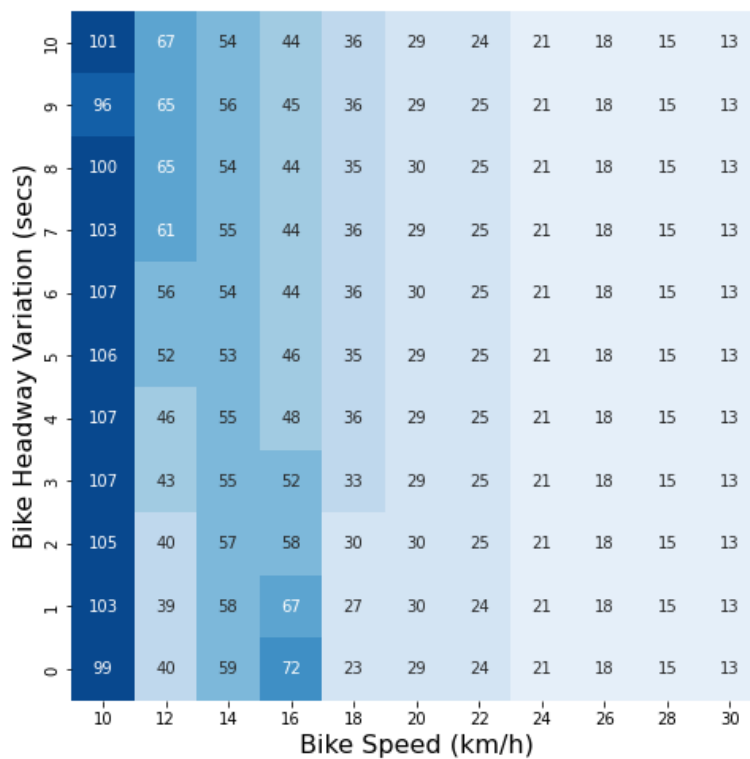


Figure 4-11 Total delay of mixed traffic flow with different bike headway variation (*veh * sec*)

Figure 4-11 suggests that the bike headway variation has the most relative impact on delay when the bike speeds are low. This is especially true since the lower bike speeds result in interactions of shockwaves that are highly sensitive to the arrival time of the bikes. However, interestingly bike headway variation does not always have the same impact on delays – for some bike speeds large bike headway variations reduce the total delay while for other bike speeds the opposite is true. Further, for large bike headway variations, the total delay values converge to the analytical delay as shown in Figure 4-8a (when the car speed is 45 km/h). This suggests that the headway variation could diminish the impact of the interaction of different shockwaves when determining the total delay. Hence, the analytical calculations can be used to approximate the delays in real-world scenarios.

4.2.2.2 Sensitivity to the bike speed variation

The bike speed is one of the factors that determine the required passing maneuver time for the following cars. Previous results revealed that the bike speed can significantly influence the total delay. Hence, here scenarios in which bikes have different speeds, i.e., bike speed variation, are considered. It is assumed that the bike speed is normally distributed, with a mean varied from 10 to 30 *km/h*, and the standard deviation is varied from 0 *km/h* to 5 *km/h*. The experiment is repeated 100 times and the average total delay for each bike speed and bike speed variation combination is calculated and demonstrated in Figure 4-12.

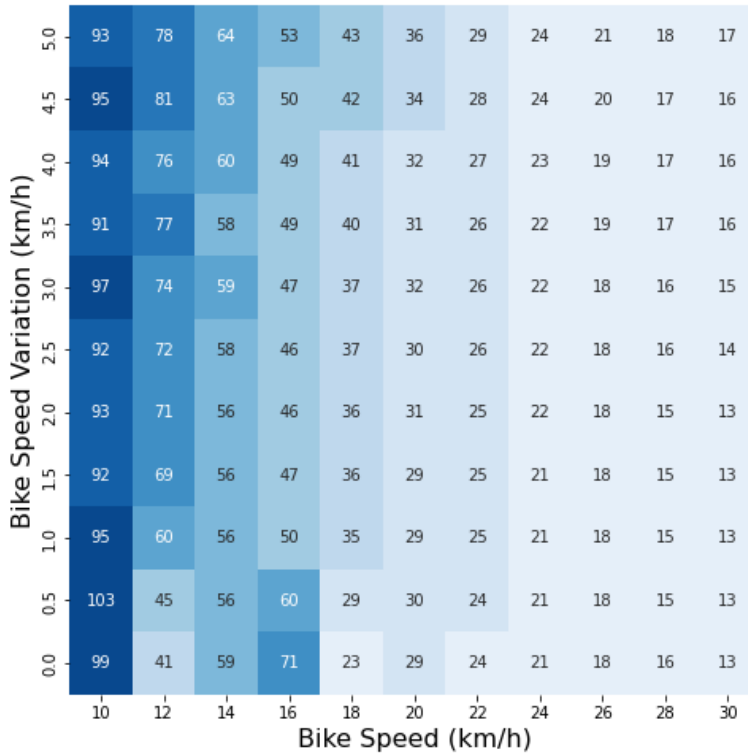


Figure 4-12 Total delay of mixed traffic flow with different bike speed variation (*veh * sec*)

The results suggest that generally increasing bike speed variation leads to larger delays, except for a few specific bike speeds. This is likely due to the bike speed variation changing the shockwave interactions. However, variations in bike speeds not only change the relative position of each bottleneck but also change the duration of the blocking time. Therefore, when the bike speed variation is high, the total delay is slightly higher than the analytical calculation of total delay (which assumes that each bottleneck behind the bike can be treated independently).

4.2.2.3 Sensitivity to the opposing car spacing variation

The distribution of cars in the opposing lane dictates when cars traveling behind a bike can pass the bike. Hence, variation in the spacing of the opposing cars will influence the blocking times and the interactions among the shockwaves of queues created behind the bike. Here, it is assumed that the

opposing car spacing is normally distributed with an average of 300 m (which corresponds to an opposing traffic flow of 150 veh/h) and standard deviation ranging from 0 m to 100 m. The experiment is repeated 100 times for different combinations of bike speed and opposing car spacing variation, and the average total delay for each combination is calculated and demonstrated in Figure 4-13.

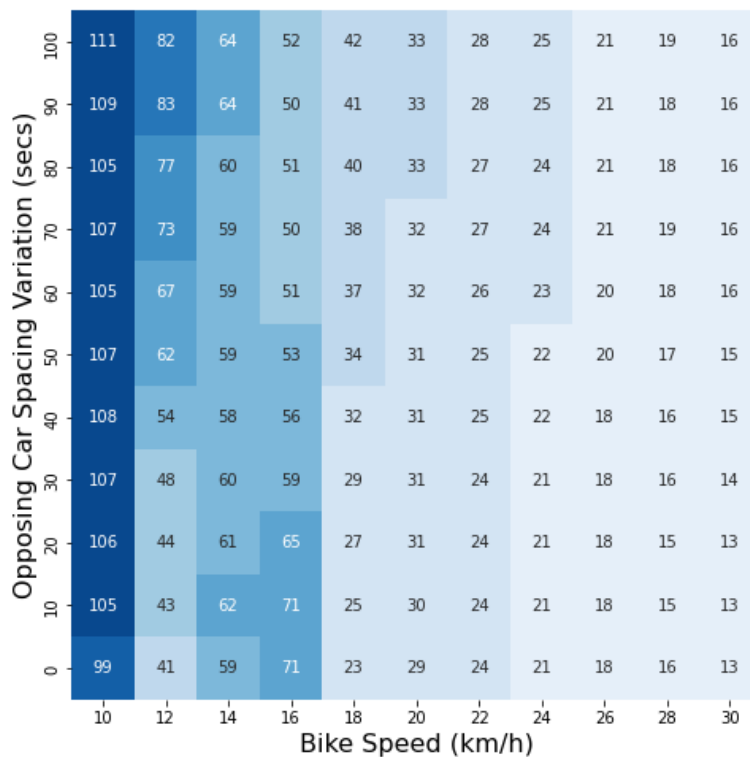


Figure 4-13 Total delay of mixed traffic flow with different opposing car spacing variation (veh * sec)

The results suggest that in general, increasing variation in opposing car spacing leads to larger delays. This is likely because uneven headways in the opposing direction can lead to a bike encountering multiple opposing cars consecutively which can lead to a long blocking time for cars following the bike. This long blocking time can then lead to larger delays for the cars traveling behind the bike. Therefore, when the opposing car spacing variation is high (100 m), the total traffic delays are higher

than the analytical calculations of total delay as shown in Figure 4-8a (when car speed is 45 *km/h*)

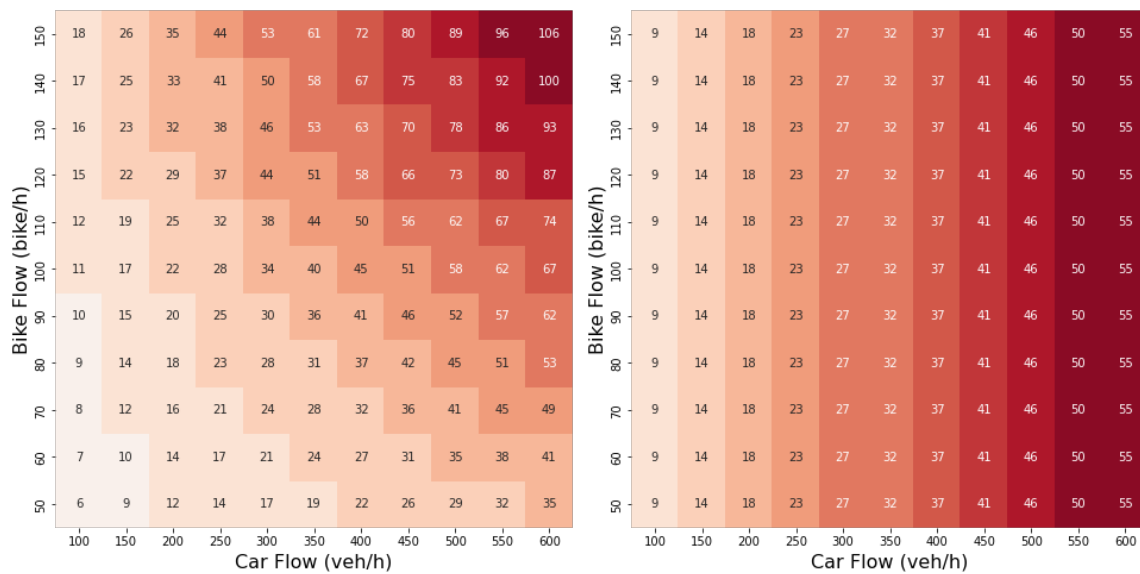
4.2.3 Domains of application of dedicated bike lanes

One option to improve the efficiency and safety of mixed traffic is to dedicate space for bikes on the roadways. The bike lane can be set either by a solid white line on the pavement or a strip garden that completely separates it from the main traffic. By separating the car flow and bike flow, the interaction between them is minimized and cars have a safer space to pass the slow-moving bikes without interacting with the opposing traffic. However, separating a bike lane from the existing roadway usually leads to a narrow roadway for cars leading to lower free-flow speeds (National Academies of Sciences, 2022). Further, bikes usually do not have enough space to pass each other, so the bicycle passing maneuver is assumed to not be allowed on the dedicated lane

Here, two components of delay are considered when dedicated bike lanes are established: delay due to reduced car speed and delay due to bike queues. Contrary to the car delay used in previous sections, person delay is used as the analysis metric in this section. A converge ratio δ between car delay and person delay is used to transfer the car delay to the person delay making them comparable. This converge ratio is determined by the average car passenger occupancy and time value ratio between motorized car drivers and cyclists. In this study, the time value of drivers and cyclists is assumed to be same. Therefore, the total delay is calculated by the sum of the bike delay and car delay multiplied by the average car passenger occupancy, which is 1.59 person/vehicle (Center for Sustainable Systems, University of Michigan, 2021).

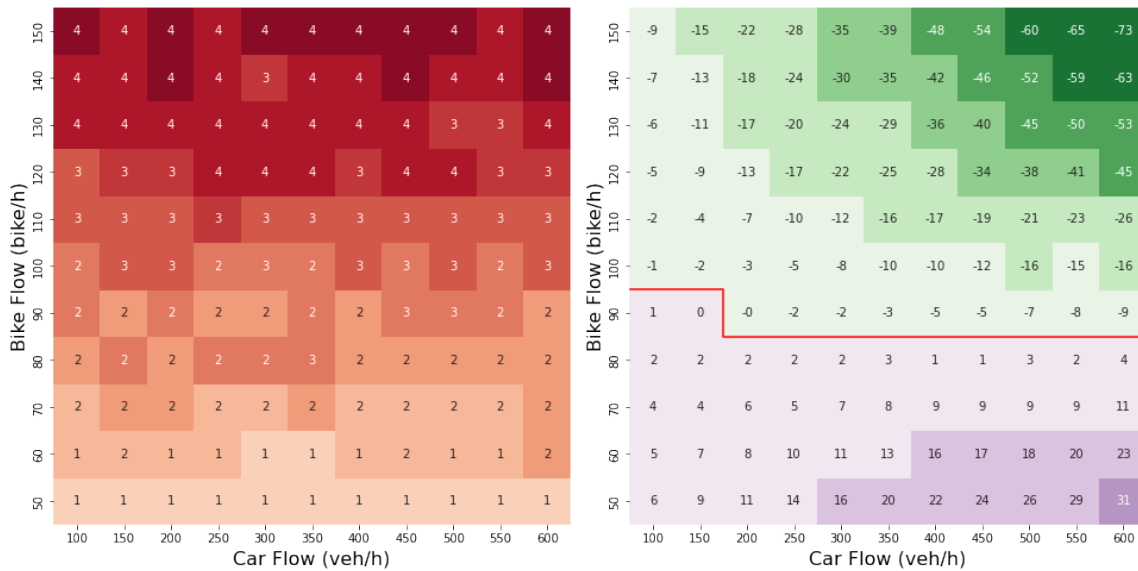
The same parameters as before are used for the analysis: car flow in the analysis direction, $Q_A = 250 \text{ veh/h}$; car flow in the opposing direction, 150 *veh/h*; and car free-flow speed (both directions), $v_c = 45 \text{ km/h}$. The bikes speed is assumed to have an average of 20 *km/h* and standard variation of 3 *km/h*. The arrival time of bikes on the upstream of the segment are uniformly distributed across the time window. The car speed is assumed to reduce by 1.2 *km/h* when a dedicated bike is added.

Since the bikes speed and time to enter this segment are randomly distributed, the average delay of 100 times repetition for each case is used as the results. The total person delay from mixed traffic flow without bike lane (where the person delay is the same as the car delay since bikes do not experience any additional delays), the car and bike delays when a lane is dedicated for bikes, along with the difference in the person delay between mixed-flow and dedicated bike lane scenarios are shown in Figure 4-14.



a). Total delay of mixed traffic flow without bike lane

b). Car delay with bike lane



c). Bike delay with bike lane **d). Changes of total delay when adding a bike lane**

Figure 4-14 Delays of the basic experiment in unit of *person * sec*

Figure 4-14a shows the total person delay when bikes randomly arrive at the upstream of the segment with exponentially distributed headways at speeds drawn from a normal distribution $N(20,3)$. Note, this is different from Figure 4-10 which assumes that bike flow is homogeneous. The randomness in the bike speed changes the results such that the total delay increases both with bike flow and car flow. Figure 4-14b shows the car delay due to the lower free flow speed when dedicating space for bikes, and as expected is independent of the bike flow. Figure 4-14c shows the bike delay when bikes cannot pass each other within the narrow dedicated bike lane, and is independent of the car flow (note that the slight variations are due to the randomness). The change in total person delay after adding a dedicated bike lane is shown in Figure 4-14d, where negative numbers represent a delay reduction with bike lane and positive numbers represent higher delay with bike lane. The solid line shows the border above which setting a dedicated bike lane can reduce the total delay. The results indicate as the bike or car flow increase, bike lanes can be useful to reduce overall person delay. Note that both the bike and car flow are important in determining the domains of application.

4.2.3.1 Impact of speed reductions on the domains of application of bike lanes

Figure 4-15 shows the changes in domains of application if the implementation of dedicated bike lane results in different car free-flow speed reductions.

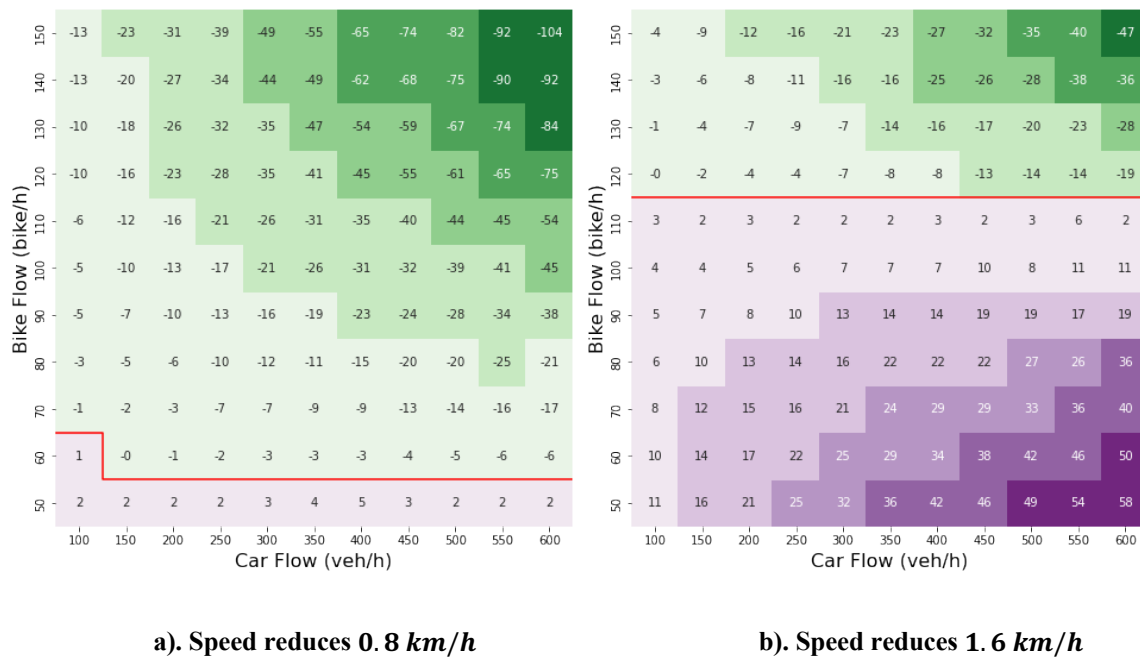


Figure 4-15 Change in total delay due to the presence of a dedicated bike lane for different car speeds reductions (*person * sec*)

The domains of application of bike lane in Figure 4-15 suggests that if dedicating a lane for bikes reduces the car speed more, the domains of application of a bike lane become smaller, i.e., bike lanes can reduce total person delay only for higher bike flows or car flows. This is expected since in these scenarios the increase to car delay from dedicating a lane for bikes is larger than the car delay when cars have to follow bikes and only occasionally encounter them. For example, for a roadway with a car flow of around 250 *veh/h*, a dedicated bike lane would reduce the overall person delay as long as there are more than 50 *bike/h* if the bike lane only reduces car speeds by 0.8 *km/h*. This threshold is increases to about 110 *bike/h* if the bike lane reduces car speeds by 1.6 *km/h*.

4.2.3.2 Impact of traffic flow on the domains of application of bike lanes

The opposing traffic flow is a critical criterion in deciding to dedicate a lane for bikes, since cars rely on the gaps in the opposing flow to pass the bike. Here, two different opposing traffic flows which are higher and lower than the base experiment, respectively, are tested and the domains of application of bike lanes are shown in Figure 4-16.

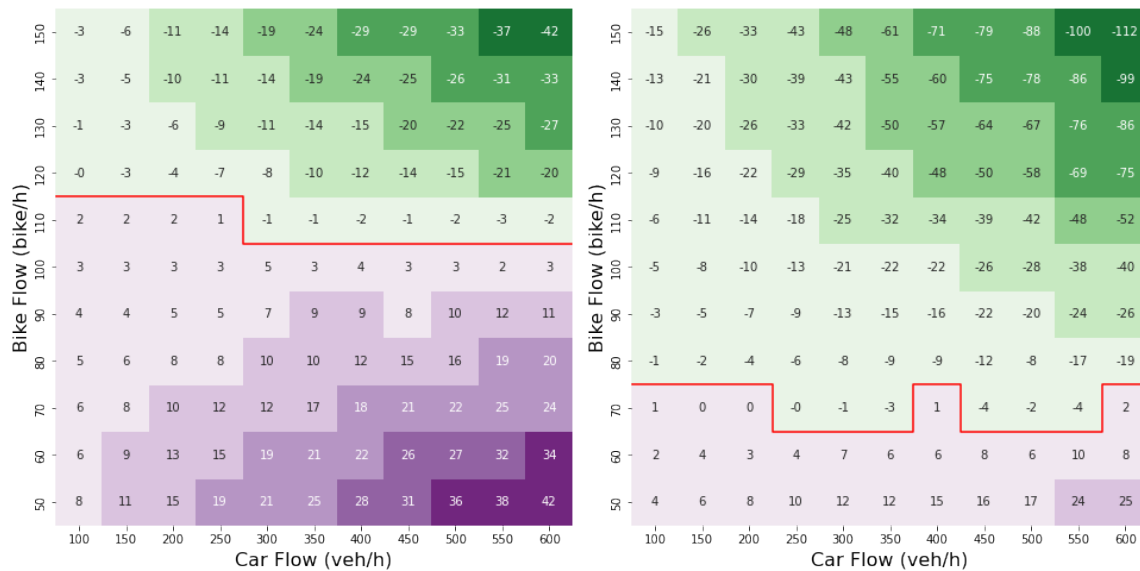
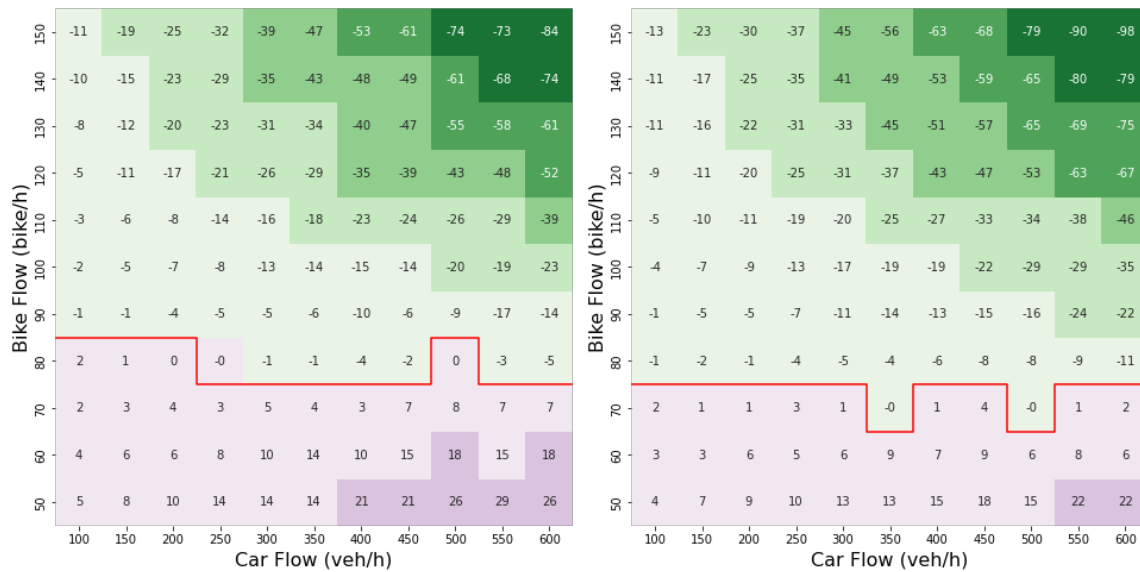


Figure 4-16 Change in total delay due to the presence of a dedicated bike lane for different opposing traffic flow (person * sec)

Figure 4-16 suggests that the domains of application of bike lanes become narrower as the opposing flow becomes lower (120 veh/h). This is because as the opposing flow decreases, cars have more chances to pass the bikes using the opposing traffic lane so that there is higher tolerance of bikes in the mixed traffic flow. Correspondingly, the domains of application increase for a larger opposing traffic flow (180 veh/h). Note, the benefit of bike lanes is more pronounced when the traffic flows in two directions are not the same and the analysis direction has higher traffic flow.

4.2.3.3 Impact of opposing traffic spacing variation on the domains of application of bike lanes



b). Opposing spacing variation 50 veh/h **c). Opposing spacing variation 100 veh/h**

Figure 4-17 Change in total delay due to the presence of a dedicated bike lane for different opposing spacing variation (person * sec)

Variation in the spacing of the opposing traffic will impact the total delay when there is no bike lane, however, does not impact delay when there is a bike lane. In the above experiments, the opposing spacing is constant, e.g., 300 m with standard variation of 0 m. Figure 4-17 shows the domains of application of bike lanes considering different opposing spacing standard variations, e.g., 50 m and 100 m. It can be seen that as the opposing spacing variation increases, the domains of application of bike lanes become slightly larger. This is expected since Figure 4-13 shows that the total delay increases as the opposing spacing variation increases in a mixed flow scenario when the average bike speed is 20 km/h.

4.2.4 Discussion

This study developed the car-bike mixed traffic flow model based on several assumption. However,

some of the derived conclusions may still stand and can be easily adopted when the assumptions are relaxed.

The proposed car-bike mixed traffic flow is based on two-way undivided roadways with one lane in each direction. When there are multi-lanes in each direction, there usually have enough space to set a separating bike lane if there is significant bike flow along the road. In the situation of no bike lane on multi-lane roadways, it can be modeled in scenarios. Assuming there is n lanes on the road, when the car flow is lower than the capacity of $n - 1$ lane, the bikes will ride on the rightmost lane and the cars that need to pass a bike can change to the left lanes to finish the passing maneuver. In this case, there will be no delay occur. When the car flow is higher than the capacity of $n - 1$ lane, the roadway can be separated into two system—a. left $n - 1$ lane and b. the rightmost lane. The cars will tend to use the left lanes to avoid intercourse with bikes. Therefore, in the left $n - 1$ lane, there will be saturating traffic flow on all of the lanes. The rest traffic flow will be distributed to the rightmost lane and the proposed car-bike mixed traffic flow may apply to this situation. The only necessary modification will be the way to decide when a car can pass the bike. In the current study, it is based on whether a car occupying the opposing lane or not; in the multi-lane situation, it will be decided based on if the traffic flow up to the capacity of $n - 1$ lanes. This modification could extend the implementing scenario into more broad cases.

This study only considers one directional bike flow along the analyzing direction. However, the proposed model can be adopted to address two directional bike flow as well. The major difference will be that traffic flow in both directions will not be homogeneous. This will cause the length of each moving bottlenecks cannot be explicitly calculated. It will require to dynamically calculate the time-space diagram time step by time step. In each time step, the traffic status, e.g., traffic density, traffic speed, and traffic flow, need to be updated for both directions, and based on the updated traffic status, whether a bike will generate a moving bottleneck can be decided for both directions. The traffic delay can be calculated in the same way when the entire time-space diagram is derived. More

detailed research to address two directional bike flow is worthy to conduct in the future.

The evaluation of the best domain of dedicated bike lane application is only based on the traffic efficiency in this study. However, the implementation of bike lane is complex decision intertwined with multi-aspects concerns, for example, the traffic safety. The passing maneuver between bikes and cars posed a high risk to the cyclists, which can be moderated by setting up a dedicated bike lane as well. The proposed car-bike mixed traffic flow model could also provide a tool to measure the traffic safety impact of dedicated bike lane. For example, when there is no bike lane, the car and bike shared the same road; the associated passing times, relative speed, time length of each passing maneuvers, time length of cars queue behind a bike, etc., can be derived from the proposed model, which can be valuable variables to analysis the crash risk.

4.3 SUMMARY

This study proposed a car-bike mixed traffic flow model based on the Lax-Hopf equations. The car is not allowed to pass a bike when there is not enough space in the opposing lane provided by the opposing traffic stream and hence a bike generates a series of moving bottlenecks to following cars as it encounters cars from the opposing direction. Such bottlenecks are modeled as internal conditions in the Lax-Hopf equations. The total delay of the mixed traffic flow is calculated and some of the critical traffic criteria are derived. The results suggest that in order to borrow the opposing lane to pass a bike, the traffic flow in the opposing direction needs to be less than 300 *veh/h*. Numerical tests reveal that as the difference between car and bike speeds gets smaller, lower opposing traffic flow is required to avoid significant congestion. The total delay, in general, is expected to increase as the speed difference between car and bike increases or car flow and bike flow increase. However, the interaction among bottlenecks can either increase or decrease the total delay, due to multiple interacting queues, or starving downstream bottlenecks of flow, respectively. Further, the impact of interaction among bottlenecks is significant to the total delay in a homogenous traffic flow but will

be diminished when variations are added to the traffic system, such as non-homogeneous bike headway, varying bike speed, and non-homogeneous opposing car spacing. Numerical results suggest that when bike headway variation is large, the results converge to the analytical calculation, while large bike speed variation or opposing car spacing variation leads to a total delay slightly higher than the analytical total delay. This study also examined the application domain of dedication bike lane to achieve minimum total person delay. The results suggests that it bike lanes can reduce overall delay when bike flow and car flow increase, the opposing traffic flow is high, or the spacing variation of opposing traffic flow is high. These conclusions can help traffic operators determine how to utilize the available space when both cars and bikes operate in the same space.

In this study, only one directional bikes are considered. The two-directional bike flow and the impact of intersection signal control on the mixed traffic flow will be studied in future work.

5 CONCLUSIONS

This study first examined the correlation between shared bike modes and public transit based on the trip datasets where the results of the individual characteristics analysis indicated that e-scooters are more likely to be used for leisure trips than public transit and shared bikes based on the temporal distribution of usage and the ratio of the morning peak to the afternoon peak trips. The Transit-GWR model showed that, while e-scooter and shared bike usage is highly positive correlated with each other, e-scooters tend to complement public transit, particularly in the west areas, the shared bikes are more likely to complement public transit and promote it in the east of the city. These differences suggest that shared bike and shared e-scooter have different roles in relation to public transit considering the built environment and the demographics. It can be concluded from the results that implementing a single mode strategically could help improve access to public transit, but when two of the micromobility modes co-exist, residents could have a preference for one over the other based on the specific regional environment. Since shared bikes and e-scooters attract users with different travel purposes, the implementation of these modes to complement public transportation can be chosen based on land use, demographic, and terrain characteristics in different areas.

This study then explored the potential impact of shared bike trips on traffic safety by developing a series of crash frequency models. Comparison of the models with different kernel functions indicated that the Hoerl function can improve the performance of the crash prediction model, and the POIs that are related to necessary trips are more tightly related to the crash frequency of all modes compared to using all POI trips. Spatial variation in the relationship was not observed in crashes involving cyclists; however, such variation was observed in crashes involving pedestrians and motorists. The model results indicate that an increase in shared bikes can increase the crash risk of cyclists and pedestrians but does not influence motorist crash risk. Other variables that are related to crash risk include the density of the roadway network, which reduces cyclist crash risk; the number of intersections, which increases crash risk for all modes; and younger population and lower family

income, which increases crash risk.

Finally, this study proposed a mixed car-bike traffic flow model based on the Lax-Hopf equations. The results suggest that in order to borrow the opposing lane to pass a bike, the traffic flow in the opposing direction needs to be less than 300 *veh/h*. As the difference between car and bike speeds gets smaller, lower opposing traffic flow is required to avoid significant congestion. The total delay, in general, is expected to increase as the speed difference between car and bike increases or car flow and bike flow increase. However, the interaction among bottlenecks can either increase or decrease the total delay, due to the multi-times queue or invalidating some of the blocks downstream, respectively. Further, the impact of interaction among bottlenecks is significant to the total delay in a homogenous traffic flow but will be diminished when variations are added to the traffic system, such as non-homogeneous bike headway, varying bike speed, and non-homogeneous opposing car spacing. This study also examined the application domain of dedication bike lane to achieve the minimum total delay. The results suggests that it is much recommended to set a bike lane when the car speed reduction caused by separating a bike lane from the main road is small, the opposing traffic flow is high, or the spacing variation of opposing traffic flow is high.

5.1 ENGINEERING SIGNIFICANCE

This research aims to enhance the understanding of the performance of micromobility and help design a multi-modal transportation system in a more cooperative, safer, and efficient manner. The conclusions from the comparison among micromobility modes and public transit clarify the similarities and differences of shared bike, shared e-scooter and public transit, so that the role of each in the multi-modal transportation system can be defined. Based on the trip characteristics of each traffic mode, shared micromobility can be designed to complement the public transit system and improve connectivity and accessibility of the public transit network. More complementary effects can be achieved by distributing the shared micromobility system to the most demanding time period

and spatial area. This also provides a guide for cities that operate or plan to operate more than one type of micromobility mode in the same area. The traffic safety modeling research helps understand the safety impacts of bike trips in big cities to reveal factors that most tightly correlate to different types of crash frequency. The safety cost of implementing shared micromobility systems thus can be quantified and by modeling the influences of different factors on each type of crash frequency specific operating regulations can be designed to reduce the crash possibility, such as separating the bicycle flow from the motorized traffic flow. The proposed model for mixed car-bike traffic flow helps to reveal the domains of application of dedicated bike lanes to achieve a system-wide efficiency that considers both bike and motorized vehicle traffic. This can help achieve a better multi-modal transportation system and minimize the average delay per person.

5.2 LIMITATIONS

While the goal of this thesis is to develop a comprehensive understanding of shared micromobility in a multi-modal transportation system, several limitations still exist. For the traffic demand analysis, the actual trip purposes of the users of different modes were unknown. A trip purpose survey in the future could help to examine the relationship between micromobility usage and geographical information to derive more detailed reasons for the different correlation distributions. Even though a large dataset was assembled to analyze the impacts of micromobility on traffic safety, not all factors that contribute to crashes were captured. Mainly, the motorist volumes were missing from the models which likely contributed to the lower predictive accuracy of the models. The model performance could be further improved with novel datasets once available in the future. While the car-bike mixed traffic flow is simplified in this study to draw basic conclusions, some improvements that make the model more realistic can be made in the future, such as considering two-directional bike flow and the impact of intersection signal control on the mixed traffic flow. These models can further be validated using empirical data.

REFERENCES

- Abdulhafedh, A., 2016. Crash Frequency Analysis. *Journal of Transportation Technologies* 06, 169–180. <https://doi.org/10.4236/jtts.2016.64017>
- Agarwal, A., Zilske, M., Rao, K.R., Nagel, K., 2015. An Elegant and Computationally Efficient Approach for Heterogeneous Traffic Modelling Using Agent Based Simulation. *Procedia Computer Science, The 6th International Conference on Ambient Systems, Networks and Technologies (ANT-2015), the 5th International Conference on Sustainable Energy Information Technology (SEIT-2015)* 52, 962–967. <https://doi.org/10.1016/j.procs.2015.05.173>
- Akaike, H., 1974. A new look at the statistical model identification. *IEEE Transactions on Automatic Control* 19, 716–723. <https://doi.org/10.1109/TAC.1974.1100705>
- Almanna, M.H., Ashqar, H.I., Elhenawy, M., Masoud, M., Rakotonirainy, A., Rakha, H., 2021. A comparative analysis of e-scooter and e-bike usage patterns: Findings from the City of Austin, TX. *International Journal of Sustainable Transportation* 15, 571–579. <https://doi.org/10.1080/15568318.2020.1833117>
- American Public Transportation Association, 2020. 2020 Public Transportation Fact Book, 71st Edition. ed.
- Anastasopoulos, P.C., Mannering, F.L., Shankar, V.N., Haddock, J.E., 2012. A study of factors affecting highway accident rates using the random-parameters tobit model. *Accid Anal Prev* 45, 628–633. <https://doi.org/10.1016/j.aap.2011.09.015>
- AustinTexas.gov [WWW Document], 2022. URL <https://www.austintexas.gov/sharedmobility> (accessed 4.1.22).

- Bai, S., Jiao, J., 2020. Dockless E-scooter usage patterns and urban built Environments: A comparison study of Austin, TX, and Minneapolis, MN. *Travel Behaviour and Society* 20, 264–272. <https://doi.org/10.1016/j.tbs.2020.04.005>
- Bao, J., Liu, P., Ukkusuri, S.V., 2019. A spatiotemporal deep learning approach for citywide short-term crash risk prediction with multi-source data. *Accid Anal Prev* 122, 239–254. <https://doi.org/10.1016/j.aap.2018.10.015>
- Barwaldt, M.S., Franzin, R. de F., Casarin, V.A., Santos, A.V. dos, 2014. Using the Theory of Graphs on the Implementation of Bike Lane in Small Towns. *Procedia - Social and Behavioral Sciences*, XVIII Congreso Panamericano de Ingeniería de Tránsito, Transporte y Logística (PANAM 2014) 162, 350–358. <https://doi.org/10.1016/j.sbspro.2014.12.216>
- Bordagaray, M., dell'Olio, L., Fonzone, A., Ibeas, Á., 2016. Capturing the conditions that introduce systematic variation in bike-sharing travel behavior using data mining techniques. *Transportation Research Part C: Emerging Technologies* 71, 231–248. <https://doi.org/10.1016/j.trc.2016.07.009>
- Brunsdon, C., Fotheringham, S., Charlton, M., 1998. Geographically Weighted Regression. *Journal of the Royal Statistical Society: Series D (The Statistician)* 47, 431–443. <https://doi.org/10.1111/1467-9884.00145>
- Buck, D., Buehler, R., Happ, P., Rawls, B., Chung, P., Borecki, N., 2013. Are Bikeshare Users Different from Regular Cyclists?: A First Look at Short-Term Users, Annual Members, and Area Cyclists in the Washington, D.C., Region. *Transportation Research Record* 2387, 112–119. <https://doi.org/10.3141/2387-13>
- Caliendo, C., De Guglielmo, M.L., Guida, M., 2013. A crash-prediction model for road tunnels. *Accid Anal Prev* 55, 107–115. <https://doi.org/10.1016/j.aap.2013.02.024>

- Campbell, A.A., Cherry, C.R., Ryerson, M.S., Yang, X., 2016. Factors influencing the choice of shared bicycles and shared electric bikes in Beijing. *Transportation Research Part C: Emerging Technologies* 67, 399–414. <https://doi.org/10.1016/j.trc.2016.03.004>
- Campbell, K.B., Brakewood, C., 2017. Sharing riders: How bikesharing impacts bus ridership in New York City. *Transportation Research Part A: Policy and Practice* 100, 264–282. <https://doi.org/10.1016/j.tra.2017.04.017>
- Capital Metro, 2020. Open Data Portal [WWW Document]. State of Texas Open Data Portal. URL https://data.texas.gov/browse?Dataset-Category_Agency=Capital+Metropolitan+Transportation+Authority&category=Transportation&utf8=%E2%9C%93 (accessed 6.6.20).
- Chen, Q., Wang, Y., 2016. A cellular automata (CA) model for motorized vehicle flows influenced by bicycles along the roadside. *Journal of Advanced Transportation* 50, 949–966. <https://doi.org/10.1002/atr.1382>
- Cheng, W., Gill, G.S., Sakrani, T., Dasu, M., Zhou, J., 2017. Predicting motorcycle crash injury severity using weather data and alternative Bayesian multivariate crash frequency models. *Accid Anal Prev* 108, 172–180. <https://doi.org/10.1016/j.aap.2017.08.032>
- Chunchu, M., Kuzhiyamkunnath, B.B., 2014. Analysis of the effect of two-wheeler lane-sharing behavior on macroscopic traffic flow modeling. *Transport* 29, 146–153. <https://doi.org/10.3846/16484142.2014.928788>
- Citi Bike, 2022. Citi Bike System Data [WWW Document]. Citi Bike NYC. URL <http://ride.citibikenyc.com/system-data> (accessed 8.22.22).
- Claudel, C.G., Bayen, A.M., 2010a. Lax–Hopf Based Incorporation of Internal Boundary Conditions Into Hamilton-Jacobi Equation. Part II: Computational Methods. *IEEE*

Transactions on Automatic Control 55, 1158–1174.

<https://doi.org/10.1109/TAC.2010.2045439>

Claudel, C.G., Bayen, A.M., 2010b. Lax–Hopf Based Incorporation of Internal Boundary Conditions Into Hamilton–Jacobi Equation. Part I: Theory. IEEE Transactions on Automatic Control 55, 1142–1157. <https://doi.org/10.1109/TAC.2010.2041976>

Clewlow, R.R., 2019. The Micro-Mobility Revolution: The Introduction and Adoption of Electric Scooters in the United States.

Costeseque, G., 2018. Lax-Hopf formula for ARZ traffic flow model.

da Silva, A.R., Rodrigues, T.C.V., 2014. Geographically Weighted Negative Binomial Regression—incorporating overdispersion. Stat Comput 24, 769–783. <https://doi.org/10.1007/s11222-013-9401-9>

Daganzo, C.F., 2005. A variational formulation of kinematic waves: basic theory and complex boundary conditions. Transportation Research Part B: Methodological 39, 187–196. <https://doi.org/10.1016/j.trb.2004.04.003>

Daganzo, C.F., 1994. The cell transmission model: A dynamic representation of highway traffic consistent with the hydrodynamic theory. Transportation Research Part B: Methodological 28, 269–287. [https://doi.org/10.1016/0191-2615\(94\)90002-7](https://doi.org/10.1016/0191-2615(94)90002-7)

David, G., 2016. Will helmet law kill Seattle’s new bike-share program? [WWW Document]. The Seattle Times. URL <https://www.seattletimes.com/seattle-news/transportation/will-helmet-law-kill-seattles-new-bike-share-program/> (accessed 8.22.22).

Economic Development, 2018. Dockless Vehicle Trips: Dockless Vehicle Trips 3KB. <https://doi.org/10.26000/030.000003>

- Gayah, V.V., Donnell, E.T., 2021. Estimating safety performance functions for two-lane rural roads using an alternative functional form for traffic volume. *Accid Anal Prev* 157, 106173.
<https://doi.org/10.1016/j.aap.2021.106173>
- Gitelman, V., Pesahov, F., Carmel, R., Chen, S., 2017. Exploring the characteristics of potential and current users of mobility scooters, among older people in Israel. *Transportation Research Part F: Traffic Psychology and Behaviour, Special Issue on Special Road safety as reflected by empirical non-crash data* 46, 373–389.
<https://doi.org/10.1016/j.trf.2016.07.010>
- Godunov, S.K., Bohachevsky, I., 1959. Finite difference method for numerical computation of discontinuous solutions of the equations of fluid dynamics. *Matematičeskij sbornik* 47(89), 271–306.
- Gooch, J.P., Gayah, V.V., Donnell, E.T., 2018. Safety performance functions for horizontal curves and tangents on two lane, two way rural roads. *Accid Anal Prev* 120, 28–37.
<https://doi.org/10.1016/j.aap.2018.07.030>
- Gooch, J.P., Gayah, V.V., Donnell, E.T., 2016. Quantifying the safety effects of horizontal curves on two-way, two-lane rural roads. *Accid Anal Prev* 92, 71–81.
<https://doi.org/10.1016/j.aap.2016.03.024>
- Gu, T., Kim, I., Currie, G., 2019. Measuring immediate impacts of a new mass transit system on an existing bike-share system in China. *Transportation Research Part A: Policy and Practice* 124, 20–39. <https://doi.org/10.1016/j.tra.2019.03.003>
- Guidon, S., Becker, H., Dediu, H., Axhausen, K.W., 2019. Electric Bicycle-Sharing: A New Competitor in the Urban Transportation Market? An Empirical Analysis of Transaction Data. *Transportation Research Record* 2673, 15–26.
<https://doi.org/10.1177/0361198119836762>

- Guler, S.I., Grembek, O., 2016. Use of different exposure metrics for understanding multi-modal travel injury risk. *International Journal of Transportation Science and Technology, Future Road Transport Technology* 5, 28–37. <https://doi.org/10.1016/j.ijst.2016.06.004>
- Hardt, C., Bogenberger, K., 2019. Usage of e-Scooters in Urban Environments. *Transportation Research Procedia, 21st EURO Working Group on Transportation Meeting, EWGT 2018, 17th – 19th September 2018, Braunschweig, Germany* 37, 155–162. <https://doi.org/10.1016/j.trpro.2018.12.178>
- Hauer, E., 2015. *The Art of Regression Modeling in Road Safety*. Springer International Publishing. <https://doi.org/10.1007/978-3-319-12529-9>
- He, T., Bao, J., Ruan, S., Li, R., Li, Y., He, H., Zheng, Y., 2020. Interactive Bike Lane Planning Using Sharing Bikes' Trajectories. *IEEE Transactions on Knowledge and Data Engineering* 32, 1529–1542. <https://doi.org/10.1109/TKDE.2019.2907091>
- Hilbe, J.M., 2011. *Negative Binomial Regression*. Cambridge University Press.
- Hopf, E., 1969. On the Right Weak Solution of the Cauchy Problem for a Quasilinear Equation of First Order. *Journal of Mathematics and Mechanics* 19, 483–487.
- Islam, A.S.M.M., Shirazi, M., Lord, D., 2022. Finite mixture Negative Binomial-Lindley for modeling heterogeneous crash data with many zero observations. *Accident Analysis & Prevention* 175, 106765. <https://doi.org/10.1016/j.aap.2022.106765>
- Jacobsen, P.L., 2003. Safety in numbers: more walkers and bicyclists, safer walking and bicycling. *Injury Prevention* 9, 205–209. <https://doi.org/10.1136/ip.9.3.205>
- Jia, R., Khadka, A., Kim, I., 2018. Traffic crash analysis with point-of-interest spatial clustering. *Accid Anal Prev* 121, 223–230. <https://doi.org/10.1016/j.aap.2018.09.018>
- Jordehi, B.A., Rose, G., Thompson, R.G., 2013. Motorcycle and Motor Scooter Use in Victoria, Australia. *Transportation Research Record* 2388, 61–70. <https://doi.org/10.3141/2388-09>

- Jung, S., Jang, K., Yoon, Y., Kang, S., 2014. Contributing factors to vehicle to vehicle crash frequency and severity under rainfall. *J Safety Res* 50, 1–10.
<https://doi.org/10.1016/j.jsr.2014.01.001>
- Khan, S.I., Maini, P., 1999. Modeling Heterogeneous Traffic Flow. *Transportation Research Record* 1678, 234–241. <https://doi.org/10.3141/1678-28>
- Kong, H., Jin, S., Sui, D., 2020. Deciphering the relationship between bikesharing and public transit: Modal substitution, integration, and complementation. *Transportation Research Part D: Transport and Environment* 85, 102392. <https://doi.org/10.1016/j.trd.2020.102392>
- La, Q.N., Lee, A.H., Meuleners, L.B., Van Duong, D., 2013. Prevalence and factors associated with road traffic crash among taxi drivers in Hanoi, Vietnam. *Accid Anal Prev* 50, 451–455.
<https://doi.org/10.1016/j.aap.2012.05.022>
- Laa, B., Leth, U., 2020. Survey of E-scooter users in Vienna: Who they are and how they ride. *Journal of Transport Geography* 89, 102874.
<https://doi.org/10.1016/j.jtrangeo.2020.102874>
- Lax, P.D., 1957. Hyperbolic systems of conservation laws II. *Communications on Pure and Applied Mathematics* 10, 537–566. <https://doi.org/10.1002/cpa.3160100406>
- Levy, N., Golani, C., Ben-Elia, E., 2019. An exploratory study of spatial patterns of cycling in Tel Aviv using passively generated bike-sharing data. *Journal of Transport Geography* 76, 325–334. <https://doi.org/10.1016/j.jtrangeo.2017.10.005>
- Li, Y., Sun, D., 2012. Microscopic car-following model for the traffic flow: the state of the art. *J. Control Theory Appl.* 10, 133–143. <https://doi.org/10.1007/s11768-012-9221-z>
- Lighthill, M.J., Whitham, G.B., 1955. On kinematic waves II. A theory of traffic flow on long crowded roads. *Proceedings of the Royal Society of London. Series A. Mathematical and Physical Sciences* 229, 317–345. <https://doi.org/10.1098/rspa.1955.0089>

- Liu, S., Shen, Z.-J.M., Ji, X., 2021. Urban Bike Lane Planning with Bike Trajectories: Models, Algorithms, and a Real-World Case Study. *M&SOM*.
<https://doi.org/10.1287/msom.2021.1023>
- Ma, T., Liu, C., Erdoğan, S., 2015. Bicycle Sharing and Public Transit: Does Capital Bikeshare Affect Metrorail Ridership in Washington, D.C.? *World Transit Research*.
- Ma, X., Zhang, X., Li, X., Wang, X., Zhao, X., 2019. Impacts of free-floating bikesharing system on public transit ridership. *Transportation Research Part D: Transport and Environment* 76, 100–110. <https://doi.org/10.1016/j.trd.2019.09.014>
- Mahmud, A., Gayah, V.V., 2021. Estimation of crash type frequencies on individual collector roadway segments. *Accident Analysis & Prevention* 161, 106345.
<https://doi.org/10.1016/j.aap.2021.106345>
- Martin, E.W., Shaheen, S.A., 2014. Evaluating public transit modal shift dynamics in response to bikesharing: a tale of two U.S. cities. *Journal of Transport Geography* 41, 315–324.
<https://doi.org/10.1016/j.jtrangeo.2014.06.026>
- Mathew, J.K., Liu, M., Seeder, S., Li, H., Bullock, D.M., 2019. Analysis of E-Scooter Trips and Their Temporal Usage Patterns. *Institute of Transportation Engineers. ITE Journal* 89, 44–49.
- Mazzanti, S., 2020. Boruta explained the way I wish someone explained it to me [WWW Document]. *Medium*. URL <https://towardsdatascience.com/boruta-explained-the-way-i-wish-someone-explained-it-to-me-4489d70e154a> (accessed 10.20.20).
- McAndrews, C., 2012. A Regional Perspective on the Relative Risk of Travel Modes: Geographic, Demographic, and Infrastructure Characteristics of Road Safety Disparities in San Francisco Bay Area.

- McKenzie, G., 2019. Spatiotemporal comparative analysis of scooter-share and bike-share usage patterns in Washington, D.C. *Journal of Transport Geography* 78, 19–28.
<https://doi.org/10.1016/j.jtrangeo.2019.05.007>
- Médard de Chardon, C., 2019. The contradictions of bike-share benefits, purposes and outcomes. *Transportation Research Part A: Policy and Practice* 121, 401–419.
<https://doi.org/10.1016/j.tra.2019.01.031>
- Mesbah, M., Thompson, R., 2011. Optimal design of bike lane facilities in an urban network. Presented at the Australasian Transport Research Forum (ATRF), 34th, 2011, Adelaide, South Australia, Australia.
- Metkari, M., Budhkar, A., Maurya, A.K., 2013. Development of Simulation Model for Heterogeneous Traffic with no Lane Discipline. *Procedia - Social and Behavioral Sciences*, 2nd Conference of Transportation Research Group of India (2nd CTRG) 104, 360–369.
<https://doi.org/10.1016/j.sbspro.2013.11.129>
- Munira, S., Sener, I.N., 2020. A geographically weighted regression model to examine the spatial variation of the socioeconomic and land-use factors associated with Strava bike activity in Austin, Texas. *Journal of Transport Geography* 88, 102865.
<https://doi.org/10.1016/j.jtrangeo.2020.102865>
- NACTO, 2020. Bike Share and Shared Micromobility Initiative [WWW Document]. National Association of City Transportation Officials. URL <https://nacto.org/program/bike-share-initiative/> (accessed 12.12.20).
- National Academies of Sciences, E., 2022. Highway Capacity Manual 7th Edition: A Guide for Multimodal Mobility Analysis. <https://doi.org/10.17226/26432>
- NYC Taxi and Limousine Commission, 2022. TLC Trip Record Data - TLC [WWW Document]. URL <https://www1.nyc.gov/site/tlc/about/tlc-trip-record-data.page> (accessed 8.22.22).

- Oshan, T.M., Li, Z., Kang, W., Wolf, L.J., Fotheringham, A.S., 2018. mgwr: A Python implementation of multiscale geographically weighted regression for investigating process spatial heterogeneity and scale (No. bphw9), OSF Preprints, OSF Preprints. Center for Open Science.
- Pan, H., Shen, Q., Xue, S., 2010. Intermodal Transfer Between Bicycles and Rail Transit in Shanghai, China. World Transit Research.
- Police Department (NYPD), 2022. Motor Vehicle Collisions - Crashes | NYC Open Data [WWW Document]. URL <https://data.cityofnewyork.us/Public-Safety/Motor-Vehicle-Collisions-Crashes/h9gi-nx95> (accessed 8.22.22).
- Pucher, J., Dijkstra, L., 2003. Promoting safe walking and cycling to improve public health: lessons from The Netherlands and Germany. *Am J Public Health* 93, 1509–1516.
<https://doi.org/10.2105/ajph.93.9.1509>
- Qu, Z., Cao, N., Chen, Y., Zhao, L., Bai, Q., Luo, R., 2017. Modeling electric bike–car mixed flow via social force model. *Advances in Mechanical Engineering* 9, 1687814017719641.
<https://doi.org/10.1177/1687814017719641>
- Reck, D.J., Haitao, H., Guidon, S., Axhausen, K.W., 2021. Explaining shared micromobility usage, competition and mode choice by modelling empirical data from Zurich, Switzerland. *Transportation Research Part C: Emerging Technologies* 124, 102947.
<https://doi.org/10.1016/j.trc.2020.102947>
- Renub Research, 2020. Electric Scooter Market Global Forecast by Country, Product, Battery Type, Company [WWW Document]. URL <https://www.researchandmarkets.com/reports/4912095/electric-scooter-market-global-forecast-by> (accessed 6.5.20).

- Reynolds, C.C., Harris, M.A., Teschke, K., Cripton, P.A., Winters, M., 2009. The impact of transportation infrastructure on bicycling injuries and crashes: a review of the literature. *Environmental Health* 8, 47. <https://doi.org/10.1186/1476-069X-8-47>
- Richards, P.I., 1956. Shock Waves on the Highway. *Operations Research* 4, 42–51. <https://doi.org/10.1287/opre.4.1.42>
- Roque, C., Cardoso, J.L., 2014. Investigating the relationship between run-off-the-road crash frequency and traffic flow through different functional forms. *Accid Anal Prev* 63, 121–132. <https://doi.org/10.1016/j.aap.2013.10.034>
- SafeGraph, 2020. Places Data & Foot-Traffic Insights | SafeGraph.com [WWW Document]. URL <https://www.safegraph.com/> (accessed 1.24.21).
- Sanders, R.L., Branion-Calles, M., Nelson, T.A., 2020. To scoot or not to scoot: Findings from a recent survey about the benefits and barriers of using E-scooters for riders and non-riders. *Transportation Research Part A: Policy and Practice* 139, 217–227. <https://doi.org/10.1016/j.tra.2020.07.009>
- Sengupta, A., Gayah, V.V., Donnell, E.T., 2021. Examining the impacts of crash data aggregation on SPF estimation. *Accident Analysis & Prevention* 160, 106313. <https://doi.org/10.1016/j.aap.2021.106313>
- Shaheen, S.A., Cohen, A.P., Martin, E.W., 2013. Public Bikesharing in North America: Early Operator Understanding and Emerging Trends. *Transportation Research Record* 2387, 83–92. <https://doi.org/10.3141/2387-10>
- Simoni, M.D., Claudel, C.G., 2020. A Fast Lax–Hopf Algorithm to Solve the Lighthill–Whitham–Richards Traffic Flow Model on Networks. *Transportation Science* 54, 1516–1534. <https://doi.org/10.1287/trsc.2019.0951>

- Singleton, P.A., Clifton, K.J., 2014. Exploring Synergy in Bicycle and Transit Use: Empirical Evidence at Two Scales. *Transportation Research Record* 2417, 92–102.
<https://doi.org/10.3141/2417-10>
- Song, M., Wang, K., Zhang, Y., Li, M., Qi, H., Zhang, Y., 2020. Impact Evaluation of Bike-Sharing on Bicycling Accessibility. *Sustainability* 12, 6124.
<https://doi.org/10.3390/su12156124>
- Syed, S., Khan, A., 2000. Factor Analysis for the Study of Determinants of Public Transit Ridership. *Journal of Public Transportation* 3. <https://doi.org/10.5038/2375-0901.3.3.1>
- Taylor, B.D., Miller, D., Iseki, H., Fink, C., 2009. Nature and/or nurture? Analyzing the determinants of transit ridership across US urbanized areas. *Transportation Research Part A: Policy and Practice* 43, 60–77. <https://doi.org/10.1016/j.tra.2008.06.007>
- Vishnoi, S.C., Claudel, C.G., 2022. Variable Speed Limit and Ramp Metering Control of Highway Networks Using Lax-Hopf Method: A Mixed Integer Linear Programming Approach. *IEEE Transactions on Intelligent Transportation Systems* 23, 7441–7456.
<https://doi.org/10.1109/TITS.2021.3069971>
- Wang, K., Zhao, S., Jackson, E., 2020. Investigating exposure measures and functional forms in urban and suburban intersection safety performance functions using generalized negative binomial - P model. *Accident Analysis & Prevention* 148, 105838.
<https://doi.org/10.1016/j.aap.2020.105838>
- Wang, M., Zhou, X., 2017. Bike-sharing systems and congestion: Evidence from US cities. *Journal of Transport Geography* 65, 147–154. <https://doi.org/10.1016/j.jtrangeo.2017.10.022>
- Wang, X., Zhou, Q., Yang, J., You, S., Song, Y., Xue, M., 2019. Macro-level traffic safety analysis in Shanghai, China. *Accident Analysis & Prevention* 125, 249–256.
<https://doi.org/10.1016/j.aap.2019.02.014>

- Wierbos, M.-J., Knoop, V., Hanseler, F., Hoogendoorn, S., 2018. A Macroscopic Flow Model for Mixed Traffic using Two-Dimensional Speed Functions: Mathematics Applied in Transport and Traffic Systems. *Mathematics Applied in Transport and Traffic Systems*.
- Wu, K., Lu, M., Guler, S., 2020. Modeling and optimizing bus transit priority along an arterial: A moving bottleneck approach. *World Transit Research*.
- Ye, X., Wang, K., Zou, Y., Lord, D., 2018. A semi-nonparametric Poisson regression model for analyzing motor vehicle crash data. *PLOS ONE* 13, e0197338.
<https://doi.org/10.1371/journal.pone.0197338>
- Younes, H., Zou, Z., Wu, J., Baiocchi, G., 2020. Comparing the Temporal Determinants of Dockless Scooter-share and Station-based Bike-share in Washington, D.C. *Transportation Research Part A: Policy and Practice* 134, 308–320.
<https://doi.org/10.1016/j.tra.2020.02.021>
- Zhang, C., Zhou, H., 2019. The study of Coopetition between Public Bus and Bike Sharing based on Environmental Protection. *E3S Web of Conferences* 136, 04015.
<https://doi.org/10.1051/e3sconf/201913604015>
- Zhang, X., Waller, S.T., Jiang, P., 2020. An ensemble machine learning-based modeling framework for analysis of traffic crash frequency. *Computer-Aided Civil and Infrastructure Engineering* 35, 258–276. <https://doi.org/10.1111/mice.12485>
- Zhang, Y., Mi, Z., 2018. Environmental benefits of bike sharing: A big data-based analysis. *Applied Energy* 220, 296–301. <https://doi.org/10.1016/j.apenergy.2018.03.101>
- Zheng, L., Li, Y., 2020. The Development, Characteristics and Impact of Bike Sharing Systems: A Literature Review. *International Review for Spatial Planning and Sustainable Development* 8, 37–52. https://doi.org/10.14246/irspsd.8.2_37

Zou, Z., Younes, H., Erdoğan, S., Wu, J., 2020. Exploratory Analysis of Real-Time E-Scooter Trip Data in Washington, D.C. Transportation Research Record 2674, 285–299.

<https://doi.org/10.1177/0361198120919760>

Curriculum Vitae

Educational Background

The Pennsylvania State University (PhD student) 09/2019 – 12/2022
Major: Civil Engineering GPA: 3.82/4.0
Dissertation: Modeling and Evaluation of Micromobility in A Multi-Modal Transportation System

Research Interests

Multi-modal urban transportation
Infrastructure management

Publications

- [1] **M. Lu**, J. Hydock, A. Radlińska, and S. I. Guler, “Reliability Analysis of a Bridge Deck Utilizing Generalized Gamma Distribution,” *Journal of Bridge Engineering*, Vol. 27, No. 4, 2022, p. 04022006. [https://doi.org/10.1061/\(ASCE\)BE.1943-5592.0001842](https://doi.org/10.1061/(ASCE)BE.1943-5592.0001842).
- [2] **M. Lu**, S. I. Guler, “Comparison of Random Survival Forest with Accelerated Failure Time-Weibull Model for Bridge Deck Deterioration,” *Transportation Research Record*. March 2022. doi:10.1177/03611981221078281.
- [3] K. Wu, **M. Lu**, and S. I. Guler, “Modeling and Optimizing Bus Transit Priority along An Arterial: A Moving Bottleneck Approach,” *Transportation Research Part C: Emerging Technologies*, vol. 121, p. 102873, Dec. 2020, doi: 10.1016/j.trc.2020.102873.
- [4] E. Yao, **M. Lu**, Y. Liu, L. Yuan, “Electric Bus Area Driving Plan Preparation Considering Charging Constraints”, *Journal of South China University of Technology (Natural Science Edition)*, vol. 47, no. 9, p. 68, Sep. 2019, doi: 10.12141/j.issn.1000-565X.180553.
- [5] Y. Liu, E. Yao, **M. Lu**, and L. Yuan, “Regional Electric Bus Driving Plan Optimization Algorithm considering Charging Time Window,” *Math. Probl. Eng.*, vol. 2019, p. e7863290, Oct. 2019..
- [6] **M. Lu**, S. I. Guler, V. V. Gayah, “*Multi-Objective Optimization of Maintenance, Rehabilitation and Reconstruction Decision Making Considering Safety*,” *Transp. Res. Rec.*, Accepted.
- [7] **M. Lu**, E. J. Traut, S. I. Guler, “*Comparing the Roles of Shared E-scooters and Bikes When Complementing and Competing with Public Transit*,” *J. Intell. Transp. Syst.*, Accepted.
- [8] **M. Lu**, H. Liu, S. I. Guler, “*Impact Of Bikes on Traffic Efficiency Based on A Car-Bike Mixed Traffic Flow Modeling*,” *TRANSPORT RES B-METH*, Under review.
- [9] **M. Lu**, V. V. Gayah, S. I. Guler, “*Does the Shared Bike Cause More Crashes to Pedestrian and Cyclists? A case study based on New York City*,” *Transp. Res. Rec.*, Under review.

Awards & Honors

2021 Mark E. and Claire L. Alpert Graduate Fellowship
2021 Leo P. Russell Graduate Fellowship
2021 18th Annual College of Engineering Research Symposium – *Best Oral Presentation Award*