

University of Massachusetts Amherst

ScholarWorks@UMass Amherst

Doctoral Dissertations

Dissertations and Theses

March 2022

ASSESSING THE IMPACT OF BICYCLE TREATMENTS ON BICYCLE SAFETY: A MULTI-METHODS APPROACH

Aikaterini Deliali

University of Massachusetts Amherst

Follow this and additional works at: https://scholarworks.umass.edu/dissertations_2



Part of the [Transportation Engineering Commons](#)

Recommended Citation

Deliali, Aikaterini, "ASSESSING THE IMPACT OF BICYCLE TREATMENTS ON BICYCLE SAFETY: A MULTI-METHODS APPROACH" (2022). *Doctoral Dissertations*. 2459.

<https://doi.org/10.7275/27202789> https://scholarworks.umass.edu/dissertations_2/2459

This Open Access Dissertation is brought to you for free and open access by the Dissertations and Theses at ScholarWorks@UMass Amherst. It has been accepted for inclusion in Doctoral Dissertations by an authorized administrator of ScholarWorks@UMass Amherst. For more information, please contact scholarworks@library.umass.edu.

**ASSESSING THE IMPACT OF BICYCLE TREATMENTS ON
BICYCLE SAFETY: A MULTI-METHODS APPROACH**

A Dissertation Presented

by

AIKATERINI DELIALI

Submitted to the Graduate School of the
University of Massachusetts Amherst in partial fulfillment
of the requirements for the degree of

DOCTOR OF PHILOSOPHY

February 2022

Civil & Environmental Engineering

©Copyright by Aikaterini Deliali 2022

All Rights Reserved

**ASSESSING THE IMPACT OF BICYCLE TREATMENTS ON
BICYCLE SAFETY: A MULTI-METHODS APPROACH**

A Dissertation Presented

by

AIKATERINI DELIALI

Approved as to style and content by:

Eleni Christofa , Chair

Michael Knodler, Member

Chengbo Ai, Member

Shannon C. Roberts, Member

John Tobiason, Department Head
Civil & Environmental Engineering

ACKNOWLEDGMENTS

There is an end to every journey. My Ph.D. journey at UMass Amherst helped me grow as a professional and individual and many people from the UMass community have played a role in that.

First, I do not have words to describe how grateful I am for working under the supervision of Dr. Eleni Christofa. Eleni, thank you for your continuous encouragement and support, for believing in my abilities, and for mentoring me on how to be a better researcher and professional. I would like to thank Dr. Michael Knodler for always making time to discuss and give feedback for my research. I am also thankful for his kind personality that has shaped the way transportation faculty, staff, and students coexist. I am thankful for having the opportunity to work with Dr. Chengbo Ai in two of the studies in my dissertation. His input and feedback improved my research skills. Lastly, I would like to thank Dr. Shannon Roberts for her detailed recommendations on my dissertation.

In addition to having the opportunity of working with exceptional professors and mentors, I was lucky to share an office with very nice people that I could call friends. Interacting with all of them made improved my day-to-day life. My academic siblings, Dr. Farnoush Khalighi and Dr. Nick Fournier, have always been there when I needed a piece of advice or a chat - I am very thankful to both of you. The presence of Dr. Haris Sipetas in the office made everything look positive, even when working after midnight. Having Dr. Alyssa Ryan right next to me to quickly discuss our research or chat for a while is something that I already miss. I am thankful for Dr. Francis Tainter and Qing Hu for helping me with my data collection, during extreme weather conditions. I am happy for meeting Dr. Yalda Ebadi, Pai Ganesh Manglore, Dr. Sayeeda Ayaz, Dr. Xinlian Yu, Dr. Marie Louis, Aamani Gautam, Rana Eslami Fard, and Hossein Ghafouriyan.

Nothing in journey would have been the same without my lovely Greek friends. Panos, Fani, George, Iro, Haris, Antonis, Ana-Maria, Konstantinos, Androklis, you know how important you were in my life in Amherst. Special thanks go to Panos for being a great roommate and for making me part of his circle when I arrived in Amherst. Many, many thanks go to Iro for always being there.

I am also very happy for the friends I made outside the office and the Greek-community. The kindest person happened to be my roommate, Rodrigo (now Dr. Rodrigo Mercado-Fernandez). Dr. Deniz Besik and Dr. Pritha Dutta and Mojtaba Salarpour are interesting, smart, and funny and it was also great to work with them in the INFORMS student chapter.

Many thanks go to Jodi Ozdarski, Kris Stetson, and Kassandra Rounds for answering all of my questions on time and with a smile on their faces.

Lastly, I would like to thank my mother, Annie, my three brothers Dimitris, Parmention, and George, and my sweet grandmother for their encouragement throughout this process!

ABSTRACT

ASSESSING THE IMPACT OF BICYCLE TREATMENTS ON BICYCLE
SAFETY: A MULTI-METHODS APPROACH

FEBRUARY 2022

AIKATERINI DELIALI

B.Sc. and M.Eng., NATIONAL TECHNICAL UNIVERSITY OF ATHENS
M.S., UNIVERSITY OF MASSACHUSETTS AMHERST

Ph.D., UNIVERSITY OF MASSACHUSETTS AMHERST

Directed by: Dr. Eleni Christofa

Compared to other modes, bicyclists are disproportionately affected by crashes considering their low mode share. There is evidence that crashes between bicyclists and motorized vehicle take place at road segments and signalized intersections where bicycle treatments (e.g., bike lanes) are present, urging for in-dept analysis of the safety impact of the various bicycle treatment types. Additionally, it is important to identify sensor types that have the potential to advance field data collection and traffic monitoring in multi-modal road environments. In this dissertation, three approaches, namely crash analysis, traffic conflict analysis, and analysis of driver speeding and glancing behavior, were implemented to investigate the safety impact of bicycle treatments at the segment- and the intersection-levels on bicycle safety. Prediction models were developed to predict bicycle-motorized vehicle crashes at road segments and signalized intersections, and traffic conflicts between straight-going bicyclists and right-turning vehicles at signalized intersections. Driver speeding and glancing behavior was analysed for the segment and the intersection levels. A mode classification framework to classify trajectories recorded using a radar-based sensor

was developed to test the feasibility of using radar-based sensors in field studies. The findings of this dissertation contribute to bicycle safety research in terms of quantifying the safety impact of various bicycle treatment types and how to assess and also, by showing how to assess bicycle safety. The findings of this research have the potential to stand as a valuable tool for transportation policymakers and officials in charge of establishing safe bicycle networks.

TABLE OF CONTENTS

Acknowledgments	iv
Abstract	v
List of Tables	xii
List of Figures	xvi
1 Introduction	1
1.1 Motivation	1
1.2 Research questions	5
1.3 Research contributions	5
1.4 Dissertation organization	9
2 Literature Review	11
2.1 Crash-based safety metrics	12
2.1.1 Injury severity	13
2.1.2 Crash rates	18
2.1.3 Crash prediction models	24
2.2 Surrogate safety metrics	32
2.2.1 Traffic conflict indicators	34
2.2.2 Categories of surrogate safety studies	40
2.3 Safety impact of bicycle infrastructure treatments	47
2.3.1 Segment-level bicycle infrastructure treatments	47
2.3.2 Intersection-level bicycle infrastructure treatments	50
2.4 Summary of literature review	54

3	Investigating the safety impact of segment- and intersection-level bicycle treatment on bicycle-motorized vehicle crashes.....	56
3.1	Introduction	56
3.2	Literature review	58
3.2.1	Bicycle safety on road segments	58
3.2.2	Bicycle safety at signalized intersections	60
3.3	Data	62
3.3.1	Crash data	63
3.3.2	Road network	63
3.3.3	Bicycle exposure data	63
3.3.4	Traffic exposure data	67
3.3.5	Bicycle treatments	67
3.3.6	Final dataset	68
3.4	Methodology	68
3.4.1	Development of signalized intersection crash prediction models	70
3.4.2	Development of road segment crash prediction models	72
3.5	Results	72
3.5.1	Crash prediction model for road segments	72
3.5.2	Crash prediction models for the intersection	73
3.6	Discussion	76
3.6.1	Crash prediction model for road segments	76
3.6.2	Crash prediction models for signalized intersections	77
3.7	Conclusions and future extensions	79
4	Right-hook traffic conflicts between motorists and bicyclists at signalized intersections.....	83

4.1	Introduction	84
4.2	Literature review	87
4.2.1	Segment-level treatments at intersections	87
4.2.2	Intersection-level treatments at intersections	89
4.2.3	Summary of the literature	90
4.2.4	Site selection and video data collection	92
4.2.5	Traffic conflict extraction	94
4.2.6	Model formulation	98
4.3	Results	100
4.3.1	Traffic conflict model	100
4.3.2	User sequence and PET values	103
4.4	Discussion	105
4.5	Conclusions and future extensions	107
5	Assessing driver speeding and glancing behavior in the presence of protected bicycle treatments.....	110
5.1	Introduction	111
5.2	Literature review	112
5.2.1	Protected bike lanes	113
5.2.2	Protected Intersections	115
5.2.3	Summary of the literature	116
5.3	Methodology	117
5.3.1	Apparatus	117
5.3.2	Research hypotheses	119
5.3.3	Scenario development	120
5.3.4	Dependent and independent variables	123

5.3.5	Sample size and participants	124
5.3.6	Experimental procedure	126
5.4	Results	127
5.4.1	Glance data	127
5.4.2	Speed data	131
5.5	Discussion	134
5.5.1	Glances	134
5.5.2	Speed	136
5.6	Conclusions and future extensions	136
6	A framework for mode classification in multimodal environments using radar-based sensors.....	139
6.1	Introduction	140
6.2	Literature Review	142
6.2.1	Vision-based traffic monitoring	142
6.2.2	Radar-based traffic monitoring	145
6.2.3	Summary of the literature	146
6.3	Radar-based sensor configuration and data format	148
6.4	Mode Classification Framework	149
6.4.1	Support Vector Machine Algorithm for mode classification	149
6.4.2	Data labeling	153
6.4.3	Feature vector formulation	154
6.4.4	SVM algorithm implementation	156
6.5	Experimental Tests	158
6.6	Test sites	159
6.6.1	Test site 1: Unsignalized intersection	159

6.6.2	Test site 2: Signalized intersection	160
6.6.3	Data preparation	161
6.6.4	Results	164
6.7	Conclusions and future extensions	169
7	Conclusions, practical implications, and future extensions.....	172
7.1	The impact of bicycle treatment type on bicycle safety	174
7.1.1	Segment-level bicycle treatments at road segments	174
7.1.2	Segment-level bicycle treatments at signalized intersections	175
7.1.3	Intersection-level treatments at signalized intersections	177
7.1.4	Contributions	178
7.2	Methodological findings	179
7.2.1	Strengths and limitations of the three safety assessment approaches	179
7.2.2	Sensors for conducting traffic monitoring studies	182
7.2.3	Safety-in-numbers effect	182
7.2.4	Contributions	183
7.3	Practical implications and methodology transferability	184
7.4	Future extensions	186
	REFERENCES.....	192

List of Tables

1.1	Summary of studied bicycle treatments	7
3.1	Summary of research crash-based studies assessing the presence and/or type of segment-level bicycle treatments on intersection bicycle safety	61
3.2	Linear regression model for the the network-wide AADB estimation .	66
3.3	Logistic regression model for the segment crashes	73
3.4	Intersection crash prediction model - Impact of segment-level treatments	75
3.5	Intersection crash prediction model - Impact of segment-level treatments	76
4.1	Data collection sites	95
4.2	Base model	101
4.3	Traffic conflicts model with bicycle treatment type	102
4.4	Kruskal-Wallis test results for different PET values and user sequence	104
5.1	Scenario Design	122
5.2	Mean and variance for the dependent variables used in the power analysis	125
5.3	Logistic regression model for glances at the bicyclist (segment)	129
5.4	Final logistic regression model for glances at the intersection Zones 1 and 2	131
5.5	Repeated measures ANOVA for segment speed	132
6.1	Summary of data collected per test site	160
6.2	C and γ values for the eight SVM models	164

List of Figures

2.1	Safety pyramid (adopted from Hyden, 1987)	35
2.2	Diagram for classifying conflict severity, adopted from [104]	36
2.3	Conventional definition of PET [7]	38
2.4	Illustration of relationship between Delta-V and probability of a severe injury (adopted from [80]).	39
2.5	Different configurations of conventional bike lanes	227
2.6	Buffered bike lanes	228
2.7	Different configurations of protected bike lanes	229
2.8	Bicycle treatments that enhanced bicyclists to use the full traffic lane along with motorized vehicles	230
2.9	Bike box	231
2.10	Intersection-crossing pavement marking and turning queue box	231
2.11	Protected or “Dutch” intersection	231
3.1	Segments with Ride app data (in total annual rides) and bicycle counter locations in the City of Portland	232
3.2	Scatter plot of the 23 road segments for which is available both Ride app and counter data.	233
3.3	Bicycle-motorized vehicle crashes at road segments.	233
3.4	Bicycle-motorized vehicle crashes at signalized intersections.	234
3.5	Distribution of motorized vehicle and bicycle demand for the studied road segments	235
4.1	Right-hook collision between a bicycle and a right-turning vehicle (Adopted from [75])	236
4.2	Example of green-colored intersection crossing markings (Seattle, WA)	236

4.3	Example of bike box (Cambridge, MA)	237
4.4	Cambridge Street at Springfield Street (Cambridge, MA). [Segment: sharrows; Intersection: None]	237
4.5	Binney street (Cambridge, MA). Bicycle treatment type: protected bike lanes with green-colored intersection crossing markings	238
4.6	Western Ave at Memorial Drive (Cambridge, MA). [Segment: protected bike lanes; Intersection: None]	238
4.7	Massachusetts Avenue at... (Cambridge, MA). Segment: conventional bike lane; Intersection: crossing Markings]	239
4.8	Cambridge Street at Sudbury Street (Boston, MA). [Segment: conven- tional bike lane; Intersection: bike box]	239
4.9	Massachusetts Avenue at Beacon Street (Boston, MA). [Segment: pro- tected bike lane; Intersection: bike box and crossing markings]	240
4.10	Massachusetts Avenue at Commonwealth Avenue (Boston, MA). [Seg- ment: conventional bike lane; Intersection: bike box and crossing markings]	240
4.11	Beacon Street at Street (Somerville, MA). [Segment: conventional bike lane; Intersection: bike box and crossing markings (not visible on Google Maps imagery due to recent installation)]	240
4.12	Post Encroachment Time graphical representation; adopted from [7] .	241
4.13	Conflict rates per bicycle treatment type	241

4.14	Heatmaps for the percentage of the number of traffic conflicts per PET value and per site over the total conflicts	242
5.1	Different configurations of protected bike lanes (PBL); these bicycle treatments are also known as separated bike lanes or cycle tracks. . .	243
5.2	University of Massachusetts Amherst Human Performance Lab driving simulator and eye-tracking device	243
5.3	Bicycle infrastructure treatment combinations	244
5.4	Simulated conventional and protected intersections	244
5.5	Scenario geometric configuration. The orange arrow shows the participants' driving path. The drive parts AB and CB denote the areas for which speed and glance data were collected and analyzed.	245
5.6	Intersection areas of interest.	245
5.7	Driver is glancing at the bicyclist	245
5.8	Areas of interest for right glances	246
5.9	Percentage of drives within each scenario that participants glanced right at the intersection at least once (Zone 1 or 2)	247
5.10	Speed violin plots for segment AB (Blue: no bicyclist is present; Brown: a bicyclist is present)	247
5.11	Interaction between bicycle infrastructure treatment at the segment and bicyclist presence	248
6.1	Radar-based sensor data visualization	248
6.2	Data processing interface in MATLAB	249

6.3	Test site 1: Unsignalized intersection	249
6.4	Test site 2: Signalized intersection	250
6.5	Motorized vehicle speed profiles	250
6.6	Motorized vehicle acceleration profiles	251
6.7	Pedestrian speed profiles	251
6.8	Pedestrian acceleration profiles	251
6.9	Bicycle speed and acceleration profiles	252
6.10	Precision for the motorized vehicle, bicycle, and pedestrian classification	252
6.11	Recall for the motorized vehicle, bicycle, and pedestrian classification	253
6.12	Accuracy of the classifier across all eight models for all three modes .	253

1 Introduction

1.1 Motivation

In an effort to set the ground for a “more sustainable future for all”, the United Nations (U.N.) established the Sustainable Development Goals in 2015. These represent a set of seventeen interconnected goals that aim to address major societal and environmental issues. In particular, Goal 11 “Sustainable Cities and Communities” highlights the role of bicycling, walking and use of public transportation, to the creation of sustainable, livable, and resilient cities and communities. Increased bicycle mode share is recognized as a means of improving public health and well-being, ensuring access to multiple activities and amenities in an emission-free, energy-saving, and affordable manner.

Governments seek to implement policies that simultaneously address bicycle mobility, i.e., policies to establish bicycle as a viable mode of transportation, and safety [221]. The most common strategy is the allocation of roadway space to bicyclists, so that they have a designated space to ride. Several studies confirm that the existence of bicycle infrastructure treatments affects the number of people bicycling [22]. Extensive research has been conducted to understand people’s perceptions with respect to bicycling and identify factors that deter people from or attract people to bicycling. Bicyclists feel safer when bicycle treatments are present [19, 213, 224, 232, 247]. At the same time, the majority of drivers are in favor of bicycle treatments and designated roadway space for bicyclists, which can improve driver awareness of bicyclists and bicyclist movement predictability. [213].

In addition to understanding bicyclist and driver preferences with respect to the types of bicycle treatments (or bicycle infrastructure treatments or bicycle facilities), it

is of high importance to identify the factors that affect bicyclist safety, i.e., factors that are responsible for crashes between bicyclists and motorists, and use that information to develop countermeasures. In the U.S. bicyclist fatalities caused by bicycle-motorized vehicle crashes saw a 6.3% increase from 2017 to 2018; during the same period the overall motorized vehicle crashes decreased by 2.3% [170]. Bicycle safety improves at locations where bicycle treatments are present [54]; yet, crashes still take place when bicycle treatments are present, highlighting the need to further investigate the safety benefits of the various bicycle infrastructure treatment types.

Bicycle treatments can be broadly separated into two categories; those for the segment-level, such as bike lanes, and those for the intersection-level, such as intersection-crossing pavement markings. Segment-level treatments indicate where bicyclists should ride while at the segment and provide no information on how bicyclists should navigate an intersection, despite the fact that segment-level treatments are often present at the intersection area (i.e., upstream and downstream an intersection). On the other hand, intersection-level treatments are present only at the intersection. It should be noted that compared to segment-level treatments, intersection-level treatments are not frequently and consistently implemented [176]. Road user movements and interactions at the segment-level differ from those at the intersection-level. In addition, intersection behavior could be affected by upstream segment conditions and infrastructure. As a result, there is a need to study the impact of segment-level bicycle treatments at both the segment and intersection. For example, protected bike lanes provide separation between bicyclists and motorists along the segment; however, this separation is not maintained at the intersection which might affect drivers' awareness of bicyclist presence during potential interactions of both user types at the intersection. Moreover, it is unclear whether there are safety benefits related to specific combinations of segment-level and intersection-level treatments.

Traditionally, traffic safety research employs crash analysis to associate crash outcomes with various crash contributing factors and identify high-risk sites, i.e., locations that are more prone to crashes. However, crash analysis is not always the most appropriate way to assess safety; limited bicycle crash data availability or reliability (e.g., crash location or severity is not precisely recorded) could limit the analysis. At the same time, crashes are not the best proxy for assessing road user behavior and this applies to both bicyclists and motorists. For example, crash analysis cannot reveal which road user started braking during a collision course. Therefore, additional quantitative methods have been developed to assess traffic safety. These methods focus on road user behavior and interactions and mainly aim to associate unsafe behaviors and/or interactions with various aspects such as traffic conditions, road environment characteristics, etc. Road user behavior and interactions are usually measured in terms of kinematics, e.g., speed or acceleration data, or proximity metrics, e.g., the distance between two road users or a road user and an object. Proximity metrics are often coupled with time-based information, to indicate how close two road users approached in time and space.

In bicycle safety research methods that assess road user behavior and interactions are particularly beneficial as they provide insights on unsafe situations that might lead to a crash. Despite the issue of under-reporting, focusing on a particular crash type reduces the number of crashes to be analyzed, and in turn, the analysis might be inconclusive. For example, there is evidence that right-hook crashes at urban intersections are frequent [25, 73, 103]. Those crashes involve a right-turning vehicle and a straight-going bicyclist at an intersection. The vehicle driver might have not seen the bicyclist or have assumed that the bicyclist will yield, and so he/she continues turning right and collides with the bicyclist. As road user behavior and interactions data can be obtained more easily, especially nowadays with the current

technological advancements, e.g., video cameras, GPS data, etc., it is more feasible to study right-hook conflicts instead of right-hook crashes.

Road user behavior and interactions data can be obtained from (1) field or naturalistic studies, where various sensing technologies are used to record road users in real time while are traveling along a segment or an intersection, and (2) driving simulator studies, where human drivers are exposed to simulated road environments. Both study types allow for the acquisition of kinematic data (e.g., speed and acceleration) as well as vehicle position data. Field studies aimed at capturing and analyzing road user interactions are primarily performed with video cameras that collect traffic data. There are other types of field studies, e.g., analysis of vehicle trajectories, but those are not appropriate for studying interactions, as data from only some vehicles/ road users are obtained. In field studies one can assess multiple types of unsafe behaviors, such as speeding, red-light crossing, etc. For bicycle safety purposes, bicyclist behavior when navigating an intersection and interacting with motorized vehicles and the infrastructure (e.g., bicycle treatments and traffic signals) can be analyzed and has the potential to reveal and quantify unsafe traffic events. In driving simulator studies, the researcher can study all the above when exposing human drivers in a simulated environment. Additionally, in driver simulator studies the researcher can also be benefited by using additional equipment such as an eye-tracking device that captures driver glancing patterns; this is a way of evaluating driver attention and response in road-related stimuli.

The main difference between field studies and driving simulator studies is that the former capture road users in real time and thus, in realistic traffic conditions while driver simulator studies expose human drivers to a set of controlled and predefined conditions. Depending on the case, these differences may appear beneficial for studying a particular problem. Field studies are advantageous in that they can capture multiple bicyclists

interacting with motorized vehicles in real traffic conditions. Driving simulator experiments allow the development of multiple scenarios where only one variable changes per scenario and so, the individual impact of that variable on driver behavior is captured.

In the case of field studies conducted with video cameras there might be a few additional disadvantages. Video cameras underperform in adverse lighting and weather conditions and thus, their ability to detect and track road users is reduced. Additionally, detection and tracking of non-motorized users can be problematic due to the dimension, shape, and color of those road users. Therefore, considering alternative data collection sensor types, such as radar-based sensors, that remain unaffected from lighting and weather conditions has the potential to improve traffic monitoring and in turn, non-motorized user safety. Lastly, radar-based sensors are not related with data privacy issues as video cameras, and are expected to be more favored from the public.

1.2 Research questions

This dissertation aims to answer the following research questions with respect to bicycle safety and alternative technologies for traffic monitoring:

1. How do various bicycle treatment types contribute to bicycle safety at the segment- and the intersection-levels?
2. Are radar-based sensors a feasible data collection tool for traffic monitoring purposes in multimodal environments?

1.3 Research contributions

The two research questions are answered through four studies. The first three studies (Chapters 3-5) are related to the bicycle safety aspect expressed by Questions

1 while Question 2 is addressed by the study presented in Chapter 6. With respect to bicycle safety, more than one type of studies are needed to better understand and quantify the safety impact of various bicycle treatments at the segment- and the intersection-levels. This is because each study is associated with a number of limitations and therefore, another approach is needed to address those limitations as explained below.

In Chapter 3 a crash analysis explores whether crashes between bicyclists and motorized vehicles at the segment and the intersection are associated with the type of the bicycle treatment. Crash analysis has certain limitations, e.g., does not provide insights regarding the underlying crash mechanism, and therefore, the field study in Chapter 4 as well as the driver simulator study in Chapter 5 have the potential to bridge this gap as they allow to assess road user interactions and capture those unsafe interactions that can possibly lead to crashes. The driver simulator study (Chapter 5) assesses driver behavior in the presence of several treatments that have also been studied in Chapters 3 and 4, but allows to assess innovative bicycle treatments that are not widely implemented and thus, cannot be assessed through crash records or a field study. The studied bicycle treatments as well as the utilized method to assess safety are summarized in Table 1.1. Chapter 6 is connected with the rest of the dissertation as it demonstrates the capabilities of an alternative to video cameras sensor type, i.e., radar-based sensor, for collecting data for field studies.

The contributions of each study are presented below:

1. Crash prediction models for road segments and signalized intersections

The objective of this study is to develop crash prediction models for midblock road segments (i.e., between two intersections) and signalized intersections, considering

Table 1.1: Summary of studied bicycle treatments

Performance metric	Site	Bicycle treatments for	
		Segments	Intersections
Crash frequency (Chapter 3)	Segment	CBL ¹ , PBL ² , Sharrows	n/a*
	Signalized int.	CBL, PBL, Sharrows	BB ³ , crossings ⁴
Traffic conflict frequency (Chapter 4)	Signalized int.	CBL, PBL, Sharrows	BB ³ , crossings
	Segment	CBL, PBL	n/a
Driver speed and glance (Chapter 5)	Signalized int.	CBL, PBL	PI ⁵ , crossings

¹conventional bike lanes; ²protected bike lanes;

³bike box; ⁴intersection-crossing pavement markings; ⁵protected intersection

*not applied

the presence of bicycle treatments for the segment-level such as conventional bike lanes, protected bike lanes, and sharrows as well as intersection-level treatments such as bike boxes and intersection-crossing pavement markings. Up to date there is no research that connects all three different types of bicycle treatments to crash outcomes on midblock road segments. For signalized intersections there are several studies that have considered the presence and type of bicycle treatments, however, none of them have compared all three treatment types. Moreover, there is no study that has emphasized on the combination of treatments, e.g., an intersection where one road has conventional bike lanes and the other has protected bike lanes versus an intersection where both streets have conventional bike lanes. Lastly, intersection-level bicycle treatments, such as bike boxes and intersection-crossing pavement markings, have not been incorporated in any of the existing crash prediction models for intersections. *This work will be the first one to relate the presence of various bicycle treatments to crash frequency at both the segment- and (signalized) intersection-levels. Specifically for signalized intersections, this is the first study to relate different combinations of segment-level treatments at the intersection with crash frequency and explore the effect of intersection treatments such as bike boxes and intersection-crossing pavement*

markings on crash frequency.

2. Right-hook traffic conflicts between motorists and bicyclists at signalized intersections

The objective of this study is to develop prediction models for right-hook conflicts at signalized intersections where one of the following treatments: conventional bike lanes, protected bike lanes, or sharrows, is present at the segment upstream the intersection and bike boxes or intersection-crossing pavement markings are present at the intersection. Right-hook crashes are a common type of crash between bicyclists and motorists at signalized intersections. Studying traffic conflicts between these road users has the potential to reveal the underlying mechanism of right-hook crashes and help develop countermeasures to prevent those crashes. Existing research on right-hook conflicts has only compared intersections with protected bike lanes and those with no bicycle treatments, without considering the presence or type of intersection treatments such as bike boxes or intersection-crossing pavement markings. *This dissertation contributes to the literature by comparing the impact of the five bicycle treatment types on the frequency of right-hook conflicts between motorized vehicles and bicyclists at signalized intersections.*

3. Assessing driver speeding and glancing behavior in the presence of protected bicycle treatments

A driving simulator experiment has been conducted to investigate driver speeding and glancing behavior at the segment-level while driving next to conventional and protected bike lanes and at the intersection-level while making a right turn when intersection-crossing pavement markings or protected intersection design elements have been implemented. Driver behavior has not been assessed while traveling next to conventional and protected bike lanes in order to allow for a comparison between the

different treatments. Moreover, there is limited research on protected intersections and no previous study has examined whether the segment-level treatment type upstream the protected intersection affects driver behavior. *This is the first study to capture and compare driver behavior in the presence of protected and not protected segment- and intersection-level treatments, while at the same time accounting for the different combinations of segment- and intersection-levels treatments.*

4. A framework for mode classification in multimodal environments using radar-based sensors

The objective of this chapter is to develop a mode classification framework that can be used to assign mode class to trajectories recorded by radar-based sensors. This is the first step in the deployment of radar-based sensors for traffic monitoring in multimodal road environments, i.e., where motorized and non-motorized road users are present. Previous efforts have demonstrated the feasibility of utilizing radar-based sensors in motorized user-only road environment. Therefore, it is uncertain whether these sensor are capable of detecting and recording non-motorized user trajectories in addition to motorized ones. Then, it is unclear whether it is feasible to assign the correct mode class to the recorded trajectories. *The proposed mode classification framework is the first one to classify trajectories that have been recorded by a radar-based sensor to three different classes, namely: pedestrians, bicyclists, and motorized vehicles. The proposed approach achieves high accuracy and can perform well in various traffic scenes that vary in terms of traffic control.*

1.4 Dissertation organization

The second chapter of this dissertation reviews and synthesizes the literature related to bicycle safety. First, the chapter presents different methods that are used

to assess traffic safety. Strengths and limitations of each approach are discussed, focusing specifically on bicycle safety. The different bicycle treatment types for the segment- and the intersection-levels are presented followed by a synthesis of the existing literature regarding their safety benefits. Lastly, literature on traffic sensing technologies is presented, focusing on those technologies that can be used in studies that aim to assess bicycle safety.

Chapters 3-5 present the three bicycle safety-related studies. Chapter 3 centers on the development of crash prediction models as a means of identifying the impact of bicycle treatments on crash occurrence. Chapter 4 presents an analysis of right-hook traffic conflicts between bicyclists and motorized vehicles at signalized intersections where bicycle treatments are present. Chapter 5 refers to a driving simulator experiment, designed to capture and assess driver speeding and glancing behavior in the presence of bicycle treatments while turning right at a signalized intersection. Chapter 6 presents the mode classification framework. These four chapters are organized in (scientific) paper-format meaning that each one of them consists of sections that correspond to introduction, literature review, methodology, results, discussion, and conclusions.

Lastly, Chapter 7 summarizes the findings and contributions of this dissertation. Practical implications as well as future extensions are also discussed.

2 Literature Review

Traffic safety can be objectively quantified as long as one or more of the following data types are available: injuries, crashes, or conflicts [202]. In addition to injuries, crashes, or conflicts, the present literature review includes studies that have emphasized on other measures such as speed, acceleration, deceleration, glancing behavior, vertical or lateral distance between two road users or one road user and a fixed object, to assess traffic safety as these are objective measures that have been associated with unsafe road behavior. Studies that have been excluded from the present literature review are the ones that focus on road user preference and perception, which are subjective measures. For example, bicycle route choice studies offer invaluable input with respect to bicyclists (or potential bicyclists) preferences on road environments that appear safe and convenient, however, from these studies we cannot extract input on countermeasures. Lastly, it should be clarified that the reviewed bicycle safety literature is about bicyclists and motorized vehicles; bicyclists might also be involved in collisions with fixed objects, animals, pedestrians, and other bicyclists however, these collision types have different consequences for bicyclists in terms of injury severity and are subject to different countermeasures.

Regarding the objective approach of studying safety, injury data is essentially a sub-category of crash data as explained in more detail in the following sections, and therefore, traffic safety studies can be grouped in two categories based on the needed data: those that utilize crash records and those that rely on road user interaction to assess safety. Crash-based studies are reviewed in Section 2.1 and non-crash-based studies are reviewed in Section 2.2. The different bicycle infrastructure treatment types are reviewed with respect to their safety impact in Section 2.3.

2.1 Crash-based safety metrics

In the past, the term accident was used to describe the collision (i.e., crash) between two (or more) vehicles or road users. However, it has been argued that this term infers that the event was attributed to chance, while a crash is the output of an array of factors that can be identified and in turn, modified [202]. This concept aligns with Vision Zero principles in stating that traffic fatalities and severe injuries are preventable [257]. Therefore, the term crash and/or collision can be used instead of the term “accident” [202]. When one or more of the involved users is injured, the event can also be referred as injury. For this dissertation the terms collision and crash are used interchangeably and refer to an event that involves (at least one) bicycle and a motorized vehicle, unless it is otherwise specified. The term injury used to refer to a crash between a bicyclist and a motorized vehicles that resulted in the injury of the bicyclist.

Crashes that involve bicyclists and motorized vehicles can be recorded by the police but also from medical professors in hospitals. Compared to crashes between motorized vehicles that are primarily recorded by the police and underreporting is not significant, crashes between bicyclists and motorists may be underreported. Several studies have compared records from (a) hospital admissions related to bicycling crashes and (b) police registrations of such crashes and have found that police records covered 10-50% of the crashes that were recorded in hospitals [52, 57, 195, 237].

The following subsections summarize the research on bicycle crashes that occurred with the involvement of a motorized vehicle. Some studies may be exclusive focus on crashes in which the bicyclist was injured and/or killed, while other studies might include non-injury crashes as well. In addition to differentiate between the injury severity level, the reviewed studies may have as dependent variable the number of

events (e.g., crashes or injuries), the rate of an event, the type of crash in terms of injury severity level, or the probability of the latter. The following subsections have been structured in a way that reflects the different ways that the dependent variable, i.e., crash outcome, has been modeled in the literature and the objective is to identify which is the most effective way to study the impact of bicycle treatments on crash occurrence.

2.1.1 Injury severity

On a worldwide level different scales have been developed in the medical field to assess injury severity. Practitioners and researchers in the traffic safety field are in turn using these scales when it comes to assess the level of injury that was caused to a person that sustained a crash. The scales that most common in practice are: the Abbreviated Injury Scale (AIS), the Injury Severity Score (ISS), and the US-based KABCO Injury Classification Scale [69]. It should be highlighted that AIS and ISS are used by medical professionals to assess injuries caused by a variety of factors, while KABCO is only used by law enforcement professionals in traffic-related injuries. The following paragraphs present these scales, findings from studies that have utilized either scale, and lastly, findings from studies that examined the underlying causes behind severely injured bicyclists.

Methods to measure injury severity

AIS was initially developed during the 70s by the American Association for the Automotive Medicine [177]. This scale is being used on a global level and is defined as: "an anatomically-based, consensus-derived, global severity scoring system that classifies each injury by body region according to its relative importance on a 6 point ordinal scale", [177]. Injury is rated on a scale from 1 to 6, where 1 is minor injury, 2

is a moderate injury, 3 is a serious injury, 4 is a severe injury, 5 is a critical injury, and 6 is a maximum injury. Injuries that fall under the sixth category are the ones that result in a fatality. This scoring is determined for eight body regions namely head, face, neck, thorax, abdomen, spine, and upper and lower extremity. Essentially, AIS includes information on trauma location and extent of injury and then assigns a score between one and six, with one corresponding to the most severe injuries. Since its creation AIS has undergone through several updates with the most recent one finalized in 2015. These updates usually aimed in improving either the content or the way information is either gathered or presented. For example, the 1980 revision acknowledged and incorporated the need to account for consciousness [187].

According to AIS a person with a head injury falls under a different category versus a person with a thorax injury. ISS was developed [14] to allow for these comparisons; it combines the separate AIS ratings for each of the person's injuries into a single measure of overall injury severity. ISS represents the sum of squares of the highest values in each of the three most severely injured body regions. As it is inferred in the previous sentence ISS relies on the AIS scale. Since the maximum AIS score is 6, then the maximum ISS number assigned is 108 (6^2 by 3).

In various parts of the world, research has been conducted to understand how bicyclists involved in a crash are being injured. The majority of these studies has concluded that bicyclists are more likely to hit their head and therefore, helmet use would really reduce the chances of getting severely injured [85, 139, 181, 263]. Apart from head injuries, a recent study that analyzed bicyclist-motorized vehicle crashes from the Netherlands and Sweden found that upper and lower extremities are as likely to be affected after a crash [139].

Both AIS and ISS were developed and are being used by medical professionals. When it comes to traffic safety and related injuries, these scales are used once people

affected by a crash are being hospitalized. This is why these methods are more detailed in evaluating the effects of an injury and thus, understanding its severity. For traffic-related injuries in the US law enforcement professionals (e.g., the police) classify crashes based on their severity using the KABCO Injury Classification Scale [69]. The scale was developed in the US in the late 1960s by the National Safety Council (NSC) [10] and denotes five different categories for injury [47]: K – corresponds to a fatal injury; A – corresponds to an incapacitating injury; B – corresponds to a non-incapacitating injury; C – corresponds to a possible injury; and O – corresponds to Property Damage Only (PDO) or no injury. The KABCO scale defines injury as bodily harm to a person, however several conditions, e.g., stroke or heart attack, are excluded. It should be highlighted that KABCO scale scoring is not applied in a consistent manner among the 50 states, in the sense that there are different definitions per state on what is considered an incapacitating injury for example [34] therefore, severity analysis between different states is not necessarily comparable.

As it was clarified earlier, the most important difference between the two scales is that in practice they are used by different disciplines. KABCO is used by law enforcement officers while AIS/ISS by medical practitioners. Therefore, it is expected the latter is considered as more accurate. Several researchers have elaborated on the differences and similarities of each scale, focusing on crashes between motorized vehicles [23, 34]. Their objective was to understand the relationship between them. As AIS is a scale used by medical practitioners they tend to be more precise when assigning a score to the patient. Additionally, KABCO scoring is given at the site of the injury and the true effects are not known yet. Other disadvantage regarding KABCO is that it overestimates more severe events such as crashes between bicyclists and drivers that have not resulted in neither property damage nor injury are under-reported [173]. However, it should be noted that there is no research relating these

two scales for crashes that involve bicyclists.

Factors that increase injury severity

Regardless the method used to classify bicycle-motorized vehicle crashes, a range of studies has tried to link the occurrence of severe crashes with driver, bicyclist, or road characteristics, weather and environmental conditions by using econometric models, regression models, or Bayesian analysis [4, 5, 64, 114, 122, 126, 163, 233]. Severe crashes are the ones that result in severe injuries or fatalities and are the ones that need to be eliminated according to Vision Zero framework that has been adopted by many cities globally [257]. Therefore, these studies shed light to factors that have been found associated with these crash types.

Factors that are impacting severe crashes can be grouped in the following categories: driver and bicyclist characteristics (e.g., gender, age, alcohol or drug use), motorized traffic characteristics (e.g., percentage of heavy vehicles), roadway (e.g., presence of control at intersections, horizontal curves, intersection or road segment) and environment characteristics (e.g., lighting and weather conditions). It was found that older bicyclists are more likely to be involved in crashes that results in fatalities or severe injuries compared to younger ones [64, 122, 163]. Children are also vulnerable to severe crashes [114]. Alcohol or drug consumption either by the bicyclist or the motorist have been positively associated with higher injury severity levels [114, 163]. Higher volume of turning vehicles as well as higher percentages of heavy vehicles were found to increase injury severity levels [4, 163]. Regarding the road geometry, grades and horizontal curves tend to elevate injury severity levels [163]. Lastly, lighting and weather conditions have been associated with bicyclist injury severity as lack of light and adverse weather increase the frequency of severe crashes [4, 126, 233]. Risk of severe injuries can be lower when bicycle infrastructure is present, the crash has taken

place in urban zones (but not at major roads within these zones), and when traffic calming measures have been implemented such a speed limit of 30 kph [4].

These studies offer valuable insights regarding the factors that increase either the total number of severe crashes at a site or the injury severity levels and have to a certain extent influenced policies that aim to prevent severe crashes; for example, in several countries bicyclists are subject to alcohol tests or they are required to wear helmets and/or have lights during the night time. Only one of the aforementioned studies considered the presence of bicycle treatments [4].

There is limited research on the association of the occurrence of injury crashes and the bicycle treatment types. Three studies overall have focused on this topic, two of them took place in Canada [94, 244] and one was conducted in the U.S. [40]. The U.S. study followed the study design of the Canadian studies in terms of participants recruitment, dependent and independent variables and modeling. The crashes that were included in all of the studies correspond to bicyclists that were injured during bicycling and hospitalized. Therefore, not all crashes were included and specifically, non-injury and well as fatal crashes were excluded. Most importantly though among the included crashes there were cases that the bicyclist was injured without interacting with or trying to avoid a motorized vehicle; e.g., participants fell and injured because of an animal or another bicyclist. Despite these limitations, these three studies are among a very limited number of studies that have considered the impact of the different bicycle treatment types. At road segments, much lower injury risk was found when protected protected bike lanes or traffic diverters are present compared to bike lanes, bike paths, or the use of signage [94]. On the contrary, the U.S. study found that one- or two-way protected bike lanes pose either the same or higher risk for bicyclists compared to major streets with no bicycle treatments [40]. These differences might be related to the different design policies each country is following with respect to bicycle

treatments, e.g., where they choose to implement protected bike lanes or regarding the width of these treatments.

Aldred et al. (2018) were the first ones to consider bicycle and motorized vehicle demand as independent variables in addition to other factors related to user, roadway environment, etc. with an injury severity level outcome. [5]. This addition allows to more effectively compare case and control sites (i.e., sites where a crash has taken place vs a site used for comparison purposes) and in turn, understand the contribution of the rest factors. The need for accounting for traffic and bicycle demand when assessing bicycle crashes is discussed in the following subsection.

2.1.2 Crash rates

Bicyclists may be involved in a crash with a motorized vehicle, which in turn may or may not result in an injury. The previous section reviewed studies that investigated whether crashes of various injury severity levels can be associated with user, roadway environment, traffic, etc., characteristics. One main limitation of these studies is that they provide no information on how likely is for a bicyclist to be involved in such a crash, e.g., a fatal crash, at a specific roadway environment.

Existing research on motorized vehicle crashes has already shown that crash frequency can be misleading as a metric of traffic safety. A highway segment is used by thousands of people on a daily basis, in contrast to a residential street that is probably used only by the people who live in the neighborhood. Therefore, a highway driver is exposed to more cars compared to the residential street driver and is in turn, more likely to be involved in a crash. We might also say that for the highway driver driving is more risky.

Risk expresses the probability of a crash to occur given exposure to potential crash events. Essentially, risk consists of two factors: crash frequency and exposure

to crashes. More broadly, risk of an event is the frequency of that event over the exposure to it. This ratio is known as crash rate. Exposure acts as normalization factor, i.e., denominator, to equalize for differences in the quantity of potential crash events in different road environments [88].

In crash rates the form of the nominator, i.e., crash frequency, is straightforward and it is always in the form of crashes over a time period at a specific location. Exposure has to correspond to the exact same time period and location. As crashes are rare events a reasonable time period for the analysis should be three to five years. More years could also be used provided that major changes in the area of interest have not taken place. The location component of crashes is the one that affects the exposure. Does the analysis focus on crashes in an area, such as a county, a city, or a block? Or does it focus on specific sites such as intersections or segments? The geographical scale of the analysis will determine the exposure metrics as well. Geographical scales can be split in three groups [249]: 1. regional level, that includes an area like a county, city, metropolitan statistical area; 2. block level, that includes a traffic analysis zone, Census tract, or Census block group; 3. site-level that includes specific parts of the road network like segments or intersections. The following subsections discuss the differences among these three categories.

Region-level exposure metrics - Case 1

The scale of geographical resolution defines the data needs to account for exposure. For the first two cases, i.e., regional and block level, travel or mobility data is used as exposure. This is because when we try to answer the question “how likely is for a bicyclist to be involved in a crash in a given area, e.g., a census block” we need to know how much the average bicyclist is exposed in that specific census block. Appropriate exposure metrics for this case express the area population travel patterns and they may

be a) average distance travelled per day by mode (bicycle in this example), average hours travelled per day per mode, c) percentage of people commuting per mode, or d) number of trips taken per mode.

These metrics can be derived from different types of travel surveys. In the U.S. the following travel surveys are used and the resolution is census block or tract:

1. National Household Travel Survey (NHTS)
2. American Community Survey (ACS)
3. Regional household travel survey

Each survey collects different types of information, which may be restricting for the exposure data. For example, ACS does not have information for non-commute trips. There have been cases that more than one sources were combined in order to create a more complete set of information [249], [16].

Travel data as derived from the aforementioned sources allows for estimating crash risk for the entire population of an area. It is more common to deploy distance-related exposure metrics such as number of crashes per 100 kilometres or miles [16, 65, 195]. However, Blaizot et al. (2013) estimated exposure using other metrics such as average number of hours traveled per day as well as average number of trips taken per day. The authors concluded that overall, time-based exposure is more objective in the sense that people tend to spend similar time traveling per day, while miles covered vary depending on the mode [16]. However, bicycle advocates have expressed an opposite opinion [88]; they have argued that people drive their cars for longer distances compared to bicycles but car trips tend to be shorter in time as speeds are higher (assuming non-congested conditions). Consequently, when distance is used as normalizer of crash frequency can show that bicycle is less safe [88]. Overall, there is not a consensus in the literature on which metric better reflects exposure for an area.

Region-level crash (or injury) analysis provided a high-level information on either which areas concentrate more crashes or which user types are more prone to be involved in a crash. They are limited in the sense that they do not indicate which specific parts of a road network, i.e., segments and intersections, should be prioritized for improvement as they are related with higher number of crashes.

Region-level exposure metrics - Case 2

In addition to household travel surveys, other forms of surveys have been developed to retrieve travel data for bicyclists and allow for region-level crash analysis. Though these surveys are similar to household travel surveys in the sense that respondents report their travel patterns, they are different in the following ways. First, respondents choose to participate in a research study, instead of going through a mandatory process initiated by the government (authorized? I am saying that you have by the authorities to participate in HH surveys). Therefore, the objective of the study is to collect data for traffic safety analysis instead of collecting data that could also be used in a traffic safety study. Second, researchers collect travel diaries, i.e., a multi-days series of travel-related information, including distance of travel, trip purpose, chosen routes, and most importantly, the times during the study period that a participant was involved in a crash with a motorized vehicle [51, 98, 193]. It is important to note that the analyzed crashes (or near misses) involve the study participants.

The safety metric used in those studies is number of crashes per 100 km traveled. The detailed list of information allows to study a variety of factors related to bicycle safety. Researchers have route-related information such as bicycle infrastructure treatments that participants tend to encounter during their bicycle trip and so, they can test for relationships between treatments and crashes. Additionally, researchers can obtain participant demographics and assess which population groups are more

vulnerable to crashes.

Instead of travel diaries, technology has allowed for different ways for conducting similar studies. Recruited participants (they are bicyclists) are equipped with video cameras and through them they create logs of their bicycle trips. These studies are analyzed in the surrogate safety section (2.2) as up to date different performance metrics rather than crash rates are used to assess safety.

Overall, assessing safety through travel diaries pertains a main shortcoming. It only represents part of the bicycling population. All these studies recruit adults while teenagers and young kids are also actively cycling. Second, this is a group of relatively frequent bicyclists, that might be more familiar with a route and have adapted to the roadway environment.

Site-level exposure metrics

Site-level crash analysis focuses on either segments, intersections, or corridors. Appropriate exposure metrics should capture the vehicle flows from these locations during the same period of time as the crashes. For motorized vehicle traffic, FHWA proposes two formulas to estimate motorized vehicle crash rates at segments and intersections, respectively [84]. Both ratios have as nominator crash frequency at the given site. Segment exposure is the product of two terms (a) an estimation of the segment's traffic volume and (b) the length of it. Intersection exposure is only the traffic volume. Given that crash frequency is expressed in average number of crashes per year (i.e., 365 days), traffic volume should be expressed in Annual Average Daily Traffic (AADT). Note that for intersections, the denominator of crash rates is the total AADT, i.e., the sum of AADT of all intersecting corridors.

While for motorized vehicle crashes exposure is straightforward, i.e., AADT, this has not been the case for bicycle-motorized vehicle crashes. Exposure for bicycle-

motorized vehicle crashes has taken the form of Average Daily Traffic (ADT) [197], Annual Average Daily Bicycles (AADB) [149] but also volume metrics in periods of time other than the annual average day. For example, hourly volume data has been used to estimate crash rates across segments [159, 206]. In both cases, crash rates are estimated as annual crash frequency over hourly volume and the authors note that this methodology is not ideal, but they chose to follow it as exposure should be considered and they lacked more adequate and accurate bicycle demand estimation. In both studies it is mentioned that hourly volumes should be collected in a consistent manner from each site in order to allow for reliable comparisons. In particular, Minikel (2012) collected bicycle demand data only between 4:00-6:00 PM on weekdays for a period of January to March in Berkeley, CA [159]. Rothenberg et al. (2016) analyzed separately sites for which they had (a) peak hour bicycle volume and (b) sites for which the average hourly bicycle volume was available [206].

As it mentioned earlier, for the case of motorized vehicle crashes that occur at a site understanding what is an appropriate exposure metric is straightforward. motorized vehicles are exposed to motorized vehicles and higher volumes have been related to higher crash risk [146]. It is not so straightforward in the case of bicycle-motorized vehicle crashes. Bicycle demand as the single exposure metric for crashes between bicycle and motorized vehicles is not representing the actual risk as it does not account for motorized vehicle traffic, that correspond to the second component of a bicycle-motorized vehicle crash. Such a crash could not occur in a bicycle trail for example, where automobiles have no access. Therefore, bicycle demand is not a complete exposure metric. Similarly, motorized vehicle demand is not a complete exposure metric on its own either. A segment or an intersection that are parts of popular bicycle routes are more likely to have more bicycle-motorized vehicle crashes as more bicyclists are present. Not accounting for bicycle demand has been reported

as a main problem in before-after studies that test for example whether cycle tracks have improved bicycle safety; several studies concluded that cycle tracks were not safe while what they were missing is the fact that more people used a route where cycle tracks were implemented [246]. Crash rates for bicycle-motorized vehicles should account for both bicycle and automobile demand [76]; otherwise, crash risk is either underestimated or overestimated.

Overall, crash rates stand as a straightforward method for assessing bicycle safety at a site. In contrast to crash frequency, crash rates allow for comparisons between different sites and time periods provided that they have been correctly estimated [76]. However, this method of is associated with several limitations. First, it assumes that there is a linear relationship between number of crashes and traffic volume, but researchers have found that this relationship is not linear [173]. Additionally, crash rates cannot account for different factors that might be associated with crash, e.g. speed limit and number of lanes. Actually FHWA suggests that crash rates method should be used to compare sites with similar volumes and operation (e.g., control type) and design (e.g., number of intersection legs) [167]. The latter note refers to motorized vehicle crashes but is should be interpreted as a general principle, and so considered for the case of bicycle-motorized vehicle crashes.

2.1.3 Crash prediction models

In the traffic safety literature it was found important to develop statistical models that can estimate the number of expected crashes at a site, i.e., road segment or intersection, given a set of predictors. This can be considered as a proactive safety approach in the sense that it can inform road authorities, designers, and road safety practitioners on the potential crash occurrence at a site before any of those crashes took place. Crash prediction models are either stand-alone and multivariate, i.e.,

have been developed for a specific, relatively small area and contain multiple road- and traffic-related parameters, or they are based on the U.S. Highway Safety Manual Predictive method.

In the U.S. there is a specific methodology on how to develop crash prediction models and then, how to account for additional road elements that affect traffic safety. The American Association of Highway Transportation Officials (AASHTO) has developed the Highway Safety Manual (HSM) where Parts C and D are dedicated on the prediction of expected crashes. The models are known as Safety Performance Functions (SPFs) and predict the expected number of crashes based on the site's traffic volume and segment length (for segments only); this process is described in Part C of the manual. SPFs exist for road segments and intersections and describe the base conditions ($SPF_{s_{base}}$). The following equation is the SPF for rural two-lane highways [9]:

$$N_i = e^{\beta_0} AADT_i^{\beta_1} length_i^{\beta_2} \quad (1)$$

where N is the number of crashes per year, $AADT_i$ is the motorized vehicle traffic for segment i in Annual Average Daily Traffic values, $length_i$ is the length of segment i , and $\beta_n X_n$ is a vector of explanatory variables other than traffic volume that might affect the number of crashes. The β values are estimated using regression modeling.

Crash Modification Factors (CMFs) have been developed to allow for the incorporation of additional variables (i.e., more than traffic volume and segment length); the available CMFs are presented in Part D of the manual [9]. CMFs are developed using before-after or case-control analysis and capture the impact of a specific road element on crash frequency (see Equation 2).

$$CMF = \frac{\textit{Predicted crashes WITH treatment}}{\textit{Predicted crashes WITHOUT treatment}} \quad (2)$$

where in each case the number predicted crashes have been estimated through a SPF. The difference between the two SPFs is on the extra term that corresponds to the treatment. If CMF is lower than 1 it means that the treatment has safety benefits for the site, otherwise it increases the crash risk.

In practice CMFs are used to multiple the SPF_{base} so that there is a more detailed prediction of crashes:

$$N_{pred,i} = SPF_{base,i} * CMF_1 * CMF_2 * \dots * CMF_n * C \quad (3)$$

where $N_{pred,i}$ is the number of predicted crashes at site i and is the product of the base SPF developed for the same type of sites (e.g., rural two-lane highway segments), e set of CMFs, and the factor calibration factor C. The later is used to adjust a HSM-based SPF and the respective CMFs to local conditions.

The HSM predictive method can be used for three discrete purposes [171, 231], namely: (1) Network Screening where the objective is to determine the most unsafe locations across a road network, (2) Expected Changes at the Project Level, where different SPFs will be developed for the “base” condition and then for the alternative in an effort to estimate the change in expected crash frequency, and (3) Evaluating the Effect of Engineering Treatments, where before-after studies take place to determine the effect of safety countermeasures that have taken place at a location. While the HSM has multiple SPFs for road segment and junctions types and at the same time there are CMFs for those facilities, the manual has no capability of predicting bicycle crashes.

In addition to the HSM Predictive Method, there are other efforts in the existing

literature to develop mathematical functions that predict crash frequency at a site. These models are different from those in the HSM in several ways. First, the SPF is multivariate instead of having traffic volume variables only. Therefore, the impact of road elements is considered as model variable instead of a factor that multiplies the model. HSM models are developed using data from five States to enhance transferability and generalizability. Existing multivariate crash prediction models are (in most cases) developed for certain states, regions, or cities. In contrast to HSM that provides models and CMFs for motorized vehicles only, multivariate crash prediction models have been developed for bicycle safety. The following subsections present the mathematical framework for crash prediction models and then, a summary of the literature on the existing crash prediction models for bicycle-motorized vehicle crashes.

Count data models

Traditionally, these statistical models fall under the category of count data models however few more frameworks have been proposed in the literature as listed in the review of Lord & Mannering [146]. Counts are non-negative integers and represent the number of occurrences of an event within a fixed period. As such certain regression frameworks, such as linear or logistic regression that have myriads of application are not appropriate for count data. Modeling count data with linear regression would allow the dependent variable to take negative as well as non-integer values which has not physical meaning. Additionally, linear regression assumes that the dependent variable is normally distributed a condition that has found to be false for the case of crash data [115]. Logistic regression is suggested for cases where the dependent variable can only take values between zero and one.

There is another category of models known as “count data models” that are ideal for modeling count data. These models rely on the theory behind a Bernoulli trial

and thus, the binomial distribution. Assuming that there are N independent trials, where in the field of traffic safety “trials” refer to the number of vehicles crossing a site, p denotes the probability of successes, or the probability of crash occurrence in traffic safety. Consecutive Bernoulli trials are assumed to be characterized by the same failure process. The binomial distribution represent the probability model and is given by the following formula 4:

$$P(X = n) = \binom{N}{n} p^n (1 - p)^{N-n} \quad (4)$$

where $n = 1, 2, 3, \dots, N$, and corresponds to the number of successes (or crashes). Binomial distribution has mean and variance are $E(X) = Np$ and $VAR(X) = Np(1 - p)$ respectively.

Crashes are events with low probability to occur, and it is appropriate to consider very large values for the parameter N while at the same time p is small. A Poisson distribution is can better approximate the probability of crashes, but essentially Poisson distribution can be derived from binomial if we set $p = \lambda/N$ where λ is the mean of Poisson distribution.

$$P(X = n) = \binom{N}{n} \frac{\lambda^n}{N^n} (1 - \frac{\lambda}{N})^{N-n} \approx \frac{\lambda^n}{n!} \exp(-\lambda) \quad (5)$$

where, $n = 0, 1, 2, \dots, N$ and λ is the mean of a Poisson distribution.

In a Poisson regression model, the probability of site i having n_i crashes per year is given by:

$$P(n_i) = \frac{\lambda_i^{n_i}}{n_i!} \exp(-\lambda_i) \quad (6)$$

In these models the expected frequency λ_i is specified as a function of the explanatory variables (i.e., predictors). The most common relationship between the

explanatory variables and the Poisson parameter is the log-linear model shown in 7, as it has the ability to produce non-negative values. Given that λ_i expresses the expected frequency, i.e., mean value that fluctuates based on the variance, it is allowed to take non-integer values; it is the number of crashes per site per year that cannot be anything but positive integers or zero due to the physical meaning.

$$\lambda_i = \exp(\beta X_i) \tag{7}$$

where X_i is a vector of explanatory variables and β is a vector of estimable parameters; the latter are estimated by standard Maximum Likelihood methods.

Poisson regression assumes equality of the variance and mean of the dependent variable, a condition that does not hold for all cases when dealing with crash frequency data. It is quite common to have the variance of the data substantially higher than the mean [146], a phenomenon called “overdispersion.” This leads to invalid t-tests of the estimated parameters and therefore, it produces inaccurate models [115]. The restriction regarding the relationship between the variance and mean imposed by the Poisson model can be overcome with the use of Negative Binomial (NB) regression, which allows the variance of the dependent variable to be larger than the mean. The NB model can be derived by simply rewriting Equation 8:

$$\lambda_i = \exp(\beta X_i + \varepsilon_i) \tag{8}$$

where the error term $EXP(\varepsilon_i)$ is a gamma-distributed with mean 1 and variance equal to α^2 . The latter is known as the overdispersion parameter. The addition of this error term in 8 allows variance to be greater than the mean as shown below in equation 9:

$$VAR[n_i] = E[n_i][1 + \alpha E[n_i]] = E[n_i] + \alpha E[n_i]^2 \quad (9)$$

The NB distribution has the form:

$$P(n_i) = \frac{\Gamma(1/\alpha + n_i)}{\Gamma(1/\alpha)n_i!} \left(\frac{1/\alpha}{(1/\alpha) + \lambda_i}\right)^{1/\alpha} \left(\frac{\lambda_i}{1/\alpha + \lambda_i}\right)^{n_i} \quad (10)$$

where $\Gamma(\cdot)$ is a gamma function.

Essentially the relationship between the variance and the mean will determine whether Poisson or NB regression should be used to model crash frequency. In the recent traffic safety literature, there is another category of models that can be well applied in either the Poisson or NB regression. These fairly recent models have been developed to deal with datasets where there is a predominance of zeros, or in other words, datasets that are zero-inflated. Hence, the models have been named Zero-Inflated models and depending on the mean-variance criterion, we may have Zero-Inflated Poisson (ZIP) or Zero-Inflated NB (ZINB) models [146].

Zero-inflated models assume that there are two states that produce zeros in a dataset. Some zeros might be due to lack, while in other cases which actually include the majority of zeros, there is a different underlying process. A common example often given in transportation-related research papers is that of a commuter and their mode choice over a week [262]. Within a week a commuter did not use the transit at all, i.e., $use = 0$. This can be either because they never use transit as they commute with car, or because that specific week they commuted by bike. In the first case we see a generic state while in the second case a random behavior. These are the zero-count and normal count states, respectively [262]. It is essential for each study to be clear whether there are two states related to the zeros in the dataset or not. The following equations present the ZIP (11) and ZINB models:

$$\begin{aligned}
n_i = 0 & \quad \text{with probability} \quad p_i + (1 - p_i) \exp(-\lambda_i) \\
n_i = n & \quad \text{with probability} \quad \frac{(1 - p_i) \exp(-\lambda_i) \lambda_i^{n_i}}{n!}
\end{aligned} \tag{11}$$

$$\begin{aligned}
n_i = 0 & \quad \text{with probability} \quad p_i + (1 - p_i) \left(\frac{1/\alpha}{(1/\alpha) + \lambda_i} \right)^{1/\alpha} \\
n_i = n & \quad \text{with probability} \quad (1 - p_i) \left[\frac{\Gamma(1/\alpha + n)}{\Gamma(1/\alpha) n!} \left(\frac{1/\alpha}{(1/\alpha) + \lambda_i} \right)^{1/\alpha} \left(\frac{\lambda_i}{1/\alpha + \lambda_i} \right)^{n_i} \right]
\end{aligned} \tag{12}$$

Safety Performance Functions and crash prediction models for bicycles

The development of SPF for motorized vehicle-bicycle crashes is rather premature compared to the significant work that has been accomplished to estimate the respective functions for motorized vehicle crashes. Actually there are only one paper [173] that followed the methodology described in HSM and so the developed model can be considered SPF. The rest of the literature includes multivariate crash prediction models that relate crash frequency to several factors.

Nordback et al. created the first bicycle-specific SPF in the US for the intersections in Boulder, Colorado [173]. The proposed model follows a NB distribution and the independent variables are intersection AADT and AADB. This work emphasizes the need to incorporate bicycle volume when estimating crashes and confirms the safety-in-number effect, i.e., when more bicyclists are present then, the per bicyclist crash risk is lower [65]. One limitation of this work is that does not associate crashes with any factor related to the road environment.

On the segment level, Park et al. assess the effect of bike lanes on urban arterials by developing before and after SPF and estimating the respective CMF [183]. Overall it was shown that bike lanes are associated with improved safety for road segments.

A more recent work was also focus on the development of CMF for various roadway components in addition to motorized vehicle and bicycle demand [200].

Other studies have conducted crash analysis by developing Poisson or NB models. For intersection crashes there have been U.S. studies [29, 211], European studies [144, 220], Australian [250], and Korean ones [178]. They all into account various factors besides bicycle and motorized vehicle volumes, such as number and width of traffic lanes, presence of bicycle infrastructure treatments, presence and width of sidewalk, direction of the motorized vehicle traffic, intersection size and number of legs. Some of these studies are dedicated to signalized intersections [200, 249] while Schepers et al. studied intersections without traffic signals [220]. Findings indicate that when speed limit, number of traffic lanes, number of intersection legs, and intersection size increase then more crashes are expected. Similarly, at signalized intersections number of crashes are expected to be higher than the ones with no control. On the contrary, presence of bike lanes was found to decrease the number of crashes.

A main shortcoming of these studies is the lack of detail regarding bicycle infrastructure treatments. They all consider the presence of bicycle infrastructure treatments, either on the segment or the intersection approach, however they do not differentiate on their type, e.g., buffered versus protected bike lanes. Additionally, for the intersection-specific SPFs, other infrastructure treatments, such as bike box or bike signal might impact the intersection safety. Nonetheless such level of detail is ignored.

2.2 Surrogate safety metrics

Surrogate safety metrics (SSM) stand as an alternative approach to the traditional crash-based traffic safety assessment. They were developed to address several shortcomings related to that approach. First, crash-based approach requires several

years of crash data per location in order to get representative crash frequency and draw conclusions regarding traffic safety. Second, crash data cannot be used to assess potential benefits of non-existing facilities, therefore other metrics were needed [82].

It is assumed that there is a relationship between the frequency and severity of traffic events. In this case, the term “traffic event” denotes an interaction between road users. While there is a high number of interactions, only portion of them ends up in crashes. Hyden represented this relationship with a pyramid, as shown in Figure 2.1 [104]. Crashes (or accidents, as denoted in the figure) represent less frequent and less severe events.

There are two noteworthy categories regarding traffic events: 1) traffic conflicts that are defined as “an observable situation in which two or more road users approach each other in time and space to such an extent that there is risk of collision if their movements remain unchanged” [11], and 2) near misses that were defined by Laureshyn as: “situation when two road users unintentionally pass each other with a very small margin, so that the general feeling is that a collision was “near”, [138]. Surrogate safety methods claim that by studying the more frequent traffic events, i.e., traffic conflicts and near misses, one may understand the factors that lead to severe events. There are certain metrics that have been developed to identify traffic conflicts in an objective manner and assess their severity; they are known as Traffic Conflict Indicators and were initially implemented in Sweden in the the 70s [11]. As technology has allowed for automating the process of collecting and processing field data traffic conflict analysis has become more popular.

In addition to traffic conflicts, FHWA denotes several surrogate safety metrics such as STOP bar encroachments, red-light violations, percentage of left turns, speed and deceleration distributions, etc. [82]. All the aforementioned metrics are used in surrogate safety studies as they have the ability to indicate unsafe road user behavior.

This section reviews the literature on surrogate safety emphasizing on the use of objective traffic conflict indicators; the most frequently used ones are compared in terms of benefits and limitations. Additionally, different studies centered on bicycle safety that fall under the category of surrogate safety are presented and the advantages and disadvantages of each study type are discussed.

2.2.1 Traffic conflict indicators

Traffic conflict indicators are metrics that identify conflicts and assess their severity. They have been developed under the umbrella of traffic conflict techniques (TCTs). Essentially, a TCT refers to a set of procedures that have been established to classify traffic events based on their severity. Different countries have developed their own TCTs and have published manuals or handbooks to guide this process. The most popular ones are: the Swedish Traffic Conflict Technique (STCT) [104], Dutch Traffic Conflict Technique (DOCTOR) [130], and U.S. Traffic Conflict Technique (USTCT) [184]. The focus, the observational method, and the way to define a conflict are different for these standard techniques, however, there are similarities among the used indicators and therefore they can be grouped regarding their context.

There are two main groups of TCIs: 1) Time-to-Collision (TTC) and 2) the Post-Encroachment Time (PET) indicators. The order that the indicators are listed corresponds to the frequency they have been appeared in the literature as identified by de Ceunynck [50]. Additionally, there are other metrics that assess the severity of a traffic conflict, however the latter needs first to be identified using TTC or PET, based on the kinematics between the involved road users. These metrics rely on 1) Deceleration and 2) Speed differential evaluation and essentially, associate braking rate or speed with severity [50]. However, Deceleration metrics can also be used in studies that do not focus on user interaction as shown in Strauss et al. [235].

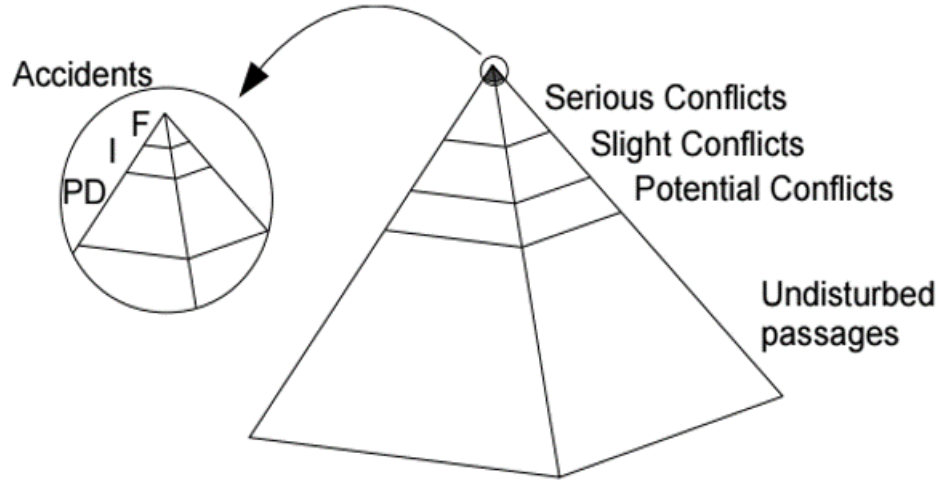


Figure 2.1: Safety pyramid (adopted from Hyden, 1987)

Time-to-Collision

TTC is the time required for two users (i.e., motorized vehicles, bicycles, etc.) to collide if they both maintain their present speed and path [95]. While it is not always clarified in the literature, both the Dutch and the Swedish TCTs share the same definition for TTC [136]. However, there is a difference in the phrasing. In the Swedish traffic conflict technique, TTC corresponds to the time difference between t_0 , which is the moment that an evasive maneuver takes place, and the estimated collision moment, t_1 [104]. The latter is estimated assuming that the two users maintain their speed and path [104], and is the time that the second user would arrive at the collision point, and so, a collision would occur. In the Swedish TCT, the actual name for TTC is Time-to-Accident (TA). The Dutch TCT uses the minimum value of the Time-to-Collision (TTC_{\min}) during an encounter [130–132]. If a collision had occurred, TTC_{\min} would be equal to zero; therefore, we can see that the Dutch TCT relies as well in the existence of an evasive maneuver. After one of the users takes an evasive maneuver they stop being in a collision course, stop approaching each other in time and space, and TTC can no longer be estimated. Therefore, we have again the time

difference between time moments t_0, t_1 .

A minimum TTC during an interaction (TTC_{\min}) of 1.5 seconds or less is considered as critical [137]. In bicycle-motorized vehicle interactions researchers deal with events that found to have a TTC less or equal to 4 seconds [198].

TTC identifies traffic conflict in an objective way, but at the same has the ability to assess their severity. The smaller the TTC value the higher the severity of a potential crash would be. However, this comparison pays no attention to other information of the road user movements, such as the speed. For the same TTC value higher speeds would result in higher crash severity; the proximity to a crash is only one dimension of “severity” [50]. The aspect of speed is captured through the Swedish conflict technique, where TA values are associated with speed in order to estimate (potential) crash severity (Figure 2.2). The other TCTs do not consider speed when assessing the severity of a potential crash and they only rely on the time proximity.

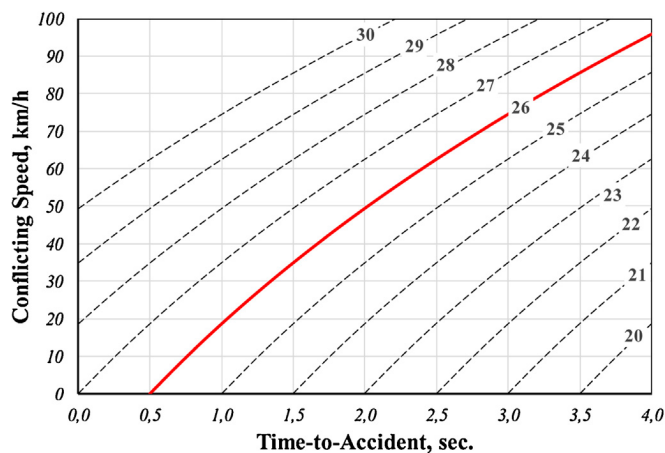


Figure 2.2: Diagram for classifying conflict severity, adopted from [104]

Many other indicators have been derived from the TTC concept. These include, for instance, Time-to-Line crossing, Time-to-Zebra, reciprocal of TTC, T_2 , Time Exposed TTC, and Time Integrated TTC. Most of them have rarely if ever been applied in practice and have not been applied in bicycle-related research [50].

Traditional TTC along with the indicators of the same family, i.e., the ones mentioned in the previous paragraph, assumes that all involved road users maintain their speed and path while being on a collision course. However, this may not be realistic. According to Saunier & Sayed that defined the probabilistic framework: “The collision probability for a given interaction between two road users can be computed at a given instant by summing the collision probability over all possible motions that lead to a collision given the road users’ states.” [218]. For each user and for each instance all possible motion patterns in terms of speed and path are generated; then, it is estimated which of all these points overlap for the different road users. The probabilistic framework has in applied in the bicycle safety research to study behaviour and safety of bicyclists at locations with bicycling infrastructure treatment discontinuities, compared to control sites [172]. The work seems promising and the authors have fully automated all video data processing, however, the paper is mostly focused on identifying the impact of the discontinuities rather than demonstrating the probabilistic framework.

Post-Encroachment Time

Provided that two road users are on a collision course, TTC may be determined. However, if none of the users changes their path or speed and given that they did not crash, TTC estimation is infeasible. A study observed that in a number of crashes no evasive action was present, or at least it was not identified [7], resulting in no TTC estimation. Therefore, there is the need for another metric that has the ability to capture events that either are not on collision course or ones that are on collision course but evasive maneuvers did not identified. For these cases, the proximity in time and space of two users can be captured by another TCI, known as PET.

PET describes the time difference between moment that the first user leaves the

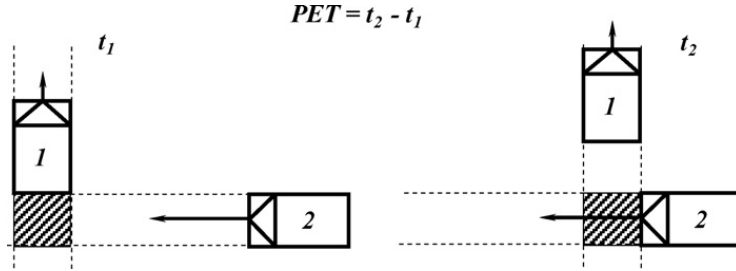


Figure 2.3: Conventional definition of PET [7]

path of the second road user and the moment when the second user reaches the path of the first road user [7]. In other words, PET measures what time margin two road users miss each other [50]. Figure 2.3 illustrates the PET concept. Smaller PET values correspond to more severe events. Some studies suggest to only consider events with PET less of equal than 4 seconds [137] while other analyze events with a PET of 5 or less seconds [129, 274]. The events at the threshold, i.e. 4 or 5 seconds, are considered to be of mild severity. Besides the time proximity there has not been proposed any other criterion to further assess the severity of the conflict.

Along with PET, several similar indicators have been developed. Time-Advance (TAdv) predicts the PET value when all involved road users keep moving on the same path maintaining the same speed [138]. Another indicator is the Time Headway, which expresses the difference between the moment that the front part of the leading vehicle enters the region and the time that the following vehicle enters the same region [258]. A similar concept is expressed through Time Gap, however, this indicator is expressed in distance instead of time. These indicators have been applied to study automobile interactions and these indicators neither have been validated for this vehicle type interactions nor have they been used to study bicycle-automobile interactions. On the other hand, the traditional PET has been validated to relate traffic conflicts with crashes for motorized traffic [7] and bicycle-automobile traffic [137, 274].

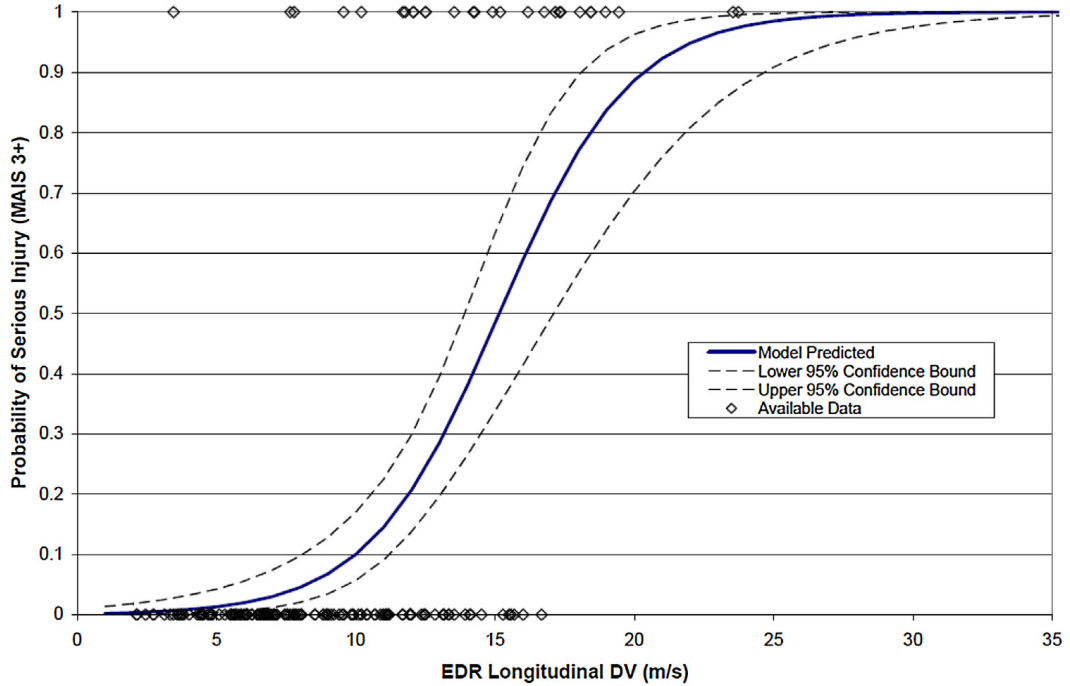


Figure 2.4: Illustration of relationship between Delta-V and probability of a severe injury (adopted from [80]).

Non-time based surrogate safety metrics

Two more categories of surrogate safety metrics have been developed: 1) Speed Differential, or Delta-V as it is usually called, and 2) Deceleration Rate (DR). These metrics are not used to identify traffic conflicts as PET and TTC. Delta-V is used to assess both the severity of an existing crash but also of a potential one, while DR is used to assess (only) the severity of crashes had they occurred.

Delta-V (ΔV) is a notation often used in physics to denote an object's change of velocity. In the context of crashes, Delta-V refers to the change of a velocity vector, i.e., magnitude and the direction of the speed, experienced by a road user during a crash. Essentially, a rapid and large change implies a crash with a high injury severity outcome. The relationship between Delta-V and the probability of a serious injury is visualised by a logistic regression curve (Figure 2.4).

For crashed that have not occurred, Delta-V estimates crash severity assuming

that the road users would collide having the speeds at the moment of the evasive maneuver [136]. This is not so realistic in the sense that before a crash, users are more likely to decelerate and therefore, their speed at the time of the collision are lower compared to the moment they realized they are about to collide. Aiming to depict this reduction, Laureshyn et al. developed the Extended Delta-V [136]. Neither of these metrics has applications that involve bicycle-automobile interactions [50].

DR is used to indicate areas or sites that could be potentially prone to crashes due to frequent and intense deceleration events. Strauss et al. used GPS data to extract bicyclists DR from intersections and segments [236]. Existing injury data for the different sites was found to be correlated with the respective DR. Specifically, the ranks of expected injuries and dangerous decelerations were found to have a Spearman correlation of 0.60 at signalized intersections, 0.53 at non-signalized intersections and 0.57 at segments [236]. Through GPS devices DR can be collected for an entire network relatively fast, although that would be data from users that have smartphone or other GPS-enabled devices and are willing to share it. One limitation of DR as a surrogate safety metric is that it does not necessarily corresponds to user interactions; a user might decelerate for various reasons and not only to avoid another road user.

2.2.2 Categories of surrogate safety studies

(beforehand) of their participation in a research project. It Surrogate safety studies observe interactions between road users and extract various metrics that can be used to quantify safety. Such studies have been conducted in various environments and two main groups may be denoted: 1) controlled and 2) not controlled studies. Studies that happen in a control environment are essentially these that involve simulation. It can be either microsimulation, i.e., a software that simulates the movements of road users, or driving simulator, where humans are exposed to a virtual environment. Under

the bicycle-safety umbrella, in addition to driving simulator studies experiments have utilized bicycle simulator. In this case, participants ride a stationary bike. On the other hand, there are studies that rely on field data that are also known as naturalistic. Two subcategories can be denoted with respect the data collection in naturalistic studies: a) fixed-point data collection, and b) network-wide data collection. In a) a sensor (e.g., a video camera) has been set up at a specific point of the road network, e.g., an intersection, and it records all the road users that appear in its range during the data collection period. Of course, this set up can be repeated to cover multiple points across the road network. In b) data is simultaneously collected from road users that are spread across the road network. This section reviews studies that have been conducted in any of the aforementioned environments and have identified and analyzed traffic conflicts.

Microsimulation studies

Simulation is a resource-efficient way to model traffic flows in a network. Simulation models can be multimodal in nature and are used to model the interactions between different modes on a transportation network. New roadway designs and facilities ranging from roundabouts to traffic signal timing can be evaluated in terms of delay, safety, etc., through simulation before they are actually built.

With respect to traffic conflict analysis several studies have been conducted in microsimulation environment. For automobiles the most commonly used tool is the Surrogate Safety Assessment Model (SSAM) [82]. It is a post-processing module compatible with various microsimulation packages, such as VISSIM, Aimsun, and Paramics. The user has the ability to define threshold values for TTC and PET, or use the predefined one that is 1.5 seconds for both indicators. Events that are found to have TTC or PET values bellow the thresholds are defined as conflicts. Then

considering the exact value reported by the time-based indicator as well as the speed of the vehicle, information can be extracted regarding the severity of each event. This model has not been expanded though to capture bicycle-vehicle interactions.

Up to date bicycle-vehicle interactions have not been studied through microsimulation. There is a considerable effort that tries to create realistic bicycle behavioral models that can be utilized in microsimulation packages [252], however only few of them have been validated with real world data. Actually, simulating bicyclists behavior is hard as these road users have a great flexibility regarding the location of the street that they use; it can be a bike lane and then, the sidewalk, or ride on a bike lane but on the opposite direction [251]. Up to date microsimulation studies that involve bicyclists have only focused on delay estimation [79, 129].

Driving simulator studies

Similar to microsimulation, driving simulator offers the possibility to study various infrastructure designs. In addition to kinematic-related metrics such as speed, acceleration, lane changing, driver simulator allows the use of alternative ways of assessing driver response to various infrastructure designs or stimuli from the roadway environment. For example, mounting an eye-tracking device on driver's heads is a way to get participant eye-gaze data that can in turn be used to assess attention allocation, impact of visual stimuli, etc. Overall, research on validating the technology of the driver simulators has found that driving in high fidelity simulators is similar to driving in the field in terms of speed, acceleration, braking, lateral position, and drivers reaction time [201].

An array of studies aimed to capture driver behavior in response to interactions with bicyclist(s) and/or bicycle infrastructure treatments [37, 53, 61, 75, 83, 93, 108, 127, 204, 261]. All of these studies have implemented various surrogate safety metrics

to assess driver behavior. For example, vehicle-bicycle lateral distance during overpassing events [93], driver speed [53, 127] and placement on the roadway with respect to bicycle infrastructure [37, 75], driver braking patterns when bicyclists are present [59], driver glances at the bicyclist [204] or the bicycle infrastructure treatments [53], etc. Part of the aforementioned studies have assessed driver behavior from the human factor's point of view. This consists of analyzing solely metrics regarding driver attention [204], situational awareness [108], distraction [61], etc. All of the previously listed surrogate safety metrics provide safety-related information and they have the potential to guide policy regarding the implementation and design of bicycle infrastructure treatments. However, their relationship with crashes has not been validated as in the case of traffic conflict indicators such as TTC and PET.

Only the study by Warner et al., conducted a traffic conflict analysis using traffic conflict indicators in a driver simulator environment [261]. In their experiment, drivers were prompted to make right turns in the presence of a variety of treatments such as bicycle signs and intersection-crossing pavement markings. Bicyclists were present riding at a bike lane on the right of the driver. TTC was estimated for the cases that a right-turning driver almost or actually hit a straight-going bicyclist. Results showed that 57% of the time TTC was equal to or less than 1.5 s [261]. In the driver simulator environment conflicting behavior is dynamic only from the driver's perspective. Bicyclists (or in general, other virtual road users) have a consistent behavior as it has been coded. On the one hand, this allows to isolate driver response in a given bicyclist behavior, on the other hand, in real world bicyclist behavior is not known either, and thus more complex.

Naturalistic studies

Naturalistic studies deploy real vehicles and the experiment involves participant driving

these vehicles in real-world (i.e., natural) conditions. Kinematic data, i.e., speed, acceleration, and position, can be collected in a naturalistic study. Two categories can be defined with respect to the area of the study. There are studies that deploy vehicles that travel through the entire road network, e.g. a city or neighborhood, and there are studies that data is collected at one or more specific sites, e.g. an intersection, and record all road users that are present in that site. While both types of studies can capture movements of bicycles and motorized vehicles in real-world conditions, there are substantial differences regarding the data collection process. In the first case, kinematic data is derived via sensors (e.g., GPS devices) that are deployed in the vehicles (either automobiles or bicycles) that study participants are utilizing. In the second case, kinematic data from all vehicles that pass through the study site can be collected. Usually this is feasible with the deployment of one of the following sensors: video camera, RADAR, or LiDAR device mounted in a fixed location.

Network-wide naturalistic studies deploy vehicles that travel throughout an area that is usually a city. In order to collect kinematic data vehicles are equipped with a GPS device. Coupled with other types of equipment, additional data can be collected during a naturalistic study conducted at a network level. For example, participants may be asked to wear an eye tracking device so that their eye movement is captured [121], bicyclists have been asked to wear helmet with a mounted camera to record route infrastructure characteristics [111], or wear devices that record pulses in order to assess stress level [27]. Moreover, as participants are recruited for these studies, socioeconomic and demographic data may also be acquired.

Many aspects regarding bicycle safety have been studied in this context. Researchers have attempted to capture the mechanisms of cars overtaking bicyclists by deploying an instrumented bicycle [58], and driver behavior while making right-turns in streets with bicycle infrastructure treatments [121]. Other studies have estimated non-time

based conflict indicators across a network indicating points with higher deceleration rates via GPS devices [236]. The significance of the latter study lies in the fact that the authors tried to associate locations with high deceleration rates with their existing crash history. The results showed a promising relationship however, they indicated that more experiments should be conducted [236].

These studies provide insightful information regarding drivers behavior in terms of lateral distance, speed, and eye-movement, however, none of them can identify traffic conflicts either with PET or TTC indicators. Network wide naturalistic studies collect data in such a way that does not allow to estimate PET or TTC values. These indicators require kinematic data from both interacting vehicles.

In addition to network-wide naturalistic studies, there are those studies that collect real world traffic data in fixed locations over a period of time. A sensor in a fixed location is used to collect vehicle kinematic data. For the bicycle-automobile interactions only video cameras have been used up to-date.

Various safety surrogate metrics have been used to analyze interactions as recorded from video data. Researchers have estimated drivers-bicyclists rear distance or drivers encroachment to bicycle infrastructure [56, 147], have assessed drivers and bicyclists yielding behavior at unsignalized intersections [228], and have also aimed to capture bicyclists and drivers understanding and compliance with left-turn regulations [26].

Apart from these surrogate metrics, site-based naturalistic studies favor the identification of traffic conflicts between bicyclists and drivers with one of the TCIs method. Many researchers have developed methodologies to accurately extract PET and TTC values and apply these methods to study bicycle-automobile interactions in various roadway designs. Both indicators have been equally used to study users interactions at intersections with the presence of bicycle infrastructure treatments [116, 153, 198, 219, 253, 274].

Some studies focus on a specific site and provide countermeasures for reducing the observed conflicts [198, 219]. The outcomes of this approach are not necessarily transferable to other sites. Also, in these studies data collection duration is only couple of hours, therefore, other different behaviors could have been observed in other time periods. There are studies that emphasize either on the video data processing framework [116] or on the used traffic conflict framework [137]. The latter category does not provide safety assessment of specific bicycle infrastructure treatments.

Only two studies collected data from different sites and aiming to compare different designs or treatments [153, 172, 274]. This is an approach that can inform about the safety benefits of certain infrastructure treatments. Madsen and Lahrmann [153] compared the TTC values from five different intersection treatments at signalized intersections where protected bike lanes were present on the segment. Zangenehpour et al. [274] collected data from 23 signalized intersections with or without protected bike lanes on the segment and used PET to identify and assess conflicts. Both studies emphasized on the importance of traffic and bicycle volumes when studying traffic conflicts and created rates in order to allow for meaningful comparison among different sites.

In traffic conflict analysis studies data collection comes from a very limited amount of time. Actually, Madsen and Lahrmann is the only study found to report data collection duration in days and not in (several) hours. Collecting data either for a longer period or in different time/day might reveal different behaviors, and so result in different number or severity of traffic conflicts. Another concern is that in either study, there is no information regarding the age of each treatment; it is expected that users eventually adjust their behavior to new elements. Therefore, data collection for traffic conflict analysis should take place during similar days and times, e.g., on weekdays during AM peak period and same weather conditions. Additionally, sites of

similar age should be compared with each other in order to avoid behaviors due to unfamiliar users.

2.3 Safety impact of bicycle infrastructure treatments

For site-level analysis, i.e. segments and intersections, multiple studies have been conducted to assess the safety impact of bicycle infrastructure treatments. The vast majority of them utilized crash records, while there are several video-based and simulator-based traffic conflict analysis studies. A study is usually dedicated to either segment- or intersection level bicycle treatments.

Various design exist for both segment- and intersection level bicycle treatments and unfortunately, researchers, practitioners, etc., are not always consistent with the terminology when referring to a certain treatment. Throughout this study, NACTO Urban Bikeway Design Guide [175] terminology and definitions have been adopted.

2.3.1 Segment-level bicycle infrastructure treatments

Definitions

According to NACTO Urban Bikeway Design Guide on the segment there are three categories of treatments: bike lanes, cycle tracks, and bicycle boulevards [175].

A *bike lane* is a portion of the roadway separated by white striping from the traffic lanes. Pavement marking and signage should be accompany bike lanes [175]. There are four types of bike lanes, namely: conventional bike lanes, buffered bike lanes, contra-flow bike lanes, and left-side bike lanes. Except of left-side bike lanes, the other types are on the right of the right-most traffic lane. Colored bike lanes, usually green in the US and Canada, red or blue in European countries, may be used either through the whole length of bike lanes or specific parts of them that are considered critical, e.g., intersection approach. Coloring is considered to enhance bicyclists visibility and

indicate vehicle priority [102]. Conventional bike lanes can commonly be found in three different configurations, as shown in Figure 2.5.

Buffered bike lanes are quite different from the conventional bike lanes as they have a buffer between the bike lane and the right-most traffic lane to ensure that bicyclists and motorists maintain a certain lateral distance between them and therefore, bicyclists are more protected. This type of bike lanes is illustrated in Figure 2.6.

A *cycle track* or *protected bike lane* is a bicycle facility physically separated from motorized traffic and distinct from the sidewalk [175]. While protected bike lanes have different designs, e.g., one or two ways, on the street or sidewalk level, raised medians or on-street parking as separators, etc., they all share common elements: they are separated from traffic and parking lanes as well as the sidewalk. NACTO defines three types of cycle tracks (i.e., protected bike lanes): one-way protected cycle tracks, raised cycle tracks, and two-way cycle tracks; see Figure 2.7. Similar to bike lanes, cycle tracks may be placed on the right or left side of the street and utilize pavement color markings.

A *bicycle boulevard* is a bicycle treatment for the segment-level where both motorized traffic and speeds are low, and bicyclists are expected to use the full traffic lane along with the motorized vehicles. There are signs and pavement markings that indicate that a street is a bicycle boulevard, while at the same time speed and traffic calming measures are implemented to discourage through trips by motorized vehicles.

There is another treatment for the segment-level that enhances mixed-traffic, i.e., allows bicyclists to use the full traffic lane, namely a *sharrow*. In contrast to bicycle boulevards, streets with sharrows do not have traffic calming measures. In terms of pavement marking and accompanied signage, this treatment is similar to bicycle boulevards, however they are not assigned to similar roadway environments. Bicycle boulevards and sharrows are shown in Figure 2.8.

Research findings

Several crash-based studies have compared streets with bike lanes to street without bike lanes, and demonstrated the safety impact of this treatment [93, 128, 197, 244]. Implementing bike lanes can reduce crash frequency [128], crash risk [93], injury severity [244]. The work of Pulugurtha and Thankur (2015) [197] indicates the importance of proper exposure metrics when estimating crash rates. When the exposure metric was the average daily traffic (ADT), streets with bike lanes were not find to differ significantly from streets with bike facilities. On the contrary, when the authors used the length of the street as exposure metric, streets with bike lanes were found to be safer for bicycling. Besides Kondo et al. [128], the other studies do not mentioned the design of the bike lanes that they studied. However, even Kondo et al., consider as bike lanes streets with sharrows [128]. Bike lanes have also been evaluated in driving simulator environment, but the focus was mostly on drivers attention [37, 261]. There are no studies on buffered, contra-flow, or colored bike lanes.

Mixed-traffic facilities mainly include sharrows and bicycle boulevards; the former treatments can be implemented with a variety of pavement markings, for example with or without green-colored squares. There is one study that compared crash rates in bicycle boulevards to near-by arterials were no treatments had been implemented [159]. The findings suggested that bicycle boulevards provide can significantly reduce the crash risk for bicyclists.

In terms of sharrows, existing studies involve a crash-based analysis [92], two driving simulator experiments [75, 93], and several field studies [74, 117, 118, 254]. The crash-based study while included sharrows they were grouped with bike lanes, therefore there is not much to be said on the sharrows safety performance. Both field studies and driving simulator experiments emphasized on the bicyclists and motorists lateral position on the roadway on segments where sharrows had been implemented.

This is because the placements of sharrows aims in placing the bicyclists in the middle of the traffic lanes instead of having them riding in the sidewalk, shoulder, or right side of the traffic lane; the latter placement is particularly dangerous in streets with parking lanes as bicyclists are likely to be involved in dooring crashes. Two of the four field studies took place in Canada and one in California, U.S., and found that bicyclists would shift their position towards the middle of the traffic lane when sharrow where present [74, 117, 118]. The last field study that took place in Norway did not observe this trend [254]. Motorists seem to be reducing their overtaking behavior and instead follow the bicyclists [74, 117, 118]. In the driver simulator settings, driver were found to have lower speed [75] and higher overtaking lateral distance [93] when in streets with sharrows.

Separated facilities such cycle tracks have extensively been studied in the North America context in the recent years, as US and Canada realized to follow the North European example of bicyclists-motorized vehicles separation [196]. The majority of these studies is based on crash records.

2.3.2 Intersection-level bicycle infrastructure treatments

According to NACTO Urban Bikeway Design Guide on the intersection there are seven categories of treatments, namely: bike boxes, intersection crossing markings, two-stage turn queue boxes, median refuge island, through bike lane, combined bike lane/turn lane, and cycle track intersection approach [175].

Definitions

A *bike box* is a designated area at the head of a traffic lane and is connected to a bike lane. Bike boxes are implemented only at signalized intersections and aim to place bicyclists in front of the queuing motorized vehicles during red signal phase.

Therefore, bicyclists visible from drivers and this enhances their safety. This treatment is particularly efficient in eliminating right-hook conflicts and in turn crashes between through-bicyclists and right-turning motorized vehicles. Figure 2.9 illustrates a typical bike box design.

Intersection-crossing pavement markings indicate the area where bicyclists should move while crossing the intersection. Usually they connect bike lanes (or other type of segment treatment) upstream and downstream of the intersection. Crossings may be green colored to enhance drivers visibility. This treatment should not be confused with *through bike lanes* that represent the part of a bike lane that vehicular traffic is supposed to cross in order to change lane, e.g., approach right-lane. Through bike lanes are placed at the intersection approach, while intersection crossing marking are on the intersection.

A *two-stage turn queue box* is an area on the intersection that offer bicyclists a place to wait before making a left turn. They can be placed at both signalized and unsignalized intersections where right side bike lanes or cycle tracks are present on the segment upstream of the intersection. Figure 2.10 illustrates the intersection-crossing pavement markings as well as a turning-queue box.

A *combined lane/turn lane* is a treatment that merges bike lanes with traffic lanes aiming to allow bicyclists to turn. It is a treatment that is placed upstream of an intersection.

Lastly, a *cycle track intersection approach* includes several designs that smoothly merge bicyclists and vehicles at the intersection approach. Depending on the cycle track (or protected bike lanes) design, one may decide upon the appropriate design for cycle track intersection approach. For example, in the case of raised cycle tracks it is common that the intersection approach brings bicyclists to the level of the vehicular traffic.

A *median refuge island* are placed in the center of two-ways streets and allow bikes and pedestrians to cross one direction at a time while crossing the street. They are configured as protected places, i.e., there a physical separation between these islands and the pavement, for example, islands can be build with an elevation.

Some additional treatments are included in the recently published NACTO Guide for intersections [176], that are not included in the previous one [175]. These are: protected intersections, dedicated intersections, and signal phasing strategies.

A *protected intersection*, also known as setback or offset intersection, offers bicyclists a dedicated path through the intersection and have the right-of-the-way over turning motorized traffic. This design incorporates elements such as corner islands and also create a setback between motorized traffic and bike lanes (either protected or conventional) and this results in drivers increased visibility; see Figure 2.11. In cases of busy streets with limited space, *dedicated intersections* can be used instead of protected intersections to accommodate bicyclist crossing movements and reduce conflicts with motorized vehicles. A dedicated intersection has a similar design with the protected one and their difference includes smaller turning radius and bicyclists have a smaller place to wait prior crossing the intersection.

Research findings

With a few exemptions of studies dedicated to roundabouts, there are no crash-based studies assessing the safety impact of intersection infrastructure treatments. Existing studies on intersections consider the placement of bike lanes, which is a segment-level treatment (e.g., cite [211]). Even though, intersections with bike lanes have been found to negatively associated with the number of intersection crashes [29, 211]. Moreover, several treatments, such as two-stages turn queue boxes and median refuge island, have not been evaluated in terms of safety in the literature. This section summarizes

research findings on the safety evaluation of intersection treatments including bike boxes, cycle track intersection approach designs, intersection crossing markings, and finally, bike signals.

Several studies have assessed safety impact of bike boxes in the US [56, 102, 147] and in Europe [8, 205]. Due to the lack of data, crash-based studies could not draw firm conclusions [8, 205]. Therefore, behavioral studies were found to be the most appropriate approach. Studies agree on an overall improvement in safety, in terms of bicyclists visibility [147], reduced number of conflicts between bicyclists and motorists and increased yielding behavior from motorists when encountering bicyclists [56]. Moreover, a study found that green-colored bike boxes led to improvements in bicyclists behavior, in terms of placement [147]. However, there are few shortcomings related to the aforementioned findings. Loskorn et al. [147], focused on motorists encroachment to bike boxes, a behavior that has not been related to crash risk in the literature. Additionally, in their work Dill et al. [56] deployed human observers to identify conflicts, such as braking event or change in direction, rather than using automated techniques. The latter are more accurate and do not suffer from human bias.

Cycle track intersection approach designs [153, 162] as well as intersections with cycle tracks [274] have been studied by video-recorded field data. Zangenehpour et al. [274] concluded that intersections where cycle tracks are present are safer than those without cycle tracks, as lower number of conflicts was identified. Protected intersection design was found to be the most effective in reducing conflicts between drivers and bicyclists at signalized intersection, compared to other cycle track intersection approach designs, such as truncated cycle track [153]. Monsere et al [162]. study was mainly focused to drivers and bicyclists understanding and compliance with various designs, included truncated cycle track and creation of mixed-traffic lane, e.g., whether users

were able to understand the right-of-the-way.

A driving simulator experiment evaluated driver behavior in various intersection designs, i.e., varying curb radius, colored and simple intersection crossing markings, protected intersection, and presence of bike signage [261]. In all the developed scenarios, conventional bike lanes existed on the segment upstream the intersection. The used performance measures were drivers glancing behavior, developed speed, and number of conflicts with (simulated or virtual) bicyclist. Signage was found to increase drivers' situational awareness in the sense that it informed drivers of bicyclists presence, however, there was no effect in observed conflicts. Protected intersection design made drivers to develop lower speeds, which can result in less severe crashes, however, it did not reduce the number of conflicts.

2.4 Summary of literature review

In the recent years there is a growing body of the literature dedicated to bicycle safety research. Attention has been given to the different bicycle treatment types and their safety benefits have been studied mainly using crash-based analysis, although various surrogate safety methods have been used as well. Overall, in the existing literature it is mostly common to compare sites that have a given bicycle treatment to sites that do not have that treatment. This approach allows to understand the safety impact of the studied treatment, however, it does not allow for a comparison between different treatment types. Such a comparison cannot be valid as a different safety approach may have been used (e.g., crash analysis vs traffic conflict analysis), but even when the same safety approach has been used the studies vary in terms of location and incorporation of independent variables. Therefore, the relationship between segment-level treatments such as sharrows, conventional bike lanes and protected bike lanes is unclear. Several studies have attempted to include some different treatment types, e.g.,

conventional bike lanes and sharrows or conventional bike lanes and protected bike lanes. Again, as not all treatments are compared within the same study, it remains unclear which treatment is safer. Similarly, it is unclear how the different intersection-level treatments such as bike boxes, intersection-crossing pavement markings, and protected intersections affect bicycle safety at the intersection.

To that end, it should be added that intersection treatments have received relatively lower attention in the literature compared to ones for the segment-level. This can be partially attributed to the fact that intersection treatment are less frequently implemented, although several studies conclude that this trend should change. There is very limited research on certain intersection treatments such as intersection-crossing pavement markings, bike boxes, and protected intersections.

Existing studies on bicycle treatment types focus on either road segments or intersections; there is no research that has simultaneously studied the impact of different bicycle treatment types both at the segment and the intersection level. For example, protected bike lanes have been found to eliminate crashes at the segment level but more crashes take place at intersections. Up to date, similar analysis has not taken place for other treatment types such as sharrows and conventional bike lanes.

3 Investigating the safety impact of segment- and intersection-level bicycle treatment on bicycle-motorized vehicle crashes

3.1 Introduction

Bicycling is a sustainable alternative to car, yet it has not reached a critical mass as a regular mode of transport in most countries around the world. This is certainly the case for the United States. The implementation of bicycle treatments, such as bike lanes, cycle tracks, bike boxes, etc., can improve bicyclists' safety and comfort level and have the potential to attract more bicyclists and legitimize bicycles as a mode of transport. However, it is evident that crashes still take place at locations that feature bicycle treatments. This could mean that bicycle treatments might not be appropriate for any location. There is a need to understand whether certain bicycle treatment types are more effective in reducing crashes compared to other types and determine the factors affecting the effectiveness of such treatments.

Bicycle treatments are typically present along a roadway segment (e.g., bike lanes, sharrows) or at the intersection (bike boxes, protected intersection elements, etc.). Segment-level treatments are intended to provide space for the bicyclists while they are traveling along the segment; however, they provide no information to bicyclists on how to navigate through an intersection. Intersection-level treatments are implemented to mitigate potential conflicts at the intersection, by providing designated space for bicyclists to wait during the red signal interval, therefore, improving their visibility, or altering the paths of vehicles and bicycles to avoid conflicts. Lastly, more research is needed with respect to intersection-level treatment, e.g., bike boxes, that tend to be omitted [176] as their safety benefits have not been well studied.

The objective of this study is to analyze bicycle-motorized vehicle crashes for

road segments and signalized intersections aiming to understand the impact the various bicycle treatment types have on crash frequency. This is done through the development of two separate crash prediction models for (a) road segments and (b) signalized intersections using data from the City of Portland, Oregon. In particular, this study has focused on protected and conventional bike lanes and sharrows, that are segment-level bicycle treatments, and two intersection-level bicycle treatments, namely bike boxes and intersection-crossing pavement markings. At intersections, the presence of segment-level treatments, e.g., intersections with conventional vs protected bike lanes is also assessed. In the U.S. conventional bike lanes and then, sharrows are the most frequently implemented treatments, a trend that is eventually changes as cities are also implementing protected bike lanes. Intersection-level treatments are not frequently applied [176], however, intersection-crossing pavement markings and bike boxes are quite common.

In both (i.e., segments and signalized intersections), the models express crash frequency as a function of motorized vehicle demand, bicycle demand, and bicycle treatment type. The analysis presented in this chapter is novel and advantageous in several ways. First, it utilizes crowdsource app data as the base for estimating bicycle demand across a road network. Previous studies on bicycle-motorized vehicle crashes and the presence of bicycle treatments have been limited to a few sites across a network as bicycle demand data were only available for those. The use of crowd sourced data allows for expanding the dataset of segments and intersections that can be studied. Second, the impact of bicycle treatment types is considered both at the segment- and the intersection-level. As a result, this analysis can provide guidance to policy makers and transportation planners and engineers on the most effective in terms of reducing vehicle-bicycle crashes, segment- and intersection-level treatments.

The rest of this chapter is organized as follows: First, literature related to the

safety impact of bicycle infrastructure treatments for roadway segments as well as signalized intersections is reviewed and synthesized to highlight existing research gaps (section 3.2). Next, the data sets used for the development of crash prediction models are presented followed by the steps taken to filter these data and develop the mathematical models (section 3.4). Section 3.5 presents the results of this study. A discussion of the results can be found in section 3.6. The final section provides concluding remarks and suggestions for future research (section 3.7).

3.2 Literature review

The reviewed literature has focused on (a) studies investigating the safety impact of bicycle treatments for the segment-level and (b) studies that have assessed the safety of intersection-level bicycle treatments. The reviewed studies have been limited to those relying on crash occurrence analysis. Particular emphasis is given on the segment- and intersection-levels, therefore, reviewed studies have been classified as either “segment-level” or “intersection-level” based on whether they focus on the safety at the segment or at the intersection. Studies that did not specify whether the crashes took place at the segment or the intersection and rather conducted a corridor-level analysis, e.g., [148], were not further considered for this review.

3.2.1 Bicycle safety on road segments

Segment-level treatments appear to improve bicycle safety at the segment, as indicated by the great majority of existing studies [159, 197, 246]. There are multiple studies focusing on sharrows, bicycle boulevards, conventional bike lanes, protected bike lanes, but there are no crash-based studies on buffered bike lanes.

In terms of shared bicycle-motorized vehicle treatments, Minikel (2012) compared bicycle boulevards, i.e., streets with traffic calming measures where bicyclists are

allowed to move as motorized vehicles do, to parallel arterials where no bicycle treatments are present [159]. The findings suggest that bicycle boulevards are associated with two to eight times lower crash rates compared to the arterials. Ferencak and Marshall [71] conducted a before-after analysis of census blocks after the installation of sharrows and observed an increase in bicyclist injury rates. The authors clarify that the increase was observed for all crash types including dooring crashes, that are supposed to be eliminated when sharrows are present. This study design cannot directly assess the impact of sharrows (or any other treatment) as it uses census blocks as the analysis unit, instead of a site (e.g., segment).

The literature on the safety impact of conventional bike lanes is inconclusive. When comparing streets with conventional bike lanes to streets with no bicycle treatment, a study observed a reduction in bicyclist crash risk [197], while another study did not find any significant difference in crash risk [15]. The work of Morisson et al. (2019) tested the effect of different types of conventional bike lanes that are implemented in Melbourne (Australia) and found that bike lanes located between parking lanes and traffic lanes are associated with reduced crash risk compared to all other types (i.e., shared bicycle and parking lane and wide-curb side lane) while curbside bicycle lanes are the least safe [164].

The review of Thomas and DeRobertis concluded that when traveling between intersections, i.e., segment-level, protected bike lanes eliminate the interaction between bicyclists and motorized vehicles [246]. In Toronto (Canada) Ling et al. conducted a before-after analysis to assess the effect of protected bike lanes and found that protected bike lanes reduced crashes at the segment but increased crashes at the intersection [143]. All of the above studies concerned the evaluation of a single bicycle treatment (e.g., sharrows). Limited studies have compared different types of treatments at the segment-level. Two Canadian studies considered multiple treatments,

such as protected bike lanes, multi-use trails, and bike lanes and found that overall injury risk at protected bike lanes is reduced for bicyclists compared to other segment types [94, 244]. A more recent U.S. study followed the same methodology as [94, 244], and concluded that protected bike lanes can be safer than conventional bike lanes or streets with no treatments, depending on the type of separation [41]. More specifically, continuous barriers or grade and horizontal separation was found to reduce collision risk compared to a separation facilitated via parked cars, posts, or low curb separation [41]. However, all three studies focused on hospitalized bicyclists, they also included bicyclists that were not injured during an interaction with motorized vehicles, such as bicyclists that had fallen during their ride or collided with objects and animals. Jensen (2008) conducted a before-after study to assess the safety effects at streets where (a) bike lanes and (b) protected bike lanes were implemented and where no bicycle treatments existed before [110]. This study highlights that both bike lanes and protected bike lanes can improve the safety at the segment-level [110].

Overall, each of the segment-level bicycle treatments has been primarily evaluated in comparison to the no treatment case. However the respective literature is inconclusive and it is overall unclear whether each treatment improves bicycle safety or not. In addition, most of the reviewed studies do not account for both bicycle and motorized vehicle exposure that has been shown to be important [76].

3.2.2 Bicycle safety at signalized intersections

Table 3.1 summarizes the methodology and research findings from existing studies that have analyzed bicycle-motorized vehicle crashes that took place at intersections and considered the presence and/or the type of bicycle treatment type. Studies that did not include signalized intersections (e.g., [220]) were excluded from the table as the focus of this study is on signalized intersections.

Table 3.1: Summary of research crash-based studies assessing the presence and/or type of segment-level bicycle treatments on intersection bicycle safety

Study (Country)	Treatment	Traffic control	Findings
Hamann & Peek-Asa, 2013 [92] (U.S.)	CBL ^a , sharrows	Signalized & Unsignalized	Reduced crash risk at intersections with treatments
Turner et al., 2011 [250] (Australia, N.Zealand)	CBL (various types)	Signalized	The presence of treatments reduced crash risk
Strauss et al., 2013 [234] (Canada)	CBL, PBL ^b	Signalized	Not statistically significant differences between the treatment types
Saad et al., 2019 [211] (U.S.)	not specified	Signalized & Unsignalized	Reduced crash risk at intersections with bike lanes
Liu & Marker, 2020 [144] (Germany)	CBL, PBL	Signalized	CBL reduces crash risk compared to PBL
Chen et al., 2020 [31] (U.S.)	Not specified	Signalized & Unsignalized	Bicycle treatment at the minor road was associated with increased crash risk

^a Conventional bike lane, ^b Protected bike lane

The existing literature is inconclusive with respect to the impact of segment- and intersection-level treatments on bicycle safety at signalized intersections. Findings on the impact of conventional bike lanes are contradicting [31, 211]. Other studies have included sites with different bicycle treatments; however, treatment types were coded as a binary variables, i.e., whether a treatment was present or not, instead of categorical ones [92, 128]. This approach does not allow for a comparison between the different treatment types. From the remaining studies, i.e., that differentiated between the treatment type, there contradicting findings when comparing the safety impact of conventional and protected bike lanes [144, 234]. In a nutshell, further research is required to assess the impact of various bicycle treatment types on bicycle-motorized vehicle crash frequency at signalized intersections.

In addition, no studies to-date have performed a comparative analysis of the impact of various intersection-level bicycle treatments such as bike boxes and intersection pavement markings on intersection safety. The effectiveness of bike boxes in improving safety, have been assessed with the use of surrogate safety metrics rather than crash data [26, 56, 147]. As a result, the impact of bike boxes on crash frequency is still not known. On the contrary crash records have been used to understand the safety benefits of colored intersection-crossing pavement markings [109]; a study found that the number of crossings at an intersection affects the number of crashes and generally, one crossing reduces the number of crashes while more than one crossings increase crash frequency [109].

3.3 Data

Multiple sources are needed for the development of crash prediction models. Crash data, is the dependent variable and exposure data (traffic count and bicycle count data) are the most commonly used independent variables. For this study, bicycle

infrastructure data is used as independent variables, too. Road network data as all data types are matched to particular spatial units; in this study those are road segments and signalized intersections.

3.3.1 Crash data

Crashes for the City of Portland were obtained from Oregon Department of Transportation database, listing all crashes that took place in Oregon between 2012-2019 [180]. Data were filtered to include only bicycle-motorized vehicle crashes for the City of Portland. Crashes are geocoded and associated with a unique ID. In the database, crashes are categorized as intersection or segment level and for the intersection ones, the control type is also reported.

3.3.2 Road network

Road network and speed limit information is publicly available in Geographic Information System (GIS) layers provided by the City of Portland [42]. A GIS layer containing all signalized intersections is available as well and was used in the study. Segment crashes were linked to the closest road segment based on their proximity and name, while intersection crashes were linked to the closest signalized intersection node.

3.3.3 Bicycle exposure data

Bicycle demand was used as an exposure metric, in addition to traffic volume data. Bicycle demand might be collected during various time intervals and could fluctuate over the year. For crash analysis, it needed to be converted to Annual Average Daily Bicycles (AADB) format. Two sources were utilized to achieve this: (a) a sample of segment-based aggregated trips and (b) point-based continuous counters. The combination of both allowed for estimating demand across the majority of the road

network in Portland.

Aggregated segment-based bicycle trips were available through the Ride mobile phone application (app) [189, 203]. The app was initially launched for the City of Portland, with the main objective to collect bicycle count data and bicyclist experience along bicycle routes across the city [191]. When enabled, the app records the user's position throughout their route. At the end of the trip users can rate the route. This rating applies to all of the segments that formed the particular route. Those ratings were not used in this study.

Mobile phone apps allowing collection of crowd sourced data are seen as a new and efficient way of collecting bicycle related information that could not have been possible otherwise. Fitness apps (e.g., STRAVA) can be used to collect bicycle trip characteristic data such as routes and trip duration and have also been used in bicycle safety studies, as a means of bicycle demand [24, 31, 211]. However, fitness apps are limited in that bicyclists using them might be more likely to be riding for recreational purposes and tend to choose routes more suitable this activity. Given that these routes tend to be consisting of off-road paths and trails and in turn, places where there is no interaction with motorized vehicles, using fitness-based crowdsourcing apps might not be ideal for assessing bicycle safety. Overall, there has been no study thus far relying on other than fitness crowdsourced apps for bicycle exposure data.

Bicycle trips per segment per month obtained from the Ride app were provided by the firm. For every segment aggregated trips per month were provided for the period of September 2016-August 2017. The total annual trips per segment were estimated and matched to the actual road network of Portland. As shown in Figure 3.1, Ride data covers a substantial portion of the Portland road network, both in the downtown area as well as in the surroundings of the city. The numbers shown are total annual rides for the 12-month period mentioned above.

Ride app data are plentiful spatially and cover the great majority of the Portland road network (Figure 3.1), but not all bicycling population uses the app. The app offers a sample of the bicycling population that must be adjusted to the AADB level. Fortunately, the City of Portland has placed counters that collect continuous bicycle count data for both short-term and long-term periods; varying from day-, week-, month-, and year-long samples. These data have been archived by the Portland State University (PSU) in an effort to establish a non-motorized counts database [192].

The PSU's database contains count data from 87 locations in Portland for the period of 2013-2016, with 23 counters having at least one full year of counts for the 2014-2016 period. These 23 multi-year continuous counters were used to scale-up the Ride app trips to obtain the AADB values. All the other short-term counts were not sufficient for this scaling process, as they would need adjustment themselves.

It is assumed that Ride trips demonstrate the bicycle demand allocation across Portland. Even if the Ride app data comes from a year, it is assumed that it remained the same during the 2014-2017 period. While factors such as land use, bicycle treatments, etc. have been found to influence bicycle demand and route choice, for the studied time period of four years land use as well as mode share trends are not likely to have changed. In terms of bicycle treatments only a very small part of the bicycle network was modified during this period and thus, it is reasonable to assume that there have not been significant changes with respect to bicyclist route choice. Additionally, it is assumed that this trend has not changed over the past few years, as there has been no change in the bicycle volumes as measured by the continuous counters during the 2014-2017 period.

For some road segments, the actual bicycle demand (i.e., total bicycle trips per year) is known by the continuous counter data. In order to estimate AADB for all network segments, all 23 continuous counters were utilized to develop a scaling model.

Table 3.2: Linear regression model for the the network-wide AADB estimation

	Coef.	Std Error	p-value	Conf. Intervals (95%)
Ride AADB	194.458	13.409	0.000 ***	[167.263, 221.652]
Pop. Density	7.821	2.418	0.003 ***	[2.916, 12.726]

*,**,*** *Statistically significant at the 90%, 95%, and 99% level*

Given the linear relationship between the counter AADB and the Ride app AADB (see Figure 3.2), a linear regression model was estimated to predict the final AADB for each segment based on the existing Ride AADB.

In addition to the Ride app AADB, one more variable was added into the linear regression model, namely the population density of the census tract, where the continuous counter is located. This was necessary as a proxy to bicycle demand given the number of segments with a crash but no Ride trips. Population density was chosen as it is expected that lower density suburban areas will have fewer rides compared to areas with higher population density. Population density was estimated as the total number of people per 1000 per square kilometre for each census tract. The results are presented in Table 3.2.

Both independent variables are statistically significant at the 95% confidence level. The model intercept was removed since it was not statistically significant, an action that improved the model in terms of R^2 (82.4% increased to 88.6%), Akaike information criterion (AIC) (523.7 decreased to 523.3), and Bayesian information criterion (BIC) (528.6 decreased to 526.5). In this implementation, the intercept might also be redundant in the sense that when population density and Ride AADB are zero, then there are no bicyclists. With respect to population density, our assumption that more dense census tracts will concentrate more rides was also confirmed as indicated by the respective coefficient signs. This model was used to predict the AADB for every network link.

3.3.4 Traffic exposure data

Traffic demand was also used as an exposure metric and was available through traffic counts performed by the City of Portland spanning the period from 2014 to 2018 [190]. Each counter collected data during three weekdays and therefore, the data was expressed in Average Daily Traffic (ADT). Additional traffic demand data in Annual Average Daily Traffic (AADT) was available for the City of Portland through ODOT; for road segments that both ADT and AADT data were available the later was selected.

3.3.5 Bicycle treatments

Segment-level bicycle treatments for Portland have been mapped and are publicly available in GIS format. Given that many changes were planned for the Portland bike network for 2018-2019, this study focused on the period between 2014 and 2017. The fact that demand data was available from two different sources from 2014 up to 2017 further motivated the selection of that period.

Only links labeled as "ACTIVE", i.e., a bicycle treatment currently exists, during the 2014-2017 period were included in the analysis. Road segments where changes took place during the 2014-2017 period were filtered out. Active bicycle infrastructure consists of the following treatments: conventional bike lanes (BL), buffered-bike lanes (BBL), enhanced shared roadways (ESR), neighborhood green ways (NG), protected bike lanes (PBL) and, off-street paths or trails (TRL). ESR and NG are marked with the sign of sharrows. Off-street paths and trails were excluded from the analysis due to not being associated with bicycle-motorized vehicle interactions. During 2014-2017, Portland had approximately 260 km of conventional bike lanes, 12 km of buffered bike lanes, 3.3 km of protected bike lanes, 132 km of mixed-traffic segments, and 135 km

of trails.

For the intersection-level treatments, Google Maps imagery was solely used to obtain the treatment type, if any, at the signalized intersections across Portland. Specifically, bike boxes and crossings were identified. Signalized intersections that had been modified in terms of bicycle treatment during the study period were excluded from the analysis.

3.3.6 Final dataset

Data availability determined the crash locations and time period that were used in this study. As mentioned earlier, the Portland bike network experienced several changes during the 2018 and 2019 while at the same time bicycle demand data was available for 2014 to 2017. Since traffic demand data was also available for that period, it was selected as the study period. The final number of segments and signalized intersections with crashes during 2014-2017 was determined by (a) the traffic demand data availability, i.e., locations with no count data were excluded, and (b) the bicycle treatment status, meaning that sites with modified bicycle treatments were excluded. Overall, 188 segments and 179 signalized intersections with crashes were used for this analysis. Additional sites, i.e., road segments and signalized intersections where no crash had taken place during the study period were randomly selected to be used for the development of the crash predictions models. The road segments and signalized intersections used for this analysis in addition to the Ride app segments and the bicycle treatment types are shown in Figure 3.3 and Figure 3.4, respectively.

3.4 Methodology

Crashes have been traditionally modeled using count data models [146]. Approaches that involve least-squares regression are not appropriate as the response variable, i.e.,

the number of crashes observed per site during a given time period, is positive integer and therefore, the functional form of the chosen model should comply with this non-negativity constraint. In practice, crashes are most commonly modeled using Poisson or Negative Binomial distribution models, however, additional modeling approaches exist [146]. Ultimately, the relationship between the mean crash frequency and the variance (across all sites from the sample) can determine whether a Poisson or Negative Binomial distribution is more appropriate; when the mean and the variance are equal, the Poisson distribution is chosen.

When modeling crash frequency with Poisson regression it is assumed that the probability of having y_i crashes at site i during a certain time period is:

$$P(y_i) = \frac{\exp(-\lambda_i)\lambda_i^{y_i}}{y_i!} \quad (13)$$

where $P(y_i)$ is the probability that site i has y_i crashes over a certain time period; λ_i is the Poisson parameter for the i -site and it expresses the average expected crash frequency per year for that site. The expected crash frequency, λ_i , is specified as a function of a set of independent variables. The Poisson regression model is appropriate for datasets where the dependent variable mean value and variance across all observations are equal. The NB regression models have the ability to deal with datasets where the variance is greater than mean.

A Poisson regression model was found to be the appropriate one for modeling crashes at signalized intersections, as it complied with the mean-variance criterion ($0.72 \approx 0.69$); as a result, no further models were considered.

The most common functional form of count data models is $\lambda_i = \exp bX_i$. However, specifically for modeling crash frequency other forms can also be considered to primarily incorporate exposure-related terms into the model and differentiate them from the rest

of the independent variables. For example, Federal Highway Administration (FHWA) proposes the following form for road segment crashes for motorized vehicles [9]:

$$N_i = e^{\beta_0} AADT_i^{\beta_1} length_i^{\beta_2} e^{\sum_{j=3}^n \beta_j X_{ij}} \quad (14)$$

where N_i is the number of crashes at road segment i per a certain time period; $AADT_i$ is Annual Average Daily Traffic in vehicles per day for segment i , $length_i$ is the length of segment i , and $\beta_n X_n$ is a vector of explanatory variables other than traffic volume that might affect the crash occurrence. Road characteristics, such as traffic volume and segment length, that are considered as exposure metrics are not in exponential format. Essentially, in the $\lambda_i = \exp bX_i$ functional form, the natural logarithm of the exposure metrics is used instead of the actual value. This transformation allows to create a relationship where the left-hand side of in Equation (14) becomes zero when either of the exposure terms are equal to zero.

3.4.1 Development of signalized intersection crash prediction models

Similar functional form as the one given by Equation 14 has been used to model bicycle intersection crashes [31, 174, 200, 211, 220]; it is clarified though that these frameworks did not have the same regression model as crashes may follow a different distribution.

For the signalized intersection crashes two functional forms are considered each one for a different model. Essentially, there are two objectives: (1) to understand the impact of the segment-level treatment type and combinations at the intersection area while not accounting for intersection treatments, see Equation (15), and (2) understand the intersection treatment impact by isolating the respective variables, see Equation (16).

$$N_i = e^{\beta_0} AADT_i^{\beta_1} AADB_i^{\beta_2} e^{\sum_{j=3}^n \beta_j X_{ij}} \quad (15)$$

where N_i is the number of crashes at the signalized intersection i during the 2014-2017 period; $AADT_i$ is the total Annual Average Daily Traffic in vehicles per day for the signalized intersection i , $AADB_i$ is the total Annual Average Daily Bicycles in bicycles per day for the signalized intersection i , and $\beta_n X_n$ is the vector of binary explanatory variables that express a signalized intersection where: (1) one of the intersecting roads has a conventional bike lane and no other treatment is present, (2) one of the intersecting roads has a protected bike lane and no other treatment is present, (3) one of the intersecting roads has a sharrow and no other treatment is present, (4) two of the intersecting roads have conventional bike lanes and no other treatment is present, (5) one of the intersecting roads has a conventional bike lanes and the other has a protected bike lane and no other treatment is present, and (6) one of the intersecting roads has a conventional bike lanes and the other has a sharrow and no other treatment is present. It is clarified that in the studied signalized intersections of Portland no other combinations were found, e.g., sharrow and protected bike lanes.

The second model for the signalized intersections has the following functional form:

$$N_i = e^{\beta_0} AADT_i^{\beta_1} AADB_i^{\beta_2} e^{\sum_{j=3}^n \beta_j X_{ij}} \quad (16)$$

where N_i is the number of crashes at the signalized intersection i during the 2014-2017 period; $AADT_i$ is the total Annual Average Daily Traffic in vehicles per day for the signalized intersection i , $AADB_i$ is the total Annual Average Daily Bicycles in bicycles per day for the signalized intersection i , and $\beta_n X_n$ is the vector of binary explanatory variables that express a signalized intersection where: (1) at least one of the approaches has intersection crossing pavement markings and (2) at least one of

the approaches has a bike box.

3.4.2 Development of road segment crash prediction models

For the road segment crashes it was observed that during the studied period the final crash dataset, i.e., after removing segments for which AADT was not available and/or there have been changes with respect to the bicycle treatment type, about 87% percent of the crash segments had one crash; two segments had three crashes and seven segments had two crashes. Therefore, this is a dataset with very low sample mean. In the review of Lord & Mannering, it is highlighted that count data models will produce models with false coefficient values when the sample mean is considerably lower than one [146]. In this case it a common practice in the literature to model crash probability instead of crash frequency. In other words, the developed crash prediction model will predict what is the probability that a segment has a crash or not during the studied period. Binary logistic regression models are ideal for this purpose.

3.5 Results

3.5.1 Crash prediction model for road segments

A logistic regression model was developed to express the crash probability as a function of AADT, AADB, and segment bicycle treatment type, i.e., no treatment, sharrows, conventional bike lane, and protected bike lane. The results are shown in Table 3.3.

Odds ratio (OR) are reported for each one of the independent variables. Values higher than one indicate a positive association between the dependent and the specific independent variable while values lower than one indicate a negative association. However, values close to one indicate a poor association. AADT is positively correlated

Table 3.3: Logistic regression model for the segment crashes

	Odds Ratio (OR)	Std Error	<i>p</i> -value	Conf. Intervals (95%)
Intercept	1.062	0.278	0.828	[0.616, 1.830]
AADT (10^{-3})	1.056	0.017	0.002**	[1.022, 1.092]
AADB	1.001	0.001	0.104	[1.000, 1.002]
¹ Treatment [Conv.]	0.552	0.293	0.043**	[0.311, 0.980]
¹ Treatment [Protected]	0.032	1.085	0.002**	[0.004, 0.272]
¹ Treatment [Shared]	0.211	0.495	0.002**	[0.080, 0.558]

*,**,*** Statistically significant at the 90%, 95%, and 99% level

¹No treatment is the base for the comparison

with the crash probability, meaning that streets with higher AADT are riskier for bicyclists. The different level of the “Treatment” variable are compared against to the no-treatment case and showed that protected bike lanes are safer than conventional bike lanes and sharrows, which are also safer than conventional bike lanes. The only independent variable that was not found statistically significant, although with a positive association to the crash occurrence, is the AADB.

The model reported a pseudo R^2 equal to 47.8%, which is relatively high considering that the response describes crash probability and crashes are affected by various factors such as weather, human factors, safety culture, etc. A chi-squared test was used to assess the model’s Goodness-of-Fit (GoF) and reported a *p*-value that was approximately 0.94. This value is greater than the target 0.05 significance level allowing us to accept the null hypothesis, stating that observed and estimated values are the same.

3.5.2 Crash prediction models for the intersection

For the intersection models, the objective was to assess both the segment-level treatments (e.g., conventional bike lanes) that they are present at least one of the intersecting streets, and also the intersection-level treatments such as bike boxes

and intersection-crossing pavement markings (or simply crossings). The impacts of segment- and intersection-level treatments were assessed developing two different models. First, intersection-level treatments are present only if the segment-level ones exist, therefore an interaction modeling approach would be more appropriate. However, the great majority of Portland bicycle network is made up by conventional bike lanes meaning that there are relatively few intersections with sharrows or protected bike lanes. For example, there is only one intersection with sharrows and a bike box and no intersection with sharrows and crossings, and therefore the developed regression model has limited information to “learn” the case of sharrows and bike box, compared to conventional bike lanes and bike box.

A Poisson regression model was developed considering the following independent variables: a) *AADT*, b) the natural logarithm of *AADB* that are prerequisites for bicycle crash prediction models [31, 173], and then a set of bicycle treatment-related binary variables: c) only-CBL that is equal to 1 if the intersection has a road with conventional bike lanes and no other treatment, and zero otherwise; d) CBL-PBL that is equal to 1 if the intersection has a road with conventional bike lanes and a road with protected bike lanes, and zero otherwise; d) CBL-CBL that is equal to 1 if the intersection has two roads both with conventional bike lanes, and zero otherwise. It is noted that independent variables to represent the sharrow-only, protected bike lane-only, and conventional bike lane and sharrow cases were also in the initial model but were not found statistical significant and were removed. Table 3.5 presents the results.

The motorized traffic and bicycle traffic terms are significant, although *AADT* is significant at the 90% level. Intersections with conventional bike lanes, either at one or both of the intersecting roads, were found to be safer compared to the intersections

Table 3.4: Intersection crash prediction model - Impact of segment-level treatments

	Coefficient	Std Error	p-value	Confidence Intervals (95%)
Intercept	-1.3839	0.269	0.000***	[-1.911, -0.857]
AADT (10^{-3})	0.0104	0.006	0.092*	[-0.002, 0.022]
ln(AADB)	0.0857	0.032	0.007**	[0.023, 0.148]
only-CBL	0.6663	0.161	0.000***	[0.350, 0.983]
CBL-CBL	0.7266	0.268	0.007**	[0.201, 1.252]
CBL-PBL	1.2178	0.382	0.001**	[0.469, 1.966]

*,**,*** Statistically significant at the 90%, 95%, and 99% levels

¹ Only-CBL: a CBL exists at only one road of the intersection

¹ CBL-CBL: CBLs exists at both intersecting roads

¹ CBL-PBL: one intersecting road has a CBL, and one has a PBL

with conventional and protected bike lanes as indicated by the coefficient values. This model indicates that everything else kept equal, signalized intersections where the aforementioned combinations of segment-level treatments are present are more likely to experience bicycle-motorized vehicle crashes. A GoF chi-square test ($stat. = 220.665$, $p\text{-value}=0.987$) confirmed that the data fits the Poisson distribution.

The following CMFs can be developed from the above model, considering as the base case no bicycle treatments at the intersection approach:

- Signalized intersections with CBL: $CMF = 1.94$
- Signalized intersections with CBL-CBL: $CMF = 2.07$
- Signalized intersections with CBL-PBL: $CMF = 3.38$

A second Poisson regression model was fit in the same dataset as the previous model, however, here the considered independent variables were different. In addition to AADT and AADB variables, the binary variables to express the presence or not of bike boxes at a signalized intersection as well the presence or no of intersection-crossing pavement markings were used as explanatory factors for bicycle-motorized vehicle crashes. The following table (Table 3.5) presents the results.

Table 3.5: Intersection crash prediction model - Impact of segment-level treatments

	Coefficient	Std Error	p-value	Confidence Intervals (95%)
Intercept	-1.1155	0.258	0.000***	[-1.620, -0.611]
AADT (10^{-3})	0.0135	0.006	0.025**	[0.002, 0.025]
ln(AADB)	0.0899	0.032	0.005**	[0.027, 0.153]
Crossings	0.3292	0.190	0.082*	[-0.042, 0.701]
Bike Box	0.5677	0.246	0.021**	[0.086, 1.050]

*,**,*** Statistically significant at the 90%, 95%, and 99% level

Traffic volume and bicycle volume are statistically significant and the findings indicate that intersections with higher AADT and AADB will experience more crashes. The findings also show that everything else kept equal, intersections with crossings or bike boxes increase crash frequency. A GoF chi-square test ($stat. = 224.358$, $p\text{-value}=0.988$) confirmed that the data fits the Poisson distribution.

The following CMFs can be developed from the above model, considering as the base case no bicycle treatments at the intersection approach:

- Signalized intersections with Bike Box: $CMF = 1.76$
- Signalized intersections with Crossings: $CMF = 1.39$

3.6 Discussion

3.6.1 Crash prediction model for road segments

The logistic regression model for the segment crashes showed that overall, the presence of bicycle treatments reduces the probability of a crash. This finding supports the implementation of bicycle treatments at segments and routes that bicyclists choose to ride across a network. It would be beneficial for cities to inspect where bicyclists ride and implement treatments in those segments instead of following the “build them and they will come” approach which by the way has been found to work under certain

conditions [45]. For the case of Portland, given that two crowdsourcing apps, i.e., Ride app and STRAVA [31], are used by the bicyclists it is feasible to understand where bicyclists ride and use this information to expand or improve the bicycle network.

In terms of the different treatment types, it was found that conventional bike lanes is the least safe treatment while sharrows come next and finally, protected bike lanes is the safest one. It should be highlighted that generally Portland has implemented sharrows in low-volume streets, mostly in residential neighborhoods, where the speed limit tends to be 20 mph, but also in some downtown locations. During 2014-2017 protected bike lanes were implemented in a small number of segments; specifically, across two downtown corridors and then, at some of the bridges. All of these cases have relatively high bicycle and motorized vehicle demand compared to the rest of the road network. Generally, there is a consistency regarding the segment types where sharrows and protected bike lanes have been placed in Portland. On the contrary, conventional bike lanes exist in segments that vary quite a lot in terms of bicycle and motorized vehicle demand, road classification, and speed limit. Figure 3.5 illustrates the AADT as well as the Ride app trips across the different treatment types for the segments with and without a crash.

Overall, these findings quantify the impact of three bicycle treatment types, namely, protected bike lanes, conventional bike lanes, and sharrows on crash occurrence. Cities that have developed their bicycle network similarly to Portland, i.e., mostly use conventional bike lanes, should consider switching to other treatment types such as protected bike lanes and sharrows.

3.6.2 Crash prediction models for signalized intersections

The findings from the two crash prediction models for signalized intersections should be carefully interpreted. Both models indicate that when everything else is

equal, signalized intersections where certain bicycle treatments (either those for the segment-level or those for the intersection-level treatments) are present will experience a higher crash frequency. Because of the way bicycle crashes are coded and reported by ODOT, it is unclear how the crash took place and at which intersection approach; thus, it is unclear whether a crash can be attributed to the bike box or the protected bike lane. Essentially, these signalized intersections should be prioritized for safety improvements. Additional countermeasures have the potential to assure that bicyclists and motorized vehicles interact safely and in turn, the number of crashes is reduced. However, the countermeasures for each intersection should be selected in a way that serves the specific intersection. Literature from the Netherlands [222] as well recently published guidelines in the U.S. [176] highlight the need for intersection treatments in addition to the segment-level ones. Such countermeasures can be related to control strategies, e.g., bike signals, or more innovative intersection treatments that eliminate the conflicting points between bicyclists and motorists such as the protected intersection [53].

The existing intersection treatments, i.e., bike boxes and crossing markings, are placed to increase driver awareness of bicyclists presence. However, they only have the ability to improve intersection safety with respect certain interactions between bicyclists and motorized vehicles. For example, bike boxes are effective in placing bicyclists in the front of motorized vehicles during the red phase so that bicyclists can remain in the front when the green starts [175]. Hence, bike boxes can be effective for the beginning of the green phase (i.e., right after the red phase) and not necessarily for the later green phase. Additionally, at the studied intersections bike boxes are placed at one or two approaches while crashes might occur in the other ones as well (the same holds for the intersection-crossing pavement markings). This discussion aims to clarify that bike boxes and/or intersection crossing-pavement markings should continue consider as potential countermeasures for signalized intersections; although at the

same time, both treatments might need to be couple with additional countermeasures, e.g., dedicated right-turning phase or bicycle signal phasing that separates bicyclists from motorized vehicles.

3.7 Conclusions and future extensions

This work developed bicycle crash prediction models for road segments and signalized intersections in an effort to associate the presence and type of different bicycle infrastructure treatments with crash frequency. The analysis took place in the City of Portland, Oregon for which a crowdsourcing app data, i.e., Ride app, was available and provided a the base for estimating network-wide bicycle demand. This allowed bicycle demand to be estimated for the great majority of Portland segments and intersections and therefore, the crash analysis was not limited to a few sites. The probability of a crash to occur at a road segment was found to be a function of the segment's AADT, AADB, and bicycle treatment type, i.e., no treatment, sharrows, conventional bike lane or protected bike lane. Then for signalized intersections, two crash prediction models were developed: one to associate crash frequency with the intersection's AADT, AADB, and segment-level treatment type, and the second associated crash frequency with the intersection's AADT, AADB, and intersection-level treatments.

The contribution of this work can be seen in multiple levels. First, it provided a quantification of the safety risk associated with three bicycle treatment types for the segment-level. While there is some evidence in the existing literature that segment-level bicycle treatments reduce bicycle-motorized vehicle crashes at the segment, it is not feasible to compare the impact of the three different bicycle treatments, namely: sharrows, conventional bike lanes, and protected bike lanes. In Portland, Oregon protected bike lanes and then sharrows are associated with lower crash risk compared to conventional bike lanes, although the placement of any treatment was found safer

compared to the no-treatment case. Using the no-treatment case as the base for the model, it was found that sharrows have an OR equal to 0.211, conventional bike lanes have an OR equal to 0.552, and for protected bike lanes OR is equal to 0.032. Second, this work explored the impact of the presence and the type of the aforementioned treatments as well their combinations, e.g., conventional and protected bike lanes, at signalized intersections. Up to date, the treatment presence and type has not been studied both for segments and intersections while exposure in terms of AADT and AADB are incorporated. Signalized intersections with certain treatment combinations, e.g., conventional and protected bike lanes or only conventional bike lanes, were found to be positively related to higher crash frequency. Lastly, the present study also associated the presence of intersection treatments, i.e., intersection-crossing pavement markings and bike boxes, with crash frequency; signalized intersections with either bike boxes and intersection-crossing pavement markings experience a higher number of crashes.

While these results are specific to the City of Portland, Oregon there are several key takeaways that can inform transportation engineers and policy makers with respect to the safety of the bicycle treatment types. At the segment, bicycle treatments should be placed to elevate motorist awareness and designate space for bicyclists. The majority of segment crashes in Portland took place in streets with no treatments therefore, implementing bicycle treatments where bicyclists ride is crucial. Then depending on the AADT and AADB values as well as the presence of other traffic calming measures, a certain treatment should be chosen over another. The current practice, where conventional bike lanes represent the majority of the city's bicycle treatment, appears to be less safe for bicyclists. At intersections, it is shown that the presence of segment-level treatment does not provide safe navigation for bicyclists. Therefore, additional treatments should be considered for the intersections that separate bicyclists

and motorized vehicles in space and time, such as protected intersections and bicycle signals. To understand the appropriate intersection-level treatments, on-site analysis should be conducted (e.g., Road Safety Inspection or traffic conflict analysis), as the existing crash reports contain limited information regarding the crash mechanism, exact location as well as the site's unsafe conditions. Essentially on-site analysis can focus on each intersection approach and specific movements and so, reveal more details on how road users interact and what is the role of the bicycle treatments in those interactions.

Future work should expand the developed models for other cities in the U.S. that are similar to Portland in terms of bicycling culture and road network/ land use, weather, etc. (e.g., Seattle, WA) to explore whether similar trends will be observed. The treatment types should also be explored for rural settings as well as cities and states that do have a strong bicycling culture. Additional treatments should also be studied, such as bicycle signals, merging zones, etc, and it would be a very useful input for engineers to know which combinations of segment- and intersection-level treatments work better for given roadway environments, bicycle and motorized traffic volumes.

In addition to testing the transferability of the developed models, future work should develop Safety Performance Functions (SPFs) as well as the respective Crash Modification Factors (CMFs) to effectively predict the expected number of crashes at road segments and signalized intersections using the HSM methodology. This work developed models and associated the different bicycle treatment type to crash-related outcomes and so, it has demonstrated the importance of differentiating on the bicycle treatment type, however, as the developed models are (a) for one city and (b) multivariate, cannot be considered SPFs as defined in HSM. To this end, crash prediction models need to have the following characteristics to be considered SPFs:

- Be developed for facilities of similar characteristics, i.e., arterials, local roads, three-leg vs four-leg signalized intersections, etc., so that they represent similar base conditions.
- Be developed using data from multiple states to ensure transferability and generalizability of the findings. Data from different states will capture local factors such as weather, bicycle culture, safety culture, road design and other road network characteristics, urban planning characteristics, etc.
- Employ Empirical Bayes theorems to remove regression-to-the-mean bias.

4 Right–hook traffic conflicts between motorists and bicyclists at signalized intersections

In the previous chapter, Chapter 3, crash prediction models were developed to relate bicycle-motorized vehicle crashes at segments and signalized intersections with exposure metrics and the type of different bicycle treatments. Crash analysis is the traditional approach of assessing safety. Analyzing crash data along with other data types that relate to road environment, road users, weather, etc., allows to understand which of those factors contribute to crashes. However, depending on availability (e.g., all crashes have been reported), level of detail (i.e., all information regarding the crash have been completed), and accuracy (i.e., the information regarding the crash is correct, e.g., precise crash location) the final outcome of the crash analysis can be affected.

For the case of bicycle-motorized vehicle crashes, there is a lot evidence that they remain underreported especially when no property damage or injury have occurred [227]. Therefore, bicycle crash data is analyzed, this only includes a portion of the crashes that can be as low as 10% of all crashes [227]. Particularly for the Oregon State and the City of Portland, where the crash analysis in Chapter 3 took place, the crashes are self-reported and the police investigates only crashes that involve the presence of an ambulance [43].

Low crash data availability may limit crash-based analysis as explained below. In the previous chapter, the models developed for signalized intersections need to be carefully interpreted. For the example, bike box presence was associated with higher crash frequency but it is hard to understand the actual reasons for this association. Ideally, crash-based analysis should differentiate based on the crash type, e.g., right-hook crashes and the safety impact of each treatment should be assessed per crash

type. However, there are not many crash records for such detailed analysis.

This Chapter studies the interactions between bicyclists and motorists at signalized intersections emphasizing on a specific interaction type, the right-hook conflicts. Traffic conflict analysis in this case is used to capture unsafe interactions between bicyclists and motorists that may lead to a crash. Interactions are categorized as safe or unsafe based on the Post Encroachment Time (PET), with smaller PET values corresponding to less safe cases as this indicates that a bicyclist and a motorist approach in space and time. Compared to crash records, traffic conflicts occur more frequently and so, certain traffic conflict types (e.g., right-hook conflicts) can be studied to assess the safety impact of a treatment.

4.1 Introduction

Increasing bicycling mode share is gaining popularity as it has been associated with environmental and multiple public health benefits, while also being an effective way to address congestion in highly crowded areas. The implementation of bicycle treatments has the potential to increase bicycle mode share by improving safety and convenience. While there is evidence that bicycle treatments improve bicyclist safety, bicycle-motorized vehicle crashes still occur at locations where treatments are present [54, 246]. This highlights the need to further explore the safety impacts of bicycle treatment types and provide guidance on the most appropriate treatment for different roadway environments and vehicles and bicycle demands.

Crash statistics from the United States (U.S.) suggest that a significant portion of bicycle-motorized vehicle crashes take place at urban intersections [186]. Furthermore, some evidence from countries around the globe like the U.S., Canada, and Germany suggests that signalized intersections are associated with a higher risk for bicyclists compared to unsignalized intersections [103, 144, 235]. A common crash type between

bicyclists and motorized vehicles at intersections is the “right-hook” crash [25, 73, 103], where a right-turning vehicle collides with a through-going bicycle; Figure 4.1 illustrates this collision type. Right-hook conflicts are also very common compared to other conflicting interactions [25]. Right-hook conflicts (and in turn, crashes) occur when right-turns are allowed during the green phase at intersections where bicyclists and motorists coexist. While both right-turning vehicles and through-bicycles are traveling as intended, their paths can cross. Given that the placement of the bicycle with respect to the motorized vehicle affects both the occurrence and severity of right-hook crashes [108, 261], it is critical to study the specifics of the interactions between bicyclists and motorists performing the aforementioned movements in the presence of bicycle treatments.

Segment-level treatments such as protected and conventional bike lanes and sharrows affect bicyclist placement with respect to motorized vehicles. Protected and conventional bike lanes separate bicyclists from motorized vehicles. In the former case bicyclists and motorized vehicles are physically separated via various objects (e.g., bollards, parked vehicles, etc.) while in the case of conventional bike lanes pavement marking is used to indicate where bicyclists should ride. Sharrows on the other hand, allow mixed-traffic conditions and bicyclists can share the exact same road space as motorized vehicles. When protected or conventional bike lanes are implemented, bicyclists are usually on the right of motorists. Due to the physical separation that protected bike lanes offer, drivers might be driving closer to protected bike lanes compared to when driving next to conventional bike lanes. When sharrows are implemented bicycles could be found to the right or left of motorized vehicles at intersections.

In addition to segment-level treatments, it is becoming common for cities to implement intersection-level treatments, such as intersection crossing markings and

bike boxes. Intersection crossing markings indicate how bicyclists should navigate through an intersection and inform drivers about the potential of a bicyclist being present within that space. Intersection crossing markings are placed in both signalized and unsignalized intersections. Bike boxes are a dedicated to bicyclists area located just upstream of signalized intersections where bicyclists can get ahead of the car queue and wait during a red signal phase [87, 147?]. This improves bicyclist visibility and provides some level of priority to bicyclists. Various combinations of intersection- and segment-level treatments exist in the real world; for example, intersection-crossing markings can be combined with protected and conventional bike lanes. While existing research has focused on assessing the safety impact of various intersection-level treatment types, it has not assessed the safety impacts of those treatments when combined with various segment-level treatments.

This study contributes to existing literature by using field data to assess and compare the safety impact of the following bicycle treatment types at signalized intersections: (i) three segment-level treatments, namely protected bike lanes, conventional bike lanes, and sharrows, and (ii) two intersection-level treatments, namely intersection crossing markings and bike boxes. The analysis focuses on right-hook conflicts between through-going bicyclists and right-turning motorists assessed using surrogate safety metrics. The primary objective of this work is to determine whether there is correlation between the frequency of traffic conflicts and the bicycle treatment type while accounting for exposure metrics, in particular right-turning vehicle and through-bicycle volumes. Emphasis is given on user sequence, e.g., a bicyclist followed by a motorized vehicle and vice versa, to assess whether it has an impact on the chosen threshold used to determine existence of a conflict. The following section presents literature on the safety impacts of bicycle treatments. Next, the methodological framework is presented focusing on data collection, traffic conflict definition and extraction,

as well as the choice and development of the appropriate regression models. The models are presented next and the obtained insights and implications for real-world implementations are discussed. The final section includes the conclusions of this study as well as recommendations for future work.

4.2 Literature review

Bicycle treatments can be separated into segment-level (e.g., bike lanes and protected bike lanes) and intersection-level (e.g., bike box and intersection crossing markings). In North America, segment-level treatments can be broadly separated into two categories: those that separate bicyclists and motorized vehicles, e.g., conventional, buffered, and protected bike lanes, and those that do not, such as sharrows and bicycle boulevards [175]. Subsection 4.2.1 focuses on studies that have evaluated segment-level treatments at intersections. Subsection 4.2.2 presents safety-related findings from studies that have focused on bike boxes and intersections crossing markings. Finally, subsection 4.2.3 summarizes the findings from the literature and highlights existing research gaps.

4.2.1 Segment-level treatments at intersections

While an array of studies has assessed the safety impact of segment-level treatments, the focus of this section is to summarize studies that have evaluated the safety impact of these treatments at the intersection level, e.g., intersections where bike lanes (protected or conventional) or sharrows are present just upstream of the intersection. Table ?? summarizes the bicycle treatment types that have been studied, the type of intersection control, and the findings of studies that have assessed the impact of segment-level treatments on intersection bicycle safety. Note that findings are expressed as the outcome that was observed when such treatments were in place (i.e., reduced crash

risk implies a reduction in crash risk when the noted bicycle treatment(s) was in place versus when not unless otherwise specified).

Several aspects were taken into consideration while reviewing the relevant literature; the following decisions were made regarding the inclusion or not of relevant studies. The listed studies (Table 3.1, Chapter 3) are consistent in that they all account for bicycle and motorized vehicle demand as exposure terms; studies that did not account for either type of exposure [128, 164] were excluded. This is because recent findings on bicycle research have shown the need of incorporating both types of exposure in bicycle safety analysis [76, 173]. Additionally, studies that only considered unsignalized intersections (e.g., [220]) were excluded as paper's focus is on signalized intersections. It is also worth noting that due to the fact that these studies were conducted in various countries, there is some inconsistency in the treatment type and configuration. European and Canadian studies consider protected and conventional bike lanes, while the ones in the U.S. mostly include conventional bike lanes and sharrows, as protected bike lanes had not been widely implemented until recently. In Australia and New Zealand, there are several different configurations of conventional bike lanes (e.g., using a dotted versus a continuous line to denote the bike lane) as demonstrated in [164, 250]. All studies except for one [274] used crash records to evaluate bicycle safety; Zangenehpour et al. (2016) used field data of right-hook traffic conflicts between bicyclists and motorized vehicles.

Overall, existing literature is inconclusive with respect to the impact of different treatments on bicycle safety at signalized intersections when accounting for exposure. Even though studies included test sites with different bicycle treatments the developed models only considered the presence of such treatments as a binary variable (i.e., whether a treatment was present or not) [30, 92, 128, 211]. This approach does not allow for a comparison between the different treatment types, e.g., are intersection approaches

with conventional bike lanes versus with sharrows safer? From the remaining studies, i.e., those that differentiated among treatment types, one concluded that there is no difference between the presence of conventional bike lanes versus protected bike lanes [234], while another one found that conventional bike lanes enhance safety compared to protected bike lanes [144]. In addition to the fact that the results of the aforementioned studies do not agree, neither of them considered sharrows in their analysis. As a result, there is no comparative analysis on the safety impacts of conventional bike lanes, protected bike lanes, and sharrows.

4.2.2 Intersection-level treatments at intersections

Several studies have assessed the impact of intersection-level treatment types on bicyclist safety by analyzing historic crash records or field data or through driver simulation experiments. The present section reviews those studies that include intersection crossing markings and/or bike boxes (see Figure 4.2 and Figure 4.3).

Bike boxes is a treatment time that is implemented only at signalized intersections [168]; therefore, the following studies include by default only such intersections. Research findings related to the impact of bike boxes are conflicting. Two different studies have analyzed bicycle-motorized vehicle crash records before and after the installation of bike boxes in Portland, OR [54] and in New Zealand [?]; in Portland, OR there was a 50% increase to the number of crashes in the after period while in New Zealand there was a reduction in the observed crashes. Increased yielding rates to bicyclists from motorists turning right and reduced conflicts (and avoidance maneuvers by bicyclists) have also been reported when bike boxes are implemented in the U.S. [56, 147] and the United Kingdom (UK) [8?]. In one study colored pavement appeared to have a negative impact on the avoidance maneuvers compared to the white outlined bike box [147]. Finally, a driver simulator study concluded that drivers

who are also bicyclists are more likely to behave as intended when encountering a bike box (i.e., stopping behind the stop line and not encroaching on the bike box) [75]. While there is some agreement in the research findings related to the occurrence of non-crash events at intersections where bike boxes are present, it is unclear how “conflicts” and “maneuvers” are defined and detected in the aforementioned studies, therefore, limiting the ability to compare findings across different studies.

Intersection crossing markings indicate the area where bicyclists should move while crossing the intersection. Usually they connect bike lanes upstream and downstream of the intersection [168] and they are placed in both signalized and unsignalized intersections. Crossings may be green colored to enhance drivers visibility.

Research on the impact of intersection crossing markings is limited compared to bike boxes. Intersection crossing markings’ impact on bicyclist safety has been evaluated in various contexts and thus, existing findings cannot be synthesized in a comprehensive and conclusive manner. A driver simulator experiment found that intersection crossing markings designed as white-dotted line markings outperformed the ones that included green-colored crossing markings in terms of increasing drivers’ ability to detect bicyclists while approaching an intersection to turn right [261]. A Danish study correlated the number of intersection approaches per intersection where blue-colored intersection crossing markings had been installed with crash frequency; in particular, intersections with one approach with blue-colored crossings reduces the number of intersection crashes, while more approaches with crossings increase that number [109].

4.2.3 Summary of the literature

Findings that related to the impacts of segment-level treatment on bicyclist safety at the intersection are inconclusive; the impact can be either positive or negative.

Additionally, there is no study up to date assessing all three types of treatments, namely, conventional bike lanes, protected bike lanes, and sharrows. Research on bike boxes does not explicitly show whether these treatments impact crash occurrence, while in cases where non-crash events have been studied the definition of metrics used, e.g., conflict, is not clear or consistent across all studies. Finally, research that simultaneously assesses the impact of segment-level and intersection-level treatments on bicyclist safety is limited.

The great majority of the studies assessing the impact of segment-level treatments that were reviewed rely on crash records while this is the case for a significant number of the studies focusing on the safety impacts of intersection-level treatments. Crashes are rare and random events, therefore, often limiting the ability to perform statistical analysis. The fact that they are rare events is even more apparent in bicycle safety research. Bicycle-motorized vehicle crashes tend to be underreported especially when they do not result in an injury or property damage [52, 57, 195, 237]. As a result, crash analysis might be an ineffective method in the sense that it requires many years of data to establish a representative crash frequency estimate for a site. Data availability is more limiting when a specific crash type is of interest, e.g., right-hook crashes. Consequently, alternatives to crash-based analyses have been developed; these approaches are denoted as surrogate safety methods.

Several studies have used surrogate safety methods to assess bicycle safety at intersections. One area of surrogate safety studies relies on the definition and use of objectively defined and identified safety performance metrics, known as surrogate safety indicators. Such indicators describe how close two or more road users approach in time and space, whether a collision is likely to occur and lastly, what would the injury severity of that collision be. An array of field studies has been conducted with the objective to assess the effectiveness of intersection-level bicycle treatments [153, 219]

and control strategies such as Leading Bicycle Interval [129, 208], as well the effect of discontinuities in the bicycle network on bicycle safety [172]. However, there is no study that differentiates on the different treatment types such as conventional and protected bike lanes and sharrows that simultaneously accounts for intersection-level treatments such as bike boxes and intersection crossing markings.

This section first presents the experimental design in terms of site selection and video data collection. Video data processing to extract relevant interactions between through-bicyclists and right-turning motorists is explained next. Statistical models are then developed to relate the observed traffic conflicts with bicycle and motorized vehicle demand as well as with the bicycle treatment type.

4.2.4 Site selection and video data collection

Video data were collected from ten signalized intersection approaches located in Boston (3), Cambridge (6), and Somerville (1), Massachusetts to investigate and compare the safety impact of the three segment-level and two intersection-level bicycle treatments on right-hook conflicts. Data collection took place in November and October of 2019 for the Cambridge sites and October and November 2020 for the Boston and Somerville sites. Cambridge data were recorded using a GoPro Hero7 camera mounted on a tripod while video data collection for Boston and Somerville was facilitated by cameras provided by Street Simplified, which were mounted on traffic or light poles. At each site the camera was placed to capture the studied approach and, in particular, the area containing potential crossing paths of through-bicyclists and right-turning vehicles.

Table 4.1 provides details on the data collection sites. The column “Period (Hours)” describes the peak period of the day for which data was collected and analyzed (total hours of data collection). The “Segment” and “Intersectio” columns contain

information on the segment and intersection bicycle treatments. The extraction of traffic conflicts (i.e., column “Conflicts”) is explained in the following subsection. The data collection sites are also illustrated in Figures 4.4-4.11. The bike path (whether it is on a protected or conventional bike lane or shared with motorized vehicles) is noted with a yellow arrow. The path of right-turning vehicles is noted with a red arrow. Finally, the light blue rectangular area marks where the traffic conflicts between right-turning vehicles and through-bikes might occur (i.e., where the aforementioned paths are crossing).

In Cambridge, data were collected during weekdays and specifically on Tuesdays, Wednesdays, and Thursdays of October and November 2019 during clear weather conditions (e.g., no snow or rain). For each bicycle treatment type, i.e., sharrows (one intersection), conventional (two intersections) and protected bike lanes (three intersections), data were collected approximately between 8:30-10:30 AM and 5:00-7:00 PM, resulting in a total of about four hours of data per treatment type. The selected intersections have consistency in terms of design: (1) when present conventional and protected bike lanes are located to the right of the traffic lanes, (2) there are no bicycle signals, and (3) turning right on red is not permitted. The latter is important as it prohibits drivers from moving during red and enter the location where bicyclists wait to cross the intersection. Overall, a total of four intersection approaches were observed during the morning peak hours in Cambridge (one with a sharrow, two with protected bike lanes, and one with a conventional bike lane) and three during the evening peak hours (one with a sharrow, one with a conventional bike lane and one with a protected bike lane). Of the seven intersection approaches four had intersection crossing markings as their intersection-level treatment and the rest had none. The decision to select different intersection approaches for the AM and PM data collection stems from the need to ensure sufficient bicycle and car demand was present, allowing for more

interactions between bicyclists and motorists to be observed. These approaches were typically not located at the same intersection as that did not always feature the same bicycle treatments; instead intersection approaches along the same main corridor were considered. The only exception was the intersection that featured sharrows as there was no other intersection in Cambridge (during the data collection period) where a sharrow was present. Lastly, due to very low bicycle demand during the AM period at Binney Street, additional data was collected from the Western Avenue and Memorial Drive intersection on a different day during the AM period.

Video recordings from Boston and Somerville sites were collected during weekdays in November 2020. Morning peak and afternoon peak periods were analyzed for the scope of this study. With respect to bike boxes, data was collected from the sites during both time periods (AM and PM) as it was not always possible to find similar sites and use one for the AM period and one for the PM period. In total, data were collected at three intersection approaches that featured conventional bike lanes upstream the intersection and one with a protected bike lane upstream the intersection. Two of the conventional bike lane sites and one of the protected bike lane ones also presented a bike box at the intersection approach and two of those were in combination with bike boxes. Finally, one intersection approach in Boston presented a combination of conventional bike lanes and a bike box at the intersection.

4.2.5 Traffic conflict extraction

Surrogate safety methods focus on the interactions between two road users, i.e., a right-turning vehicle and a through-bicycle in this case. More specifically, interactions between two road users should align with one of the following definitions in order to be considered as conflicts: “an observable situation in which two or more road

Table 4.1: Data collection sites

Site (City)	Period (Hours)	Segment	Intersection	RT Vehicles ^a	Thru-bikes ^b	Conflicts
Cambridge & Springfield St (Cambridge)	AM, PM (3)	Sharrow	None	186	126	36
Binney & First St (Cambridge)	AM (2)	PBL ^c	Crossings	109	13	3
Binney & Third St (Cambridge)	PM (1.5)	PBL	Crossings	79	36	10
Western Ave & Memorial Dr (Cambridge)	AM (1.5)	PBL	None	122	40	9
Massachusetts Ave & Albany St (Cambridge)	PM (1.5)	CBL ^d	Crossings	162	241	43 ^a
Massachusetts Ave & Sidney St (Cambridge)	AM (2)	CBL	Crossings	114	257	30
Cambridge St & Sudbury St (Boston)	AM, PM (2)	CBL	Bike box	142	41	5
Massachusetts Ave & Beacon St (Boston)	AM, PM (4)	PBL	Bike box & Crossings	514	238	69
Massachusetts Ave & Commonwealth Ave (Boston)	AM, PM (2)	CBL	Crossings	103	171	22
Beacon St & Park St (Somerville)	AM, PM (2)	CBL	Bike box & Crossings	147	60	10

Right-turning vehicle, ^b Through-bikes, ^c Protected bike lane, ^d Conventional bike lane

users approach each other in time and space to such an extent that there is a risk of collision if their movements remain unchanged” [11], or a “situation when two road users unintentionally pass each other with a very small margin, so that the general feeling is that a collision was “near”, [138].

Different time-based indicators have been developed to objectively quantify the proximity aspect of two interacting users; the most commonly used ones are the Time to Collision (TTC) and the Post Encroachment Time (PET) [50]. TTC is appropriate when users are in a collision course, meaning that one user needs to change their path or speed to avoid the collision. Essentially, TTC can be detected only when such action (i.e., change in speed or path or in other words evasive maneuver) is observed and it is estimated as the time difference between the moment of the evasive maneuver until the time one of them would reach the collision point. On the other hand PET, which is defined as the time difference between the moment that the first user leaves the path of the second road user and the moment when the second user reaches the path of the first road user [7], is appropriate for cases where the user paths are crossing (or in other words, are perpendicular) by default and so a user does not aim at changing their path to avoid another one. Since in the present context of right-hook conflicts right-turning vehicles cross paths with through-going bicyclists, PET is the appropriate as it can be estimated without the existence of evasive maneuvers. Figure 4.12 graphically illustrates the definition of the PET indicator.

Surrogate safety methods and in particular, the time-based indicators have a proven association with traffic safety. As summarized in the review of Johnsson et al. (2018), eight studies have correlated the number of observed traffic conflicts that have been identified using the PET definition with crash occurrence [112]. In bicycle safety research several studies have used PET to assess the impact of bicycle treatments on right-hook conflicts between bicyclists and right-turning vehicles [129, 208, 274].

The recorded videos were reviewed to manually extract the following information. A 15-minute interval was used as the time unit for the number of: (1) right-turning motorized vehicles, (2) through-bicyclists, (3) traffic conflicts, at each intersection approach. This interval was considered appropriate since it is usually used in volume studies. Smaller intervals such as 5 min. would have greater variability to the number of conflicts and recorded volumes depending on the occurrence of red phase per 5 min. As mentioned earlier, traffic conflicts are identified using the PET as the surrogate safety indicator. The number of traffic conflicts was further grouped by (a) PET value (i.e., 1, 2, 3, and 4 seconds), and (b) the road user sequence in terms of who is arriving first at the conflict area, i.e., a bicycle arrives first and is followed by a motorized vehicle or vice versa.

For each site the conflict area was defined as the area where the through-bicycle and right-turning vehicle paths were crossing. This area is illustrated with a blue rectangle in Figures 4.4-4.11. The total number of observed traffic conflicts during the data collection period were grouped per treatment type. The total number of detected traffic conflicts along with the respective volumes are shown in Table 4.1.

Smaller PET values indicate a closer chance of collision in the sense that users have approached each other closer in time and space. This is because a slight increase in the second user's speed (see Figure 4.12) would result in collision. Different time thresholds have been proposed in the literature to (a) define which interactions are severe enough to be considered as traffic conflicts, and (b) differentiate these interactions to severe and less severe ones. Some studies suggest to only consider events with PET values lower or equal to 4 seconds [137] while others analyzed events with a PET threshold of 5 seconds [129, 274]. Events reporting a PET value equal to the threshold, i.e., 4 or 5 seconds, are considered to be of mild severity. For this study, very few interactions were observed where PET was equal to 5 seconds. In addition, video analysis revealed

that these interactions did not appear to be unsafe. As a result, the upper PET value used for the study was 4 seconds.

User sequence was also obtained for each PET value smaller than 4 seconds. In particular, conflicts that occurred when a bicyclist was the first user or the second user, i.e., the first or the second one to arrive at the conflict area, were counted separately. The focus on the user sequence might reveal some further information with respect to the user behavior and allow for assessing whether bicyclists and motorists have different preferences with respect to the gap they leave between themselves and the leading vehicle.

4.2.6 Model formulation

The objective of this study is to correlate the number of conflicts per 15 minutes with the types of bicycle treatments that are present (both at the segment- and the intersection-levels), as well as the right-turning motorized vehicle and through-bicycle volume for the respective time period.

The dependent variable, i.e., the number of traffic conflicts per 15 minutes, is a positive integer and conflicts are random events. As a result, count data models, which can model discrete outcomes, are the appropriate family of models to consider [146]. In traffic safety literature, count data models have been used to model crash frequency data, however, more recently, they have also been used to model traffic conflicts for motorized vehicles [67], between bicycles or pedestrians and motorized vehicles [56, 112, 129, 208].

The Poisson distribution is appropriate for a set of observations where its mean and variance are approximately equal [146]; for the observed traffic conflicts this relation holds although variance is slightly higher than the mean (the mean and variance are 2.69 and 3.56 respectively). On the other hand, the Negative Binomial (NB)

distribution is flexible, since it can represent observations where the variance exceeds the mean. Poisson, is a subcategory of NB, when the error term has zero variance. N

According to the NB distribution, the average expected number of events λ_i (e.g., traffic conflicts) is given by the following equation:

$$\lambda_i = \exp(\beta X_i + \varepsilon_i) \quad (17)$$

where X_i is a vector of explanatory variables for the i^{th} -observation and β is a vector of estimable parameters. The term ε_i is the error term that follows the gamma distribution with *mean* = 1 and *variance* = α , where α is the dispersion parameter. The addition of the gamma-distributed error term allows the observations' variance to be greater than the mean; the physical meaning is that some sites experience quite higher or lower events (e.g., traffic conflicts or crashes) compared to the mean across all sites. Equation 17 represents the average expected frequency of events given by a Poisson distribution.

The NB probability distribution has the following form as determined by Long ((year?):

$$P(y_i|X_i) = \frac{\Gamma(1/\alpha + y_i)}{\Gamma(1/\alpha)y_i!} \left(\frac{1/\alpha}{(1/\alpha) + \lambda_i}\right)^{1/\alpha} \left(\frac{\lambda_i}{1/\alpha + \lambda_i}\right)^{y_i} \quad (18)$$

where y_i is the number of events (e.g., traffic conflicts) for the i^{th} -observation, $\Gamma(\cdot)$ is a gamma function, and α is the dispersion parameter.

Finally the variance of the NB probability distribution is given by:

$$Var(y_i|X_i) = \lambda_i + \frac{\lambda_i^2}{1/\alpha} \quad (19)$$

In this study given that the mean and variance are close, both Poisson and NB distributions were considered to model traffic conflict frequency. The final selection

of the distribution relies on statistical criteria; essentially, the objective is to keep the models that fit the data better. The Akaike Information Criterion (AIC) and Bayesian Information Criterion (BIC) were estimated for each model. The model that showed the lower AIC and BIC values was the one presenting a better fit for the available data. The parameters of the developed count data models were estimated by maximizing the log-likelihood function. All of the analyses were conducted by using the *statsmodel* module of the Python programming language [223].

As mentioned earlier, the recorded traffic conflicts were also categorized based on different PET values, and the road user sequence, i.e., whether a bicyclist was followed by a motorist or vice versa. Different time thresholds, e.g., PET of one versus two seconds, correspond to a higher probability of collision. These data collection would allow for a better understanding of the frequency and type of more and less severe conflicts.

4.3 Results

4.3.1 Traffic conflict model

The first model that was estimated was the “base model” which relates the number of traffic conflicts per 15 minutes to the exposure terms, i.e., right-turning motorized vehicles and through-bicycles, and excludes any other independent variable. Note that the natural logarithm of each exposure term is used for the model instead of the actual count. This transformation allows us to model the following relationship between the dependent variable and the exposure terms: when either of the exposure terms is zero, then the dependent variable is zero as well. The traffic conflicts per 15 min are given by the following equation (20):

Table 4.2: Base model

	Coefficient	Std Error	<i>p</i>-value	Confidence Intervals (95%)
Intercept	-3.5922	0.556	0.000***	[-4.682, -2.503]
Right-Turning Veh.	0.9084	0.171	0.000***	[0.573, 1.244]
Through-Bicyclists	0.7057	0.088	0.000***	[0.533, 0.879]

*** Statistically significant 99% confidence level

$$N_i = RT_i^{\beta_1} TB_i^{\beta_2} e^{\beta_0} \quad (20)$$

where N is the number of conflicts per 15 minutes observed during the i^{th} interval, RT_i is the number of right-turning motorized vehicles during the i^{th} interval, TB_i is the number of through-bicycles observed during the same interval, and X_4 is the nominal variable for the bicycle treatment type.

The model as defined by Equation 20 was estimated by fitting the Poisson and NB distributions. The Poisson distribution was found to have lower AIC and BIC values ($AIC_{Poisson} = 268$ and $BIC_{Poisson} = 275$) compared to the NB one ($AIC_{NB} = 269$ and $BIC_{NB} = 279$), meaning that the Poisson distribution is more appropriate in terms of fitting. The model specifications of the Poisson model are shown in Table 4.2.

The base model reveals that both the number of right-turning motorized vehicles and the number of through-bicycles are significantly and positively associated with the number of conflicts at signalized intersections. These findings align with existing research on bicycle-motorized vehicle collisions at signalized intersections [173]. The developed model was assessed in terms of goodness-of-fit (GoF) using the chi-square statistical test. The GoF results revealed that the Poisson distribution fits the conflict data with a chi-square value of 38.53 and a p -value of 0.273.

The base model was then extended to consider the variables for the bicycle treatment presence. Specifically, the variables indicate whether a treatment is present

in the studied intersection approach or not: the segment-level treatment type was treated as nominal variable with three levels (i.e., CBL, PBL, and sharrows); the variable Crossings indicates whether there are intersection crossing markings, and finally the Bike Box variable indicates whether a bike box is present. Both bike boxes and intersection crossing markings can be present at an approach. The model form (Equation 21) and specifications (Table 4.3) are presented below:

$$N_i = RT_i^{\beta_1} TB_i^{\beta_2} e^{\beta_0 + \beta_3 CBL + \beta_4 PBL + \beta_5 Crossings + \beta_6 BikeBox} \quad (21)$$

where N is the number of conflicts per 15 minutes observed during the i^{th} interval, RT_i is the number of right-turning motorized vehicles during the i^{th} interval, TB_i is the number of through-bicyclists observed during the same interval, and CBL, PBL, Crossings, and BikeBox are the variables for the various bicycle treatment types.

Table 4.3: Traffic conflicts model with bicycle treatment type

	Coefficient	Std Error	p-value	Confidence Intervals (95%)
Intercept	-3.0417	0.622	0.000***	[-4.260, -1.824]
Right-Turning Veh.	0.8484	0.224	0.000***	[0.409, 1.288]
Through-Bicyclists	0.6927	0.137	0.000***	[0.424, 0.962]
CBL	-0.7046	0.371	0.057*	[-1.431, 0.022]
PBL	-0.5342	0.402	0.184	[-1.322, 0.254]
Crossings	0.3851	0.308	0.210	[-0.218, 0.988]
Bike Box	-0.1313	0.166	0.429	[-0.457, 0.194]

***, ** Statisticaly significant at the 90%, 95%, and 99% confidence level

The exposure variables (i.e., right-turning vehicles and through-bicycles) and the constant are statistically significant at the 99% confidence level. However, the treatment variables, are not statistically significant at the 95% significant level or higher. This finding suggests that the studied bicycle treatment types, at the segment or the intersection, do not affect the frequency of right-hook conflicts. However, the chi-square test results show that this model fits the Poisson distribution well ($\chi^2 =$

37.49, p -value = 0.231). The only treatment that appears to have an impact, although at the 90% confidence level is the conventional bike lane. Compared to sharrows, this segment-level bicycle treatment reduces right-hook conflicts.

Conflict rates (Equation 22) were estimated for every 15 minutes interval and then, grouped by segment treatment type.

$$C_{Ri} = \frac{100 * Conflicts}{TB_i * RT_i} \quad (22)$$

where C_{Ri} is the conflict rate estimated for the i^{th} interval, RT_i is the number of right-turning motorized vehicles during the i^{th} interval, and TB_i is the number of through-bicycles observed during the same interval.

Figure 4.13 shows three violin plots displaying the conflict rates per segment-level bicycle treatment. Violin plots summarize information in a succinct manner and show the probability density of the data at different values. These violin plots reveal that lower conflict rates are associated with intersections where conventional bike lanes are present compared to intersections with protected bike lanes and sharrows.

4.3.2 User sequence and PET values

This part of the analysis examined the observed right-hook conflicts in terms of road user sequence. There are two potential user sequences: a bicyclist is user 1 and the motorist is user 2 and so, the bicycle is followed by the motorized vehicle, or the opposite. As noted earlier, while the observed conflicts were recorded, the user sequence was recorded as well. While analyzing the data it became apparent that bicyclists tend to have smaller PET values when follow motorized vehicles compared to when they are being followed.

For the two different user sequences a heatmap was created to illustrate how PET

values vary depending on the user sequence; see Figure 4.14. Each cell on the heatmap corresponds to the percentage of traffic conflicts per PET value and site over the total number of number of conflicts.

The heatmaps reveal that right-hook conflicts where a bicyclist is followed by a motorist tend to have PET values of 2 or 3 seconds, while when motorists are followed by bicyclists, PET values are more likely to be equal to 1 second. Further statistical analysis was conducted to test whether there are statistically significant differences between the occurrence of conflicts at the various PET values and the user sequence. The data per group, i.e., per PET value and per user sequence, are not normally distributed, according to the Shapiro-Wilk test [226]. Therefore, the Kruskal–Wallis test [134] was used to assess whether the two user sequence groups per each PET value, i.e., 1, 2, and 3 seconds, are significantly different. The results of the Kruskal-Wallis test are presented in Table 4.4.

Table 4.4: Kruskal-Wallis test results for different PET values and user sequence

PET value (sec)	Statistic	<i>p</i>-value
PET = 1	3.871	0.049**
PET = 2	13.055	0.000***
PET = 3	12.151	0.000***
All PET	5.436	0.020**

** *Statistically significant at the 95% level*
*** *Statistically significant at the 99% level*

Findings from this analysis showed that there is a statistically significant difference between the reported PET values between the two different groups: conflicts where a bicyclist was followed by a motorized vehicle or was following one. Bicyclists tend to maintain smaller distance between themselves and the vehicle in their front that is turning right, while motorists maintain a relatively larger distance. This finding could impact on the way PET is recorded.

4.4 Discussion

Overall, the bicycle treatment type does not appear to have a significant impact on the frequency of right-hook conflicts. Conventional bike lanes showed promising results for improving safety compared to sharrows, but more research is needed to conclude whether they are indeed capable of reducing the frequency of right-hook conflicts. With respect to the other treatment types, a few considerations are listed below in an effort to explain their lack of impact on reducing right-hook conflicts between bicyclists and motorists.

Intersection-level treatments, such as bike boxes and intersection crossing markings, can indeed be beneficial for improving bicyclist safety as concluded by previous studies [56, 75, 147, 261]. However, their safety impact is not necessarily related to reducing right-hook conflicts and it is reasonable to infer that additional countermeasures are needed to reduce (or even eliminate) right-hook conflicts at signalized intersections. Bike boxes and intersection crossing markings should be placed at intersections to improve bicycle safety and increase driver awareness on bicyclist presence in general, however, it is critical for practitioners to understand which safety aspect they intent to improve by implementing those treatments and whether additional signage or other control devices are needed to enhance the safety impact of such treatments specifically on reducing right-hook crashes.

Another reason that explains why the particular treatments at the segment and the intersection levels are not strongly impacting right-hook conflicts is bicyclist behavior when going through the intersection. Anecdotally, bicyclists often chose to wait in front of motorized vehicles and particularly right after the crosswalk (regardless the presence of intersection treatments), ensuring that they would be the first to proceed through the intersection once the light turned green, thus, eliminating any potential

conflicts. This was observed at sites with and without bike boxes and/or intersection crossing markings. Bicyclists crossing the intersection during the red signal indication was also a relatively common phenomenon. The presence of bicycle signals would be beneficial for bicyclists in terms of safety and convenience.

Finally, in the studied sites there are also considerable pedestrian volumes due to the fact that all sites were located in the downtown areas. More pedestrians and bicyclists improve safety for pedestrians and bicyclists, creating the “safety-in-numbers” effect [107] something also observed (but not recorded) during the video data analysis. Motorized vehicles turning right yield to pedestrians and during those moments bicyclists can also proceed through the intersection without interacting with the motorized vehicles. Right-turning motorized vehicles are stopped ahead of the segment-level treatments or bike boxes during this time, which in turn reduces the potential for right-hook conflicts. The “Yield to Bicycles” sign that is placed in most of the studied intersections, might also have a strong impact on driver situational awareness and consequently, driver behavior. A driving simulator study found that the presence of “Yield to Bicycles” signs attracted drivers’ glances [261]. An increase in the time drivers spent looking at their right mirror before making a right turn was also observed in the presence of such signs [261].

In addition to evaluating the safety impact of bicycle treatment types on right-hook conflict occurrence at signalized intersection, the analysis also centered on the different PET values reported for the two user sequence types. PET values for conflicts that involve bicyclists followed by motorized vehicles are lower compared to the opposite user sequence and it was found that this difference is statistically significant for all the PET intervals (1, 2, 3, and 4 seconds). It seems that bicyclists feel safer while the conflicting vehicle is in their front and so they do not consider it important to for example slow down so that they can increase their distance from the right-turning

vehicle. This finding could motivate research on the appropriate thresholds to classify the detected conflicts as more or less severe.

4.5 Conclusions and future extensions

This study aimed to assess the safety impact of five bicycle treatments at signalized intersections focusing in particular, on right-hook conflicts between right-turning motorized vehicles and through-going bicycles. Poisson regression was used to model the observed traffic conflicts while additional analysis focused on the impact of user sequence, i.e., a bicyclist arrives first at the conflict area and is followed by a motorized vehicle or vice versa, in relation to the PET values.

The developed model found a strong positive association between the observed number of conflicts and the exposure terms, i.e., right-turning vehicles and through-bicycles, but did not find a statistically significant relationship between the conflicts and the bicycle treatment type; conventional bike lanes appear to improve safety for bicyclists however, this finding was significant at the 90% confidence level.

The observations collected through the video recordings also concluded that lower PET values correspond to cases where a motorized vehicle is followed by a bicycle. This suggests that bicyclists tend to maintain a smaller distance from the motorized vehicle in their front and potentially, different PET thresholds should be defined for this user sequence in an effort to capture the potential severity of a conflict. Intuitively, it is riskier when a motorized vehicle maintains a small distance from the leading bicycle compared to the opposite.

Possible limitations of this study is the relatively small number of sites hours of the overall data collection effort; across the different sites 22 hours of data were analyzed. However, an effort was made to ensure consistency, e.g., data collection took place only on weekdays of October and November, during the peak hours of the day and

during clear weather conditions. Another limitation is the lack of considering the presence and impact of control devices. For example, this study did not include signal timing considerations such as phasing sequence and signal timings or the presence of signage that could be affecting bicyclist and motorist behavior. Bicyclist and motorist behavior could also be affected by the level of familiarity with certain treatments, some of which are fairly new in the study area.

Overall, existing literature on surrogate safety techniques and specifically on traffic conflict studies using indicators such as PET or TTC, is inconclusive regarding the amount of data is needed to accurately assess safety using such metrics. But even with this uncertainty, data collection used for surrogate safety studies is more easily acquired and more informative compared to crash record datasets; video data collection provides data for traffic conflict analysis but also allows us to study other factors, e.g., user compliance with the bicycle treatment and intersection control.

Future research should focus on several aspects to better understand the crash mechanism behind right-hook crashes by studying right-hook conflicts. First, the occurrence of traffic conflicts should be studied in relation to user compliance, i.e., given the lack of bicycle signals bicyclists tend to cross the intersection during red, which potentially eliminates the potential of conflicts. The presence of other intersection treatments, e.g., protected intersections, that are specifically placed to protect bicyclists from right-turning vehicles [53] should also be evaluated using surrogate safety methods. It is also important to consider sites where intersection treatments in addition to bicycle signals are present and develop recommendations to guide their implementation based on bicycle and vehicle demand levels and intersection geometric characteristics; there is some research on this field but it is again limited in terms of the studied treatment types and control strategies [129, 208]. Finally, future research should focus on the user sequence when assessing conflicts between bicycles

and motorized vehicles, leading to recommendations on the appropriate thresholds needed to determine safe and unsafe interactions between motorists and bicyclists.

5 Assessing driver speeding and glancing behavior in the presence of protected bicycle treatments

Chapter 4 emphasized on right-hook conflicts between right-turning motorized vehicles and through-going bicyclists at signalized intersections. The objective was to explore whether (i) the bicycle treatment type upstream the signalized intersection and (ii) intersection-level treatments affect the frequency of right-hook traffic conflicts at signalized intersections. The studied treatment types were: conventional bike lanes, protected bike lanes, and sharrows for the segment-level and bike boxes and intersection-crossing pavement markings for the intersection-level.

Traffic conflict techniques are one alternative approach to crash-based analysis to assess traffic safety and they fall under the umbrella of surrogate safety methods. Their focus is on traffic conflicts as these are events where different road users approach each other in time and space and so, replicate the crash mechanism. Traffic conflict analysis based on field data which can be used in various analyses in addition to traffic conflicts (e.g., demand studies, compliance studies). Field data and in turn traffic conflict analysis are limited in that they can only capture kinematic-related metrics, such as position, speed, acceleration. Metrics related to road user mental processes such as attention, situational awareness, perception, etc. cannot be assessed.

One additional restricting aspect related to field studies, is that through those studies only existing treatments can be assessed. “Dutch” or protected intersection design is not commonplace in the US. Therefore evaluating the safety impact of this treatment is not feasible using field data. A driving simulator experiment is another surrogate safety approach, appropriate for assessing bicycle-motorized vehicles interactions in the presence of bicycle treatments for treatments not frequently found in the field and it can benefit from the use of additional equipment such as eye-tracking

device that captures driver glancing behavior.

5.1 Introduction

Urban intersections account for one third of crashes between bicyclists and motor vehicles [186] and 43% of bicyclist fatalities in the United States [176]. Right-hook crashes, i.e., crashes involving a right-turning vehicle and a through-going bicyclist, are particularly common at urban intersections as indicated in various studies [25, 73, 103].

Intersection complexity has been found to increase driver workload and reduce driver ability to detect potential hazards related to non-motorized users [108, 125, 264]. Drivers navigating complex intersections might omit searching for or detecting bicyclists in their proximity, a condition that raise crash risk. Additionally, in bicyclist-driver interactions there is the known “looked-but-failed-to-see” phenomenon [21], which occurs when drivers look at a target, yet they fail to process that information. Finally, research has shown that incorrect driver or bicyclist expectations related to bicyclist visibility accounts for some bike-car crashes [265].

In an effort to improve safety for bicyclists, many cities around the world have committed to improving and expanding their bicycle network by implementing separated or protected bicycle infrastructure, i.e., protected (separated) bike lanes and protected (Dutch) intersections. Such treatments offer dedicated space to bicyclists, which can lead to improved comfort and perceived safety for bicyclists [27, 44, 161]. Finally, protected bike lanes have been associated with increased diversity in the bicycling population, making bicycling more inclusive in terms of gender and age [3, 6, 55, 160]. However, the impact of such bicycle infrastructure treatments on driver behavior and bicyclist safety has not been adequately studied.

Protected bike lanes have the potential to reduce drivers’ ability to detect bicyclists or bicycle treatments, which may act as a warning that bicyclists may be present.

This is due to the distance they add between motorists and bicyclists, that in turn reduces the probability of a driver detecting a bicyclist. The distance between a travel lane and the protected bike lane can be as low as 60 centimeters in the case of a buffer zone and as high as the width of a car when a parking lane is used to separate bicyclists and motor vehicles [155]. Particular protected bike lane configurations, such as the presence of a parking lane, can actually block the drivers' view to the protected bike lane and limit their ability to detect bicyclists. Even when bicyclists are detected by a driver, there are still concerns that after a period of separation drivers might not anticipate interacting with bicyclists as the two merge back at the intersection. This could raise the risk of right-hook crashes between drivers and bicyclists. The placement of protected or Dutch intersection features after a protected bike lane is a potential solution to right-hook crashes. This is because protected intersection elements allow for physical separation between motorists and drivers at the intersection. This in turn encourages drivers to make wider angle turns and results in drivers encountering bicyclists in front of them, reducing the risk for right-hook collisions.

The objective of this research is to understand whether driver behavior at right turns is affected by the presence of protected bicycle infrastructure treatments, namely protected bike lanes and protected intersections. A driving simulator experiment was designed to capture driver right-turning behavior, in terms of speed and glances, under the presence of protected bicycle infrastructure treatments.

5.2 Literature review

This section summarizes bicycle safety research on protected bicycle treatments. The safety impact of protected bike lanes both at the segment and the intersection levels are presented in the first subsection (subsection 5.2.1), which is followed by research related to protected intersections (subsection 5.2.2). The final section

summarizes research gaps that are addressed through the research presented in this paper.

5.2.1 Protected bike lanes

Protected bike lanes, also known as separated bike lanes or cycle tracks, provide a physical separation between motorists and bicyclists as shown in Figure 5.1. As a result, they eliminate the possibility of collisions between bicyclists and motorized vehicles along roadway segments.

During the 1980s-1990s research on the safety aspect of protected bike lanes had focused on Northern European countries, as those were the first to implement such treatments. The majority of European studies have revealed positive safety impacts at the segment level when cycle tracks are implemented versus when no bicycle treatment is present [246]. Similar positive benefits were observed in Copenhagen, Denmark, when protected bike lanes were compared against conventional bike lanes [113].

Recent implementations of protected bike lanes in North America have motivated studies that compare the safety performance of road segments with protected bike lanes to ones where other bicycle infrastructure treatments or no treatments at all are present. Crash record studies reveal that both two-way and one-way protected bike lanes are associated with reduced crash rates and injury risk when compared to segments where no bicycle infrastructure is present [143, 148]. Naturalistic studies also support these findings [27]. However, all of these studies have focused on comparisons between segments with protected bike lanes and those with no bicycle treatments. Additionally, none of the aforementioned studies have studied the impact of protected bike lanes on intersection level safety; essentially, they do not differentiate on whether a crash has taken place at the segment (i.e., between two intersections) or at the the intersection. Lastly, they have not studied driver behavior and situational awareness

of drivers traveling next to protected bike lanes.

Given that bicyclists and drivers merge back at the intersection, it is critical to investigate the impact protected bike lanes have on drivers' ability to detect bicyclists and behavior of drivers at intersections. Recent research suggests that more than 50% of drivers turning right omit to scan right for bicyclists at intersections after traveling next to protected bike lanes [121].

There are several crash-based analyses on the impact of protected bike lanes at intersections. Several studies found an increase in the number of intersection bicycle-motorized vehicle crashes after the installation of protected bike lanes [2, 65, 110]. However, some of these studies did not control for exposure (i.e., bicycle and motorized vehicle demand) [2, 65], which can result in over- or under-estimation of crash risk [76]. On the contrary, studies that have accounted for bicycle and motorized vehicle demand as the exposure metric, reported conflicting findings regarding the presence of protected bike lanes at signalized intersections. When compared against no treatment, protected bike lanes have been found to reduce crash risk [240] but when compared to conventional bike lanes, protected bike lanes have been found to increase crash risk [144]. Besides the conflicting findings, none of these crash-based studies differentiate between the type of crashes that occur at intersections, e.g., rear-end versus right-hook crashes. Therefore, it remains unclear if and how protected bike lanes impact right-hook crashes at intersections.

A more accurate understanding can be obtained through field studies and conflict analyses. Analysis of conflicts between right-turning drivers and straight-going bicyclists using data from two types of signalized intersections: with one- or two-way cycle tracks or with no bicycle treatment revealed lower conflict rates when cycle tracks were present [274]. However, other field studies revealed that drivers performing right turns at intersections following protected bike lanes are focused on the left-approaching

traffic and are less likely to scan right prior to a right turn [121, 238].

Several crash-based analyses have suggested the need for interventions at intersections when protected bike lanes are placed upstream of those intersections to reduce the risk of right-hook crashes [220, 246]. Examples of intersection treatments that have been studied include the evaluation of traffic signal phases dedicated to bicyclists [78, 240] and various intersection-designs such as mixing zones, i.e., configurations that mix the traffic upstream of the intersection [153, 162, 240], raised bicycle intersection-crossings [81], and protected intersections [81, 261]. Most notably, mixing was found to be appropriate only in cases where the speed limit is lower than 30 km/hour [222], while the understanding of the impacts of combining protected bike lanes with protected intersections on driver behavior and bicycle safety are limited as explained in the next subsection.

5.2.2 Protected Intersections

Protected or Dutch intersections are intersections consisting of design elements such as corner refuge islands, curb extensions, and setback bicycle crossings [176]. These design elements alter the placement of drivers and bicyclists at the intersection and right-turning drivers encounter the bicyclist in front of them, not to the right of them which is the case for non-protected (i.e., conventional) intersections. Finally, vehicles and bicyclists are physically separated with a corner refuge island that encourages the driver to make the turn at a wider angle.

Overall, even though limited, existing studies agree on the positive safety implications of protected intersections. A before-after study in Utah evaluated the safety benefits of protected intersections by observing bicyclist, pedestrian, and driver behavior; several unsafe behaviors were reduced after the installation of the protected intersection and in particular, the number of bicyclists using the crosswalk and the

number of bicyclists waiting to cross at the wrong place of the intersection (e.g., bicyclists waiting on the crosswalk) [?]. A bicyclist simulation study exposed bicyclists to protected and non-protected intersections (upstream of protected bike lanes) where they interacted with a virtual right-turning vehicle [188]. The presence of protected intersection increases the distance between the bicyclist and the turning vehicle at the intersection while at the same time in the case of protected intersections bicyclists were found to arrive at the crossing point with a lower speed compared to non-protected intersections. Other field and driving simulator studies have revealed reduced conflicts between bicyclists and right-turning motorists when protected lanes are present and protected bike lanes [153], or conventional ones [261] are implemented upstream the intersection. However, the results from both of these studies lacked statistical significance due to limited sample sizes. More recently, another driving simulator study, investigated the impact of protected intersection elements and design specifics on driver behavior [53], revealing that corner refuge islands with larger width as well as the presence of intersection-crossing pavement markings can reduce right-turning vehicle speed. However, this study did not compare scenarios where no protected intersection elements were present or scenarios featuring protected bike lanes upstream of the protected intersection. Overall, none of the aforementioned studies compared scenarios that included different combinations of protected/non-protected intersections and protected/conventional bike lanes upstream those intersections.

5.2.3 Summary of the literature

While research on the safety impacts of protected bike lanes has flourished over the past few years as more and more urban areas have been implementing them, research on protected intersections and their safety impacts is limited, raising the need to understand their safety benefits and design implementation guidelines. Additionally,

existing studies either focus on the segment treatments (e.g., away from the intersection) or only evaluate intersection treatments, ignoring the benefits that could result from combining segment and intersection bicycle infrastructure treatments. The goal of this study is to fill this gap by assessing how drivers behave when traveling on segments and making right turns through intersections with protected bicycle treatments.

5.3 Methodology

A driving simulator experiment was designed to assess driver behavior along segments consisting of conventional or protected bike lanes while performing right turns at protected and non-protected intersections. Driver behavior was captured via glances at the infrastructure and bicyclists that were present in the simulated scenarios, as well as through driver speed.

The following subsections describe the apparatus, followed by the research hypotheses and the experimental design which presents the scenarios, dependent and independent variables. Participant recruitment as well as the experimental procedure are discussed at the end.

5.3.1 Apparatus

Driving simulator

The experiment took place at the University of Massachusetts Amherst Human Performance Lab (HPL) where a high fidelity driving simulator is housed. The HPL driving simulator is a 2013 model Ford Fusion Sedan (see Figure 5.2) offering similar functionalities as real cars, such as accelerating, braking, steering, etc. The car is surrounded by screens creating a 330-degree field-of-view, where scenarios are displayed. The scenes are provided by six high resolution projectors. The front five projectors provide a resolution of 1920 x 1200 pixels, while the rear projector provides a resolution

of 1400 x 1050 pixels. The rear scene can be seen through the in-cab rear-view mirror while the side-view mirrors display the simulated world. The vehicle is also equipped with a virtual dash and 17-inch touch screen center stack. With regards to sound, the HPL driving simulator has two separate systems, one external and one internal to the cab, to imitate environmental sounds and engine noise, respectively as in the real world. A 5.1 channel audio system that is external to the car cab provides the environmental sounds such as traffic, passing vehicles, and road noise, an internal audio system provides engine sounds and vibrations; those two elements contribute further to providing a driving experience that resembles real-world conditions. Vehicle data such as position, speed, acceleration, etc. are continuously collected at a frequency of 96 Hz.

Eye-tracking device

An eye-tracking device was used to record participant eye movements. HPL is equipped with an Applied Science Laboratory MobileEye tracker, shown in Figure 5.2, which is used to capture participants' right pupil movement through video cameras located on a set of goggles ; the camera works at 30 Hz. The device is equipped with an additional camera that is used to record participant's view of the road.

Prior to the experiment, the eye-tracking device is calibrated for every participant to ensure that the device is capable of following the participant's right pupil and correctly record its movement. During this process, the participant is in the driver's seat, has already adjusted the seat and rear mirror, and is wearing a seat belt; so, the participant's position before the calibration needs to be settled. The front screen of the simulator displays a board with letters A, B, C, D, E, F, G, H, and I placed on a 3x3 grid. The participant is asked to view the middle letter (E) so that the device's camera can detect their right pupil. After this step is completed, the participant is asked to view all letters, one-by-one. Through this phase, it is ensured that the device

follows the participant's pupil across all letters on the board. The molecular needs to be adjusted up to the point that all letters are viewed.

5.3.2 Research hypotheses

This experiment tested various hypotheses related to driver glancing and speeding while traveling along the segment and while entering the intersection. There are three hypotheses related to driver glancing behavior:

1. Drivers traveling along a segment configured with protected bike lanes (located to the right of a parking lane) are less likely to glance at a bicyclist that is riding on the protected bike lane, compared to the case of conventional bike lanes; protected bike lanes increase the lateral distance between drivers and bicyclists while parked cars have the potential to block drivers' view of the protected bike lane.
2. Drivers turning right at protected intersections are more likely to glance to their right compared to when entering a non-protected intersection. Protected intersections incorporate design elements that can capture driver attention.
3. Bicyclist presence at the segment is expected to increase the occurrence of right glances at the intersection.

Additionally, there are three hypotheses related to driver speeding behavior:

1. Drivers are expected to develop higher speeds when traveling along a segment where protected bike lanes are present compared to when traveling next to conventional bike lanes, as the potential to interact with bicyclists is lower in the former case.

2. Drivers turning right at protected intersections are expected to develop lower speeds compared to non-protected intersections.
3. Drivers are expected to drive slower along the segment if a bicyclist is present, with the objective to be more careful.

5.3.3 Scenario development

In order to capture the impact of protected bicycle infrastructure treatments, i.e., protected bike lanes and protected intersections, both treatments were compared with conventional bike lane and non-protected intersection designs, respectively. Two segment- and two intersection-level bicycle infrastructure treatments were considered, leading to four roadway environments (Figure 5.3).

The roadway environment was a straight road segment leading to an intersection, configured as followed based on the bicycle treatment types:

1. a straight segment with a protected bike lane leading to a protected intersection
2. a straight segment with a protected bike lane leading to a non-protected intersection
3. a straight segment with a conventional bike lane leading to a protected intersection
4. a straight segment with a conventional bike lane leading to a non-protected intersection

At intersections, it is common to indicate the bicyclist path, which is ultimately the extension of the bike lane through the intersection referred to as intersection-crossing pavement markings. In North America the design of a bicycle intersection-crossing

path varies, e.g., solid line versus green-colored markings [168]. In this study the dotted extension design has been chosen for the conventional non-protected intersection design (Figure 5.4a), while the protected intersection design includes green-colored intersection-crossing pavement markings in addition to a corner refuge island for the intersection approach of interest (Figure 5.4b).

In each roadway environment, the cross-sectional design of the road per direction of traffic is: two traffic lanes (i.e., for motorized vehicles), one parking lane, and one bike lane. The relative placement of the bike and parking lanes was dependent on the type of the segment-level bicycle treatment. More specifically, a conventional bike lane is placed between the right-most traffic lane and the parking lane, while a protected bike lane is placed between the parking lane and the sidewalk. The posted speed limit is 56 km per hour (35 miles per hour).

Eight scenarios were developed in total, and each of these four roadway environments appeared twice; see Table 5.1. The development of eight instead of four scenarios aimed at the inclusion of an additional variable, that of the bicyclist presence. The later variable aimed to assess whether driver glancing and speeding behavior at the intersection depends on bicyclist presence along the segment. In the real world, drivers would ideally be able to detect bicycle infrastructure treatments along the roadway and translate this stimuli as an indication of potential bicyclist presence. In an attempt to replicate reality, participants drove the same roadway environment twice; once with a bicyclist riding on the segment (upstream the intersection) and once with no bicyclist present.

A bicyclist was programmed to ride on the bike lane (conventional and protected) along the segment denoted by letters A and B in Figure 5.5. This was achieved by placing a sensor along the driver's path, 10 meters before point A. The sensor activated the bicyclist's motion as soon as the driver crossed it. The bicyclist was programmed

Table 5.1: Scenario Design

Scenarios	Segment	Intersection	Bicyclist
1	Protected Bike Lane	Protected	YES
2	Protected Bike Lane	Non-Protected	YES
3	Conventional Bike Lane	Non-Protected	YES
4	Conventional Bike Lane	Protected	YES
5	Protected Bike Lane	Protected	NO
5	Protected Bike Lane	Non-Protected	NO
7	Conventional Bike Lane	Non-Protected	NO
8	Conventional Bike Lane	Protected	NO

to be moving at a constant speed of 16 kilometers per hour (km/h). The bicyclist's starting and ending points, as well as speed remained unchanged across all scenarios.

The bicyclist presence combines both top-down and bottom-up attention processing elements in the design. Along the segment, a bottom-up design has been adopted where the driver may or may not glance at the bicyclist, while at the intersection the design is top-down; it is aimed to assess whether drivers will check right based on their previously gained knowledge (i.e., glance at the bicyclist, if one was present) or based on their overall knowledge related to safe driving.

A within-subject design was adopted exposing each participant to all eight scenarios. An eight by eight Latin square matrix was used to randomize the order of the scenario presentation to the participants. At each scenario participants drove a straight roadway segment of 310 meters that led to an intersection. In advance of the intersection participants were given the instruction to turn right. The signals were programmed to always show a green indication so that the driver would not have to stop at the intersection. Each scenario had two programmed cars moving in the opposite direction, along the segment and away from the intersection. The scenario terminated 50 meters after the participant completed the right turn. The outline of the geometric design of the scenarios is shown in Figure 5.5.

5.3.4 Dependent and independent variables

The columns of Table 5.1 display the independent variables used in the experiment, namely the bicycle segment treatment (protected or conventional bike lane), bicycle intersection treatment (protected or non-protected intersection), and the binary variable indicating the presence of a bicyclist.

The segment-level analysis focused on the AB part of the drive (Figure 5.5), which coincides with the part of the drive where the driver would encounter the bicyclist (if driving a scenario with a bicyclist being present). The intersection analysis focused on the segment defined by points C and D. The intersection was further split into two zones; see Figure 5.6. Zone 1 is defined as the intersection approach section and captures the area corresponding to 3 seconds before the driver reaches the stop bar. The starting point for Zone 2 is the stop bar and the end point coincides with the end of the bike lane, where a potential bicyclist could enter the intersection. The intersection glance analysis is time-based instead of distance-based (e.g., as the speed analysis) to ensure that regardless their traveling speed, all participants had the opportunity to allocate the same time (i.e., 3 seconds) to glance right. A different amount of time would have resulted in different memory allocation per participant, e.g., one participant would have processed this information earlier compared to another.

Speed data were analyzed for the following parts of each drive: (1) while the driver was traveling along section AB, and (2) while the driver was traveling along section CD. For both cases, the average participant speed per scenario for these two parts of the drive was used for the analysis.

Driver eye glance analysis focused on: (1) whether the driver glanced at the bicyclist in the presence of protected and conventional bike lanes while traveling along section AB and (2) whether the driver glanced to the right at the intersection while

traveling in Zone 1 or 2.

Glances were treated as binary variables, indicating whether the participant placed a glance or not at the area of interest. The scoring process consisted of assigning a score of 1 every time the red cross shown in Figures 5.7 and 5.8 was pointing at the area of interest; otherwise, assigning a score of 0. For segment AB the scoring process was relatively straight-forward, recording 1 if the red cross was on the bicyclist at least once during the drive along the segment. The expected glance type at the intersection and in particular, while in Zone 1, varied based on the segment treatment. When protected bike lanes were present, it was anticipated that drivers would place a glance at the intersection area where a bicyclist would be expected to be in a real-world environment. In the case of conventional bike lanes a right-mirror glance while in Zone 1 was scored with 1.

5.3.5 Sample size and participants

This within-subject experimental design includes three binary, independent variables. To determine the required sample size necessary to achieve a target statistical power of 0.95, GLIMMPSE [90, 133] was used. GLIMMPSE requires the user to define certain parameters such as the desired power, the number and type of the independent variables, estimated mean value and variance for the dependent variables, in order to estimate the minimum sample size to achieve the desired statistical power. While some parameters are straight-forward based on the experimental design (e.g., the number of independent variables), some others require assumptions to be made by the researcher. For example, in order to estimate the sample size the researcher needs to import values, i.e., expected mean and variance, for the dependent variables. In this experiment this information was obtained using data collected from three participants, referred to as the “pilot study” participants. The data obtained from the pilot study

were not used for the main analysis of this study. Table 5.2 lists the mean value and variance imported into the GLIMPSE software. In addition to mean and variance for the four dependent variables, the following information was used to complete the power analysis:

1. Desired power = 0.95
2. Type I Error Rate = 0.05
3. Control for a single, normally distributed covariance
4. Four dependent variables
5. “Repeated measures” option was checked indicating that for every dependent variable eight measurements were taken per participant
6. Under the Hypothesis tab it was indicated that the objective of the study was the Main Effect
7. Within-Participant Variability: derived from the pilot study

Table 5.2: Mean and variance for the dependent variables used in the power analysis

Dependent variable	Mean	Variance
Segment speed	35.90	23.07
Intersection speed	16.79	7.77
Glance at the bicyclist	0.92	0.08
Glance right (intersection)	0.04	0.04

GLIMPSE indicated that a sample size of 31 participants is needed to achieve a 0.95 statistical power. Thirty-two participants were recruited and successfully completed the experiment.

Participants with a valid U.S. driving license for at least a year were recruited from the University of Massachusetts Amherst area. In total, 35 drivers completed the

study: three that served as pilot and 32 that were used for the entire analysis. Two more attempted to complete the study but due to simulation sickness and problems in calibrating the eye-tracking device (section 5.3.1) they had to stop.

5.3.6 Experimental procedure

The 32 participants were equally split between male and female and their age was between 18 to 36 years old with a mean of 23.7 years old, a median of 24 years old and a standard deviation of 4.5 years. When participants arrived at the HPL they were asked to fill out a consent form and a questionnaire asking about their gender, age, and driving history; this is the pre-study questionnaire. Then they were asked to enter the vehicle in the drivers' seat and were fitted with the eye tracker; the device was properly calibrated before starting the vehicle.

Participants first drove a practice drive that allowed them to familiarize themselves with the simulator and its controls. The practice drive was three to four minutes in length and exposed participants to straight segments and curves, signalized and unsignalized intersections so that they can experience different driving modes such driving with constant speed, decelerating, and accelerating. Additionally, prior to each turn, a pop-up message informed the driver whether to turn left, right, or continue through the intersection. The pop-up message related to navigation is an element present in the experiment and participants familiarized themselves with this type of communication during the practice drive. Similar to the experimental roadway environment, the practice drive featured an urban network, however, no other similarities exist with the actual scenarios. The particular practice drive has been used in other driving simulator experiments conducted in HPL (e.g., [53, 60, 209, 210]). After the practice drive, participants started the driving part of the experiment that exposed them to the eight scenarios described earlier. Lastly, after completing the

driving part of the experiment, participants completed the post-study questionnaire that gathered information related to participant's bicycling habits, if any. Overall, each participant remained in the lab for about 15 to 20 minutes.

5.4 Results

This section presents the glance and speed related results organized in subsections corresponding to results for the segment or intersection parts of the drive. In addition to the three roadway environment-related independent variables, i.e., segment treatment, intersection treatment, and bicyclist presence, the impact of participants' age was also investigated. While data on participants gender was collected it was mostly used to ensure that males and females are equally represented among the participants and was not used for any further analysis.

5.4.1 Glance data

Data collected with the eye-tracking device were analyzed for the AB segment and the intersection area, focusing on Zones 1 and 2; see Figure 5.6. Glances were treated as binary variables, indicating whether the participant glanced or not at the respective area of interest. Logistic regression models were developed aiming to associate the binary response variable with design factors (e.g., segment or intersection treatments), bicyclist presence, and participant age. The information on participants driving history, i.e., age that they got their driving license, was found correlated (0.908) with the Age variable and so, the latter was included in the analysis.

A closer investigation of the data revealed that the positive and negative responses in terms of glances (i.e., glanced or not) were not equally represented in the dataset; while at the segment the great majority of the drivers glanced at the bicyclist, most of them did not place a glance to their right when at the intersection. This imbalance

in the number of positive and negative responses can result in a model that is biased towards the most represented response, or in other words, the majority class. This bias is due to the fact that the model will have more opportunities to learn why the most represented response occurred [199]. Therefore, both datasets, i.e., glances at the bicyclist at the segment and glances at the intersection, needed to be balanced prior to the development of any logistic regression models. Note that the fact that both datasets appeared to be imbalanced can be attributed to randomness; studying a different population might have resulted in only of them being balanced.

One approach to address imbalance in datasets is to change the size of the classes in order to make them equal; this can be achieved by either oversampling the minority class or undersampling the majority class [165, 182]. In the context of the present study, undersampling the majority class would be achieved by withdrawing participant data; therefore, resulting in an insufficient sample size to achieve the target statistical power. As a result this imbalance was addressed by implemented a method, introduced by Chawla et al. (2002), named Synthetic Minority Oversampling TEchnique (SMOTE). SMOTE increases the sample size of the minority class by generating new synthetic data points. This is done by forming a convex combination of neighboring data points [28]. For example, in a two-dimensional space, SMOTE will generate a third point that lies somewhere on the straight line that connects two neighboring data points [28].

SMOTE has several applications in the field of transportation and particularly, in traffic safety research. Traffic incidents such as near misses and crashes are relatively rare compared to the total number of road user interactions occurring on roadways. SMOTE has been implemented for the development of real-time crash risk models where the dataset used to develop the model consisted of crash and no-crash traffic events with the former exceeding by far the latter [120, 185]. In driving simulator

studies SMOTE has been applied for crash prediction [63] and balancing datasets with depressed (minority class) and non-depressed participants (majority class) [119]. For the present study, this method was applied for the recorded glance datasets at the segment and the intersection and was used to create models that are capable of modeling both the minority and majority classes.

The analysis of glance data along the AB segment included only scenarios where a bicyclist was present resulting in a total of 128 observations (i.e., 32 participants across 4 scenarios). In the case of conventional bike lanes, all participants glanced at the bicyclist; in the case of protected bike lanes participants glanced at the bicyclist in 76.2% of all drives, creating the aforementioned imbalance that called for the implementation of SMOTE. After implementing SMOTE, the total number of observations increased from 128 to 156. This increase allowed us to obtain a more representative sample for the no-glancing data points.

The following independent variables were considered for the logistic regression model: segment treatment and age (see Table 5.3). Age that was treated as a continuous variable ranging from 18 to 36 years old and segment treatment is a binary variable equal to 1 for protected bike lane and 0 otherwise.

Table 5.3 presents the logistic regression model that describes glances at the bicyclist while traveling along segment AB. All independent variables were statistically significant at the 95% level of significance.

Table 5.3: Logistic regression model for glances at the bicyclist (segment)

Parameter	Coefficient	Odds Ratio (CI 95%)	<i>p</i>-value
Segment Treatment	-5.367	0.005 (0.000, 0.031)	0.000**
Age	0.203	1.224 (1.123, 1.329)	0.000**

**Statistically significant at the 95% level of significance*

*** Statistically significant at the 99% level of significance*

CI: Confidence interval

The model reported a pseudo R^2 equal to 40.4%, which is relatively high considering that the response describes human behavior that is complicated and affected by many different factors. A chi-squared test was used to assess the model's Goodness-of-Fit (GoF) and reported a p -value that was approximately 0.99. This value is greater than the target 0.05 significance level allowing us to accept the null hypothesis, stating that observed and estimated values are the same. The model's performance in terms of accuracy was assessed with the Area Under the Curve (AUC), which essentially summarizes the classifier's performance over a range of trade-offs between True Positive and False Positive error rates [241]. For this model the AUC value was found to be equal to 0.98. Higher AUC values are associated with higher model accuracy and, according to Hosmer et al. (2013), an AUC that is greater than 0.7 results in acceptable models [101].

The intersection glancing behavior analysis focused on Zone 1 and Zone 2 as shown in Figure 5.6. The glance variables indicate whether there was at least one right glance (either in Zone 1 or 2) at the intersection per drive per participant. The percentage of drives where at least one glance was recorded while either in Zone 1 or 2 is shown in Figure 5.9. Bicyclist presence as well as the protected intersection design appear to be influencing glancing behavior, despite the fact that the percentage of drives that have at least one glance per scenario rarely exceeds 25%.

Given the imbalance observed regarding the intersection glances, SMOTE was applied prior to developing a logistic regression model. In this model, the independent variables were: age (continuous), segment treatment = 1 for protected bike lane, 0 otherwise; intersection treatment = 1 for protected intersection, 0 otherwise; and bicyclist = 1 for scenarios where a bicyclist was present at the segment, 0 otherwise. A backwards elimination process was used to determine the final model; in the initial model, segment treatment was not found statistically significant at the 95% confidence

level and was removed from the final model, shown in Table 5.4 for the final logistic regression model.

Table 5.4: Final logistic regression model for glances at the intersection Zones 1 and 2

Parameter	Coefficient	Odds Ratio (CI 95%)	p-value
Intercept	-3.299	0.037 (0.007, 0.209)	0.000**
Intersection Treatment	0.690	1.994 (1.201, 3.310)	0.008**
Bicyclist	1.464	4.322 (2.526, 7.340)	0.000**
Age	0.082	1.085 (1.157, 1.157)	0.013*

*Statistically significant at the 95% level of significance

** Statistically significant at the 99% level of significance

In this model all the independent variables except from age are statistically significant at the 99% level, indicating a strong association with the response variable. The model reported a pseudo $R^2 = 0.09$. According to the odds ratio, the presence of a bicyclist was the strongest determinant of driver glancing behavior at the intersection. Intersection treatment had a positive impact on the right glances, meaning that protected intersections can increase right glances at the intersection. The goodness of fit for this model was assessed through a chi-square test, which reported a p -value equal to 0.99, indicating that observed and estimated values are similar. The AUC value for this model was equal to 0.76 suggesting that the model has quite strong predictive power [101].

5.4.2 Speed data

Driver speeding behavior was analyzed for the AB and CD parts of the drive as shown in Figure 5.5. There was an attempt to develop linear regression models both for the segment and the intersection speeds including both design independent variables (i.e., intersection and segment treatment) as well as participant age; however, these models were found to have a poor fit and resulted in no significant variables, so

they are not presented. The analysis presented here focuses instead on capturing the effect of each design variable individually on the segment and intersection speeds.

Segment speeds were assessed using violin plots; see Figure 5.10. The white dot in the middle of each violin plot corresponds to the median value and the thick black bar represents the interquartile range. The end of the thin black line (upper and lower) represents the upper (max) and lower (min) adjacent values in the data [97]. Violin plots represent the order ranking of a data set and the additional information of how frequently a certain value appears in the data set, in other words illustrating a density plot of the data.

The violin plots show that there was a relatively small variation in speeds among the four segment environments. The impact of individual factors and in particular, bicycle treatment and bicyclist presence on speed was assessed through an Analysis of Variance (ANOVA). A repeated measures ANOVA revealed that bicyclist presence (F-statistic = 7.36, p-value = 0.033) as well as the type of bicycle treatment (F-statistic = 10.16, p-value = 0.011) had an impact on the observed speeds. However, there is no interaction effect between the segment treatment and the presence of a bicyclist (F-statistic = 0.04, p-value = 0.946) as shown in Figure 5.11.

Table 5.5: Repeated measures ANOVA for segment speed

Parameter	F-statistic	p-value
Bicyclist	7.36	0.033*
Segment Treatment	10.6	0.011*
Interaction term	0.04	0.946

**Statistically significant at the 95% level of significance*

An average speed of 54.2 km/hour (mean = 54.2, variance = 9.02) was reported for the scenarios where a bicyclist was present versus 56.5 km/hour (mean = 56.5, variance = 8.02) for those with no bicyclist. For these values the Cohen's D value was 0.274 which indicates a small effect size [46], meaning that the presence of the

bicyclist has a small effect on the segment speed. When participants drove next to protected bike lanes developed on average speed of 54.6 km/hour versus 56.2 km/hour, which was the average speed across all scenarios with conventional bike lanes. The Cohen's D value was 0.183 indicating a small effect. Capturing the effect of each one of the independent variables, i.e., bicycle treatment and bicyclist presence, on speed with sufficient statistical power would require a larger sample size [46].

The intersection speed analysis focused on understanding whether the segment and intersection bicycle treatments as well as the bicyclist presence would trigger participants to select different speeds within the CD part of the drive. A repeated measures ANOVA on the intersection (CD part of the drive) speeds revealed that none of these independent variables or their interactions were statistically significant at the 95% confidence level.

Further speed analysis focused on assessing whether there is a correlation between participant glancing behavior and speed selection at the intersection (CD part of the drive). The reasoning behind this analysis is the fact that both the intersection treatment and bicyclist presence variables were found to have an impact on driver behavior. A Student t-test was conducted for the intersection speeds of the participants who glanced right at the intersection (mean = 20.9 km/hour) versus the ones who did not (mean = 24.4 km/hour). A p-value equal to 0.000021 allowed to reject the null hypothesis stating that the means across the two groups are equal. An additional test aimed to assess whether glancing at the bicyclist at the segment upstream the intersection influences participant intersection speed. A Student t-test revealed that the mean speed of participants who glanced at the bicyclist (mean = 23.7 km/hour) was significantly different (p-value = 0.00033) to the speed (mean = 25.9 km/hour) of participants who did not glance at the bicyclist.

5.5 Discussion

This research aimed at understanding whether bicycle infrastructure treatments such as protected bike lanes and protected intersections affect driver behavior when drivers are approaching an intersection to make a right turn. Driver behavior was analyzed separately for the segment and the intersection using both glance and speed data.

5.5.1 Glances

For the segment the initial hypothesis was that participants would be less likely to glance at a bicyclist when a protected versus conventional bike lane is present. The hypothesis was confirmed and in particular, protected bike lanes were found to reduce the number of drivers glancing at the bicyclist. While not glancing at the bicyclist was not a frequent behavior, which was the reason for using SMOTE to increase the sample size by generating synthetic samples, it is important to consider that: 1) the ambient traffic was very low, i.e., two vehicles per scenario were programmed to travel in the opposite direction, and 2) the parking lane was not entirely full. According to Werneke and Vollrath (2012), lower levels of motorized vehicle density were found associated with higher attention levels to non-motorized users [264]. Due to this, a low ambient traffic was chosen for the experiments to allow for capturing differences in glances between conventional and protected treatments when other potential distractions are not present. This indicates that even in a roadway environment with reduced driver workload some participants did not directly glance at the bicyclist that was traveling on the protected bike lane. Anecdotally, after completing the experiment, two participants reported that they were nervous when overpassing the bicyclist that was next to them and preferred having the bicyclist further, separated by the parked

cars. Two other participants changed lanes to overpass the bicyclist that was riding on the conventional bike lane. This observation has also been reported in recent studies on driver preferred bicycle infrastructure treatments; drivers in the Bay Area reported a stronger preference in driving next to protected/separated bike lanes compared to conventional ones [212]. Traffic signs indicating the presence of bicyclists and education could serve as potential countermeasures that would allow for increased driver awareness of potential bicyclist presence in the case of protected bike lanes.

At the intersection, a small portion of the participants glanced right. Glancing was positively correlated with the implementation of a protected intersection design and the presence of a bicyclist on the segment. The protected intersection design was effective in increasing the likelihood of participants glancing right at the intersection, indicating that the intersection design itself can indeed increase driver awareness of the potential presence of other road users at the intersection. The presence of a bicyclist upstream the intersection triggered participants to glance right at the intersection. The analysis of glances at the intersection also concluded that the segment treatment type does not have an effect on participants' right-turning behavior.

Age was found to affect participants' glancing behavior. In agreement to other studies on young drivers and their risk perception and situational awareness [17, 18], younger drivers in this study were less likely to glance at the bicyclist while at the segment and also less likely to place a glance to their right while being at the intersection. Generally speaking, young and inexperienced drivers have a reduced ability to anticipate and as a result, check for hazardous events. In the present study a potential interaction with a bicyclist can be considered as a "hazardous event."

5.5.2 Speed

The speed hypothesis that participants would develop higher average speeds when traveling on segments with protected bike lanes versus with conventional bike lanes was not confirmed. One reasoning can be that parking lanes adjacent to traffic lanes decrease the effective lane width and as a result, they motivate lower speeds. This was also the finding of an earlier driving simulator experiment, which studied the effect of on-street parking on driver speed and reaction times and found that the presence of on-street parking reduces average speed [62].

While speeds at the intersection were not statistically different across the scenarios, the results indicate that participants who glanced right at the intersection were those who had significantly lower average speeds. Considering that intersection glancing behavior was associated with the implementation of a protected intersection, it is reasonable to infer that protected intersections may have speed reduction benefits as well.

Lower average speeds were observed at the segment when a bicyclist was present. At the intersection, bicyclist presence triggered right glances, which in turn, were associated with lower turning speeds. Therefore, bicyclist presence resulted in reduced intersection turning speeds only for those participants that glanced right at the intersection. Although the speed reduction was small, it still shows that drivers adjust their behavior once they are aware of the presence of other road users.

5.6 Conclusions and future extensions

This study conducted an in-depth analysis of driver behavior when the roadway environment is configured with bicycle infrastructure treatments, namely protected or conventional bike lanes and protected or non-protected intersections. The contribution

of this study is in the comparison of various combinations of segment and intersection bicycle treatments, which allows for the development of guidelines for the appropriate implementation of such treatments and development of countermeasures that ensure safe operations and interactions for all users.

The results of this study conclude that protected bike lanes, especially those that are located between the sidewalk and the parking lane, have the potential to reduce driver ability to detect the bicyclist. At the same time, they were also found to affect driver speed selection although causing only a small change to it while traveling along the segment. Findings from this study also indicate that the protected intersection design improves driver behavior as it results in higher rates of glancing at the intersection prior to a right turn, which in turn were associated with lower average speeds. However, glances at the intersection were not affected by the type of segment-level bicycle infrastructure upstream of the intersection. Notably, the presence of a bicyclist on the bike lane along the segment was positively correlated with glances and lower speeds at the intersection.

While this study lacked scenarios with driver-bicycle interactions, it still contributes to the understanding of driver behavior in complex roadway environments such as when bicycle infrastructure treatments are present. Given that the presence of a bicyclist along the segment appeared to impact driver behavior, future studies should examine this impact for a variety of bicycle demand scenarios (both in terms of volume and location where bicyclists are present). A limitation of this study is that the ambient traffic was very low; future work could explore how drivers behave when they are exposed to higher workload conditions, e.g., higher traffic volumes. In such scenarios it would be important to also assess whether the presence of traffic signs or education improves drivers' ability to detect bicyclists but also, check for bicyclist presence prior to turns or other movements. Lastly, a more diverse population in

terms of age as well as familiarity with bicycle infrastructure treatments should also be tested in the future; existing research has demonstrated that drivers familiar with bicycle treatments behave as indented in the presence of those treatments [75].

Acknowledgements

This research was funded in part through a research project funded by the Safety Research Using Simulation (SAFER SIM) University Transportation Center (UTC) at the University of Iowa. Funding for the UTC Program is provided by the Office of Assistant Secretary for Research and Innovation (OST-R) of the United States Department of Transportation (USDOT).

6 A framework for mode classification in multimodal environments using radar-based sensors

As demonstrated in Chapters 3-5 bicycle safety can be assessed through various ways including crashes and traffic conflicts between bicyclists and motorized vehicles as well as by capturing and studying driver behavior in response to bicycle treatments and bicyclist presence. Crash analysis and driver simulator experiments have been used for traffic safety research for a relatively long time and are established methodologies; there is a well-known procedure to collect and analyze data. On the contrary, field studies are tightly related to the existing technology that is available for data collection. Essentially, the way real-time field data is collected as well its quality are affected by the technology and in turn, affect the analysis. Cambridge field data was collected using video cameras as this technology is mostly used in the bicycle and pedestrian traffic monitoring and safety studies. However, it was found that the recorded video data suffered from certain inaccuracies, regarding bicyclists and pedestrians detection and tracking. These aspects and a few more limitations regarding video data collection are discussed in the literature; video cameras underperform in adverse lighting and weather conditions.

Overall, the present study was motivated by certain limitations associated with video cameras as tools to collect traffic data. Technologies alternative to video cameras, such as radar-based sensors, can be used for traffic monitoring purposes. Radar-based sensors have not been studied in terms of mode classification in multimodal environments where pedestrians, bicyclists, and motorized vehicles are present. It is of high importance to first assure that through the data records obtained by a radar-based sensor it is feasible to differentiate between the different mode types prior to considering this sensor for traffic monitoring. This study developed a mode

classification framework able to work in varying traffic scenes where pedestrians, bicyclists, and motorized vehicles are present.

6.1 Introduction

Non-motorized transportation modes such as bicycling and walking are gaining popularity given their potential to improve public health and energy efficiency, reduce congestion, and contribute to livable communities. However, data from around the globe show that non-motorized user safety is on the line [68, 169, 267].

Enabling safe and convenient travel for bicyclists and pedestrians is critical in achieving sustainable mobility goals; however, this requires a better understanding of road users' behavior and needs as well as the impact of geometric design and other countermeasures on safety outcomes. Adequate and accurate traffic data remains pertinent in supporting the procedural decision-making of improving roadway facilities (e.g., median crossings, bike lanes, and control strategies) to more safely accommodate non-motorized users. Additionally, accurate data are essential for evaluating the effectiveness of implemented treatments.

Traditional sensing technologies, such as loop detectors and pneumatic tubes, can be implemented in multimodal (i.e., motorized vehicles and non-motorized users) environments but cannot provide detailed trajectory data that are essential for performing behavioral and safety analyses. Thus far, bicycle and pedestrian safety studies focusing on user behavior and, more specifically, interactions between users have primarily utilized global positioning systems (GPS) or computer vision technology (i.e., vision-based). Despite being able to concurrently collect data across multiple locations, GPS data has its limitations when studying the interactions between different users, given that not all users are equipped with GPS devices.

Safety analysis through user interactions is feasible with sensors that collect data

from all users that are present at a location of interest. Fixed-point sensors, such as vision- and radar-based sensors, offer this capability. Vision-based sensors utilize video cameras that capture traffic data, which is later processed and analyzed to reveal roadway user behavior and interactions. While the vision-based approach has recently gained popularity in traffic monitoring, video cameras are still constrained by adverse lighting and weather conditions. Given that many bicycle-motorized vehicle crashes occur during reduced light conditions [233, 266], it is essential to explore different sensing technologies that can provide accurate data necessary for safety studies even under adverse environmental conditions.

Radar-based sensors are unaffected by external weather and lighting conditions but have been relatively underutilized for traffic monitoring. In addition, their capabilities have not yet been demonstrated in multimodal environments. There is a need to assess the feasibility of radar-based sensors in detecting and classifying pedestrians and bicyclists in multimodal environments prior to implementing them in large traffic monitoring studies in such environments.

This work is the first to bridge this gap by developing a mode classification framework to assign mode class to trajectories recorded by radar-based sensors. The proposed mode classification framework incorporates the following aspects to ensure transferability and flexibility: (1) it deploys the Support Vector Machine (SVM) algorithm as the classifier, as it enables robustness and interoperability; (2) it is capable of accommodating the unique properties of multimodal traffic captured by radar-based sensors as: (a) it can be applied in different feature spaces (e.g., speed/length or speed/acceleration feature space), (b) it incorporates training sample balancing strategy, and (c) it uses cross-validation for determining optimal SVM implementation and deriving performance metrics; (3) it is validated using data from two multimodal environments that vary in terms of traffic conditions and control type.

The rest of the paper is organized as follows: First, literature on traffic monitoring using vision-based and radar-based sensors is summarized. Next, a description of data collected with a radar-based sensor is provided. This is followed by the proposed mode classification framework and the numerical results of its application. The paper concludes with a discussion of this study’s findings and limitations as well as considerations for future exploration.

6.2 Literature Review

Fixed-point sensors such vision-based and radar-based sensors have the ability to simultaneously collect traffic data from multiple road users that are present on a site (e.g., intersection or road segment). In turn, these data can be used to study user interactions and obtain safety-related insights. Traffic monitoring applications of these sensor types and in particular, their performance with respect to detection and mode classification are discussed next along with their limitations.

6.2.1 Vision-based traffic monitoring

Video cameras have been widely implemented for traffic monitoring studies. The data obtained with video cameras are processed to identify road users through a three-step process consisting of user (1) detection, (2) classification, and (3) tracking [48]. During object detection, a real-world object, represented as a set of pixels, is identified as different from its background. Detection is followed by assigning the object to a class, e.g., human, car, bicycle. Once this is complete, tracked observations from consecutive frames are connected to form a classified object’s trajectory.

Object detection and tracking are affected by the video image quality and the placement of the video camera. Aspects such as image resolution and rendered colors are critical for differentiating among different road users [48]. The absence of poles,

the presence of buildings or vegetation affect camera placement and therefore, the view angle. Overall, cameras should be placed so that occlusion, shadows, and light reflections are minimized.

Video cameras have been deployed for traffic monitoring in multimodal environments. While the demonstrated classification frameworks perform exceptionally in detecting and classifying motorized vehicles, they underperform in the case of pedestrians and especially, bicyclists [273]. Non-motorized users' low detection and classification accuracy has been attributed to the variability in color and shape [1]. However, pedestrian detection and classification can often be facilitated by the fact that pedestrians travel on designated facilities, i.e., sidewalks or crosswalks. Bicyclists, on the other hand, do share the roadways with motorized vehicles (e.g., in the absence of bicycle facilities), which can further degrade detection accuracy and cause partial occlusion problems. Lastly, due to their size, bicyclists and pedestrians can only be detected when they are in close proximity to the camera; for motorized vehicles this distance can be larger [141].

Another limitations of the aforementioned studies is that they took place in daylight and clear weather conditions. Adverse lighting, i.e., low or extreme illumination, and weather, e.g., fog, conditions have been found particularly untactful with respect to user detection [13, 152]. Alternative video camera technologies, such as thermal infrared cameras, can be used in adverse environmental conditions. Studies that utilized thermal infrared cameras and took place in highways during thick fog or snowy conditions have reported high detection accuracy [105, 106]. However, the implementation of infrared cameras in multimodal environments is challenging as this technology is associated with low image resolution. Non-motorized user size remains an issue regarding detection with the use of thermal infrared cameras; pedestrians appearing at a distance greater than 50 meters from the sensor cannot be captured by

such cameras [179]. Bicyclist detection is further challenging due to shape of bicycles and their irregular movement, e.g., they can be traveling within or separated from motorized vehicle traffic, [77]. Lastly, infrared cameras could also face limitations in detection during daylight conditions due to the low contrast between moving objects and their background [151]. While recent efforts have addressed issues related to low image quality, pedestrians were not extensively studied and bicyclists were not included at all in the studied environments [31, 140].

Video camera traffic monitoring applications in multimodal environments might also suffer from partial occlusion. While this type of occlusion is temporary, it could be occurring when, for example, turning-vehicles are blocking bicyclists and pedestrians and therefore, resulting in trajectory discontinuities [13, 141]. Several researchers have proposed the deployment of unmanned aerial vehicles (UAVs) equipped with video cameras to obtain a “bird’s eye view” [33, 151]. While differentiating between pedestrians and motorized vehicles was successful studies using UAVs, pedestrians could be missed as they look similar to fixed roadside objects (e.g., light poles) [33, 151].

In a nutshell, the implementation of video cameras for traffic monitoring purposes is limited by several factors such as adverse weather and lighting conditions as well as partial occlusion. These limitations are more pronounced when detecting and classifying pedestrians and bicyclists. Alternative camera technologies, i.e., thermal infrared cameras, as well as alternative placement of video cameras, e.g., on a UAV, have the potential to address some of the limitations of conventional video cameras, but they still perform poorly in multimodal environments. Moreover, thermal infrared cameras are associated with higher procurement costs [12], while UAVs are limited by battery life [72] and stability issues [151].

In recent years additional vision-based technologies have emerged, such as the Light Detection and Ranging (LiDAR) sensors. LiDAR imaging systems have been

extensively studied as autonomous vehicle (AV) equipment that assist in sensing the vehicle's surroundings as discussed in the review by Royo & Ballesta-Garcia [207]. Several studies have demonstrated that LiDAR sensors can be used for traffic monitoring through effective frameworks that have been developed to facilitate the mode classification task using such data [124, 268, 275, 276]. For this study, LiDAR sensors are not considered due to limitations related to their high purchase and deployment costs.

6.2.2 Radar-based traffic monitoring

A radar-based sensor utilizes an antenna, which transmits electromagnetic waves toward the area of interest. Once the waves hit an object they are reflected back to the sensor enabling it to calculate the distance between its location and the object and, simultaneously, to estimate the object's speed based on Doppler effects. Given the radar-based sensors' ability to provide accurate distance and speed measurements and their non-intrusive nature of the installation, they have been utilized as an alternative to loop detectors [123] and pneumatic tubes. An advantage of radar-based sensors compared to vision-based methods is that their operation remains unaffected from external conditions such as weather (e.g., fog) and lighting (e.g., extreme illumination or lack of light) [156, 157].

While radar-based sensors have been mostly utilized for speed [96, 123, 135] and traffic demand (e.g., count data collection) monitoring [96, 215], some researchers have demonstrated that these sensors can also be used for analyzing intersection delay [216] or assessing safety [214, 243]. Speed data collected by radar-based sensors have also been used for emission-related research [96, 272].

Existing literature includes several radar-based vehicle classification studies, e.g., differentiate between trucks and personal vehicles [96] or motorcycles and other types

of motorized vehicles [35]. Vehicle classification with radar-based sensors relies on vehicle kinematic; measurement data such as position, speed, and vehicle length have commonly been used to classify different vehicle types. As an example, speed data along with vehicle length obtained by radar-based sensors, have been used to differentiate between trucks and cars [96] or motorcycles and cars and trucks [35]. Only a few radar-based sensor studies have taken place at locations where road users other than motorized vehicles are present [215, 216, 272]. While none of these studies differentiated between the type of motorized vehicles (e.g., cars vs. buses or trucks), processes were developed to filter out recorded pedestrians that were traveling either on the sidewalk or the crosswalk. The classification was performed either based on the location pedestrians travel or by excluding trajectories that were parallel to the stop bar [215, 216, 272].

6.2.3 Summary of the literature

This study focuses on multimodal environments where pedestrians, bicyclists, and motorized vehicles are present. As revealed by the review of the literature on radar-based traffic monitoring, there are no studies to-date considering all three modes. The inclusion of bicyclists essentially prohibits a mode classification framework from relying on location-specific data. This is because bicyclists often use different parts of the roadway such as sidewalks or shared paths where they interact with pedestrians, bicycle facilities that can be physically separated from motorized vehicles, or they may travel on traffic lanes where they interact with motorized vehicles. Hence, even in the presence of bicycle-specific infrastructure treatments, mode classification could still be relevant and needed.

The presence of multiple road users in the same traffic scene has the potential to create imbalance in the mode classification problem, and require various features to

be considered for the classification since different mobility patterns might be present depending on traffic operating conditions. The imbalance is caused due to the fact that bicyclists and pedestrians tend to be fewer compared to motorized vehicles. Depending on the traffic scene, non-motorized users might be moving with speeds that are comparable to those of motorized vehicles, i.e., due to congestion. An additional challenge with the use of any sensor data is the measurement noise. Existing research is limited in that the developed so far mode classification frameworks have been tested for single sites and have provided limited information on their transferability and generalizability. In addition, existing studies do not deal with data imbalance. Overall, there is a need to develop a transferable mode classification framework that can handle imbalanced datasets of motorized vehicles, bicyclists, and pedestrians, while accounting for radar-based sensor measurement inaccuracies.

The proposed framework addresses all of these limitations. First it includes different road user types, a condition that has not previously been addressed in the literature. Existing studies that use a radar-based sensor for traffic monitoring have taken place at locations where motorized vehicles, pedestrians, and bikes are not all present simultaneously. The presence of the three different modes further complicates the data imbalance issue, that is also explicitly addressed in this framework, through the implementation of a weighted SVM classification. SVM classifiers are robust and in addition to being able to address imbalanced data they can account for measurement noise that is inherent to radar-based data. Lastly, one challenge related to mode classification of trajectories recorded with radar-based sensors is the impact of varying traffic conditions on recorded trajectories and therefore, their potential to be classified as one mode versus the other. A motorized vehicle recorded on the main road of an unsignalized intersection is likely to have a very different speed profile compared to that of a motorized vehicle that arrives at a signalized intersection during the red

signal indication or one traveling on a congested roadway, which could easily bias the mode classification process. The proposed framework addresses this issue by assessing whether the classification should rely on (i) speed only, (ii) speed and vehicle length, (iii) speed and acceleration, or (iv) speed, acceleration and vehicle length measurement to assign each trajectory to its corresponding mode, ensuring that this framework can be effectively implemented at any traffic scene.

6.3 Radar-based sensor configuration and data format

For the purposes of this study data collected using a generic radar-based motion and presence sensor. The device detects moving objects that are approaching it and records their position every 0.5 seconds. Each detected moving object is assigned a unique ID. Moving object position information reported by the radar refers to the middle point of the front bumper of the vehicle, i.e., the location of the radar-based sensor setup is represented by coordinate (0, 0) in a Cartesian plane and the position of the front bumper of a vehicle is reported in reference to this origin. Using the vehicle position over time, the device can estimate vehicle speed and length. The output data file records the following metrics for each detected moving object within the device's range, which is approximately 200 meters (600 feet), every 0.5 seconds: X coordinate, Y coordinate, speed, and, vehicle length, essentially allowing for trajectory construction for each of the recorded moving objects. A sample of data collected with this device is shown in Figure 6.1. Different colors and sizes represent different speed levels and vehicle lengths that were recorded for the various data points of each moving object.

While radar-based sensors are capable of operating in adverse weather and lighting conditions, they also present some limitations. Compared to video cameras radar sensors provide limited contextual information (e.g., study site characteristics such as

geometric ones that are not visible through the collected data). Data pre-processing needs are also higher for radar-based sensors. This is due to the fact that some preliminary data collection experiments and analysis are often necessary before finalizing the sensor placement to ensure that collected data with minimal noise. In the case of video cameras, one can immediately see whether one placement is better than another in terms of capturing the area of interest, minimizing occlusion, etc. Radar-based sensors can also be noisy in terms of vehicle length measurements.

6.4 Mode Classification Framework

The proposed mode classification framework assesses whether trajectories recorded by a radar-based sensor can be correctly classified based on motion (e.g., speed) and physical (e.g., vehicle length) features. The Support Vector Machine (SVM) algorithm has been chosen as the classifier since it has a set of mathematical properties that make it ideal for dealing with different roadway environments and traffic conditions. The proposed mode classification framework has been structured with the following properties: (1) trajectory normalization scheme (i.e., a process for selecting only a set of records per trajectory); (2) deployment of different feature space combinations (e.g., speed only or speed along with vehicle length); (3) deployment of different sample balancing strategies, and (4) use of cross-validation to obtain the optimal values for SVM parameters and average scores across different performance metrics.

6.4.1 Support Vector Machine Algorithm for mode classification

The varying nature of multimodal traffic creates challenges that need to be addressed by the proposed mode classification framework. We identify the following three challenges: (1) unbalanced classes: the potential of having a higher number of motorized vehicle trajectories than bicyclist and pedestrian trajectories [273]; (2)

applicability various operating environments: the proposed framework should be flexible to work in both signalized and unsignalized intersections and under varying traffic conditions over time for the same site; (3) inaccurate or “noisy” measurements: noise in radar-based sensor measurements can be attributed to cross and/or adjacent traffic, as well as sensor confusion [217]. In addition, some data points might be reported with wrong X and/or Y coordinates or a specific recorded object might have variations in its length measurements.

SVM, that falls under the Support Vector Classifiers (SVCs) category, can outperform existing classifiers, e.g., logistic regression, discriminant analysis, or trained-based methods (e.g., Neural Networks (NN), random forests (RF)) for several reasons. SVCs can be used for multi-class classification problems; this study considers three classes: motorized vehicles, pedestrians, and bicyclists. Compared to trained-based methods such as NN and RF, SVCs are semi-interpretable; the user may later interpret which features (i.e., independent variables) and observations (e.g., trajectories) improve the performance of the classifier. Other interpretable methods, like regression models, are not flexible in using all data points from a trajectory as parameters; one would have to use descriptive statistic values (e.g., mean, standard deviation) to develop a classification model. Most importantly, SVC algorithms are robust and capable of dealing with unbalanced and/or noisy datasets as explained next.

Among a set of possible hyperplanes that separate two (or more) classes in a p -dimensional space (where p is the number of features), SVCs identify the one that maximizes the distance between these classes. Compared to other classifiers “Support vectors” are points close to the class boundaries that support (i.e., inform) the definition of the hyperplane. However, support vectors are not necessarily the most extreme points of each class. The fact that class boundaries are set to maximize the distance between the classes but without relying on the most extreme points, that

could correspond to noisy measurements, allows for robust solutions to be obtained; the solutions are not sensitive to a specific set of points which could correspond to less frequently occurred or noisy values. Other classifiers such as k-nearest neighbors, simply define class boundaries based on the most extreme points. Consequently, SVCs allow for some misclassification with an overall robustness trade-off. In the case of SVC the user can determine the parameter related to misclassification to obtain a solution that better fits particular needs. This flexibility is also important for imbalanced and/or noisy datasets as the user can determine to which extent misclassification is accepted for a particular class. For the present study the ability to handle unbalanced datasets as well noisy measurements is especially important.

Compared to other SVCs, SVM can deal with non-linearly separable classes. Assuming a dataset, consisting of N observations: $(x_1, y_1), (x_2, y_2), \dots, (x_N, y_N)$, where x_i is an input vector consisting of the characteristics based on which the classification rule will be developed, and y_i denotes the class of the i^{th} observation. If the different classes (i.e., mode types) are not linearly separable in the initial feature space, a fixed feature space transformation, e.g., $\phi(x_i)$ can be defined to map each data point (x_i, y_i) into a higher dimensional feature space in which the transformed data are linearly separable. These transformations are facilitated with the use of kernel functions. A kernel function denoted as $K(x_i, x'_i)$, where x'_i is the input vector in the transformed feature space, can take various forms such as polynomial, sigmoid, etc. SVM's ability to define non-linear boundaries is essential for mode classification as in certain cases different road users have comparable, and so non-linearly separable, speed profiles. This can be the case in more congested conditions where motorized users move in lower speeds that are similar to those of non-motorized users or in shared paths where bicyclists move at lower speeds since they share the space with pedestrians.

SVM solves the following mathematical program:

$$\underset{w, \beta, \epsilon}{\text{minimize}} \frac{w^T w}{2} + C \sum_{i=1}^N \epsilon_i \quad (23)$$

subject to:

$$y_i(w^T \phi(x_i) + \beta) \geq 1 - \epsilon_i, \forall i = 1, 2, \dots, N \quad (24)$$

$$\epsilon_i \geq 0, \forall i = 1, 2, \dots, N \quad (25)$$

where w and β are parameters that define the hyperplane in p -dimension space, e.g., when $p = 2$ they represent the slope and intercept of the line. The product $w^T w$ is the square of the margin between two classes. C and ϵ_i are parameters related to misclassification. C bounds the sum of the ϵ_i 's; the latter is a slack variable that informs on the placement of the i^{th} observation with respect to the margin and the hyperplane. C expresses the severity of the violation that is tolerated. Smaller C values result in small-margin hyperplanes, while larger values create larger-margin hyperplanes. The first constraint is related to the class assignment of the i^{th} observation.

The ability to deal with non-linearly separable classes is crucial in the case of varying traffic conditions (e.g., different control type and congestion levels), as these conditions affect the motion of the different modes. Depending on traffic conditions, motorized vehicles might need to slow down and adopt a lower speed comparable to those of bicycles and even of pedestrians. Therefore, for the simple case of separating a set of speed measurements of a bicycle from those from a slow motorized vehicle, it can be argued that this cannot be facilitated with a straight line; SVM allows for non-linear lines.

In a nutshell, SVM as a classifier separates two (or more) classes by maximizing the distance between them based on points close to the class boundaries. The classes can

be separated either in a linear or non-linear way. These mathematical properties make SVM ideal for the case of multimodal classification in varying traffic conditions. The maximization of the distance between two (or more) classes based on the points close to the boundaries eliminate noise and make SVM flexible in dealing with unbalanced datasets. The ability to define non-linear boundaries between the classes favor the mode classification under varying traffic conditions.

6.4.2 Data labeling

Ground truth data were collected using a video camera that was mounted next to the radar-based sensor to be able to have the same field of view. Each object ID recorded by the radar-based sensor was compared with the video footage to assign one of the following labels: passenger car, bus/truck, pedestrian, bicyclist, or other; for the rest of the paper this task is referred to as vehicle labeling.

In order to facilitate vehicle labeling, a tool (Figure 6.2) was developed that allowed for: (1) synchronization of the radar-based sensor and video data and (2) plotting of the trajectory, speed, and length values for each object ID recorded by the radar-based sensor. The tool was able to read the first and last time stamps for each object ID recorded by the radar-based sensor and display the same time period from the video data footage. Then the user was able to view the video that corresponds to each trajectory and assign the correct mode type to it.

At the initiation of the tool's interface (Figure 6.2), the user is asked to insert the path for the video and radar-based sensor data files as well as provide the time differences between the two data sources. This way the code can synchronize the two datasets. The radar-based sensor records are grouped per object ID and listed on the left part of the interface. By selecting a recorded object ID, the following actions occur:

1. the part of the video recording that corresponds to the radar-based sensor's record appears on the right side of the interface so that the user can view it;
2. a plot (in red) of the selected vehicle trajectory appears to the left of the video; and
3. vehicle speed and length as recorded by the radar-based sensor are plotted and appear under the video recording view.

6.4.3 Feature vector formulation

The SVM algorithm classifies an observation to one of the available classes based on a set of characteristics that are provided as input, denoted as “feature vector”. For the proposed classification framework there are several features that could be considered to ensure a low error classification. These features are measurements such as speed, acceleration, and vehicle length. Acceleration, when not directly reported from the device, can be estimated using the speed and time information. Depending on the traffic scene, it is hypothesized that different combinations of feature vectors can increase classification performance, and thus, it is advised to consider and evaluate those combinations and select the most appropriate one. The grounds of this hypothesis are explained below.

Speed and acceleration have been used in the literature to differentiate between various motorized and/or non-motorized modes [70, 225, 239, 269, 271]. Speed profiles of motorized vehicles, bicyclists, and pedestrians in free-flow conditions can be quite distinct; under congested conditions different road user types might travel with comparable speed [142, 255]. Therefore, it is unlikely to use speed only for classification. Vehicle length has been used to differentiate between various types of motorized vehicles [35, 96]. Overall, depending on the specific site, one measurement type could

be more important than another. The proposed framework aims at assessing which measurement type combinations achieve the best classifier performance.

The measurement values for each feature vector combination need to be aggregated to allow for an efficient implementation of the SVM algorithm. Mean values for speed, acceleration, and vehicle length can be used to build feature vectors. However, a single averaged value cannot capture the variability that is potentially present in a trajectory. For example, a motorized vehicle that approaches an intersection and stops at a red traffic signal would be represented with a low mean speed that could be similar to a pedestrian speed.

Alternatively, multiple measurements from each trajectory can be used to demonstrate the potential variability of each measurement rather than a single mean value. However, as recorded objects move at different speeds, each of the trajectories consists of a different number of data points (e.g., for the same length a pedestrian would have more points per trajectory compared to a moving motorized vehicle). In this case the algorithm would have to receive and then process more data and in turn, learn better the pedestrian compared to the motorized vehicle. Moreover, regardless of mode, some trajectories are shorter than others, which occurs as they are only partially included in an area where data collection takes place. Again, this phenomenon would result in SVM receiving more information for some moving objects versus another. A trajectory normalization scheme is therefore, proposed to allow for obtaining the same number of data points per trajectory, without losing the integrity of the overall trajectory, in order to ensure consistency in the feature vector inputs. In particular, ten data points are extracted along each trajectory, i.e., at the 10%, 20%, etc. of its covered length, A normalization scheme by ten data points per trajectory is a balance between the complexity of the training scheme and the availability of the training data since using more data points would also increase the number of training

samples required. Depending on the sensor utilized for the analysis and in particular the recording frequency, as well as the configuration of the site, a different value (i.e., other than ten) could be considered for the trajectory normalization.

Speed and acceleration measurements are extracted and interpolated; these values are then used to form the feature vectors. The mean vehicle length for each trajectory is estimated by averaging all obtained length measurements for that trajectory. Unexpected variability in length measurements was found in many trajectories, therefore, obtaining a mean value across all points per trajectory would result in a more representative vehicle length value.

Overall, the following feature vectors consisting of speed (S_j), acceleration (A_j), and/or mean length (L_m) measurements are considered:

1. $[S_1, S_2, \dots, S_{10}]$
2. $[S_1, S_2, \dots, S_{10}, L_m]$
3. $[S_1, S_2, \dots, S_{10}, A_1, A_2, \dots, A_{10}]$
4. $[S_1, S_2, \dots, S_{10}, A_1, A_2, \dots, A_{10}, L_m]$

6.4.4 SVM algorithm implementation

Following the feature vector selection is the selection of the appropriate kernel function and values for parameters C and γ . C reflects the misclassification tolerance. The second parameter, γ , is present in SVM formulations that involve certain kernel functions, e.g., Radial Basis Function (RBF), polynomial, etc. Techniques available in multiple programming languages can determine the appropriate kernel function and the parameter values; this work proposes the implementation of the exhaustive grid search [242] that is applied for each feature vector. The grid search receives a set of kernels to

test and for each kernel a combination of C and γ values. Those combinations create a grid and the grid search algorithm will test each node of the grid (for each kernel type) and then, it will produce the solution that yields to the most accurate results. It is clarified that the optimal solution is obtained with cross-validation. Therefore, as grid search considers multiple kernels and multiple combinations of C and γ per kernel, it can be a time-consuming process. Grid search receives the feature vector, the class of each trajectory as well as ranges for C and γ parameters as input.

In mode classification problems, imbalanced classes, i.e., classes with different sizes, might occur as the non-motorized mode share is lower compared to that of motorized vehicles. This imbalance is likely to affect the performance of the classifier, as the algorithm would be well-trained in recognizing passenger cars (which is expected to be the majority class) , but would be less successful in detecting the rest of the modes (or in other words the minority classes). Assigning a weight to all objects of a class that is inversely proportional to the class size penalizes the classifier when it miss-classifies observations. The idea is to implement a greater penalty when the classifier miss-classifies minority class observations compared to majority class observations. For the rest of the paper the term “balanced” approach refers to the case where weights (other than one) have been assigned. The weights are calculated as follows:

$$w_j = \frac{n}{kn_j} \tag{26}$$

where w_j is the weight assigned to the j^{th} class; n is the total number of observations; k is the number of the different classes; n_j is the number of observations that belong to the j^{th} class.

To summarize, the proposed framework assesses eight SVM models that vary in: (1) the types of feature vectors (e.g., speed or speed and acceleration), and (2) the

implementation of weights or not to account for imbalanced classes. The combination of the two balancing approaches along with the four feature vectors (see section 6.4.3) results in a total of eight models.

All eight models need to be evaluated with respect to their predictive power. The objective is for a given traffic scene to identify the optimal model to perform the mode classification. Overall, this formulation is robust and flexible in that it can incorporate various traffic characteristics with respect to traffic conditions, control type, and mode share that may appear in a roadway environment. Lastly, for each model, the respective dataset is divided into two sets: one used to train the model (train set) and another to test the model (test set). To assess the SVM performance for every model cross-validation is recommended; use multiple pairs of train-test sets to obtain average scores across these sets. Avoiding the use of a single train-test set can eliminate the chance of randomly obtaining a very good or bad training set, and in turn creating a model with high bias or variance. Five pairs of train-test sets are generally considered adequate.

6.5 Experimental Tests

Traffic data was collected with a radar-based sensor at two urban intersections where pedestrians, bicyclists, and motorized vehicles are present to test and validate the proposed mode classification framework. Two test sites were chosen to allow for testing of the proposed framework under different traffic operating conditions. In particular, the chosen intersections in this study vary in terms of control types and the roadway configurations.

6.6 Test sites

6.6.1 Test site 1: Unsignalized intersection

Test site 1 is the intersection of North Pleasant Street and Butterfield Terrace (Figure 6.3), in Amherst, Massachusetts. Data was collected on two different weekdays between 9:30 to 11:30 AM and 10 AM to 12 PM.

North Pleasant Street is an one-way per direction street, while Butterfield Terrace is a low-volume local street. This is an uncontrolled intersection with no yield or STOP signs present. In North Pleasant Street, each direction of traffic includes a bike lane adjacent to the travel lane and a sidewalk. Generally, vehicles and non-motorized users travel on North Pleasant Street without interruptions, which allows for data collection during free-flow traffic conditions. This is important because bicyclists and motorized vehicles can develop different speed profiles depending on whether they are experiencing near free-flow or congested traffic conditions. Only a few motorized vehicles were observed stopping for crossing pedestrians or decelerating prior to turning right at Butterfield Terrace, and these conditions allowed for obtaining wider speed ranges for motorized vehicles. This in turn, allows for the algorithm's robustness to be further tested.

The radar-based sensor was mounted on a light pole, located a few meters downstream of the studied intersection to capture the northbound direction of travel (Figure 6.3). While the majority of bicyclists used the bike lane, some were observed using the right or the left sidewalk similarly to pedestrians. As a result, the roadway position of these users cannot be used as a criterion to differentiate between motorized and non-motorized users, as it was done in other studies to filter out pedestrians [215, 272].

Table 6.1 presents the number of motorized vehicles, pedestrians, and bicyclists

recorded per site and additionally, traffic flow, mean speed, mean acceleration, and mean length values for the same modes. For test site one (i.e., unsignalized approach) data was collected during two different days; the total hours of video data collection are noted in a parenthesis. Traffic flow information has been estimated from the video data by counting the number of motorized vehicles, pedestrians, and bicyclists that were present in the traffic scene and approaching the sensor over the data collection period in hours. However, due to some trajectories being incomplete, fewer motorized vehicles were used in the analysis. The majority of the recorded modes in test site one were motorized vehicles, followed by pedestrians and bicyclists. Based on Table 1, it can be seen that the two data collection days are comparable in terms of level of traffic demand. Given the low traffic flow, this segment was operating under free-flow traffic conditions.

Table 6.1: Summary of data collected per test site

Test site	Day (hours)	Mode	Flow ^a (veh/h)	No. trajectories ^b	Speed ^b (km/h)	Acceleration ^b (m/s^2)	Veh. Length (m)
Unsignalized	Day 1 (2)	Motorized veh.	283	537	44.9	0.0007	4.4
		Bicycles	4	5	21.2	0.0068	2.4
		Pedestrians	28	50	4.5	0.0079	2.2
	Day 2 (2)	Motorized veh.	291	527	48.3	0.0092	4.5
		Bicycles	16	27	22.1	-0.0009	2.8
		Pedestrians	22	37	4.4	-0.0057	3.0
Signalized	Day 3 (1)	Motorized veh.	542	304	10.5	0.0023	4.8
		Pedestrians	49	34	3.6	0.0031	3.1

^a Values estimated based on video footage.

^b Values based on the final number of trajectories used to develop the model.

6.6.2 Test site 2: Signalized intersection

The second test site is the intersection of Main and State Streets in Northampton, Massachusetts downstream of the merging point of Elm and West Streets in the

eastbound direction as shown in Figure 6.4c. One hour of data were collected during the 10:00 to 11:00 AM period on a weekday.

The radar-based sensor was mounted on the traffic light pole at the intersection of Main and State Streets, capturing the traffic approaching from West and Elm Streets (see Figure 6.4b, 6.4c). The captured approach has one right-only traffic lane, a bike lane, two through lanes, and a left-only lane. Pedestrians are walking on the sidewalk on the right of the traffic and bike lanes. Due to the sensor placement, pedestrians crossing the intersection were not captured.

The signalized approach of interest operated under a cycle length of 110 seconds, a green time for the through and right-turning vehicles of 50 sec, and a green time for left-turning vehicles of 30 sec. Given the low flow of motorized vehicles and the aforementioned signal settings, traffic conditions at this intersection approach were undersaturated. The presence of signalization, however, allowed for capturing different speed and acceleration profiles compared to the unsignalized test intersection (where users mostly travel at free-flow speed).

6.6.3 Data preparation

Vehicle labeling Using the vehicle labeling tool presented in section 6.4.2, a total of 1341 object IDs were labeled as “Passenger car”, 27 as “Bus/Truck”, 32 as “Bikes”, 121 as “Pedestrians”, and lastly, 813 were labeled as “Other”; the latter category included either broken trajectories or those with fewer than five data points. Given the small number of large vehicles, i.e., buses and trucks, compared to passenger cars, they were all merged in one category, namely “motorized vehicles”. Differentiating between motorized vehicles using radar-based sensor trajectories has already been addressed in the literature [96] and is out of the scope of this study.

The task of labeling the recorded trajectories is a relatively time-consuming step.

For the present implementation, 3-4 hours of video data were labeled in approximately 8 hours. It should be noted that the labeling process is a one-time task, essential to train and test a mode classification model for a given traffic scene; trajectories recorded later on can be automatically classified through the SVM implementation with no additional video records needed. In addition, these labeled trajectories can be used for example, for safety analysis through surrogate safety metrics obtained from trajectories as well as for energy and emission estimation.

In total 1521 trajectories were used to develop and test the framework consisting of 1368 motorized vehicles (78% in site 1 and 22% in site 2), 32 bicycles (100% is site 1), and 121 pedestrians (71.9% in site 1 and 28% in site 2). Motorized vehicles are approximately 90% of the recorded trajectories, while the respective percentages for pedestrians and bicycles are 7.9% and 2.1%. During the video data collection several more vehicles, pedestrians, and bikes were recorded however, occlusion resulted in some of the trajectories being broken and therefore, removed from the analysis. Broken trajectories were mostly found at the signalized approach and can be attributed to occlusion due to radar positioning; ideally, the sensor should be placed at the horizontal part of a traffic pole as demonstrated in the studies by Santiago-Chaparro et al. [215, 216] as other placements capture traffic scenes with an angle that may benefit occlusion.

Feature vectors Speed and acceleration for a set of motorized vehicles and pedestrians as well as for the bicyclists were visualized using whisker plots. Additionally, mean values for speed, acceleration, and vehicle length were estimated (Table 6.1). Generally, data visualization can provide some evidence regarding the appropriate kernel function as it can indicate whether the dataset is linearly separable or not. Additionally, visualizing the data may illustrate which measurements can be more critical for the classification. However, data visualization should not be used as a decision making

tool in the present framework.

Records from test site 1 (unsignalized approach) show that the three user types, i.e., motorized vehicles, bicyclists, and pedestrians, have distinct speed profiles as shown in Figures 6.5a, 6.7a, and 6.9a . Similarly, distinct speed profiles between pedestrians and motorized vehicles were observed for test site 2 (signalized approach). Essentially, bicycles (site 1) traveled with speeds comparable to those of motorized vehicles at site 2; this is an indication that speed alone might not be a good predictor of vehicle class especially between motorized vehicles and bicycles. Acceleration profiles shown in Figures 6.6, 6.8, and 6.9b are quite similar across all user types for both sites, indicating that acceleration might not be a good predictor for the given traffic scenes. Lastly, vehicle lengths vary between motorized and non-motorized users, however, the estimated mean values for pedestrians and bicyclists are quite similar as shown in Table 6.1. While visualizing the data is informative, testing all four feature vectors as well as balancing strategies should not be omitted.

Grid search output

The appropriate kernel as well as optimal values for C and γ for each one of the four feature vectors and balancing approach combinations were determined through grid search. The following kernels were tested: linear, polynomial, RBF, and sigmoid. It is highlighted that grid search for the same feature vector should be applied separately for the balanced and unbalanced approach as balancing the data affects the misclassification error which is related to the value of C . For both parameters, the range that was provided as input was $[10^{-10}, 10^{10}]$; linear kernels do not use γ parameter so in that case only a range for C was provided. The range is big enough to ensure that the algorithm will assess multiple C and γ combinations. The RBF kernel, shown in equation 27, was found to be the most appropriate for all eight SVM

models.

$$K(x_i, x'_i) = \exp(-\gamma \|x_i - x'_i\|^2) \quad (27)$$

where x_i and x'_i correspond to the input vectors for the i^{th} observation in the non-linearly and linearly separable feature spaces, respectively.

Different optimal C and γ values were identified for each feature vector as shown in Table 6.2. The differences in C values indicate that in all cases, with the exception of the speed-acceleration combination, there is a larger penalty for misclassification.

Table 6.2: C and γ values for the eight SVM models

Balancing Approach	Feature Vector	C	γ
Unbalanced	Speed	1	10
	Speed & length	1	100
	Speed & acceleration	0.1	10
	Speed, acceleration, & length	10,000	0.01
Balanced	Speed	0.1	100
	Speed & length	1	100
	Speed & acceleration	10	100
	Speed, acceleration, & length	10	10

6.6.4 Results

Various performance measures, namely: precision, recall, and accuracy, were used to compare the eight models. The reported values for precision, recall, and accuracy correspond to the average values across the five different train-test sets that were created for each one of the eight models.

Precision is the ratio of the True Positives over the sum of True and False Positives as shown in equation 28. Precision expresses the number of cases correctly classified as positive, e.g., the number of bicycles that were correctly classified as bicycles, over

the total number of cases labeled as positive, i.e., the number of recorded objects that were classified as bicycles [230].

$$Precision = \frac{True\ Positives}{True\ Positives + False\ Positives} \quad (28)$$

Recall is the ratio of True Positives over the sum of True Positives and False Negatives, as shown in equation 29; in other words, recall expresses the number of correctly classified as positive cases over the number of positive cases in the data [230]. In this study's context, recall is calculated as the number of times a recorded object was detected as a bicycle over the number of times a bicycle was actually present.

$$Recall = \frac{True\ Positives}{True\ Positives + False\ Negatives} \quad (29)$$

Accuracy, which expresses the effectiveness of the classifier, is defined as the number of correctly classified responses over the entire sample (n) of observations, as shown in equation 30.

$$Accuracy = \frac{1}{n} \sum_{i=1}^{n-1} 1(\hat{y}_i = y_i) \quad (30)$$

where y_i is the observed response and \hat{y}_i is the predicted response.

The precision and recall results for the three-mode classification are shown in Figures 6.10 and 6.11, respectively. Accuracy across all eight models is presented in Figure 6.12.

Precision for motorized vehicles for the unbalanced approach ranges from 93-96% across the four models with the lower values reported for the speed and acceleration feature vector. In the balanced approach precision 93-99%, with speed and acceleration feature vector reporting the lowest value. Overall, precision is maximized in the balanced approach with the utilization of the feature vector that includes speed and

length.

Precision for pedestrians for the unbalanced approach was overall lower compared to that of motorized vehicles, ranging between 87-94% across the four models. Similar to the motorized vehicle classification, the balanced approach either slightly improved or did not affect the precision for pedestrians. Generally, pedestrians and motorized vehicles maintain higher precision scores compared to bicyclists. When to each other, pedestrians and motorized vehicles have distinct lengths and for this specific dataset they were also found to have quite distinct speed profiles. However, as both road users have similar acceleration profiles, the inclusion of acceleration does not improve precision score for either road user. Pedestrian and bicyclist values for vehicle length were very similar, while their speeds had close values in some cases. The balanced approach penalizes the algorithm in case of miss-classification, and so performs better when differentiating between these two modes. Hence, higher True Positive rate is reported for pedestrians.

Lastly, precision for bicycles was low in both the unbalanced and balanced approaches, ranging from 20-75% and 48-75%. The balanced approach increased precision for bicycles. This is because the bicycle class is the most underrepresented class of the three in the sample and the use of weights facilitates the algorithm's "learning" of this class. However, given the relatively low precision scores, even in the balanced approach, the classifier appears to be mistakenly assigning non-bicycle trajectories to the bicycle class.

In terms of recall, motorized vehicles reported the highest recall values (higher or equal to 92%) compared to the other two modes in both the unbalanced and balanced approaches (Figure 6.11). High recall values indicate that when motorized vehicles are present, they will most likely be correctly detected. The balanced approach causes a small drop in the recall values of motorized vehicles. This is because the inclusion of

weights when implementing SVM results in a trade-off between precision and recall. When assigning weights proportional to the class size, precision is improved for the majority class due to the fact that False Positives rate drops; the algorithm is forced not to falsely assign pedestrians or bicyclists to the motorized vehicle class. Recall on the other hand drops for the majority class as False Negatives value might increase; the algorithm is less penalized if it falsely “misses” a motorized vehicle compared to missing a pedestrian/ bicyclist. For the present sites, the aforementioned changes in precision and recall are minor. Overall, for motorized vehicles vectors that contain vehicle length information maximize both precision and recall for motorized vehicles. The recall results for pedestrians are slightly influenced by the use of a balanced vs unbalanced approach and besides the case of the speed-only feature vector, the balanced approach results in lower recall values for pedestrians. The drop in pedestrians’ recall along with a simultaneous increase in bicyclists’ recall values can be easily explained; the inclusion of weights proportional to the class size force the algorithm to correctly identify the minority class. Essentially, pedestrians’ False Negative rate drops as the algorithm learns to identify bicycles instead. The fact that speed-only feature vector results in the highest Pedestrians have distinctive speed profiles (see Figure 6.7) compared to the great majority of motorized vehicles (see Figure 6.5) and bicycles (see Figure 6.9); therefore, speed alone stands as a good classifier for pedestrians. The inclusion of acceleration and vehicle length in addition to speed, improved pedestrians’ recall while that was not the case for speed and acceleration or speed and vehicle length feature vectors.

For bicycles in the unbalanced approach zero recall values were reported, indicating that in those cases the classifier could not classify any observation as a bicycle. As a result, in the case of bicycles, the balanced approach is very beneficial for the classifier’s ability to detect and classify them correctly given the small bicycle samples in the

dataset. Additionally, the inclusion of mean length in the feature vector increases the recall scores for bicycles.

Overall, as shown by the results, the feature vector consisting of speed and mean vehicle length is able to simultaneously improve recall scores for all three modes. The ability of radar-based sensors to estimate and report vehicle length is a unique feature that is important when focusing on mode classification, but is lacking from other sensing technologies (e.g., GPS devices). This advantage is not degraded due to the fact that there are some fluctuations in the reported length values, possibly caused by partial occlusion.

Accuracy values reached or exceeded 95% except for the case of speed and acceleration feature vector (Figure 6.12). In the unbalanced approach the classifier focuses on better learning the majority class (motorized vehicles in this case) as this is a simultaneously safer and easier way of increasing its correct response rate. However, there are slight differences in accuracy between balanced and unbalanced approaches; one should focus on precision and recall metrics as they can better explain where the misclassification error occurs and by analyzing precision and recall the benefit of balancing the data is more apparent.

Overall, the balanced approach outperforms the unbalanced one. The current dataset is highly unbalanced and the inclusion of weights that are proportional to class sizes improved recall and precision. These weights are imposed on the C parameters, affecting the penalty for misclassification.

Precision, recall, and accuracy are three metrics commonly used in signal detection theory; however, depending on the context of the study, one metric is often prioritized over another [230]. For example in the case of bicycles, high precision means that most detected objects detected as bicycles are actually bicycles while high recall indicates that most bicycles in the scene are detected.

When mode classification takes place in the context of traffic monitoring in multimodal environments, either for demand estimation or safety studies, recall is more critical. Higher recall means that road users that are present can be detected. As pedestrians and bicyclists are fewer compared to motorized vehicles, a classifier should be capable of always detecting non-motorized users; missing a car is trivial compared to missing a pedestrian as a pedestrian is less frequently present. Precision is important too, as it assesses the classifier’s ability to correctly assign an object to a class; however, it is considered of lower importance in this context. Lastly, with respect to accuracy in multi-class classification settings this metrics lacks in informing where the errors are attributed to.

After determining the balancing approach as well as the most appropriate performance metric, the next step is to conclude on the feature vector that improves the classifier’s performance. As shown in Figure ??, a feature vector consisting of ten speed measurements along with the mean vehicle length resulted in recall values that were high for all three mode types.

6.7 Conclusions and future extensions

Current technologies, such as cameras that are widely implemented in traffic monitoring studies are limited by external conditions, such as light and weather. At the same time, radar-based sensors have been underutilized with implementations in motorized vehicles-only traffic environments. This study utilized a radar-based motion and presence sensor to develop a mode classification framework for motorized vehicles, bicycles, and pedestrians when they are all present in multimodal environments. Overall, there are three, interconnected layers of contributions related to the use of radar-based sensors in multimodal road environments, as this study. This is the first research to explore the feasibility of using radar-based sensor data to accurately classify

recorded trajectories in road environments where multiple modes are present. Second, a robust, flexible, and transferable procedure for the classification of motorized vehicles, bicyclists, and pedestrians is provided; the framework is capable of accommodating varying traffic conditions and control types as well as imbalanced datasets. Lastly, this study provides an instance of the complete procedure to demonstrate its applicability in two traffic scenes.

The proposed classification framework uses SVM as the classifier and can accommodate the unique properties of multimodal traffic captured by radar-based sensors. This is feasible through a step-by-step procedure for the SVM implementation to obtain the most appropriate mode classification model for a certain scene. A trajectory normalization scheme is first implemented to obtain a set of measurements for each trajectory. The SVM implementation may vary in terms of feature space (speed/length or speed/acceleration) or training sample balancing strategy. A set of performance metrics as well a discussion on how to assess the results is provided to guide users on how to select the appropriate SVM implementation. Cross-validation is used to obtain the optimal SVM parameters as well as performance metric averages across multiple train-test sets.

To test and validate the proposed framework, data was collected from two test sites that varied in terms of geometry, control, and operating speeds, i.e., interrupted vs uninterrupted flow. This allowed for testing the radar-based sensor performance using trajectories collected under a variety of operating conditions. The findings suggest that the proposed mode classification framework can accurately detect and classify bicycles, pedestrians, and motorized vehicles achieving an accuracy of 95%. A desirable performance was achieved using a balanced approach with a feature vector containing ten speed measurements and the average vehicle length. The preference for the balanced approach can be attributed to the imbalance of observations by mode.

While this combination of features and approach could be specific to the dataset used in this study, the proposed framework can be applied to determine feature vectors that are appropriate for any other test site with radar-based trajectory data available.

Future work on the mode classification aspect could validate and perhaps modify the proposed framework to be able to differentiate between types of motorized vehicles, e.g., buses and trucks, as large motorized vehicles have been reported to degrade the safety of pedestrians and bicyclists [194, 256]. Additionally, future research should focus on a large scale validation with larger numbers of non-motorized users (and types, e.g., e-scooters) in a variety of traffic conditions and geometric designs (e.g., vertical or horizontal curves or under the presence of innovative bicycle and pedestrian infrastructure treatments).

Apart from classification, there is a need to develop a methodology to deal with broken trajectory reconstruction and noise elimination so that road user interactions and behavior can be assessed. Such a methodology should combine both advanced data filtering techniques and identification of optimal placing of the sensor. Essentially, the placing of the sensor can eliminate/ reduce occlusion and noise and so, reduce the need for advanced data filtering. For example, placing the radar-based sensor in the horizontal part of a traffic light pole allows for field of “view” closer to bird’s eye view, that is free of occlusion. This placement would benefit right-hook traffic conflict analysis between bicyclists and motorized vehicles, presented in the previous chapter.

7 Conclusions, practical implications, and future extensions

Following the trend of Northern European countries, many cities across the globe have come to the realization that the bicycle should become the primary transportation mode for a higher share of the population. The implementation of bicycle treatments, such as protected or conventional bike lanes and protected intersection, attracts more bicyclists as such treatments provide designated roadway space for them, which enhances perceived safety and comfort. However, the combination of higher bicycling levels and a not appropriate selection of bicycle treatment types deteriorates bicyclist safety.

This dissertation focused on the impact of a variety of segment- and intersection-level treatments when implemented individually or in combination in an effort to quantify their safety benefits and guide future implementations. While bicycle safety literature has been receiving an increasing amount of attention during the last 20 years on a global level, there is still limited research on the comparison of different bicycle treatment types. Moreover, little attention has been given in differentiating between the safety impact of a treatment at the segment- versus the intersection-level. This limitation can be attributed to data availability. Assessing the actual (i.e., not perceived) bicycle safety in conjunction to the presence and type of bicycle treatments requires bicycle demand and bicycle crash data, in addition to data on bicycle treatment location. Such detailed datasets are hard to be obtained, and when they are available, one should consider that not all bicycle treatments can be assessed via crash-based analysis.

In this dissertation, the evaluation of the different treatment types was facilitated using three different traffic safety approaches, namely, (1) crash analysis, (2) traffic

conflict analysis, and (3) driver behavior analysis (i.e., speeding and glances). Each of these approaches sheds light to some aspects regarding bicyclist safety and the presence and type of bicycle treatments. The combination of all approaches allows for a deeper understanding of how each treatment type affects bicycle safety. Crash analysis may be inconclusive when studying specific crash types (e.g., right-hook crashes) and in this case, traffic conflict analysis can be used as an alternative. Traffic conflict analysis can focus on the specific part of a site and isolate specific road user interactions that are related to a specific crash type. For example, traffic conflict analysis can focus on one intersection approach and analyze right-turning motorized vehicles and straight-going bicyclists as in these road users will be involved in a right-hook crash. Similarly, driving simulator experiments offer the ability to isolate specific parts of a site and road user interactions but at the same time, they offer additional opportunities compared to traffic conflict analysis. Driving simulator experiments benefit from the use of additional equipment to assess road users mental processes, e.g., eye-tracking device to assess driver's glancing behavior. Driving simulator can also study treatments not frequently found in the field. Therefore, using crash analysis, then traffic conflict analysis and then, driving simulator technology allowed for a complete assessment of bicycle treatment presence and type. Recognizing that effective monitoring of sites where bicyclists coexist with other non-motorized and motorized road users is essential for safety analysis, this dissertation investigated the feasibility of alternative traffic monitoring technologies, in particular, radar-based sensors. Traffic conflict analysis and generally field data collection, could also be conducted using radar-based sensors instead of video cameras.

This Chapter presents the overall findings and contributions of this dissertation. Bicycle safety-related findings, that relate the type of various bicycle treatments to a safety outcome (e.g., predicted crashes or driver speeding behavior), are presented

separately for the segment and the intersection levels, showing the main findings from Chapters 3-5. Section 7.2 discusses methodological findings with the objective to summarize new knowledge on how to evaluate bicycle safety. Section 7.3 presents a step-by-step methodology on data needs and data analysis regarding the assessment of the safety of bicycle treatments. The objective of this section is to stand as a guideline for authorities and professionals in the field of bicycle safety. The last section (7.4) discusses the limitations of this dissertation and paths for future research.

7.1 The impact of bicycle treatment type on bicycle safety

7.1.1 Segment-level bicycle treatments at road segments

This dissertation assessed the safety impact on bicycle safety of the the following three bicycle treatments for the segment-level: sharrows, protected bike lanes and conventional bike lanes.

The crash-based analysis found that road segments with bicycle treatments are overall safer compared to those segments without bicycle treatments in terms of crash probability. With respect to the bicycle treatment type, the study found that there is a ranking between segment-level treatments; protected bike lanes, then sharrows and then conventional bike lanes are safer when compared to road segments with no bicycle treatments. Essentially, a bicyclist has less chances of being involved in a crash with a motorized vehicle when traveling on a protected bike lane or a sharrow compared to a conventional bike lane. This finding can guide cities in the selection of bicycle treatments and motivate a shift from conventional bike lanes to protected ones.

Driver behavior in terms of speeding and glancing data in the presence of bicycle treatments was assessed through a driving simulator experiment. Driver speeding data

analysis showed that drivers develop a lower speed when driving next to protected versus conventional bike lanes. As the protected bike lanes require the presence of a physical barrier between the bike lane and the traffic lane, they cause a reduction of the effective (traffic) lane width and so, reduce speeds. Implementing protected bike lanes can thus be seen as a means of achieving lower speed limits in urban areas. Driver glancing data analysis suggested that drivers are less likely to glance at a bicyclist when he/she travel at a protected versus a conventional bike lane. The fact that drivers may not glance at the bicyclist raises concerns regarding their ability to perceive and detect the bicyclist. In this case, it is important that countermeasures are introduced to improve driver awareness and increase the probability of drivers' searching for bicyclists at the intersection, when the twp road users might interact. Previous research found that the presence of the "yield to bicyclists" signage placed at the intersection increased the right side mirror scanning of right-turning drivers at the intersection by 9% [261]. To address view blocking of the protected bike lane, one could terminate the parking lane before the intersection to improve driver situational awareness of bicyclist presence as they are approaching the intersection.

7.1.2 Segment-level bicycle treatments at signalized intersections

All three safety-related studies conducted for this dissertation assessed the safety benefits of segment-level treatments at signalized intersections.

The crash prediction models (i.e., crash-based analysis) found that signalized intersections where bicycle treatments are placed at more than one of the intersecting roads are associated with higher crash frequency. Specifically, the following situations were associated with increased crash frequency: (a) when the intersection has one road with conventional bike lanes and no other treatment on the other intersection road, (b) when the intersection has a road with conventional bike lanes and a road

with protected bike lanes, and (c) when the intersection has two roads both of which with conventional bike lanes. The combination of bicycle treatments described in point (a) above correspond to a safer case, in terms of crash frequency, compared to the one described in point (b), and so on. This finding suggests that when both protected and conventional bike lanes are present at a signalized intersection the crash risk is the highest compared to (signalized) intersections where no protected bike lanes are present. This finding should be carefully interpreted. The developed crash prediction models should be seen as a tool of prioritizing high risk sites for further investigation. For example, a road safety inspection or a traffic conflict analysis at signalized intersections where both protected and conventional bike lanes are present could reveal potential unsafe situations between bicyclists and motorized vehicles, and in turn, recommend appropriate countermeasures.

In contrast to the crash-based analysis that assessed the safety performance of entire signalized intersections (i.e., all of the approaches) in terms of crash frequency, the traffic conflict study as well as the driving simulator experiment focused on one approach. The focus of these two studies was on right-hook traffic conflicts that occur between a right-turning driver and a straight-going bicyclists during the green phase. The bicyclist is located on the right side of the right-turning vehicle and the driver might not notice the bicyclist and potentially collide with him/her.

The traffic conflict study concluded that the segment-level bicycle treatment type does not affect the frequency of the observed right-hook traffic conflicts at signalized intersections. This study assessed sharrows, conventional and protected bike lanes at the intersection approach. Right-hook traffic conflicts always take place downstream the stop bar and so, the (upstream) segment-level bicycle treatment type does not affect bicyclist path while crossing the intersection.

The driving simulator experiment findings were similar in the sense that driver

behavior while turning right at an intersection was not found to be affected by the bicycle treatment type upstream the intersection. The speeding and glancing analysis showed that speed differences as well differences with respect to drivers' glancing were not statistically significant when turning right in the presence of protected versus conventional bike lanes.

Overall, as the segment-level bicycle treatment type does not play a role when it comes to right-hook conflicts and driver right-turning behavior, other safety measures should be implemented to eliminate right-hook conflicts and in turn, right-hook crashes at signalized intersections. For example, the placement of a protected intersection has the potential to increase the number of drivers that glance right before making a right turn. Hence, this treatment could potentially reduce the number of right-hook traffic conflicts.

7.1.3 Intersection-level treatments at signalized intersections

With respect to the intersection-level treatments at signalized intersections, the three following treatments were studied in this dissertation: bike boxes, intersection-crossing pavement markings (or simply crossing markings) and protected intersection design.

The crash-based analysis found that intersections with bike boxes or crossing markings increase crash frequency. This finding should be carefully interpreted by transportation professionals and authorities. The developed crash prediction models should be seen as a tool for identifying high-risk sites (i.e., signalized intersections) across a network. Bike boxes and intersection-crossing pavement markings aim to increase driver awareness of bicyclist presence and provide potential paths for bicyclists while navigating an intersection. Therefore, signalized intersections with bike boxes or crossing markings at one or more approaches need be inspected as additional

countermeasures might be needed to enhance bicycle safety.

The right-hook traffic conflict analysis considered signalized intersections with bike boxes or crossing markings. The findings indicate that none of these bicycle treatments significantly affects the frequency of right-hook traffic conflicts between right-turning drivers and straight-going bicyclists. Bike boxes are useful for a signalized intersection during the red phase, as bicyclists can wait there and be in the front of drivers and so visible to them. However, by design bike boxes cannot affect bicyclist safety during the green phase; these treatments may only increase driver awareness of bicyclists.

The driving simulator experiment found that protected intersection design is safer in terms of driver glancing behavior compared to the non-protected intersection design featuring crossing-markings. Specifically, drivers turning right at a protected intersection are more likely to glance right at the intersection compared to the non-protected design. Additionally, it was found that those drivers that glanced right at the intersection, developed lower average speed while making the right turn compared to those drivers who did not glance.

7.1.4 Contributions

The work carried out in this dissertation contributes in the following ways in the bicycle safety literature:

1. At the segment-level, it was shown that it is meaningful to differentiate between the bicycle treatment type, and specifically protected bike lanes are safer than conventional bike lanes and sharrows.
2. At the intersection-level, the presence of segment-level bicycle treatments increases crash risk however, the type of segment-level treatments does not play a role for right-hook conflicts.

3. At the intersection-level, the implementation of intersection-level treatments needs to be carefully selected.

7.2 Methodological findings

7.2.1 Strengths and limitations of the three safety assessment approaches

The analysis of historical crash records is the main approach for the assessment of road safety. The outcome of this analysis is the identification of high-risk sites, i.e., sites that experience more crashes compared to a defined threshold (e.g., average crash frequency across all studied sites). Detailed crash records with respect to the crash severity level, crash type (e.g., rear-end crash), and conditions during the crash event (e.g., day vs night) help to better identify the most critical crash contributing factors per case and so, propose appropriate countermeasures.

For this dissertation, crash analysis was used to relate (i) segment-level bicycle treatments to the probability of a bicycle-motorized vehicle crash at the segment and (ii) intersection bicycle-motorized vehicle crash frequency to the presence and type of bicycle treatments. In the first case, the outcome of the analysis can fully guide future implementation of bicycle treatments for road segments; for example, protected bike lanes are safer for bicyclists compared to conventional bike lanes. In the case of signalized intersections, both developed models produce results that need to be carefully interpreted. Further analysis is needed at sites that are found more risky, e.g., signalized intersections with bike boxes.

Ideally, crash analysis for intersections should differentiate based on the crash type (e.g., right-hook crashes vs left-hook crashes) as certain crash types can be reduced by certain countermeasures. However, in the case of bicycle-motorized vehicle crashes this differentiation is not feasible as the number of crashes is low; bicycle

crashes are few because the number of bicyclists is much lower compared to other modes and at the same time, they tend to be underreported. A detailed crash dataset was not the case for signalized intersection crash analysis in this dissertation, and so to better understand the impact of bicycle treatment types at the intersection, additional analyses were needed. It should be highlighted that not having a detailed (i.e., regarding the crash type) bicycle-motorized vehicle crash dataset is not specific to Portland, and so in turn, it does not only affect the analysis of this dissertation. In the bicycle safety literature only one study has been found to differentiate between the crash type and concerns unsignalized intersections [220]. On the contrary, for motorized vehicle crashes there are multiple studies that center on specific crash types: e.g., rear-end crashes [91, 259, 270], left-turn crashes [89, 260].

In a nutshell, crash analysis can be inconclusive regarding the safety impact of a treatment. Therefore, different safety approaches can be useful for cases where crash analysis cannot guide the selection of interventions and countermeasures. Traffic conflict analysis is a non-crash based safety approach and traffic conflict occurrence has been related to crash occurrence [112]. Traffic conflict analysis requires on-site data collection (e.g., with video cameras) that capture road users' movement and interactions. For the assessment of bicycle safety, traffic conflict analysis is a powerful tool as it can assess multiple interactions between bicyclists and motorized users, e.g., right-hook conflicts, and see how those interactions are impacted by the presence and type of bicycle treatments. On-site/video data collection benefit the safety analysis are the researcher/inspector can see how bicyclists and motorists behave in the presence of bicycle treatments, e.g., do bicyclists use the bike box during and do drivers encroach the bike box area?

There are two shortcoming related to traffic conflict analysis for bicycle safety. First, there are no guidelines on the amount of data needed for traffic conflict analysis;

crash data should be at minimum of three years to address regression-to-the-mean bias. It is unclear how many sites, how many hours of data in total or per site is needed to assess a treatment or what time of the week/year should these data be recorded. Second, research on traffic conflict analysis is based on motorized vehicles and some of the existing thresholds should be adjusted to for the case of bicyclists. For example, questions can be raised regarding the the PET traffic conflict indicator and the road user sequence in the case of right-turning vehicles and straight-going bicyclists: is the conflict between a bicyclist followed by a motorized vehicle equally risky as the conflict of a motorist followed by a bicyclist? In Chapter 4 it was showed that motorists keep a greater time distance between themselves and the leading bicyclists compared to bicyclists following a motorized vehicle.

Driving simulator experiments are another non-crash-based approach to assess the impact of bicycle treatments on bicycle safety. This approach is ideal for treatments that are cannot be found in the field, such as the protected intersections that are not commonly placed in the U.S compared to other treatments (e.g., bike boxes). Driving simulator technology benefits from technological equipment that can be used in the experiment. In this driving simulator experiment driver glancing behavior was recorded via an eye-tracking device, offering the analysis to check whether drivers glance towards certain areas of interest (e.g., at a bicyclist traveling at the protected bike lane or at the protected intersection). Glancing information which is a proxy of drivers' awareness cannot be assessed in crash-based on traffic conflict-based studies.

Overall, bicycle safety analysis can be benefited from different safety approaches. Due to the low number of bicycle crashes, crash-based analysis may be limited. In this case, traffic conflict analysis is an appropriate alternative when the studied bicycle treatments can be found in the field.

7.2.2 Sensors for conducting traffic monitoring studies

Throughout this dissertation two different sensors, namely a radar-based sensor and video cameras, were used to collect field data from intersections where pedestrians, bicyclists, and motorized vehicles are present. The objective was to test the feasibility of using radar-based sensors in multi-modal traffic environments. The first step for the deployment of radar-based sensors in field studies is to ensure that recorded trajectories can be correctly classified. A mode classification framework was developed to classify recorded trajectories in three modes, namely motorized vehicle, bicycles, and pedestrians. The framework is capable of operating in multiple traffic scenes (e.g., signalized intersection vs corridor) by considering a difference balancing approach and a different set of features, such as speed and vehicle length or speed and acceleration.

7.2.3 Safety-in-numbers effect

All of the safety-related studies agree with the safety-in-numbers effect; the presence of more bicyclists improves safety for all bicyclists [107]. Road segments with more bicyclists are more likely to experience a crash compared to segments with fewer bicyclists, meaning that the per bicyclist crash risk is lower in the former case. Additionally, a single bicyclist present at the segment was found to cause a reduction in driver speed, compared to the case that no bicyclist was present at the segment. At signalized intersections, the number of bicyclists per year and per 15 minutes were associated with higher crash frequency and traffic conflict frequency, respectively. This means that signalized intersections with more bicyclists have a lower risk per bicyclist. The fact that drivers were more likely to glance right at the intersection when a bicyclist was present confirms that even one bicyclist has the potential to increase driver awareness of bicyclist presence.

There are several take-away remarks from this finding. First, cities should try to increase bicycle mode share in an effort to improve safety for all bicyclists. Campaigns, traffic calming measures, and the provision of bicycle facilities (e.g., bicycle parking) can motivate more people to shift to bicycling. Then, bicycle networks should be in a way that (1) allow drivers to detect bicyclists and so, be aware of bicyclist presence and potential routes (2) and be well-connected. Well-connected bicycle networks have the potential to attract the great majority of bicyclists to certain segments and intersections (i.e., those that are part of the bicycle network). At those segments and intersections the safety-in-numbers effects will be present. If due to poor design, bicyclists check alternative segments and intersections where drivers do expect to see them, they are at a higher risk.

7.2.4 Contributions

The work carried out in this dissertation makes the following contributions from a methodological perspective:

1. Strengths and limitations for three safety assessment methods have been demonstrated regarding their appropriateness for assessing bicycle safety.
2. The analysis of right-hook traffic conflicts between bicyclists and motorized vehicles showed that user sequence is important when studying conflict between different road users.
3. A mode classification framework was developed for classifying trajectories recorded via radar-based sensors in multi-modal environments. The framework is flexible to operating in various traffic scenes.

7.3 Practical implications and methodology transferability

The results of this research have numerous practical implications for transportation practitioners, policymakers, officials, and researchers. In brief, the findings can assist professionals with the selection of new bicycle treatment types but also provide guidelines on how to evaluate of the effectiveness of existing treatments or the overall safety of a multimodal site.

At the segment level, for high volume streets, protected bike lanes stand as the safest treatment. Their implementation has also been found to reduce driving speed, a measure that enhances safety for all road users. Shared bicycle treatments are ideal for low-traffic, low-speed streets in residential areas.

At the intersection level, the signalized intersection approach has the potential to reduce right-hook crashes between bicyclists and motorized vehicles and also turning speeds. Treatments like bike boxes and intersection-crossing pavement markings should continue be implemented as they provide an indication on which roadway space should bicyclists occupy when they present or navigating a signalized intersection. However, transportation practitioners and officials should carry out additional analysis on a per site basis to understand which other countermeasures are needed to enhance safety.

Several takeaways can be drawn from this research regarding ways to evaluate bicycle safety. Crash analysis is the most common approach, however, 3 to 5 years of crash data are needed to obtain representative crash history of the sites, traffic volume and bicycle volume data are needed to account for exposure, while detailed road-related data are also needed for understanding which sites are more prone to crashes. Agencies need to actively collect and update these data, which can be time and resource consuming. None of these data type should be omitted when conducting

crash-based analysis, while the analysis should not be considered if the data are limited or not reliable. Overall, crash analysis is not applicable for new roads or newly implemented treatments and these are cases that different approaches become essential.

For new interventions or limited/unreliable data field data collected with a video camera is a cost-effective way to assess safety. Each intersection approach can be isolated and studied in detail, while per approach different movements can also be isolated and studied, e.g., right-hook and left-hook conflicts. Segments can also be evaluated using this technology. In segments it is interesting to observe the distance between bicyclists and motorized vehicles when conventional bike lanes are present and potentially consider implementing some sort of physical separation. Video-based traffic monitoring offers an agency a rich dataset that can be used for multiple analyses in addition to safety assessment. Multiple data types are recorded (e.g., bicycle flows, pedestrian flows, signal timing, red-light runners, traveling speeds, etc.) and they can be translated into variables for an analysis. For example the developed models for the prediction of right-hook conflict can be extended to include of motorized vehicle speed, acceleration, signal timing, pedestrian flows, etc.

Video-based traffic monitoring studies are ideal for short-term data collection as one camera can be easily set-up to multiple sites (segments and intersections). For longer periods of data collection, it would be meaningful for agencies to purchase radar-based sensors that can operate during all lighting and weather conditions. In addition to their capabilities compared to video cameras, radar-based sensors require more time to set up and at the same time, more equipment to be operate. For those reasons are more appropriate for long-term implementations. The agency should also invest in developing a data processing system to analyze the recorded data.

7.4 Future extensions

This work has allowed for a better understanding of how various bicycle treatment types affect bicycle safety. Additionally, it has tested the feasibility of using alternative sensors for collecting field data for safety analysis. Future research in the field of bicycle safety should consider the following directions in terms of treatment types, methodological approaches, and incorporation of additional objectives (e.g., equity):

1. *Additional bicycle treatment types for road segments*

Segment-level treatments were grouped in three categories, namely sharrows, conventional bike lanes, and protected bike lanes. Future research should consider removing this grouping and differentiating between bike lanes and buffered bike lanes, contra-flow bike lanes, as well as the various types of protected bike lanes (see Figure 2.7). As both road space and curb space in urban areas is limited, practitioners should have strong evidence on the safety effectiveness on various bicycle treatments types as safer bicycle treatment might also take more space. Lastly, bicycle and bus interactions should be assessed in the case of mixed bike and bus lanes.

2. *Design aspects and level of congestion of segment-level bicycle treatments*

With regards to segment-level bicycle treatments, future research should also evaluate the dimensions of the various bike lane types and how they impact bicycle safety. Essentially, the width of bike lanes and its ability to serve bicyclists becomes critical in higher bicycling levels. Congested and narrow bike lanes might force bicyclists to use the adjacent traffic lane to overtake a bicyclist or in the case of congested protected bike lanes, bicyclists might choose to use the traffic lanes.

3. *Additional bicycle treatment types for intersections*

Three intersection treatments, namely: intersection-crossing pavement markings, bike boxes, and protected intersections, were studied in this dissertation. Additional treatments exist for bicyclists such as merge zones and bicycle signals accompanied by various timing strategies have not been studied in depth. The various intersection-level treatments should be examined in terms of bicycle safety in conjunction with signal strategies, with and without bicycle signals.

4. *Traffic signal timing optimization*

As indicated earlier, control strategies should be studied in addition to bicycle treatments at intersections. Intersection delay has been found to affect bicyclists' route choice who try to avoid signalized intersections [20, 99]. Hence, signal timing should be optimized to equitably accommodate all users at an intersection and extend current person-based signal timing optimization schemes [36, 38, 39].

5. *Bicycle treatments in rural areas*

This dissertation focused on urban or suburban (for the case of crash analysis) environments. There is limited research on bicycle safety in rural areas and also, as many conditions are different compared to urban roads (e.g., lack of lighting, absence of bicyclists, presence of horizontal curves, etc.), the findings might not be transferable.

6. *Bicycle safety and social equity*

There are several research studies showing that bicycle facilities (e.g., bikesharing schemes) and treatments are more frequently adopted by higher income level communities [100, 229]. This could result in making bicycling an accessible transportation mode for only certain socioeconomic and demographic groups. At the same time, lower income and marginalized populations are more likely

to experience higher crash risk while bicycling [49, 248]. However, there are no in-depth studies on the factors that contribute to crashes at those neighborhoods. For example, is it the overall poor road conditions or specifically the lack of bicycle treatments that increase crash risk in those settings? Given that bicycle is a relatively cheap mode to own and use and at the same time has the potential to increase access to opportunities [154, 166], ensuring that it is a safe and convenient mode for low income communities is of paramount importance for ensuring transportation and health equity.

7. *Transferability of the findings*

Traffic safety analysis is affected by road user characteristics, road network design, weather, vehicle characteristics, presence of enforcement, etc. Therefore, it is unlikely that the findings of one study are directly transferable to a different setting; many U.S. States for example, have developed their own Safety Performance Functions instead of calibrating the ones provided in Highway Safety Manual as they achieved a better prediction performance for the same type of roads, e.g., [86, 158]. The three safety-related studies undertaken as part of this dissertation took place in a certain spatial and temporal context. Portland, Oregon and Boston and Cambridge, Massachusetts have a strong bicycling culture [245] and this might have affected the findings. Similar studies should take place in different locations that vary in terms of bicycle and driving culture, road and bicycle network design, land use, weather, population demographics, etc. These are only a few factors that have been found to impact bicycle safety [32] and can in turn affect the transferability of this dissertation's findings.

In addition to assessing bicycle treatments in various environments (e.g., cities with more treatment types or rural areas), it is important to consider how the presence

and type of bicycle treatments affects modern mobility. New micromobility modes are emerging (e.g., e-scooters, segways, e-bikes, etc.) that demand for roadway space, while at the same on-vehicle sensors are advancing and the world is eventually shifting towards automated driving.

1. *Bicycle treatments and micromobility users*

Laws and regulations as well as attitudes related to micromobility users vary per city [150, 277]. E-bikes and usually e-scooters tend to use the bicycle infrastructure, and it remains unclear whether the presence of these road users affects bicyclist safety and convenience. Essentially, as e-bikes and e-scooters develop higher speeds compared to the average bicyclist a set of regulations might be needed to accommodate all users in bicycle treatments.

2. *On-vehicle sensors and automated driving*

In light of the automated driving, multiple on-vehicle sensors are placed on cars to enable some level of driving assistance or automation. As urban roads are more complex (e.g., stop-and go traffic, presence of pedestrians, etc.) the majority of these sensors are trained and tested in motorways or large arterials and so, the sensors have reduced capabilities in detecting and responding to bicyclists. Some advanced driving assistance systems (ADAS) are placed in the windshield of the car and so can only detect risks in the front of the vehicle (e.g., Mobileye sensor, see the work by Emami [66]). A right-turning driver, that has not checked on their right for bicyclists, cannot be alerted that is about to collide with a straight-going bicyclist. Manufacturers of on-vehicle sensors, either these sensors fall under the umbrella of ADAS or are part of automated driving technology, should consider different types of interactions between motorized vehicles and bicyclists while taking into account the presence

of bicycle treatments to develop detection and warning systems that enhance bicycle safety. For example, the sensor should be aware of the presence of a protected bike lane behind the parking lane or it should maintain a steady distance from the bike lane in case a bicyclist enters the traffic lane

Future research can center on the following methodological extensions of the current work:

1. *Development of Safety Performance Functions and Crash Modification Factors*

The developed crash prediction models for road segments and signalized intersections should motivate the development of Safety Performance Functions (SPFs) to predict the expected crash frequency for (i) road segments and (b) signalized intersections using exposure variables as predictors, i.e., AADT and AADB. Crash modification factors are also needed to complement the developed SPFs and account for the bicycle treatment present and type. While the developed models predict crashes and can be used to extract CMFs for the treatment types, they cannot be considered SPF as defined in the Highway Safety Manual as they (a) have been developed using data from one area instead of multiple and (b) predict crashes rather than the expected crash frequency; the later is estimated using the Empirical Bayes theorem. Highway Safety Manual CMFs are estimated using before-after analysis to assess the effectiveness of a treatment type.

2. *Surrogate safety analysis and field studies as proactive safety approach*

Crash analysis stands as the starting point for traffic safety studies, however, it is important that authorities and agencies take advantage of the technology and shift towards safety assessment approaches that rely on traffic monitoring data (i.e., field data). This is a proactive safety approach in the sense that remedial

actions are taken before crashes occur. Field studies have the potential to assess multiple bicycle-motorized vehicle interaction data but also, once field data is collected it can be used for different purposes (e.g., demand studies, pedestrian safety studies, or unsafe interactions at trucks and taxis pick-up/drop-off points). However, authorities and agencies are in need of guidelines on how to conduct such studies. The current literature while it has associated traffic conflict occurrence to crash occurrence and other metrics such as speed and red-light running to crash occurrence, it is limited from conclusive guidelines on: (i) the amount of data needed to support decisions on countermeasures implementation and (ii) how different types of surrogate safety measures can be combined (e.g., how can PET events combined with speed acceleration events at an intersection?).

3. *Alternative sensor types for field studies* Traffic monitoring is mainly relied on video cameras while other sensors such as radar-based sensors are used for speed enforcement. The feasibility of using radar-based sensors in field studies was explored in this dissertation and future work should exploit recorded trajectories for assessing motorized vehicle, bicyclist, and pedestrian safety.
4. *Exploiting the capabilities of driving simulator* In this dissertation driving simulator technology was used mainly because protected intersections can be found in the field. In addition to testing treatments not frequently found in real world, driving simulator experiments can be designed to assess critical events (e.g., near-misses) between drivers and simulated (i.e., virtual bicyclists) in the presence of various bicycle treatments, or to test user interface designs that alert drivers of potentially risky interactions with bicyclists (or micromobility users).

References

- [1] D. Acharya, K. Khoshelham, and S. Winter. Real-time detection and tracking of pedestrians in cctv images using a deep convolutional neural network. In *Proceedings of the 4th annual conference of research@ locate*, volume 1913, pages 31–36, 2017.
- [2] N. Agerholm, S. Caspersen, and H. Lahrman. Traffic safety on bicycle paths. *Aalborg (Dänemark)*, 2008.
- [3] R. Aldred, B. Elliott, J. Woodcock, and A. Goodman. Cycling provision separated from motor traffic: a systematic review exploring whether stated preferences vary by gender and age. *Transport reviews*, 37(1):29–55, 2017.
- [4] R. Aldred, S. García-Herrero, E. Anaya, S. Herrera, and M. Á. Mariscal. Cyclist injury severity in spain: a bayesian analysis of police road injury data focusing on involved vehicles and route environment. *International journal of environmental research and public health*, 17(1):96, 2019.
- [5] R. Aldred, A. Goodman, J. Gulliver, and J. Woodcock. Cycling injury risk in london: A case-control study exploring the impact of cycle volumes, motor vehicle volumes, and road characteristics including speed limits. *Accident Analysis & Prevention*, 117:75–84, 2018.
- [6] R. Aldred, J. Woodcock, and A. Goodman. Does more cycling mean more diversity in cycling? *Transport reviews*, 36(1):28–44, 2016.
- [7] B. L. Allen, B. T. Shin, and P. J. Cooper. Analysis of traffic conflicts and collisions. *Transportation Research Record*, 667(HS-025 846):67–74, 1978.

- [8] D. Allen, S. Bygrave, and H. Harper. Behavior at cycle advanced stop lines. report no ppr240. Technical report, Transport for London, London Road Safety Unit, London UK, 2005.
- [9] American Association of Highway and Transportation Officials (AASHTO). Highway Safety Manual, Part D, 2010.
- [10] American National Standard. Manual on classification of motor vehicle traffic accidents. Technical report, National Safety Council, Itasca, Illinois, 2007.
- [11] F. Amundsen and C. Hyden. Proceedings of first workshop on traffic conflicts. Technical report, Oslo: Institute of Transport Economics, 1977.
- [12] M. I. Asborn, C. G. Burris, and S. Hernandez. Truck body-type classification using single-beam lidar sensors. *Transportation Research Record*, page 0361198118821847, 2019.
- [13] C. Bahnsen and T. B. Moeslund. Detecting road users at intersections through changing weather using rgb-thermal video. In *International Symposium on Visual Computing*, pages 741–751. Springer, 2015.
- [14] S. P. Baker, B. o’Neill, W. Haddon Jr, and W. B. Long. The injury severity score: a method for describing patients with multiple injuries and evaluating emergency care. *Journal of Trauma and Acute Care Surgery*, 14(3):187–196, 1974.
- [15] D. Bhatia, S. A. Richmond, C. J. Loo, L. Rothman, C. Macarthur, and A. Howard. Examining the impact of cycle lanes on cyclist-motor vehicle collisions in the city of toronto. *Journal of Transport & Health*, 3(4):523–528, 2016.

- [16] S. Blaizot, F. Papon, M. M. Haddak, and E. Amoros. Injury incidence rates of cyclists compared to pedestrians, car occupants and powered two-wheeler riders, using a medical registry and mobility data, rhône county, france. *Accident Analysis & Prevention*, 58:35–45, 2013.
- [17] A. Borowsky, T. Oron-Gilad, and Y. Parmet. Age and skill differences in classifying hazardous traffic scenes. *Transportation Research Part F: Traffic Psychology and Behaviour*, 12(4):277–287, 2009.
- [18] A. Borowsky, D. Shinar, and T. Oron-Gilad. Age, skill, and hazard perception in driving. *Accident Analysis & Prevention*, 42(4):1240–1249, 2010.
- [19] J. Broach, J. Dill, and J. Gliebe. Where do cyclists ride? a route choice model developed with revealed preference gps data. *Transportation Research Part A: Policy and Practice*, 46(10):1730–1740, 2012.
- [20] J. Broach, J. Gliebe, and J. Dill. Bicycle route choice model developed using revealed preference gps data. In *90th annual meeting of the transportation research board, Washington, DC*, 2011.
- [21] I. D. Brown. A review of the ‘looked but failed to see’ accident causation factor. In *Behavioural research in road safety: Eleventh seminar*, 2002.
- [22] R. Buehler and J. Dill. Bikeway networks: A review of effects on cycling. *Transport Reviews*, 36(1):9–27, 2016.
- [23] C. Burch, L. Cook, and P. Dischinger. A comparison of kabco and ais injury severity metrics using codes linked data. *Traffic injury prevention*, 15(6):627–630, 2014.

- [24] Q. Cai, M. Abdel-Aty, and S. Castro. Explore effects of bicycle facilities and exposure on bicycle safety at intersections. *International Journal of Sustainable Transportation*, pages 1–12, 2020.
- [25] G. Cantisani, L. Moretti, and Y. De Andrade Barbosa. Safety problems in urban cycling mobility: a quantitative risk analysis at urban intersections. *Safety*, 5(1):6, 2019.
- [26] J. M. Casello, A. Fraser, A. Mereu, and P. Fard. Enhancing cycling safety at signalized intersections: analysis of observed behavior. *Transportation research record*, 2662(1):59–66, 2017.
- [27] A. Caviedes and M. Figliozzi. Modeling the impact of traffic conditions and bicycle facilities on cyclists’ on-road stress levels. *Transportation research part F: traffic psychology and behaviour*, 58:488–499, 2018.
- [28] N. V. Chawla, K. W. Bowyer, L. O. Hall, and W. P. Kegelmeyer. Smote: synthetic minority over-sampling technique. *Journal of artificial intelligence research*, 16:321–357, 2002.
- [29] C. Chen. Crowdsourcing data-driven development of bicycle safety performance functions (spfs): Microscopic and macroscopic scales. 2017.
- [30] C. Chen, H. Wang, J. Roll, K. Nordback, and Y. Wang. Using bicycle app data to develop safety performance functions (spfs) for bicyclists at intersections: a generic framework. *Transportation research part A: policy and practice*, 132:1034–1052, 2020.
- [31] J. Chen, X. Yang, L. Lu, Q. Li, Z. Li, and W. Wu. A novel infrared image enhancement based on correlation measurement of visible image for urban traffic

- surveillance systems. *Journal of Intelligent Transportation Systems*, 24(3):290–303, 2020.
- [32] P. Chen. Built environment factors in explaining the automobile-involved bicycle crash frequencies: A spatial statistic approach. *Safety science*, 79:336–343, 2015.
- [33] P. Chen, W. Zeng, G. Yu, and Y. Wang. Surrogate safety analysis of pedestrian-vehicle conflict at intersections using unmanned aerial vehicle videos. *Journal of advanced transportation*, 2017, 2017.
- [34] M. V. Chitturi, A. W. Ooms, A. R. Bill, and D. A. Noyce. Injury outcomes and costs for cross-median and median barrier crashes. *Journal of safety research*, 42(2):87–92, 2011.
- [35] H.-J. Cho and M.-T. Tseng. A support vector machine approach to cmos-based radar signal processing for vehicle classification and speed estimation. *Mathematical and Computer Modelling*, 58(1-2):438–448, 2013.
- [36] E. Christofa, K. Ampountolas, and A. Skabardonis. Arterial traffic signal optimization: A person-based approach. *Transportation Research Part C: Emerging Technologies*, 66:27–47, 2016.
- [37] E. Christofa, M. A. Knodler, et al. Development and evaluation of infrastructure strategies for safer cycling. 2017.
- [38] E. Christofa, I. Papamichail, and A. Skabardonis. Person-based traffic responsive signal control optimization. *IEEE Transactions on Intelligent Transportation Systems*, 14(3):1278–1289, 2013.
- [39] E. Christofa and A. Skabardonis. Traffic signal optimization with application

- of transit signal priority to an isolated intersection. *Transportation Research Record*, 2259(1):192–201, 2011.
- [40] J. B. Cicchino. *Not all protected bike lanes are the same: infrastructure and risk of cyclist collisions and falls leading to emergency department visits in three US cities August 2019*. PhD thesis, Department of Emergency Medicine, Oregon Health & Science University, 2019.
- [41] J. B. Cicchino, M. L. McCarthy, C. D. Newgard, S. P. Wall, C. J. DiMaggio, P. E. Kulie, B. N. Arnold, and D. S. Zuby. Not all protected bike lanes are the same: Infrastructure and risk of cyclist collisions and falls leading to emergency department visits in three us cities. *Accident Analysis & Prevention*, 141:105490, 2020.
- [42] City of Portland, Oregon. Maps, GIS & Open Data. = <https://www.portlandoregon.gov/>, 2018.
- [43] City of Portland, Oregon. How crash data works. = <https://www.portland.gov/transportation/vision-zero/crash-data>, 2021.
- [44] C. Clark, P. Mokhtarian, G. Circella, and K. Watkins. User preferences for bicycle infrastructure in communities with emerging cycling cultures. *Transportation Research Record*, page 0361198119854084, 2019.
- [45] F. Cleaveland and F. Douma. The impact of bicycling facilities on commute mode share. In *88th Annual Meeting of the Transportation Research Board, Washington, DC*, 2009.
- [46] J. Cohen. Statistical power analysis. *Current directions in psychological science*, 1(3):98–101, 1992.

- [47] N. S. Council. Manual on classification of motor vehicle traffic crashes eighth edition. Technical report, American National Standards Institute, Inc., Board of Standards Review, 2017.
- [48] S. R. E. Datondji, Y. Dupuis, P. Subirats, and P. Vasseur. A survey of vision-based traffic monitoring of road intersections. *IEEE transactions on intelligent transportation systems*, 17(10):2681–2698, 2016.
- [49] C. M. Davison, M. Torunian, P. Walsh, W. Thompson, S. McFaull, and W. Pickett. Bicycle helmet use and bicycling-related injury among young canadians: an equity analysis. *International journal for equity in health*, 12(1):48, 2013.
- [50] T. De Ceunynck. *Defining and applying surrogate safety measures and behavioural indicators through site-based observations*. Hasselt University, 2017.
- [51] B. de Geus, G. Vandenbulcke, L. I. Panis, I. Thomas, B. Degraeuwe, E. Cumps, J. Aertsens, R. Torfs, and R. Meeusen. A prospective cohort study on minor accidents involving commuter cyclists in belgium. *Accident Analysis & Prevention*, 45:683–693, 2012.
- [52] J. De Mol and P. Lammar. Half of road casualties are not included in statistics: linking hospital and police registration necessary. *TRAFFIC SPECIALIST (MECHELEN)*, 130:15–18, 2006.
- [53] A. Deliali, N. Campbell, M. Knodler, and E. Christofa. Understanding the safety impact of protected intersection design elements: A driving simulation approach. *Transportation Research Record*, page 0361198120909382, 2020.
- [54] J. DiGioia, K. E. Watkins, Y. Xu, M. Rodgers, and R. Guensler. Safety impacts of bicycle infrastructure: A critical review. *Journal of safety research*, 61:105–119, 2017.

- [55] J. Dill, T. Goddard, C. Monsere, and N. McNeil. Can protected bike lanes help close the gender gap in cycling? lessons from five cities. 2014.
- [56] J. Dill, C. M. Monsere, and N. McNeil. Evaluation of bike boxes at signalized intersections. *Accident Analysis & Prevention*, 44(1):126–134, 2012.
- [57] R. Doom and P. Derweduwen. Optimisation of road accident statistics and use in urban safety management. research contracts cp/02/392. Technical report, CP, 2005.
- [58] M. Dozza and J. Werneke. Introducing naturalistic cycling data: What factors influence bicyclists’ safety in the real world? *Transportation research part F: traffic psychology and behaviour*, 24:83–91, 2014.
- [59] J. Duan, R. Li, L. Hou, W. Wang, G. Li, S. E. Li, B. Cheng, and H. Gao. Driver braking behavior analysis to improve autonomous emergency braking systems in typical chinese vehicle-bicycle conflicts. *Accident Analysis & Prevention*, 108:74–82, 2017.
- [60] Y. Ebadi, D. L. Fisher, and S. C. Roberts. Impact of cognitive distractions on drivers’ hazard anticipation behavior in complex scenarios. *Transportation research record*, 2673(9):440–451, 2019.
- [61] Y. Ebadi, G. P. Mangalore, and S. Samuel. Impact of cognitive distractions on drivers’ anticipation behavior in vehicle-bicycle conflict situations. In *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, volume 62, pages 1898–1902. SAGE Publications Sage CA: Los Angeles, CA, 2018.
- [62] J. Edquist, C. M. Rudin-Brown, and M. G. Lenné. The effects of on-street parking and road environment visual complexity on travel speed and reaction time. *Accident Analysis & Prevention*, 45:759–765, 2012.

- [63] Z. Elamrani Abou Ellassad, H. Mousannif, and H. Al Moatassime. Class-imbalanced crash prediction based on real-time traffic and weather data: A driving simulator study. *Traffic injury prevention*, 21(3):201–208, 2020.
- [64] N. Eluru, C. R. Bhat, and D. A. Hensher. A mixed generalized ordered response model for examining pedestrian and bicyclist injury severity level in traffic crashes. *Accident Analysis & Prevention*, 40(3):1033–1054, 2008.
- [65] R. Elvik, T. Vaa, A. Høy, and M. Sørensen. *The handbook of road safety measures*. Emerald Group Publishing, 2009.
- [66] H. Emami. *Driver Reactions to Forward Collision Warnings*. PhD thesis, Morgan State University, 2019.
- [67] M. Essa and T. Sayed. Traffic conflict models to evaluate the safety of signalized intersections at the cycle level. *Transportation research part C: emerging technologies*, 89:289–302, 2018.
- [68] European Road Safety Observatory. Pedestrians and cyclists. Technical report, European Commission, 2019.
- [69] Federal Highway Administration. KABCO Injury Classification Scale and Definitions.
- [70] T. Feng and H. J. Timmermans. Transportation mode recognition using gps and accelerometer data. *Transportation Research Part C: Emerging Technologies*, 37:118–130, 2013.
- [71] N. N. Ferenchak and W. E. Marshall. Advancing healthy cities through safer cycling: an examination of shared lane markings. *International journal of transportation science and technology*, 8(2):136–145, 2019.

- [72] C. D. Fitzpatrick, M. A. Knodler, A. Ryan, C. Ai, and D. Chen. The application of unmanned aerial systems in surface transportation-volume ii-c: Evaluation of uas highway speed-sensing application. Technical report, Massachusetts Department of Transportation, Office of Transportation Planning, Boston, MA United States, December 2019.
- [73] M. Foldberg Steffensen and A. Gaardbo. Cycling and safety for cyclists - Danish experiences. Technical report, Danish Road Directorate (Vejdirektoratet), 2016.
- [74] N. Foletta, C. Nielson, J. Patton, J. Parks, and R. Rees. Green shared lane markings on urban arterial in oakland, california: evaluation of super sharrows. *Transportation research record*, 2492(1):61–68, 2015.
- [75] N. Fournier, S. Bakhtiari, K. D. Valluru, N. Campbell, E. Christofa, S. Roberts, and M. Knodler Jr. Accounting for drivers’ bicycling frequency and familiarity with bicycle infrastructure treatments when evaluating safety. *Accident Analysis & Prevention*, 137:105410, 2020.
- [76] N. Fournier, E. Christofa, and M. A. Knodler Jr. A mixed methods investigation of bicycle exposure in crash rates. *Accident Analysis & Prevention*, 2017.
- [77] T. Fu, J. Stipancic, S. Zangenehpour, L. Miranda-Moreno, and N. Saunier. Automatic traffic data collection under varying lighting and temperature conditions in multimodal environments: Thermal versus visible spectrum video-based systems. *Journal of advanced transportation*, 2017, 2017.
- [78] P. G. Furth, P. J. Koonce, Y. Miao, F. Peng, and M. Littman. Mitigating right-turn conflict with protected yet concurrent phasing for cycle track and pedestrian crossings. *Transportation Research Record*, 2438(1):81–88, 2014.

- [79] P. G. Furth and Y. Wang. Delay estimation and signal timing design techniques for multi-stage pedestrian crossings and two-stage bicycle left turns. In *94th Annual Meeting of the Transportation Research Board, Washington, DC*, 2015.
- [80] D. J. Gabauer and H. C. Gabler. Comparison of delta-v and occupant impact velocity crash severity metrics using event data recorders. In *Annual proceedings/association for the advancement of automotive medicine*, volume 50, page 57. Association for the Advancement of Automotive Medicine, 2006.
- [81] P. Gårder, L. Leden, and T. Thedéen. Safety implications of bicycle paths at signalized intersections. *Accident Analysis & Prevention*, 26(4):429–439, 1994.
- [82] D. Gettman, L. Pu, T. Sayed, S. Shelby, and I. Siemens. Surrogate safety assessment model and validation. Technical report, United States. Federal Highway Administration. Office of Safety Research and . . . , 2008.
- [83] M. Ghodrati Abadi and D. S. Hurwitz. Operational impacts of protected-permitted right-turn phasing and pavement markings on bicyclist performance during conflicts with right-turning vehicles. *Transportation Research Record*, page 0361198119837231, 2019.
- [84] G. Golembiewski and B. Chandler. Roadway departure safety: A manual for local rural road owners. Technical report, Federal Highway Administration, 2011.
- [85] S. Gomei, M. Hitosugi, K. Ikegami, and S. Tokudome. Assessing injury severity in bicyclists involved in traffic accidents to more effectively prevent fatal bicycle injuries in japan. *Medicine, Science and the Law*, 53(4):194–198, 2013.
- [86] J. P. Gooch, V. V. Gayah, and E. T. Donnell. Safety performance functions

- for horizontal curves and tangents on two lane, two way rural roads. *Accident Analysis & Prevention*, 120:28–37, 2018.
- [87] D. Goodman and D. Christopher. Separated bike lane planning and design guide. *Separated Bike Lane Planning and Design Guide*, page 5, 2015.
- [88] T. Götschi, J. Garrard, and B. Giles-Corti. Cycling as a part of daily life: a review of health perspectives. *Transport Reviews*, 36(1):45–71, 2016.
- [89] F. Guo, X. Wang, and M. A. Abdel-Aty. Modeling signalized intersection safety with corridor-level spatial correlations. *Accident Analysis & Prevention*, 42(1):84–92, 2010.
- [90] Y. Guo, H. L. Logan, D. H. Glueck, and K. E. Muller. Selecting a sample size for studies with repeated measures. *BMC medical research methodology*, 13(1):100, 2013.
- [91] K. Haleem, M. Abdel-Aty, and J. Santos. Multiple applications of multivariate adaptive regression splines technique to predict rear-end crashes at unsignalized intersections. *Transportation research record*, 2165(1):33–41, 2010.
- [92] C. Hamann and C. Peek-Asa. On-road bicycle facilities and bicycle crashes in iowa, 2007–2010. *Accident Analysis & Prevention*, 56:103–109, 2013.
- [93] C. J. Hamann, C. Schwarz, and O. Soniyi. Examination of driver behavior in response to bicyclist behaviors. 2016.
- [94] M. A. Harris, C. C. Reynolds, M. Winters, P. A. Cripton, H. Shen, M. L. Chipman, M. D. Cusimano, S. Babul, J. R. Brubacher, S. M. Friedman, et al. Comparing the effects of infrastructure on bicycling injury at intersections and

- non-intersections using a case–crossover design. *Injury prevention*, 19(5):303–310, 2013.
- [95] J. Hayward. *Near misses as a measure of safety at urban intersections*. Pennsylvania Transportation and Traffic Safety Center, 1971.
- [96] M. Hilpert, M. Johnson, M.-A. Kioumourtzoglou, A. Domingo-Relloso, A. Peters, B. Adria-Mora, D. Hernández, J. Ross, and S. N. Chillrud. A new approach for inferring traffic-related air pollution: use of radar-calibrated crowd-sourced traffic data. *Environment international*, 127:142–159, 2019.
- [97] J. L. Hintze and R. D. Nelson. Violin plots: A box plot-density trace synergism. *The American Statistician*, 52(2):181–184, 1998.
- [98] M. R. Hoffman, W. E. Lambert, E. G. Peck, and J. C. Mayberry. Bicycle commuter injury prevention: it is time to focus on the environment. *Journal of Trauma and Acute Care Surgery*, 69(5):1112–1119, 2010.
- [99] J. Hood, E. Sall, and B. Charlton. A gps-based bicycle route choice model for san francisco, california. *Transportation letters*, 3(1):63–75, 2011.
- [100] K. Hosford and M. Winters. Who are public bicycle share programs serving? an evaluation of the equity of spatial access to bicycle share service areas in canadian cities. *Transportation research record*, 2672(36):42–50, 2018.
- [101] D. W. Hosmer Jr, S. Lemeshow, and R. X. Sturdivant. *Applied logistic regression*, volume 398. John Wiley & Sons, 2013.
- [102] W. W. Hunter, D. L. Harkey, J. Richard Stewart, and M. L. Birk. Evaluation of blue bike-lane treatment in portland, oregon. *Transportation Research Record*, 1705(1):107–115, 2000.

- [103] D. Hurwitz, M. Jannat, J. Warner, C. M. Monsere, and A. Razmpa. Towards effective design treatment for right turns at intersections with bicycle traffic. 2015.
- [104] C. Hydén and L. Linderholm. The swedish traffic-conflicts technique. In *International Calibration Study of Traffic Conflict Techniques*, pages 133–139. Springer, 1984.
- [105] Y. Iwasaki, S. Kawata, and T. Nakamiya. Vehicle detection even in poor visibility conditions using infrared thermal images and its application to road traffic flow monitoring. In *Emerging Trends in Computing, Informatics, Systems Sciences, and Engineering*, pages 997–1009. Springer, 2013.
- [106] Y. Iwasaki, M. Misumi, and T. Nakamiya. Robust vehicle detection under various environmental conditions using an infrared thermal camera and its application to road traffic flow monitoring. *Sensors*, 13(6):7756–7773, 2013.
- [107] P. L. Jacobsen. Safety in numbers: more walkers and bicyclists, safer walking and bicycling. *Injury prevention*, 21(4):271–275, 2015.
- [108] M. Jannat, D. S. Hurwitz, C. Monsere, and K. H. Funk II. The role of driver’s situational awareness on right-hook bicycle-motor vehicle crashes. *Safety science*, 110:92–101, 2018.
- [109] S. U. Jensen. Safety effects of blue cycle crossings: A before-after study. *Accident Analysis & Prevention*, 40(2):742–750, 2008.
- [110] S. U. Jensen et al. Bicycle tracks and lanes: A before-after study. In *Transportation Research Board 87th Annual Meeting*, pages 1–13, 2008.

- [111] M. Johnson, J. Charlton, J. Oxley, and S. Newstead. Naturalistic cycling study: identifying risk factors for on-road commuter cyclists. In *Annals of advances in automotive medicine/annual scientific conference*, volume 54, page 275. Association for the Advancement of Automotive Medicine, 2010.
- [112] C. Johnsson, A. Laureshyn, and T. De Ceunynck. In search of surrogate safety indicators for vulnerable road users: a review of surrogate safety indicators. *Transport Reviews*, 38(6):765–785, 2018.
- [113] S. Kaplan and C. Giacomo Prato. A spatial analysis of land use and network effects on frequency and severity of cyclist–motorist crashes in the copenhagen region. *Traffic injury prevention*, 16(7):724–731, 2015.
- [114] S. Kaplan, K. Vavatsoulas, and C. G. Prato. Aggravating and mitigating factors associated with cyclist injury severity in denmark. *Journal of safety research*, 50:75–82, 2014.
- [115] M. G. Karlaftis and A. P. Tarko. Heterogeneity considerations in accident modeling. *Accident Analysis & Prevention*, 30(4):425–433, 1998.
- [116] A. Kassim, K. Ismail, and Y. Hassan. Automated measuring of cyclist–motor vehicle post encroachment time at signalized intersections. *Canadian journal of civil engineering*, 41(7):605–614, 2014.
- [117] A. Kassim, K. Ismail, and S. McGuire. Operational evaluation of central sharrows and dooring zone treatment on road user behavior in ottawa, canada. *Transportation research record*, 2672(36):136–144, 2018.
- [118] A. Kassim, K. Ismail, and S. Woo. Investigation of the effect of super sharrows on cyclist and vehicle behavior. *Transportation Research Record*, 2659(1):224–232, 2017.

- [119] C. Katrakazas, C. Antoniou, and G. Yannis. Identification of driving simulator sessions of depressed drivers: A comparison between aggregated and time-series classification. *Transportation Research Part F: Traffic Psychology and Behaviour*, 75:16–25, 2020.
- [120] C. Katrakazas, M. Quddus, and W.-H. Chen. A new integrated collision risk assessment methodology for autonomous vehicles. *Accident Analysis & Prevention*, 127:61–79, 2019.
- [121] N. Kaya, S. Ayas, C. T. Ponnambalam, and B. Donmez. Visual attention failures during turns at intersections: An on-road study. In *Proceedings of the 28th CARSP Conference*. CARSP, Victoria, BC, 2018.
- [122] J.-K. Kim, S. Kim, G. F. Ulfarsson, and L. A. Porrello. Bicyclist injury severities in bicycle–motor vehicle accidents. *Accident Analysis & Prevention*, 39(2):238–251, 2007.
- [123] S. Kim and B. Coifman. Assessing the performance of speedinfo radar traffic sensors. *Journal of Intelligent Transportation Systems*, 21(3):179–189, 2017.
- [124] T. Kim and J. Ghosh. Robust detection of non-motorized road users using deep learning on optical and lidar data. In *2016 IEEE 19th International Conference on Intelligent Transportation Systems (ITSC)*, pages 271–276. IEEE, 2016.
- [125] K. Kircher and C. Ahlström. Attentional requirements on cyclists and drivers in urban intersections. *Transportation Research Part F: Traffic Psychology and Behaviour*, 68:105–117, 2020.
- [126] J. R. Klop and A. J. Khattak. Factors influencing bicycle crash severity on two-lane, undivided roadways in north carolina. *Transportation Research Record*, 1674(1):78–85, 1999.

- [127] M. Knodler Jr and E. Christofa. Using simulation to assess conflicts between bicyclists and right-turning vehicles.
- [128] M. C. Kondo, C. Morrison, E. Guerra, E. J. Kaufman, and D. J. Wiebe. Where do bike lanes work best? a bayesian spatial model of bicycle lanes and bicycle crashes. *Safety science*, 103:225–233, 2018.
- [129] S. Kothuri, A. Kading, A. Schrope, K. White, E. Smaglik, C. Aguilar, and W. Gil. Addressing bicycle-vehicle conflicts with alternate signal control strategies. 2018.
- [130] J. Kraay, A. van der Horst, and S. Oppe. Handleiding conflictobservatietechniek doctor [manual conflict observation technique doctor](rapport izf 1986 c-6). *Leidschendam: Institute for Road Safety Research SWOV*, 1986.
- [131] J. Kraay, A. van der Horst, and S. Oppe. Manual conflict observation technique doctor (dutch objective conflict technique for operation and research). 2013.
- [132] J. H. Kraay and A. Van der Horst. The trautenfels study: A diagnosis of road safety using the dutch conflict observation technique doctor. Technical report, 1985.
- [133] S. M. Kreidler, K. E. Muller, G. K. Grunwald, B. M. Ringham, Z. T. Coker-Dukowitz, U. R. Sakhadeo, A. E. Barón, and D. H. Glueck. Glimpse: online power computation for linear models with and without a baseline covariate. *Journal of statistical software*, 54(10), 2013.
- [134] W. H. Kruskal and W. A. Wallis. Use of ranks in one-criterion variance analysis. *Journal of the American statistical Association*, 47(260):583–621, 1952.
- [135] P. Krysiak, P. Samczynski, M. Malanowski, L. Maslikowski, and K. Kulpa. Velocity

- measurement and traffic monitoring using a gsm passive radar demonstrator. *IEEE Aerospace and Electronic Systems Magazine*, 27(10):43–51, 2012.
- [136] A. Laureshyn, T. De Ceunynck, C. Karlsson, Å. Svensson, and S. Daniels. In search of the severity dimension of traffic events: Extended delta-v as a traffic conflict indicator. *Accident Analysis & Prevention*, 98:46–56, 2017.
- [137] A. Laureshyn, M. de Goede, N. Saunier, and A. Fyhri. Cross-comparison of three surrogate safety methods to diagnose cyclist safety problems at intersections in norway. *Accident Analysis & Prevention*, 105:11–20, 2017.
- [138] A. Laureshyn, Å. Svensson, and C. Hydén. Evaluation of traffic safety, based on micro-level behavioural data: Theoretical framework and first implementation. *Accident Analysis & Prevention*, 42(6):1637–1646, 2010.
- [139] C. Leo, C. Klug, M. Ohlin, N. M. Bos, R. J. Davidse, and A. Linder. Analysis of swedish and dutch accident data on cyclist injuries in cyclist-car collisions. *Traffic injury prevention*, pages 1–3, 2019.
- [140] Q. Li, W. Wu, L. Lu, Z. Li, A. Ahmad, and G. Jeon. Infrared and visible images fusion by using sparse representation and guided filter. *Journal of Intelligent Transportation Systems*, pages 1–10, 2019.
- [141] X. Li, F. Flohr, Y. Yang, H. Xiong, M. Braun, S. Pan, K. Li, and D. M. Gavrila. A new benchmark for vision-based cyclist detection. In *2016 IEEE Intelligent Vehicles Symposium (IV)*, pages 1028–1033. IEEE, 2016.
- [142] Z. Li, M. Ye, Z. Li, and M. Du. Some operational features in bicycle traffic flow: Observational study. *Transportation Research Record*, 2520(1):18–24, 2015.

- [143] R. Ling, L. Rothman, M.-S. Cloutier, C. Macarthur, and A. Howard. Cyclist-motor vehicle collisions before and after implementation of cycle tracks in toronto, canada. *Accident Analysis & Prevention*, 135:105360, 2020.
- [144] P. Liu and S. Marker. Evaluation of contributory factors’ effects on bicycle-car crash risk at signalized intersections. *Journal of Transportation Safety & Security*, 12(1):82–93, 2020.
- [145] J. S. Long. Regression models for categorical and limited dependent variables (vol. 7). *Advanced quantitative techniques in the social sciences*, page 219, 1997.
- [146] D. Lord and F. Mannering. The statistical analysis of crash-frequency data: a review and assessment of methodological alternatives. *Transportation research part A: policy and practice*, 44(5):291–305, 2010.
- [147] J. Loskorn, A. F. Mills, J. F. Brady, J. Duthie, and R. B. Machemehl. Effects of bicycle boxes on bicyclist and motorist behavior at intersections in austin, texas. *Journal of Transportation Engineering*, 139(10):1039–1046, 2013.
- [148] A. C. Lusk, P. G. Furth, P. Morency, L. F. Miranda-Moreno, W. C. Willett, and J. T. Dennerlein. Risk of injury for bicycling on cycle tracks versus in the street. *Injury prevention*, 17(2):131–135, 2011.
- [149] A. C. Lusk, P. Morency, L. F. Miranda-Moreno, W. C. Willett, and J. T. Dennerlein. Bicycle guidelines and crash rates on cycle tracks in the united states. *American journal of public health*, 103(7):1240–1248, 2013.
- [150] Q. Ma, H. Yang, Y. Ma, D. Yang, X. Hu, and K. Xie. Examining municipal guidelines for users of shared e-scooters in the united states. *Transportation research part D: transport and environment*, 92:102710, 2021.

- [151] Y. Ma, X. Wu, G. Yu, Y. Xu, and Y. Wang. Pedestrian detection and tracking from low-resolution unmanned aerial vehicle thermal imagery. *Sensors*, 16(4):446, 2016.
- [152] C. A. MacCarley, B. M. Hemme, and L. Klein. Evaluation of infrared and millimeter-wave imaging technologies applied to traffic management. Technical report, SAE Technical Paper, 2000.
- [153] T. K. O. Madsen and H. Lahrmann. Comparison of five bicycle facility designs in signalized intersections using traffic conflict studies. *Transportation research part F: traffic psychology and behaviour*, 46:438–450, 2017.
- [154] K. Martens. Role of the bicycle in the limitation of transport poverty in the netherlands. *Transportation research record*, 2387(1):20–25, 2013.
- [155] Massachusetts Department of Transportation. Massdot separated bike lane planning & design guide. Technical report, 2015.
- [156] J. C. Medina, R. F. Benekohal, and H. Ramezani. Field evaluation of smart sensor vehicle detectors at intersections—volume 1: Normal weather conditions. Technical report, 2012.
- [157] J. C. Medina, R. F. Benekohal, and H. Ramezani. Field evaluation of smart sensor vehicle detectors at intersections—volume 2: Performance under adverse weather conditions. Technical report, 2013.
- [158] G. Mehta and Y. Lou. Calibration and development of safety performance functions for alabama: Two-lane, two-way rural roads and four-lane divided highways. *Transportation research record*, 2398(1):75–82, 2013.

- [159] E. Minikel. Cyclist safety on bicycle boulevards and parallel arterial routes in berkeley, california. *Accident Analysis & Prevention*, 45:241–247, 2012.
- [160] R. Mitra and S. Nash. Can the built environment explain gender gap in cycling? an exploration of university students’ travel behavior in toronto, canada. *International journal of sustainable transportation*, 13(2):138–147, 2019.
- [161] C. Monsere, J. Dill, N. McNeil, K. J. Clifton, N. Foster, T. Goddard, M. Berkow, J. Gilpin, K. Voros, D. van Hengel, et al. Lessons from the green lanes: Evaluating protected bike lanes in the us. 2014.
- [162] C. M. Monsere, N. Foster, J. Dill, and N. McNeil. User behavior and perceptions at intersections with turning and mixing zones on protected bike lanes. *Transportation Research Record*, 2520(1):112–122, 2015.
- [163] D. N. Moore, W. H. Schneider IV, P. T. Savolainen, and M. Farzaneh. Mixed logit analysis of bicyclist injury severity resulting from motor vehicle crashes at intersection and non-intersection locations. *Accident Analysis & Prevention*, 43(3):621–630, 2011.
- [164] C. N. Morrison, J. Thompson, M. C. Kondo, and B. Beck. On-road bicycle lane types, roadway characteristics, and risks for bicycle crashes. *Accident Analysis & Prevention*, 123:123–131, 2019.
- [165] R. O. Mujalli, G. López, and L. Garach. Bayes classifiers for imbalanced traffic accidents datasets. *Accident Analysis & Prevention*, 88:37–51, 2016.
- [166] B. Murphy and A. Owen. Implementing low-stress bicycle routing in national accessibility evaluation. *Transportation research record*, 2673(5):240–249, 2019.

- [167] M. C. Myers, J. E. Atkinson, T. E. Bryer, R. Retting, J. Smithline, J. Trim, P. Wojtkiewicz, G. B. Thomas, S. P. Venglar, S. Sunkari, B. J. Malone, and P. Izadpanah. Signalized Intersections Informational Guide, Second Edition. Technical report, Federal Highway Administration, 2011.
- [168] National Association of City Transportation Officials. Urban bikeway design guide. Technical report, National Association of City Transportation Officials, 55 Water Street, New York, NY 10041, 2014.
- [169] National Highway Traffic Safety Administration. 2018 fatal motor vehicle crashes: Overview. Technical report, U.S. Department of Transportation, 2019.
- [170] National Highway Traffic Safety Administration (NHTSA). 2018 Fatal Motor Vehicle Crashes: Overview. = <https://crashstats.nhtsa.dot.gov/Api/Public/ViewPublication/812826>, 2019.
- [171] National Research Council (U.S.). Transportation Research Board. Task Force on Development of the Highway Safety Manual and Transportation Officials. Joint Task Force on the Highway Safety Manual. *Highway Safety Manual*, volume 1. AASHTO, 2010.
- [172] M. S. N. Niaki, N. Saunier, and L. F. Miranda-Moreno. Is that move safe? case study of cyclist movements at intersections with cycling discontinuities. *Transportation Research Record*, 2019.
- [173] K. Nordback, W. E. Marshall, and B. N. Janson. Bicyclist safety performance functions for a us city. *Accident Analysis & Prevention*, 65:114–122, 2014.
- [174] K. Nordback, W. E. Marshall, B. N. Janson, and E. Stolz. Estimating annual average daily bicyclists: Error and accuracy. *Transportation research record*, 2339(1):90–97, 2013.

- [175] N. A. of City Transportation Officials. *Urban bikeway design guide*. Island Press, 2014.
- [176] N. A. of City Transportation Officials. Don't give up at the intersection. designing for all ages and abilities. Technical report, National Association of City Transportation Officials, 2019.
- [177] A. of the Advancement in Automotive Medicine. Abbreviated injury scale. <https://www.aaam.org/abbreviated-injury-scale-ais/>, 2015.
- [178] J. Oh, J. Jun, E. Kim, and M. Kim. Assessing critical factors associated with bicycle collisions at urban signalized intersections. Technical report, 2008.
- [179] D. Olmeda, C. Premebida, U. Nunes, J. M. Armingol, and A. de la Escalera. Pedestrian detection in far infrared images. *Integrated Computer-Aided Engineering*, 20(4):347–360, 2013.
- [180] Oregon Department of Transportation. Crash Data System. = <https://zigzag.odot.state.or.us>, 2018.
- [181] D. Otte, M. Jänsch, and C. Haasper. Injury protection and accident causation parameters for vulnerable road users based on german in-depth accident study gidas. *Accident Analysis & Prevention*, 44(1):149–153, 2012.
- [182] M. Ozbayoglu, G. Kucukayan, and E. Dogdu. A real-time autonomous highway accident detection model based on big data processing and computational intelligence. In *2016 IEEE International Conference on Big Data (Big Data)*, pages 1807–1813. IEEE, 2016.
- [183] J. Park, M. Abdel-Aty, J. Lee, and C. Lee. Developing crash modification functions to assess safety effects of adding bike lanes for urban arterials with

- different roadway and socio-economic characteristics. *Accident Analysis & Prevention*, 74:179–191, 2015.
- [184] M. Parker Jr and C. V. Zegeer. Traffic conflict techniques for safety and operations: Observers manual. Technical report, United States. Federal Highway Administration, 1989.
- [185] A. B. Parsa, H. Taghipour, S. Derrible, and A. K. Mohammadian. Real-time accident detection: coping with imbalanced data. *Accident Analysis & Prevention*, 129:202–210, 2019.
- [186] Pedestrian and Bicycle Information Center. Safety. http://www.pedbikeinfo.org/factsfigures/facts_safety.cfm, 2019. Online; accessed 29 November 2019.
- [187] E. Petrucelli, J. D. States, and L. N. Hames. The abbreviated injury scale: Evolution, usage and future adaptability. *Accident Analysis & Prevention*, 13(1):29–35, 1981.
- [188] J. M. Plumert, J. K. Kearney, E. O’Neal, L. D. Subramanian, et al. Using simulation to assess and reduce conflicts between drivers and bicyclists. 2020.
- [189] A. K. Porter and M. Schwartz. Ride report: mobile app user guide. *British journal of sports medicine*, 52(18):e4–e4, 2018.
- [190] Portland Bureau of Transportation, PBOT. Traffic counts. = <https://www.portlandoregon.gov/transportation/article/180473>, 2019.
- [191] Portland Department of Transportation. <https://www.portlandoregon.gov/transportation/article/680761>.

- [192] Portland State University. User Documentation for BikePed Portal. <https://github.com/PSUTrec/documentation>, 2018.
- [193] R. Poulos, J. Hatfield, C. Rissel, L. Flack, S. Murphy, R. Grzebieta, and A. McIntosh. An exposure based study of crash and injury rates in a cohort of transport and recreational cyclists in new south wales, australia. *Accident Analysis & Prevention*, 78:29–38, 2015.
- [194] G. Prati, L. Pietrantonio, and F. Fraboni. Using data mining techniques to predict the severity of bicycle crashes. *Accident Analysis & Prevention*, 101:44–54, 2017.
- [195] J. Pucher and L. Dijkstra. Making walking and cycling safer: lessons from europe. *Transportation Quarterly*, 54(3):25–50, 2000.
- [196] J. R. Pucher and R. Buehler. *City cycling*, volume 11. MIT Press Cambridge, MA, 2012.
- [197] S. S. Pulugurtha and V. Thakur. Evaluating the effectiveness of on-street bicycle lane and assessing risk to bicyclists in charlotte, north carolina. *Accident Analysis & Prevention*, 76:34–41, 2015.
- [198] F. M. Puscar, T. Sayed, A. Y. Bigazzi, and M. H. Zaki. Multimodal safety assessment of an urban intersection by video analysis of bicycle, pedestrian, and motor vehicle traffic conflicts and violations. Technical report, 2018.
- [199] A. H. A. Rahim, N. A. Rashid, A. Nayan, and A.-R. Ahmad. SMOTE Approach to Imbalanced Dataset in Logistic Regression Analysis. pages 429–433. Springer.
- [200] M. A. Raihan, P. Alluri, W. Wu, and A. Gan. Estimation of bicycle crash modification factors (cmfs) on urban facilities using zero inflated negative binomial models. *Accident Analysis & Prevention*, 123:303–313, 2019.

- [201] A. Razmpa. An assessment of post-encroachment times for bicycle-vehicle interactions observed in the field, a driving simulator, and in traffic simulation models. 2016.
- [202] C. C. Reynolds, M. A. Harris, K. Teschke, P. A. Cripton, and M. Winters. The impact of transportation infrastructure on bicycling injuries and crashes: a review of the literature. *Environmental health*, 8(1):47, 2009.
- [203] RideReport. <https://www.ridereport.com/>.
- [204] C. J. Robbins and P. Chapman. Drivers' visual search behavior toward vulnerable road users at junctions as a function of cycling experience. *Human factors*, 60(7):889–901, 2018.
- [205] A. Rodgers. Asl before & after study report no. t40900. Technical report, Transport for London, Cycling Centre of Excellence, Hertfordshire, UK, 2005.
- [206] H. Rothenberg, D. Goodman, and C. Sundstrom. Separated bike lane crash analysis. Technical report, 2016.
- [207] S. Royo and M. Ballesta-Garcia. An overview of lidar imaging systems for autonomous vehicles. *Applied Sciences*, 9(19):4093, 2019.
- [208] B. Russo, S. Kothuri, E. Smaglik, C. Aguilar, E. James, N. Levenson, and P. Koonce. Exploring the impacts of intersection and traffic characteristics on the frequency and severity of bicycle-vehicle conflicts. *Advances in Transportation Studies*, 2020.
- [209] A. Ryan, E. Casola, C. Fitzpatrick, and M. Knodler Jr. Flashing yellow arrows for right turn applications: A driving simulator study and static evaluation analysis.

- Transportation research part F: traffic psychology and behaviour*, 66:324–338, 2019.
- [210] A. Ryan, C. Fitzpatrick, E. Christofa, and M. Knodler Jr. Driver performance due to small unmanned aerial system applications in the vicinity of roadways. *Transportation research part F: traffic psychology and behaviour*, 68:118–131, 2020.
- [211] M. Saad, M. Abdel-Aty, J. Lee, and Q. Cai. Bicycle safety analysis at intersections from crowdsourced data. *Transportation Research Record*, page 0361198119836764, 2019.
- [212] R. L. Sanders. Examining the cycle: how perceived and actual bicycling risk influence cycling frequency, roadway design preferences, and support for cycling among bay area residents. 2013.
- [213] R. L. Sanders. We can all get along: The alignment of driver and bicyclist roadway design preferences in the san francisco bay area. *Transportation Research Part A: Policy and Practice*, 91:120–133, 2016.
- [214] K. R. Santiago-Chaparro, M. Chitturi, A. Bill, and D. A. Noyce. Expanding the capabilities of existing vehicle detection infrastructure to monitor red light running. *International Journal of Engineering Management and Economics*, 5(3-4):196–208, 2015.
- [215] K. R. Santiago-Chaparro, M. Chitturi, A. Bill, and D. A. Noyce. Automated turning movement counts for shared lanes: leveraging vehicle detection data. *Transportation Research Record*, 2558(1):30–40, 2016.
- [216] K. R. Santiago-Chaparro, M. Chitturi, A. R. Bill, and D. A. Noyce. Real time performance measures from radar-based vehicle detection systems. In *19th*

ITS World Congress ERTICO-ITS Europe European Commission ITS America ITS Asia-Pacific, 2012.

- [217] K. R. Santiago-Chaparro and D. A. Noyce. Expanding the capabilities of radar-based vehicle detection systems: Noise characterization and removal procedures. *Transportation Research Record*, page 0361198119852607, 2019.
- [218] N. Saunier and T. Sayed. Probabilistic framework for automated analysis of exposure to road collisions. *Transportation research record*, 2083(1):96–104, 2008.
- [219] T. Sayed, M. H. Zaki, and J. Autey. Automated safety diagnosis of vehicle–bicycle interactions using computer vision analysis. *Safety science*, 59:163–172, 2013.
- [220] J. Schepers, P. Kroeze, W. Sweers, and J. Wüst. Road factors and bicycle–motor vehicle crashes at unsignalized priority intersections. *Accident Analysis & Prevention*, 43(3):853–861, 2011.
- [221] P. Schepers, M. Hagenzieker, R. Methorst, B. van Wee, and F. Wegman. A conceptual framework for road safety and mobility applied to cycling safety. *Accident Analysis & Prevention*, 62:331–340, 2014.
- [222] P. Schepers, D. Twisk, E. Fishman, A. Fyhri, and A. Jensen. The dutch road to a high level of cycling safety. *Safety science*, 92:264–273, 2017.
- [223] S. Seabold and J. Perktold. statsmodels: Econometric and statistical modeling with python. In *9th Python in Science Conference*, 2010.
- [224] I. N. Sener, N. Eluru, and C. R. Bhat. An analysis of bicycle route choice preferences in texas, us. *Transportation*, 36(5):511–539, 2009.

- [225] M. A. Shafique and E. Hato. Use of acceleration data for transportation mode prediction. *Transportation*, 42(1):163–188, 2015.
- [226] S. S. Shapiro and M. Wilk. An analysis of variance test for the exponential distribution (complete samples). *Technometrics*, 14(2):355–370, 1972.
- [227] D. Shinar, P. Valero-Mora, M. van Strijp-Houtenbos, N. Haworth, A. Schramm, G. De Bruyne, V. Cavallo, J. Chliaoutakis, J. Dias, O. E. Ferraro, et al. Under-reporting bicycle accidents to police in the cost tu1101 international survey: Cross-country comparisons and associated factors. *Accident Analysis & Prevention*, 110:177–186, 2018.
- [228] A. P. Silvano, H. N. Koutsopoulos, and X. Ma. Analysis of vehicle-bicycle interactions at unsignalized crossings: A probabilistic approach and application. *Accident Analysis & Prevention*, 97:38–48, 2016.
- [229] C. S. Smith, J.-S. Oh, and C. Lei. Exploring the equity dimensions of us bicycle sharing systems. Technical report, Western Michigan University. Transportation Research Center for Livable . . . , 2015.
- [230] M. Sokolova and G. Lapalme. A systematic analysis of performance measures for classification tasks. *Information processing & management*, 45(4):427–437, 2009.
- [231] R. Srinivasan, D. Carter, and K. Bauer. Safety performance function decision guide: Spf calibration vs spf development. *Federal Highway Administration–Office of Safety Report*, 2013.
- [232] M. A. Stinson and C. R. Bhat. Commuter bicyclist route choice: Analysis using a stated preference survey. *Transportation Research Record*, 1828(1):107–115, 2003.

- [233] M. Stone and J. Broughton. Getting off your bike: cycling accidents in great britain in 1990–1999. *Accident Analysis & Prevention*, 35(4):549–556, 2003.
- [234] J. Strauss, L. F. Miranda-Moreno, and P. Morency. Cyclist activity and injury risk analysis at signalized intersections: A bayesian modelling approach. *Accident Analysis & Prevention*, 59:9–17, 2013.
- [235] J. Strauss, L. F. Miranda-Moreno, and P. Morency. Mapping cyclist activity and injury risk in a network combining smartphone gps data and bicycle counts. *Accident Analysis & Prevention*, 83:132–142, 2015.
- [236] J. Strauss, S. Zangenehpour, L. F. Miranda-Moreno, and N. Saunier. Cyclist deceleration rate as surrogate safety measure in montreal using smartphone gps data. *Accident Analysis & Prevention*, 99:287–296, 2017.
- [237] J. C. Stutts, J. E. Williamson, T. Whitley, and F. C. Sheldon. Bicycle accidents and injuries: A pilot study comparing hospital-and police-reported data. *Accident Analysis & Prevention*, 22(1):67–78, 1990.
- [238] H. Summala, E. Pasanen, M. Räsänen, and J. Sievänen. Bicycle accidents and drivers’ visual search at left and right turns. *Accident Analysis & Prevention*, 28(2):147–153, 1996.
- [239] Z. Sun and X. J. Ban. Vehicle classification using gps data. *Transportation Research Part C: Emerging Technologies*, 37:102–117, 2013.
- [240] C. A. Sundstrom, S. M. Quinn, and R. Weld. Bicyclist crash comparison of mixing zone and fully split phase signal treatments at intersections with protected bicycle lanes in new york city. *Transportation Research Record*, page 0361198119859301, 2019.

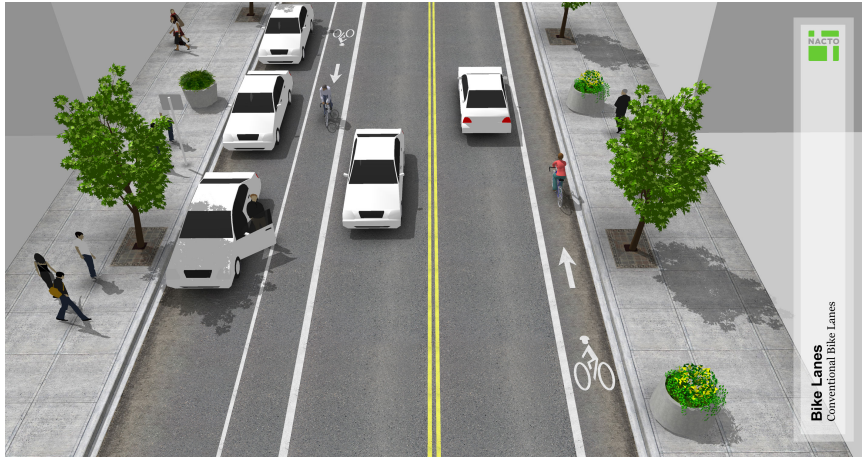
- [241] J. A. Swets. Measuring the accuracy of diagnostic systems. *Science*, 240(4857):1285–1293, 1988.
- [242] I. Syarif, A. Prugel-Bennett, and G. Wills. Svm parameter optimization using grid search and genetic algorithm to improve classification performance. *Telkomnika*, 14(4):1502, 2016.
- [243] F. Tainter, E. Christofa, and M. A. Knodler. All-red clearance intervals for use in the left-turn application of flashing yellow arrows. *Journal of Transportation Engineering, Part A: Systems*, 146(4):04020008, 2020.
- [244] K. Teschke, M. A. Harris, C. C. Reynolds, M. Winters, S. Babul, M. Chipman, M. D. Cusimano, J. R. Brubacher, G. Hunte, S. M. Friedman, et al. Route infrastructure and the risk of injuries to bicyclists: a case-crossover study. *American journal of public health*, 102(12):2336–2343, 2012.
- [245] The League of American Bicyclists. Topic III: Rates of Active Commuting. = <https://data.bikeleague.org/show-your-data/state-data/states-rates-of-active-commuting/>, 2021.
- [246] B. Thomas and M. DeRobertis. The safety of urban cycle tracks: A review of the literature. *Accident Analysis & Prevention*, 52:219–227, 2013.
- [247] N. Y. Tilahun, D. M. Levinson, and K. J. Krizek. Trails, lanes, or traffic: Valuing bicycle facilities with an adaptive stated preference survey. *Transportation Research Part A: Policy and Practice*, 41(4):287–301, 2007.
- [248] B. Tucker and K. Manaugh. Bicycle equity in brazil: Access to safe cycling routes across neighborhoods in rio de janeiro and curitiba. *International journal of sustainable transportation*, 12(1):29–38, 2018.

- [249] S. Turner, I. N. Sener, M. E. Martin, S. Das, R. C. Hampshire, K. Fitzpatrick, L. J. Molnar, M. Colety, S. Robinson, E. Shipp, et al. Synthesis of methods for estimating pedestrian and bicyclist exposure to risk at areawide levels and on specific transportation facilities. Technical report, United States. Federal Highway Administration. Office of Safety, 2017.
- [250] S. Turner, G. Wood, T. Hughes, and R. Singh. Safety performance functions for bicycle crashes in new zealand and australia. *Transportation research record*, 2236(1):66–73, 2011.
- [251] H. Twaddle. Analysis and modelling of the operational and tactical behaviour of bicyclists. In *UR: BAN Human Factors in Traffic*, pages 325–345. Springer, 2018.
- [252] H. Twaddle, T. Schendzielorz, and O. Fakler. Bicycles in urban areas: Review of existing methods for modeling behavior. *Transportation research record*, 2434(1):140–146, 2014.
- [253] W. van Haperen, S. Daniels, T. De Ceunynck, N. Saunier, T. Brijs, and G. Wets. Yielding behavior and traffic conflicts at cyclist crossing facilities on channelized right-turn lanes. *Transportation research part F: traffic psychology and behaviour*, 55:272–281, 2018.
- [254] M. Vasilev, K. Pitera, and T. Jonsson. Evaluation of bicycle sharrows within the norwegian context. *Transportation research procedia*, 27:1097–1104, 2017.
- [255] E. T. Verhoef. Speed-flow relations and cost functions for congested traffic: Theory and empirical analysis. *Transportation Research Part A: Policy and Practice*, 39(7-9):792–812, 2005.

- [256] N. Verzosa and R. Miles. Severity of road crashes involving pedestrians in metro manila, philippines. *Accident Analysis & Prevention*, 94:216–226, 2016.
- [257] Vision Zero Network. How does vision zero differ from the traditional traffic safety approach in u.s. communities? <http://visionzeronetWORK.org/wp-content/uploads/2016/03/VZN-Case-Study-1-What-makes-VZ-different.pdf>, 2019.
- [258] K. Vogel. A comparison of headway and time to collision as safety indicators. *Accident analysis & prevention*, 35(3):427–433, 2003.
- [259] X. Wang and M. Abdel-Aty. Temporal and spatial analyses of rear-end crashes at signalized intersections. *Accident Analysis & Prevention*, 38(6):1137–1150, 2006.
- [260] X. Wang and M. Abdel-Aty. Modeling left-turn crash occurrence at signalized intersections by conflicting patterns. *Accident Analysis & Prevention*, 40(1):76–88, 2008.
- [261] J. Warner, D. S. Hurwitz, C. M. Monsere, and K. Fleskes. A simulator-based analysis of engineering treatments for right-hook bicycle crashes at signalized intersections. *Accident Analysis & Prevention*, 104:46–57, 2017.
- [262] S. Washington, M. G. Karlaftis, F. Mannering, and P. Anastasopoulos. *Statistical and econometric methods for transportation data analysis*. CRC press, 2020.
- [263] W. Weijermars, N. Bos, and H. L. Stipdonk. Serious road injuries in the netherlands dissected. *Traffic injury prevention*, 17(1):73–79, 2016.
- [264] J. Werneke and M. Vollrath. What does the driver look at? the influence of

- intersection characteristics on attention allocation and driving behavior. *Accident Analysis & Prevention*, 45:610–619, 2012.
- [265] J. M. Wood, P. F. Lacherez, R. P. Marszalek, and M. J. King. Drivers’ and cyclists’ experiences of sharing the road: Incidents, attitudes and perceptions of visibility. *Accident Analysis & Prevention*, 41(4):772–776, 2009.
- [266] J. M. Wood, R. A. Tyrrell, R. Marszalek, P. Lacherez, and T. Carberry. Bicyclists overestimate their own night-time conspicuity and underestimate the benefits of retroreflective markers on the moveable joints. *Accident Analysis & Prevention*, 55:48–53, 2013.
- [267] World Health Organization. Global status report on road safety 2015. Technical report, World Health Organization, 2016.
- [268] J. Wu, H. Xu, Y. Zheng, Y. Zhang, B. Lv, and Z. Tian. Automatic vehicle classification using roadside lidar data. *Transportation Research Record*, 2673(6):153–164, 2019.
- [269] G. Xiao, Z. Juan, and C. Zhang. Travel mode detection based on gps track data and bayesian networks. *Computers, Environment and Urban Systems*, 54:14–22, 2015.
- [270] X. Yan, E. Radwan, and M. Abdel-Aty. Characteristics of rear-end accidents at signalized intersections using multiple logistic regression model. *Accident Analysis & Prevention*, 37(6):983–995, 2005.
- [271] B. Yang and Y. Lei. Vehicle detection and classification for low-speed congested traffic with anisotropic magnetoresistive sensor. *IEEE Sensors Journal*, 15(2):1132–1138, 2014.

- [272] L. Yu. *Modeling Transportation Emissions Using Radar-based Vehicle Detection Data*. The University of Wisconsin-Madison, 2016.
- [273] S. Zangenehpour, L. F. Miranda-Moreno, and N. Saunier. Automated classification based on video data at intersections with heavy pedestrian and bicycle traffic: Methodology and application. *Transportation research part C: emerging technologies*, 56:161–176, 2015.
- [274] S. Zangenehpour, J. Strauss, L. F. Miranda-Moreno, and N. Saunier. Are signalized intersections with cycle tracks safer? a case-control study based on automated surrogate safety analysis using video data. *Accident Analysis & Prevention*, 86:161–172, 2016.
- [275] J. Zhang, R. Pi, X. Ma, J. Wu, H. Li, and Z. Yang. Object classification with roadside lidar data using a probabilistic neural network. *Electronics*, 10(7):803, 2021.
- [276] J. Zhao, H. Xu, H. Liu, J. Wu, Y. Zheng, and D. Wu. Detection and tracking of pedestrians and vehicles using roadside lidar sensors. *Transportation research part C: emerging technologies*, 100:68–87, 2019.
- [277] N. Zuniga-Garcia, N. R. Juri, K. A. Perrine, and R. B. Machemehl. E-scooters in urban infrastructure: Understanding sidewalk, bike lane, and roadway usage from trajectory data. *Case Studies on Transport Policy*, 2021.



(a) Simple bike lanes



(b) Simple bike lanes with parking buffer



(c) Green-colored bike lanes

Figure 2.5: Different configurations of conventional bike lanes

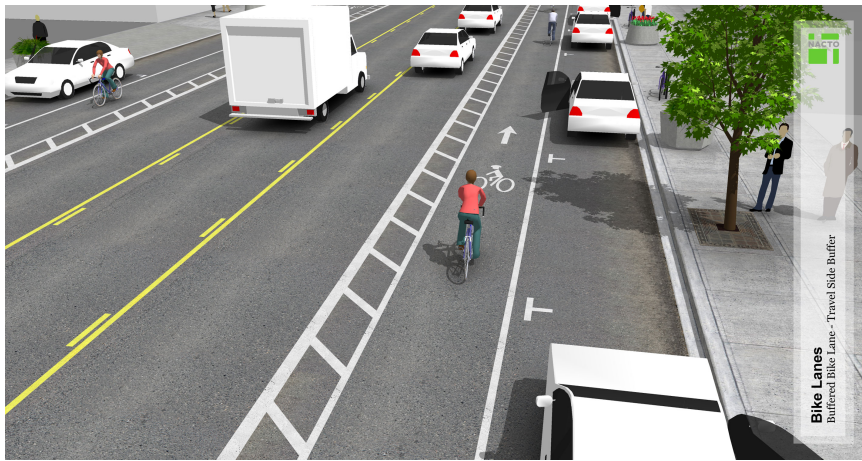
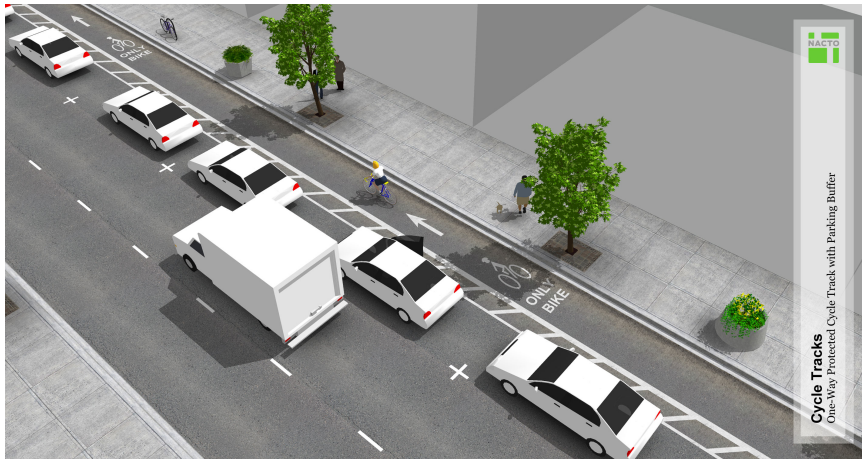


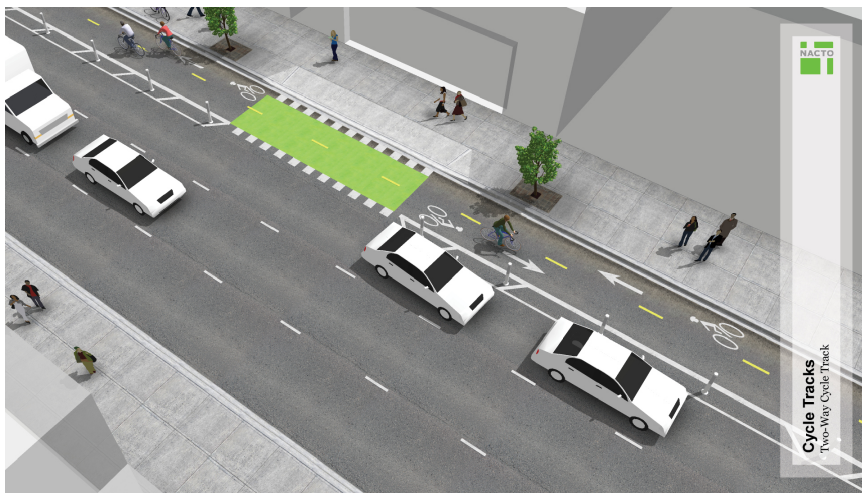
Figure 2.6: Buffered bike lanes



(a) Protected bike lane separated from traffic with a parking lane



(b) Raised protected bike lane

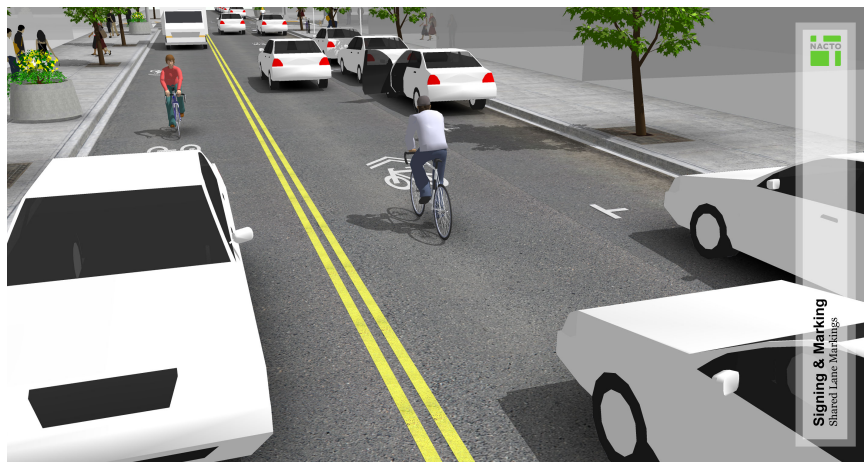


(c) Two-ways cycle track

Figure 2.7: Different configurations of protected bike lanes



(a) Bicycle boulevards



(b) Sharrows

Figure 2.8: Bicycle treatments that enhanced bicyclists to use the full traffic lane along with motorized vehicles



Figure 2.9: Bike box



Figure 2.10: Intersection-crossing pavement marking and turning queue box



Figure 2.11: Protected or “Dutch” intersection

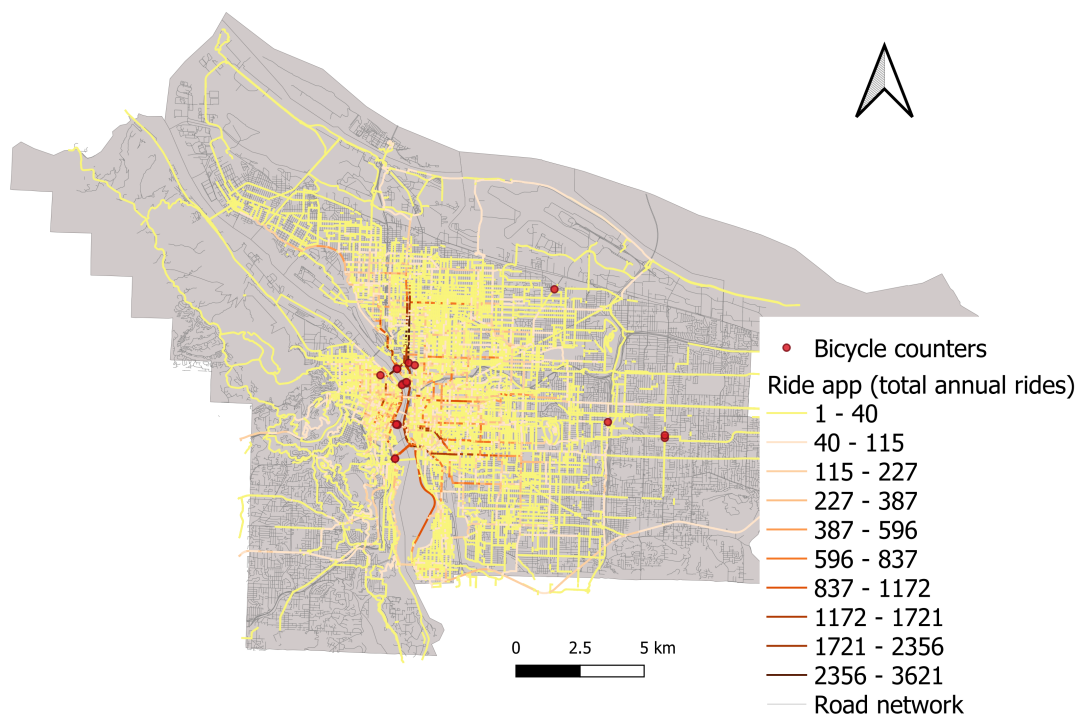
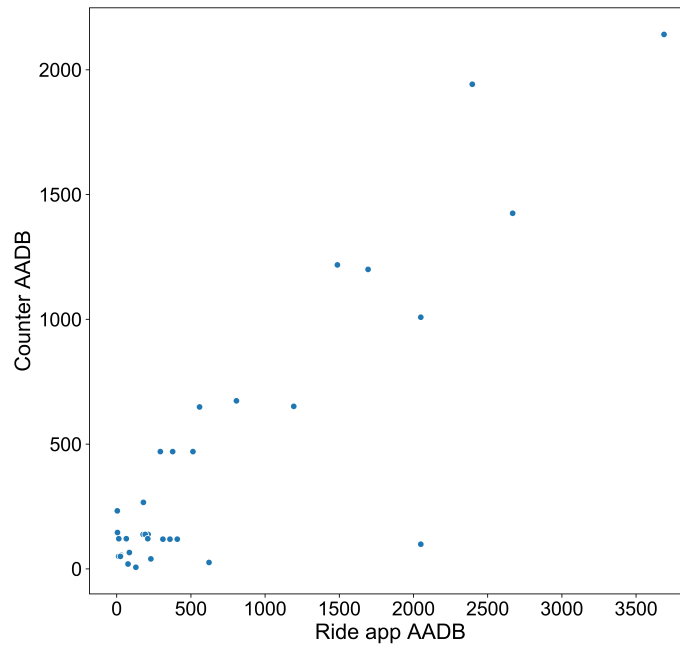


Figure 3.1: Segments with Ride app data (in total annual rides) and bicycle counter locations in the City of Portland



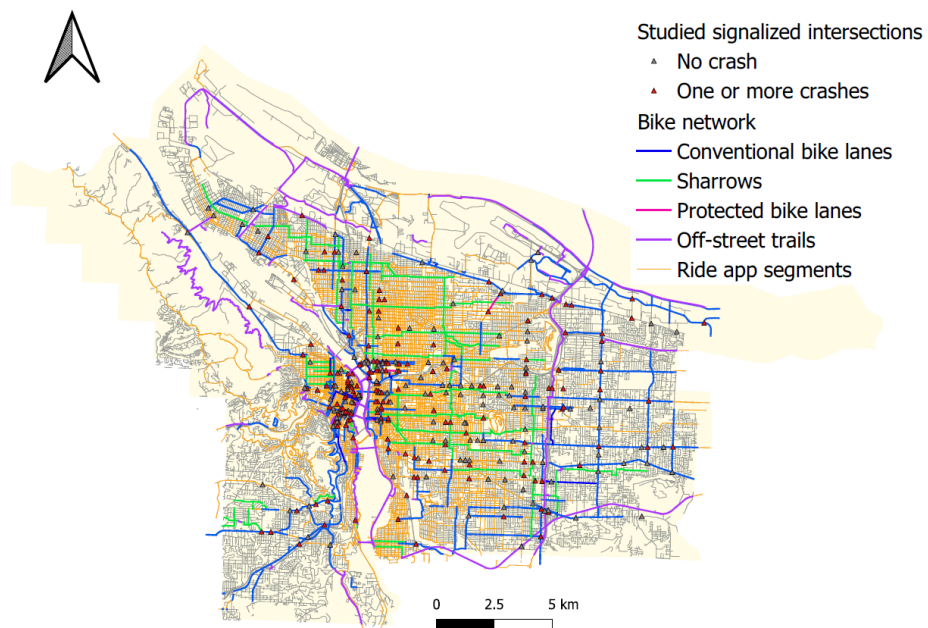
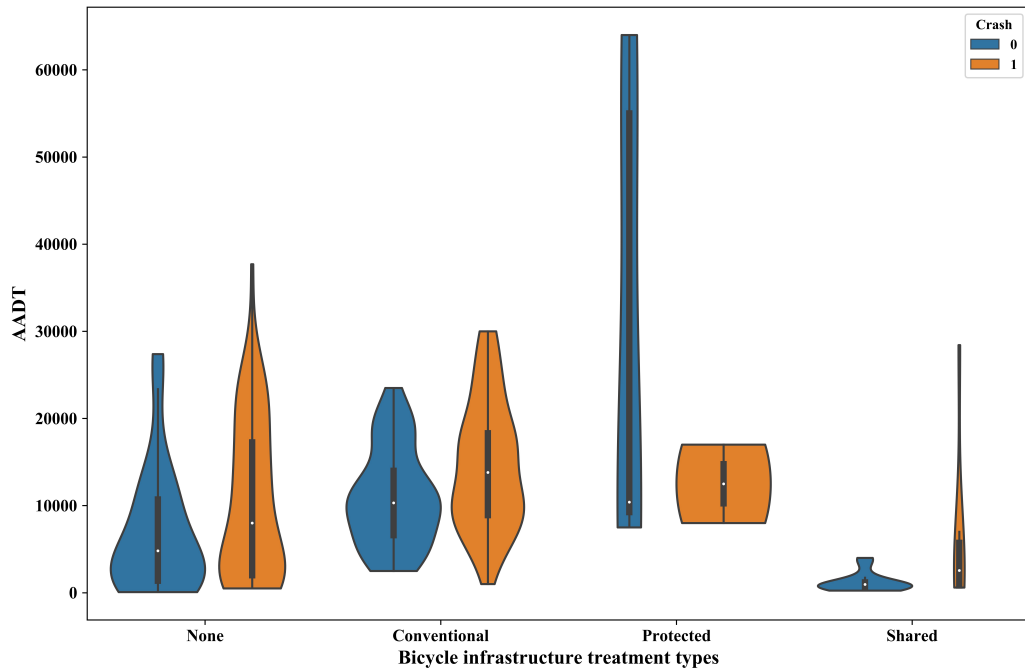
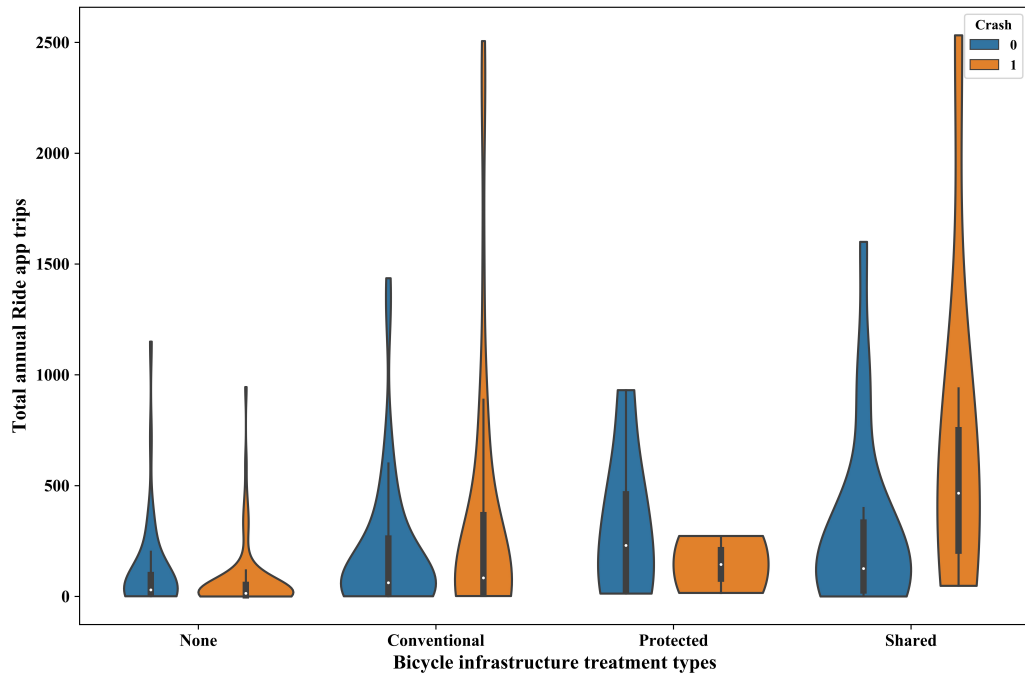


Figure 3.4: Bicycle-motorized vehicle crashes at signalized intersections.



(a) AADT and crash occurrence across the different bicycle treatment types



(b) Ride app trips and crash occurrence across the different bicycle treatment types

Figure 3.5: Distribution of motorized vehicle and bicycle demand for the studied road segments

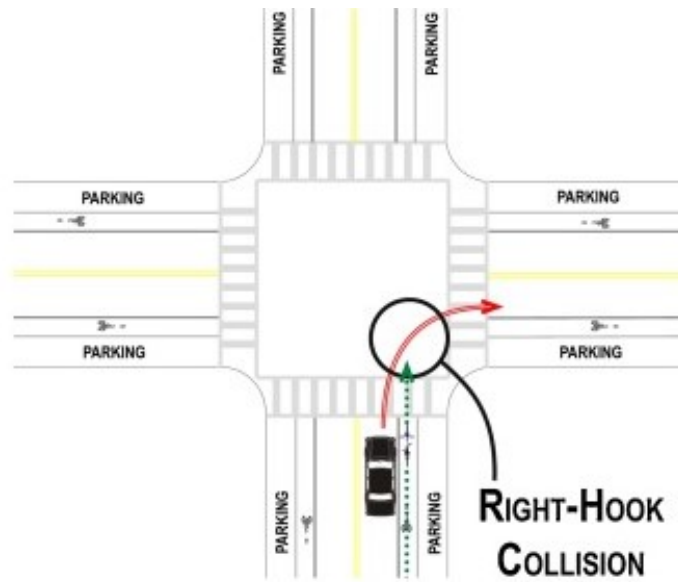


Figure 4.1: Right-hook collision between a bicycle and a right-turning vehicle (Adopted from [75])



Figure 4.2: Example of green-colored intersection crossing markings (Seattle, WA)



Figure 4.3: Example of bike box (Cambridge, MA)

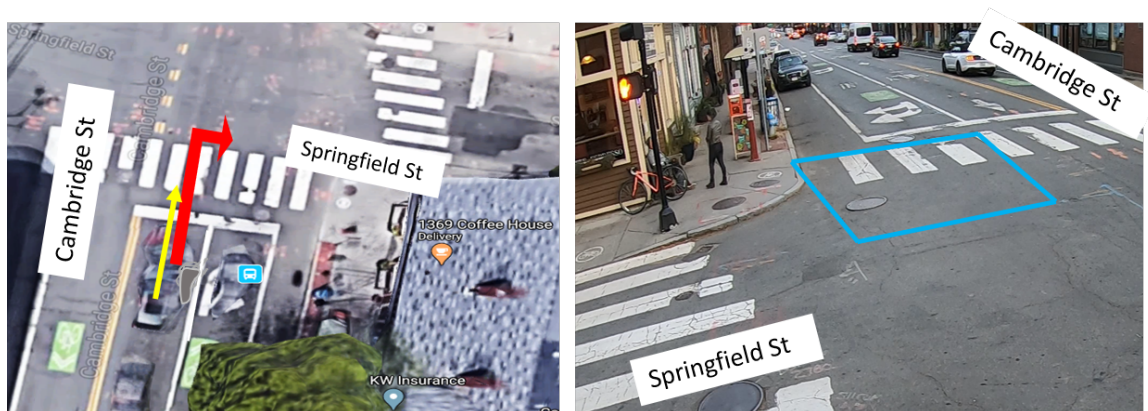
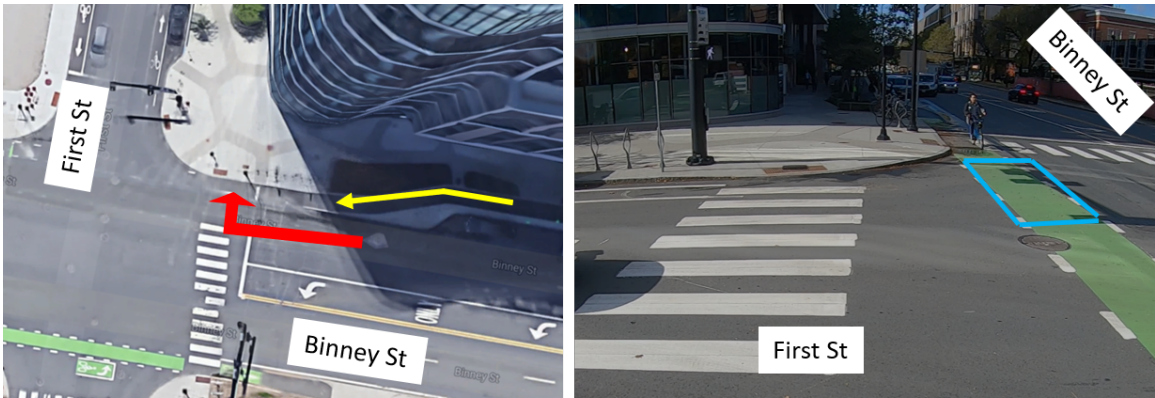


Figure 4.4: Cambridge Street at Springfield Street (Cambridge, MA). [Segment: sharrows; Intersection: None]



(a) Binney & First St

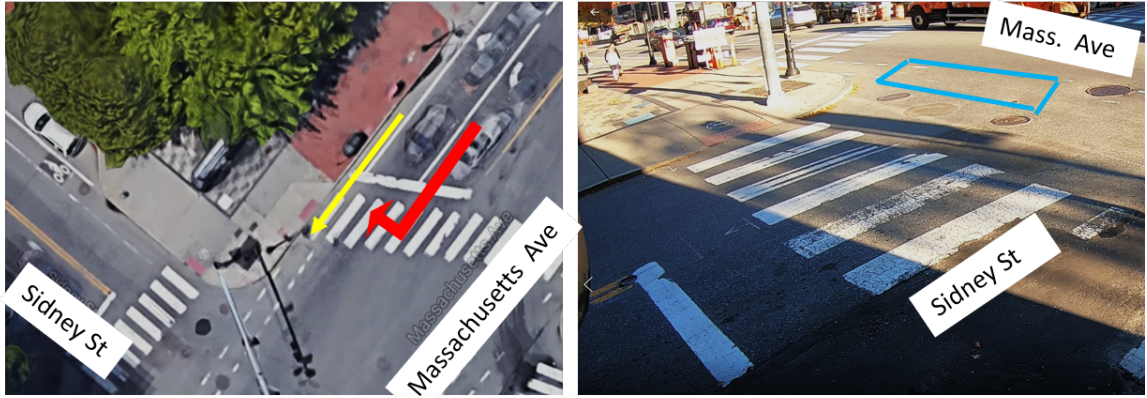


(b) Binney & Third St

Figure 4.5: Binney street (Cambridge, MA). Bicycle treatment type: protected bike lanes with green-colored intersection crossing markings



Figure 4.6: Western Ave at Memorial Drive (Cambridge, MA). [Segment: protected bike lanes; Intersection: None]



(a) Massachusetts Ave & Sidney St



(b) Massachusetts Ave & Albany

Figure 4.7: Massachusetts Avenue at... (Cambridge, MA). [Segment: conventional bike lane; Intersection: crossing Markings]



Figure 4.8: Cambridge Street at Sudbury Street (Boston, MA). [Segment: conventional bike lane; Intersection: bike box]

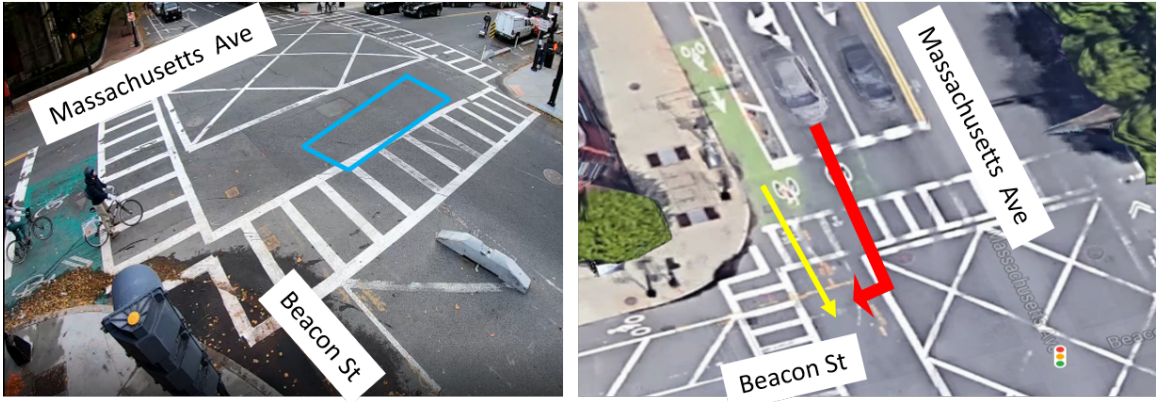


Figure 4.9: Massachusetts Avenue at Beacon Street (Boston, MA). [Segment: protected bike lane; Intersection: bike box and crossing markings]



Figure 4.10: Massachusetts Avenue at Commonwealth Avenue (Boston, MA). [Segment: conventional bike lane; Intersection: bike box and crossing markings]

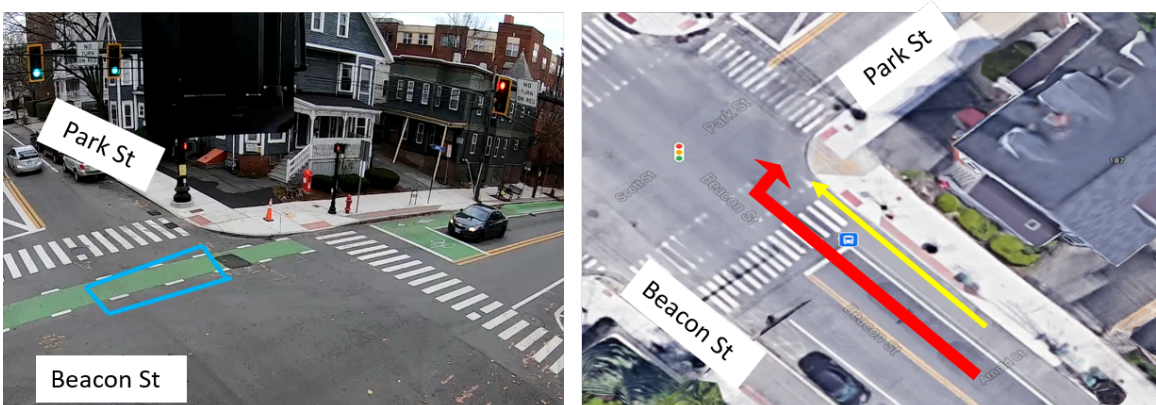


Figure 4.11: Beacon Street at Street (Somerville, MA). [Segment: conventional bike lane; Intersection: bike box and crossing markings (not visible on Google Maps imagery due to recent installation)]

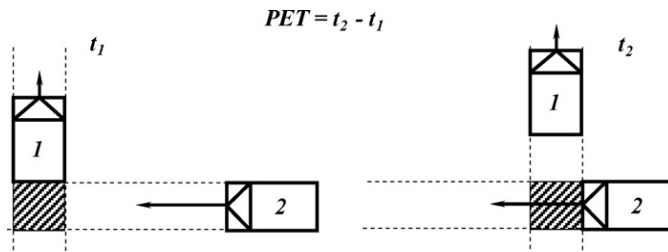


Figure 4.12: Post Encroachment Time graphical representation; adopted from [7]

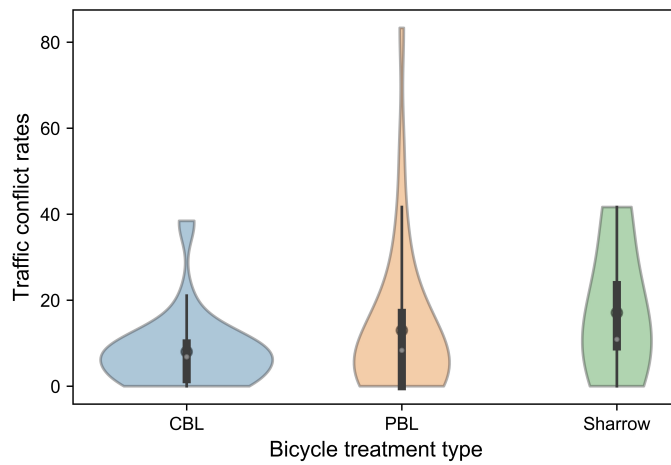
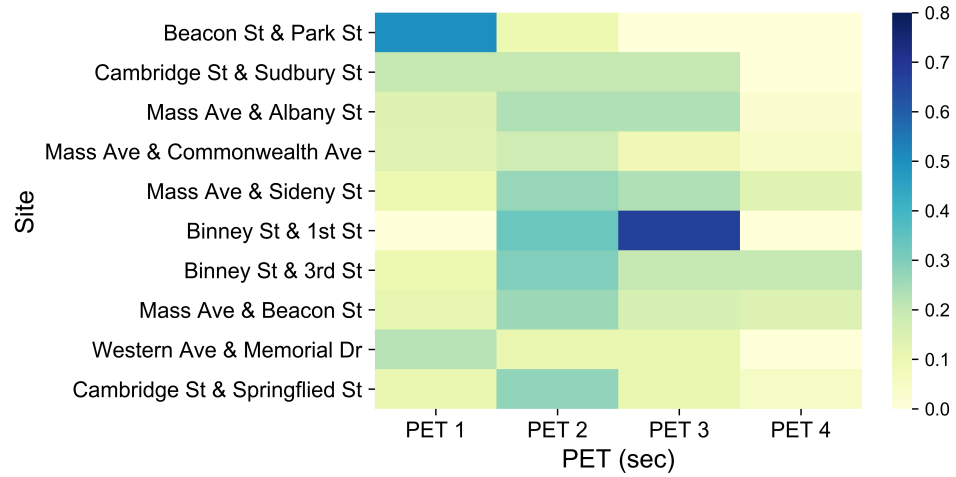
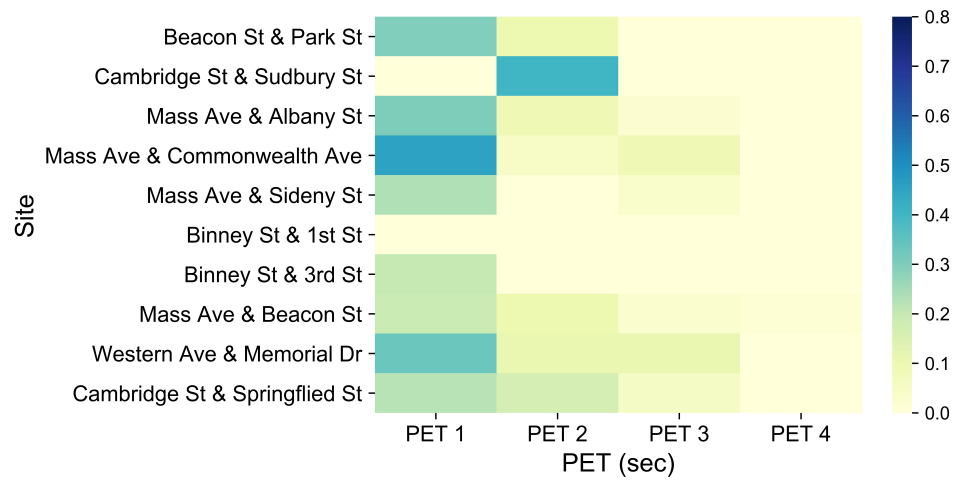


Figure 4.13: Conflict rates per bicycle treatment type



(a) Conflicts where a bicycle is followed by a motorized vehicle



(b) Conflicts where a motorized vehicle is followed by a bicycle

Figure 4.14: Heatmaps for the percentage of the number of traffic conflicts per PET value and per site over the total conflicts



(a) Protected bike lane with a fixed, raised barrier (b) Protected bike lane with parking buffer (c) Raised protected bike lane

Figure 5.1: Different configurations of protected bike lanes (PBL); these bicycle treatments are also known as separated bike lanes or cycle tracks.



Figure 5.2: University of Massachusetts Amherst Human Performance Lab driving simulator and eye-tracking device

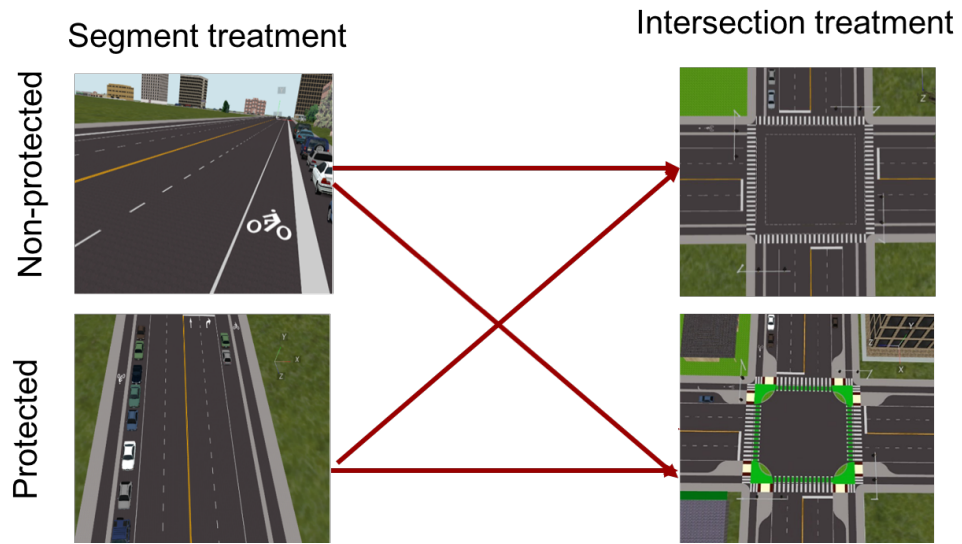


Figure 5.3: Bicycle infrastructure treatment combinations

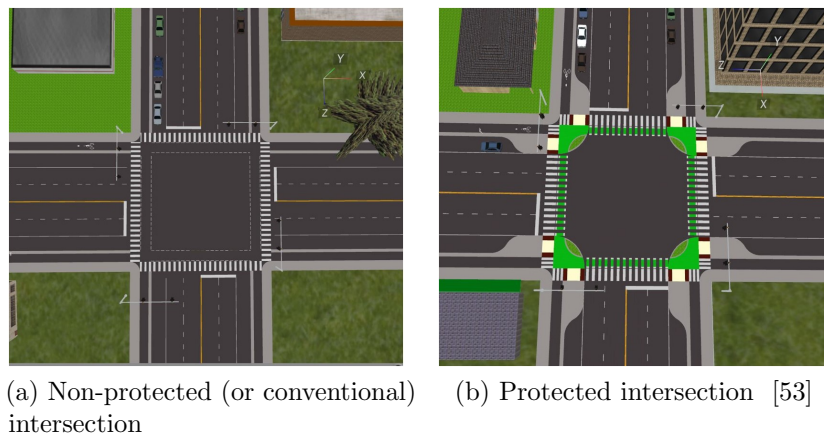


Figure 5.4: Simulated conventional and protected intersections

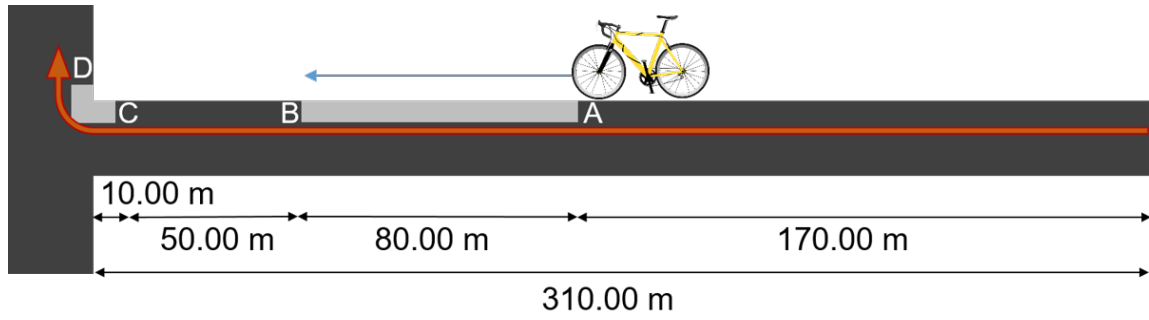


Figure 5.5: Scenario geometric configuration. The orange arrow shows the participants' driving path. The drive parts AB and CB denote the areas for which speed and glance data were collected and analyzed.

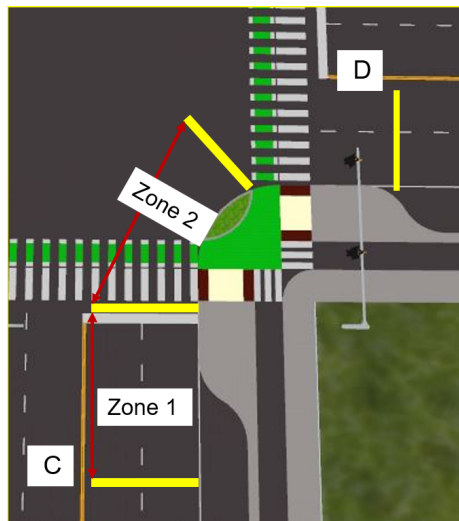


Figure 5.6: Intersection areas of interest.



Figure 5.7: Driver is glancing at the bicyclist



(a) Right glance at the intersection (non-protected intersection).

(b) Right glance at the intersection (protected intersection).



(c) Right glance at the intersection through the right side mirror.

Figure 5.8: Areas of interest for right glances

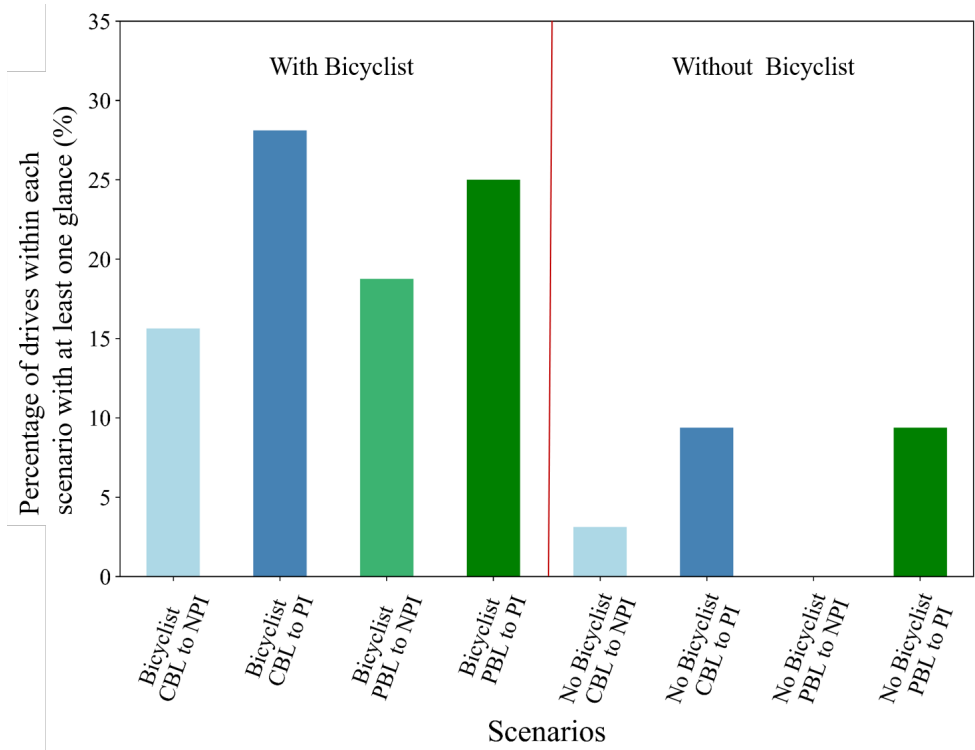


Figure 5.9: Percentage of drives within each scenario that participants glanced right at the intersection at least once (Zone 1 or 2)

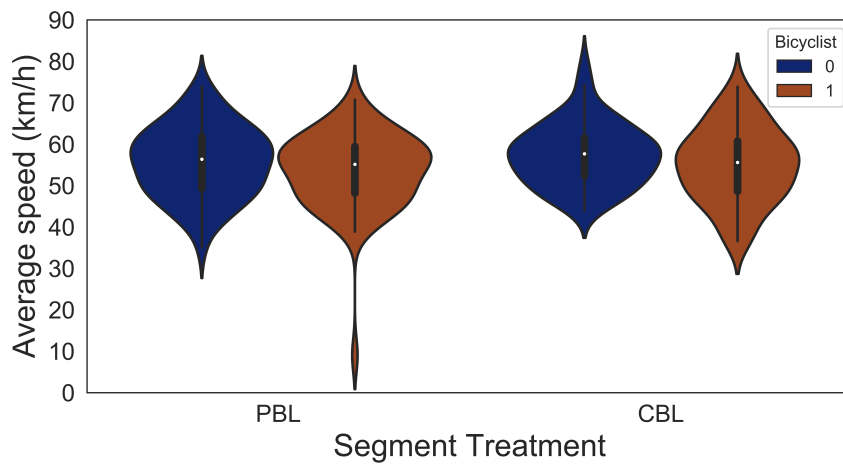


Figure 5.10: Speed violin plots for segment AB (Blue: no bicyclist is present; Brown: a bicyclist is present)

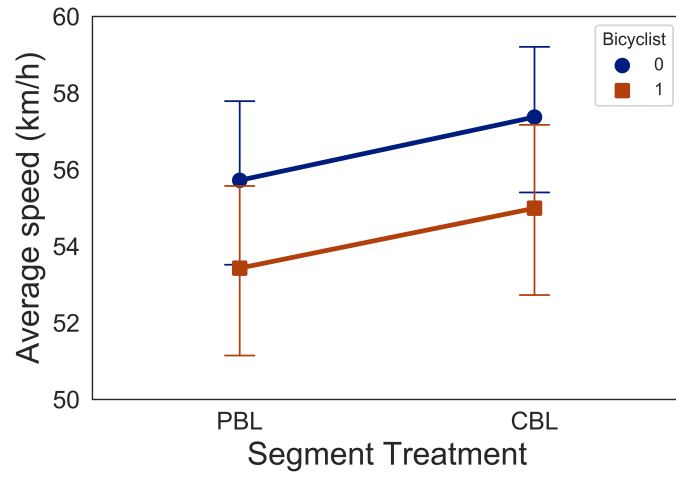


Figure 5.11: Interaction between bicycle infrastructure treatment at the segment and bicyclist presence

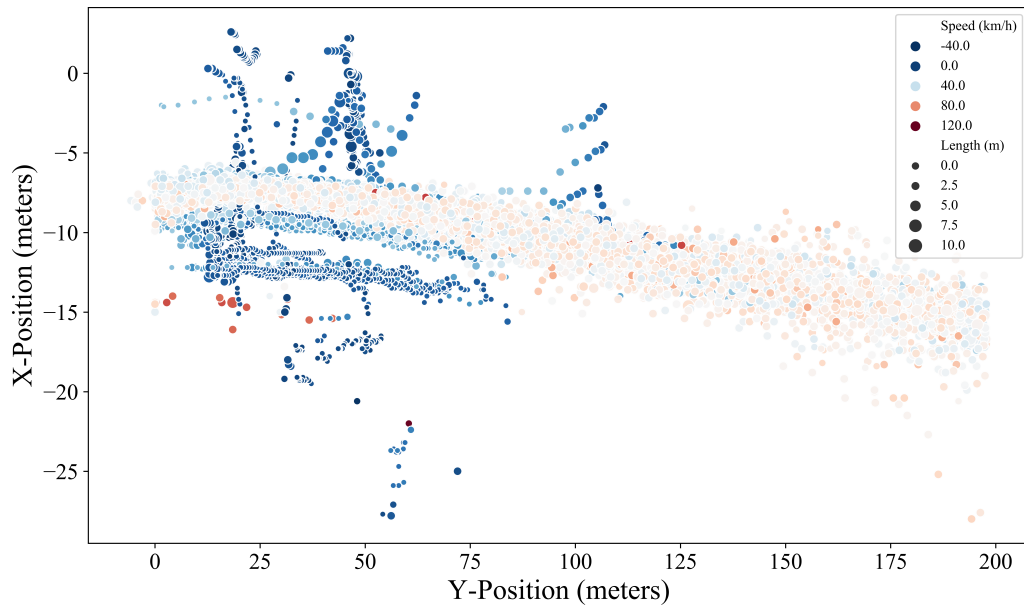


Figure 6.1: Radar-based sensor data visualization

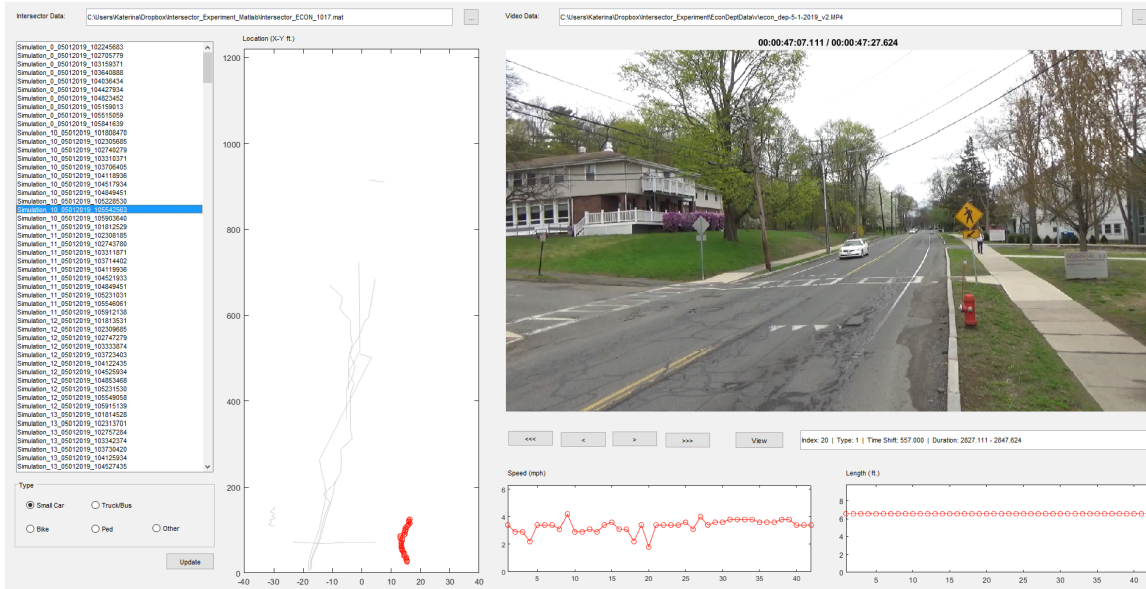


Figure 6.2: Data processing interface in MATLAB

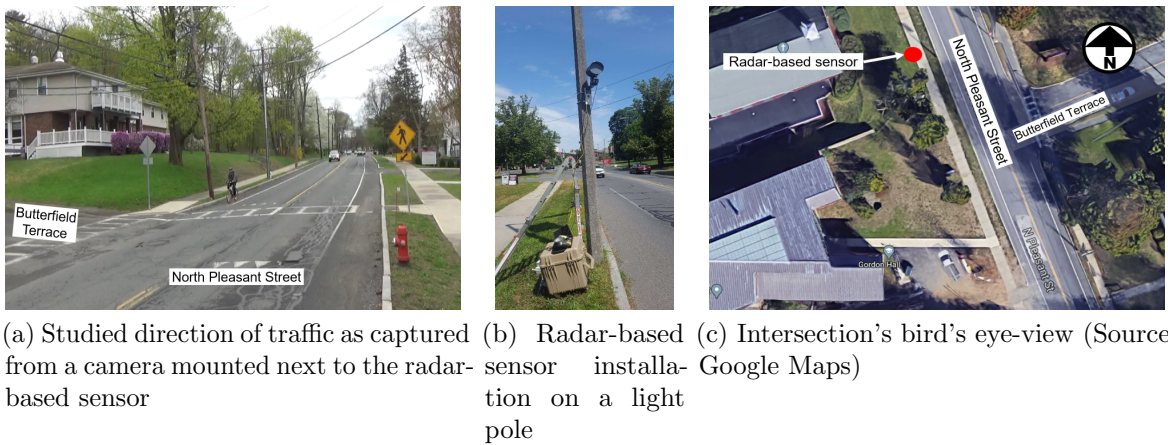
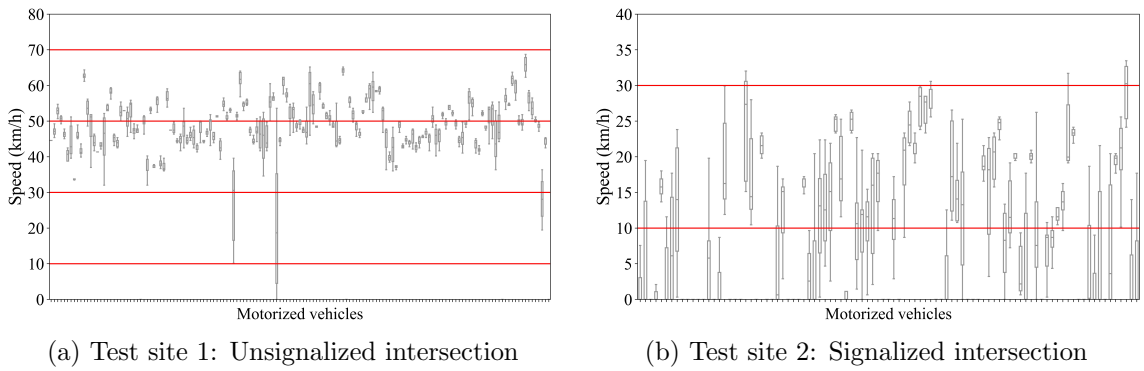


Figure 6.3: Test site 1: Unsignalized intersection



(a) Studied direction of traffic captured as approaching the signalized intersection and the radar-based sensor (b) Radar-based sensor installation (c) Intersection's bird's eye-view (Source: Google Maps)

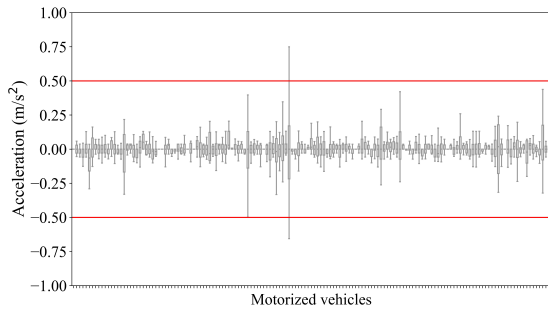
Figure 6.4: Test site 2: Signalized intersection



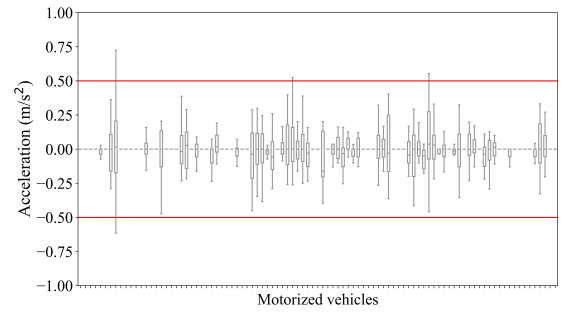
(a) Test site 1: Unsignalized intersection

(b) Test site 2: Signalized intersection

Figure 6.5: Motorized vehicle speed profiles

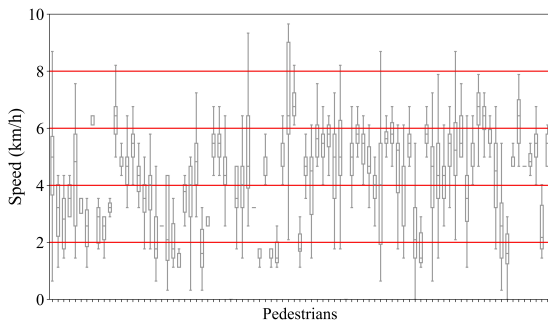


(a) Test site 1: Unsignalized intersection

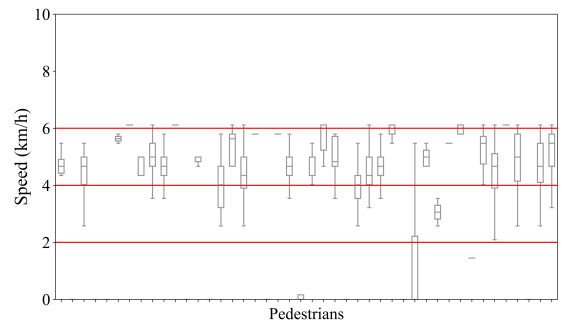


(b) Test site 2: Signalized intersection

Figure 6.6: Motorized vehicle acceleration profiles

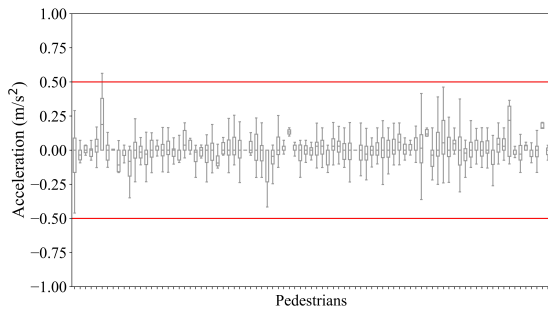


(a) Test site 1: Unsignalized intersection

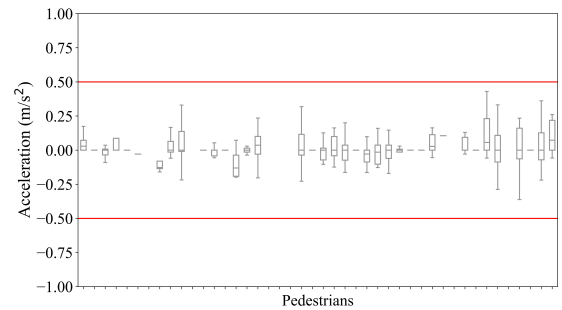


(b) Test site 2: Signalized intersection

Figure 6.7: Pedestrian speed profiles



(a) Test site 1: Unsignalized intersection



(b) Test site 2: Signalized intersection

Figure 6.8: Pedestrian acceleration profiles

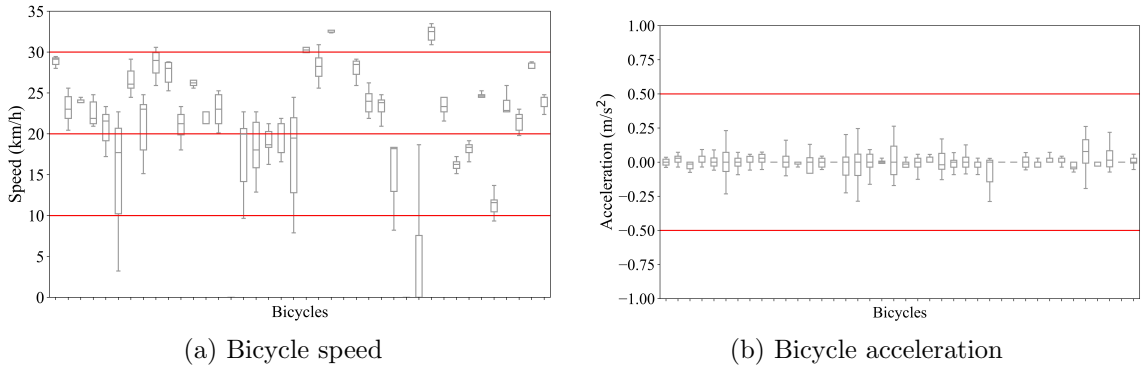
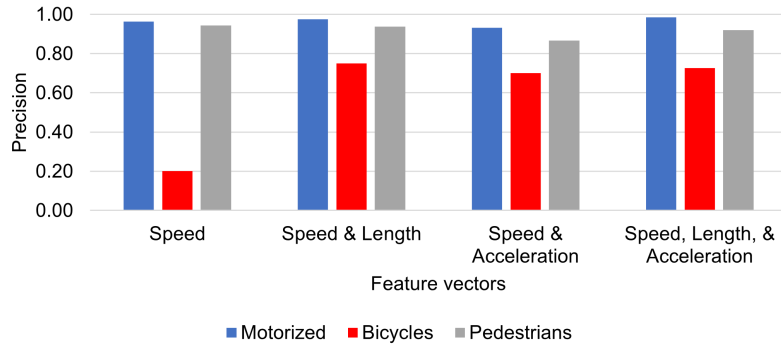
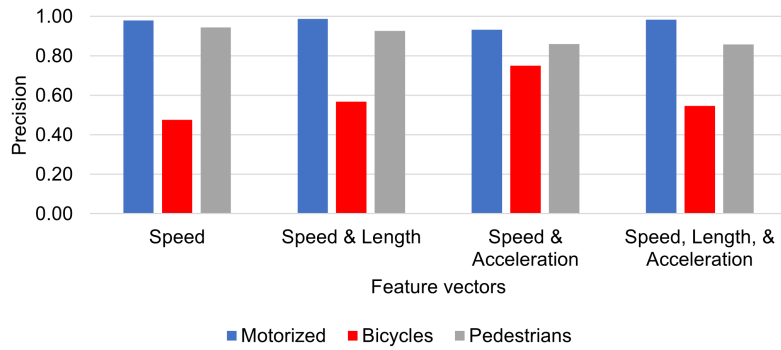


Figure 6.9: Bicycle speed and acceleration profiles

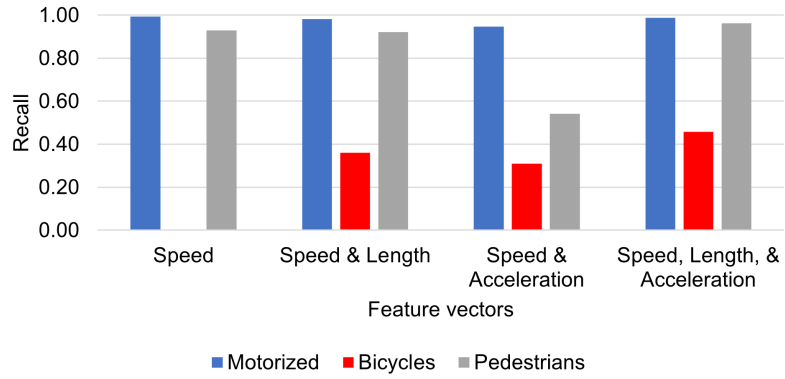


(a) Unbalanced

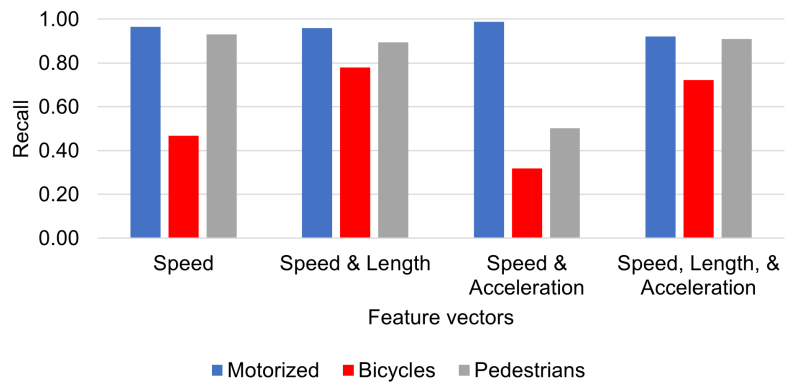


(b) Balanced

Figure 6.10: Precision for the motorized vehicle, bicycle, and pedestrian classification



(a) Unbalanced



(b) Balanced

Figure 6.11: Recall for the motorized vehicle, bicycle, and pedestrian classification

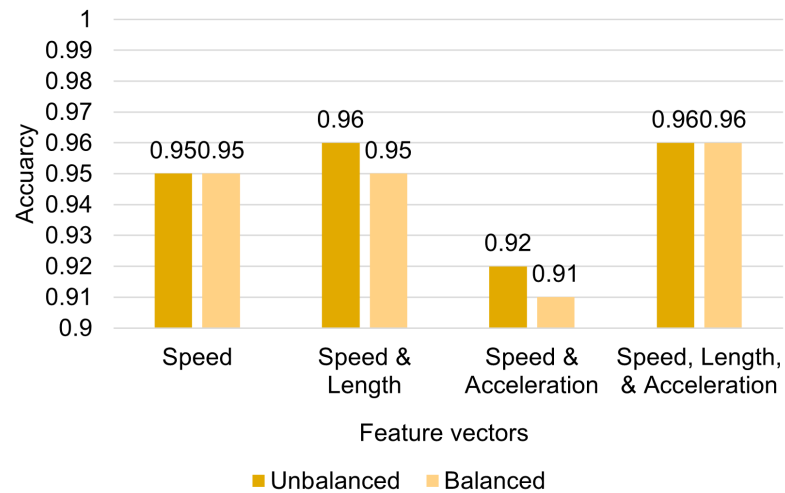


Figure 6.12: Accuracy of the classifier across all eight models for all three modes