



**University of  
Zurich**<sup>UZH</sup>

# Bicycle Accidents in Urban Zurich: An Analysis of Temporal Patterns, Influence of Network Infrastructure and Accident Severity

GEO 511 Master's Thesis

**Author**

Gregory Biland  
17-714-361

**Supervised by**

Prof. Dr. Robert Weibel  
Dr. Wernher Brucks (Wernher.Brucks@zuerich.ch)

**Faculty representative**

Prof. Dr. Robert Weibel

31.08.2023

Department of Geography, University of Zurich



**University of  
Zurich** <sup>UZH</sup>

# Bicycle Accidents in Urban Zurich: An Analysis of Temporal Patterns, Influence of Network Infrastructure and Accident Severity

---

GEO 511 - Master's Thesis

**Author**

Gregory Biland, Universität Zürich  
Mat. Nr. 17-714-361

**Supervised by**

Prof. Dr. Robert Weibel, Universität Zürich  
Dr. Wernher Brucks, Dienstabteilung  
Verkehr, Stadt Zürich

Zurich, August 30, 2023

Department of Geography, University of Zurich

## Abstract

Zurich's increasing population density has raised concerns about road safety, particularly for cyclists. In Switzerland, there were 5,287 bicycle accidents reported in 2022. Despite implementing safer road initiatives, cycling accidents continue to increase in Zurich. The COVID-19 pandemic has encouraged cycling and resulted in a surge of accidents. This study explores the relationship between car traffic, bicycle accidents, and the network infrastructure, aiming to contribute to reducing accidents, analyzing safety factors, and providing proactive solutions for traffic planning in Zurich. Regression models such as the linear mixed effect model, machine learning methods such as the random forest model, and time series analysis were utilized to analyze the spatial and temporal correlations between traffic density, bicycle accidents, and road infrastructure. To conduct this study, pseudo absence points and significant combinations of infrastructure variables were generated and utilized with crash locations to predict the location and severity of a bicycle crash. The research revealed temporal patterns in bicycle crashes and identified a positive association between motorized vehicles and bicycles. The spatial analysis pinpointed high-risk accident locations, especially within areas of high complexity traffic density, and underscored the paramount importance of bicycle infrastructure. By modeling crash severity with a focus on infrastructure variables, it was possible to predict crash outcomes and discern the role of involved vehicles and age in severe bicycle crashes. Traffic and infrastructure are central to bicycle crashes. When paired with spatial analysis, temporal patterns spotlight high-risk zones and periods. Crucial infrastructure elements, such as road width and proximity to tram tracks, profoundly impact crash severity. Addressing these components may aid in enhancing bicycle safety in urban contexts.

**Keywords:** #bicycle, #accidents, #road-safety, #infrastructure, #network, #prediction, #COVID-19, #accident severity

## Acknowledgments

I want to express my deep gratitude to Professor Dr. Robert Weibel, my supervisor, for his invaluable guidance and understanding that have significantly shaped this thesis. His patience and insightful feedback have been instrumental throughout this project.

I would also like to thank Dr. Wernher Brucks from the DAV, who co-supervised this project and provided essential support and constructive input.

Prof. Dr. Reinhard Furrer deserves recognition for his meaningful contributions to addressing my modeling-related queries. His assistance and advice were crucial to the advancement of this work.

Additionally, I am incredibly grateful to Dr. Markus Deublein from the bfu, whose invaluable advice regarding modeled parameters and constructive input has been pivotal to the success of this thesis.

I want to thank my family for their relentless support and belief in my abilities throughout my academic journey.

I am also grateful to my friends and peers, whose support and companionship have enriched my journey. Lastly, I extend my heartfelt thanks to all who have contributed, directly or indirectly, to this research. Your efforts and support have made a significant impact and are genuinely appreciated.

# Contents

<b>1</b>	<b>Introduction</b>	<b>1</b>
<b>2</b>	<b>Related Work</b>	<b>4</b>
2.1	Background . . . . .	4
2.2	Literature review . . . . .	5
2.2.1	Time-Series and Exposure . . . . .	5
2.2.2	Infrastructure and Severity . . . . .	8
2.3	Research Gaps . . . . .	10
2.4	Research Questions . . . . .	11
<b>3</b>	<b>Data</b>	<b>13</b>
3.1	Traffic volume data . . . . .	13
3.1.1	Motorized vehicle count . . . . .	14
3.1.2	Bicycle counts . . . . .	14
3.1.3	Complete traffic model (GVM-ZH) . . . . .	14
3.2	Network and transport-related features . . . . .	14
3.3	Bicycle accident data . . . . .	16
<b>4</b>	<b>Methodology</b>	<b>19</b>
4.1	Technical Setup . . . . .	20
4.2	Data Preparation . . . . .	20
4.2.1	Exploratory Analysis . . . . .	21
4.3	Statistical Analysis . . . . .	21
4.3.1	Temporal Correlation Analysis . . . . .	21
4.3.1.1	Multiple change point analysis . . . . .	21
4.3.1.2	Generalized additive models (GAM) . . . . .	22
4.3.2	Spatial Network Analysis . . . . .	23
4.3.2.1	Accident Infrastructure Analysis . . . . .	23
4.3.2.2	Accident Location Prediction . . . . .	26
4.3.2.3	Accident Severity Analysis . . . . .	27
4.3.2.4	Predictive Modelling of Accident Severity . . . . .	30
<b>5</b>	<b>Results</b>	<b>32</b>
5.1	Temporal Correlation Analysis . . . . .	32
5.1.1	Exploratory Analysis . . . . .	32
5.1.2	Statistical Analysis . . . . .	34
5.1.2.1	Multiple Change Point Analysis . . . . .	34
5.1.2.2	GAM . . . . .	36
5.2	Spatial Network Analysis . . . . .	38
5.2.1	Exploratory Analysis . . . . .	38
5.2.1.1	Average nearest neighbor analysis . . . . .	38

5.2.2	Statistical Analysis . . . . .	39
5.2.2.1	Accident Infrastructure Analysis . . . . .	39
5.2.2.2	Accident Location Predictions . . . . .	45
5.2.3	Accident Severity Analysis . . . . .	52
5.2.3.1	Infrastructure Effect . . . . .	53
5.2.3.2	Predictive Modelling of Accident Severity . . . . .	56
<b>6</b>	<b>Discussion</b>	<b>59</b>
6.1	Results . . . . .	59
6.1.1	Temporal Analysis . . . . .	59
6.1.1.1	Relationship Overview . . . . .	59
6.1.1.2	COVID-19 Impact . . . . .	60
6.1.2	Spatial Analysis . . . . .	61
6.1.2.1	Spatial Patterns and Correlations . . . . .	61
6.1.2.2	Influence on Accident Severity . . . . .	63
6.2	Methodological Approach . . . . .	65
6.2.1	Methodological Considerations: Temporal Analysis . . . . .	65
6.2.2	Model Evaluation: Spatial Analysis . . . . .	65
6.3	Data Availability . . . . .	66
6.4	Implications for Policy and Urban Design . . . . .	67
<b>7</b>	<b>Conclusion</b>	<b>69</b>
7.1	Contributions . . . . .	69
7.2	Main Findings . . . . .	70
7.3	Limitations . . . . .	70
7.4	Outlook . . . . .	71
<b>8</b>	<b>Appendix</b>	<b>77</b>

# Chapter 1

## Introduction

As Zurich's population density increases, the importance of safe and forward-thinking urban planning becomes more apparent. More people share the same road network, causing it to reach capacity limits in many areas and increasing the risk of accidents. Designing traffic areas consciously and safely is crucial to protect all road users, but holds particularly for the more vulnerable among the traffic participants, including cyclists. 2022, Switzerland experienced over 5,287 bicycle accidents, making road safety a significant concern for politicians and the public (ASTRA, 2022). The number of cyclists on the road has increased in recent years, while the risk of accidents has also risen. Despite efforts by city planners to create safe streets for all road users, the number of bicycle accidents in Zurich has continued to increase. This trend is unique to the cycling community, as accidents in other traffic groups have decreased. While there was a decrease in severe and fatal accidents in 2021, the number of light injuries has continued to rise. Thus, the decrease in accidents during the summer of 2021 and previous periods is likely due to generally rainy and cold weather conditions rather than improvements in network infrastructure (DAV, 2021).

The growing popularity of bicycles as a means of transportation has made the interaction between bicycles and cars a crucial concern for policymakers and urban planners. Unfortunately, the risk of accidents between cars and bicycles has become a significant issue in cities with limited space for bicycle infrastructure, such as Zurich. Studies have shown that reducing the number of vehicles on the road would significantly decrease the likelihood of collisions (Garrard, Rose, and Lo, 2008). Understanding the dynamics of road traffic and related accidents is essential for urban planning, policy-making, and public safety (Lord, Manar, and Vizioli, 2005). In particular, the correlation between car traffic density and bicycle accidents is a significant concern in urban environments where these two modes of transport coexist (Vandenbulcke, I. Thomas, Geus, et al., 2009). The relationship between car traffic density and bicycle accidents is complicated and influenced by various factors, including road infrastructure, traffic regulations, and the behavior of road users (Kim et al., 2007). Furthermore, this relationship is not static and changes over time due to societal trends, policy changes, and unexpected events, such as the COVID-19 pandemic (Ebrahim Shaik and Ahmed, 2022).

The COVID-19 pandemic has significantly impacted daily life and transportation worldwide. To contain the spread of the virus, governments have taken measures that have drastically reduced traffic in many areas. With people only allowed to leave their homes for essential reasons, commuting became a thing of the past, and work shifted to the home office. This shift led to unusual traffic situations in many cities worldwide, promoting a shift towards non-motorized forms of transport, such as cycling (Bucsky, 2020). During the pandemic, there was a notable rise in cyclists on the road. Unfortunately, this led to a significant increase in bicycle accidents during the first year of the pandemic compared to before the shutdown. However, there was a decrease in accidents during the second half of 2020,

which was a positive development (Tiefbauamt Stadt Zürich, 2021). This change presents a unique opportunity to analyze the relationship between traffic volume, traffic safety, and accident occurrence and identify potential approaches to how modal split changes might affect future road safety. However, the long-term impacts of the pandemic on the interplay between car traffic and bicycle accidents are yet to be thoroughly studied, marking a crucial research gap that this study aims to address.

Another essential aspect that this research delves into is the role of road infrastructure in modulating the correlation between car traffic and bicycle accidents. Road infrastructure characteristics such as intersections, roundabouts, and bicycle lanes/paths have significantly affected cyclists' safety (B. Thomas and DeRobertis, 2013). However, less is known about how these elements interact with car traffic density to influence accident rates, a topic of inquiry that will be explored in this study. Examining the data and drawing the soundest potential deductions from this situation for future roadway network planning is essential. Researchers have begun to study the past relationship between automobile traffic and bicycle accidents in an urban context. For example, Garrard and Rose (2010) found that in areas with high automobile traffic and inadequate bicycle infrastructure, bicycle accidents increase (Garrard, Rose, and Lo, 2008). In contrast, areas with less automobile traffic and more bicycle infrastructure have fewer bicycle accidents (Liu and Marker, 2020). Other studies have shown that especially bicycle lanes/paths and paths can improve safety for bicyclists and motorists (Lusk et al., 2011; Teschke et al., 2012). A Swedish study has also demonstrated that intersections and junctions are among the most frequent accident sites and that the exposure of bicycles to cars in the same lane/path drastically increases the risk of accidents (Kullgren et al., 2019).

Similarly, local and international studies have shown that infrastructure, such as the length of the bicycle lane/path network and the speed at which people ride on the road, drastically influence the likelihood and severity of accidents. Thus, the fatality of accidents is highest when the speed is higher than 50 km/h on impact and lowest when cars and bicycles travel less than 30 km/h (Kim et al., 2007). Although studies have shown the benefits of bicycle infrastructure and the dangers of traffic circles, the short time span and coverage of studies in this domain remain a significant uncertainty factor in effective planning (Reynolds et al., 2009). Particularly in Zurich, where the density of intersections in combination with various infrastructure elements is densely packed into a small space, a conclusive assessment of the effects of such combinations remains essential for increasing road safety in the city. Various studies have repeatedly shown that weather negatively impacts road safety and significantly affects the number of accidents (Qiu and Nixon, 2008). The number of accidents and, thus, the probability of accidents depends heavily on seasonal and weather-related forecasts, but equally, the weather also influences various infrastructures significantly (Pazdan, 2020). These findings show that investments in bicycle infrastructure and measures to reduce automobile traffic and exposure can substantially improve road safety for all road users (Pedroso et al., 2016). Nevertheless, there need to be more meaningful studies investigating these findings in Zurich. This thesis aims to make a significant contribution to analyzing the relationship between infrastructure and accidents in recent years and to be able to make initially targeted statements.

A comprehensive study of road infrastructure has revealed that the network's state significantly impacts the severity of accidents. Extensive research has demonstrated that various factors, including the type of road and signalization, play a crucial role in determining the outcome of accidents (Prati, Pietrantonio, and Fraboni, 2017). Additionally, accidents tend to be more severe in areas with high speeds, complex curves, and one-way streets (Jaber, Juhász, and Bálint Csonka, 2021). As reducing the number of accidents remains a top priority, there is an increased focus on minimizing their severity in road safety. To achieve this objective, it is imperative to thoroughly analyze accidents and their correlation with infrastructure (Kaplan and Giacomo Prato, 2015). Therefore, careful road infrastructure planning and design can positively affect accident outcomes and minimize their severity.

The **main objectives** of this work are twofold. One is to conduct an in-depth study of traffic accidents in Zurich in collaboration with the Traffic Division of the City of Zurich (DAV<sup>1</sup>) and analyze how the COVID-19 pandemic has affected road safety and the development of accidents over time. Secondly, spatial patterns will be examined, and initial evidence of the influence of infrastructure will be collected. In contrast to existing studies, the focus is placed on the totality of factors and is not dealt with on a situation-specific basis. By statistically investigating the relationship between bicycle accidents and Zurich's road network and topology, it is possible to evaluate potential weak points and hotspots best and estimate how much individual network characteristics influence them. This should make it possible to predict accident trends and enable future-oriented traffic planning using variables-based projections. This holistic analysis is guided by *Vision Zero*<sup>2</sup>, which aims to reduce the number of crashes involving bicyclists and ultimately result in zero fatalities and a minimum of accidents regardless of the severity. Potential solutions will be explored to improve road safety and prevent future crashes, mitigating the impact of the road network infrastructure and advancing *Vision Zero*.

---

<sup>1</sup><https://www.stadt-zuerich.ch/dav>

<sup>2</sup><https://visionzero.global>

# Chapter 2

## Related Work

### 2.1 Background

The increase in road construction since the 1980s has created numerous new traffic circles, intersections, and tram tracks, connecting Zurich’s districts with a network of more than 2000 km of tram tracks, bicycle lanes/paths, car roads, and bus lanes<sup>3</sup>. As Zurich has grown and developed new network properties, the safety landscape for all road users has significantly transformed. However, despite this transformation, bicycle traffic remains the only mode of transport in Switzerland that has not seen a decline in accidents over the past decade. In 2011, there were 294 bicycle accidents in Zurich, but in 2020 during the pandemic, there were 606—an all-time high indicated in Figure 2.1.



Figure 2.1: Bicycle accidents in the city of Zurich between 2012 and 2021, broken down by conventional bicycles (dark blue) and e-bicycles (light blue). Source: City of Zurich, Department of Safety.

Compared to car traffic and pedestrians, bicycle accidents have steadily increased in recent years (Foley et al., 2021). They have not yet shown any negative trend as pedestrian and car accidents did. Given this situation, planning and designing a safer network infrastructure for cyclists is becoming increasingly critical. The same strategies are being pursued for car drivers and pedestrians, where awareness, stricter legislation, and modern infrastructure have drastically reduced accident rates and risks in Zurich. When creating safer infrastructure for bicyclists, it is vital to consider the difference between perceived safety and statistical safety. Cyclists may mistakenly feel unsafe even if a road is built to safety standards. Therefore, it is essential to raise awareness and design the infrastructure in a forgiving way for cyclists - applying the *SERFOR*- principle (ASTRA, 2016). However, a modern planned infrastructure culture can provide a direct foundation for safer traffic. Combining all three

<sup>3</sup>[https://www.stadt-zuerich.ch/ted/de/index/taz/verkehr/webartikel/webartikel\\_netzlaenge.html](https://www.stadt-zuerich.ch/ted/de/index/taz/verkehr/webartikel/webartikel_netzlaenge.html)

approaches is necessary to create a genuinely safe and forgiving bicyclist infrastructure. In addition, inclusive planning should involve all traffic participants and a clear strategy for bringing cars, trams, buses, and pedestrians together on a network without one variable limiting the other, such as car traffic limiting bicycle traffic too much in terms of safety (Li, Shi, and D. Chen, 2011). Furthermore, there are voices in Zurich that claim the most significant risk for cyclists was caused by cars. To plan traffic sustainably and with a strong foundation, it is essential to thoroughly research this variable’s impact.

Similarly, a road may feel safe but does not meet up-to-date safety requirements and standards (Wegman, Zhang, and Dijkstra, 2012). The ultimate goal is the approach of a safe, planned infrastructure, which is also perceived as such by road users. For such a scenario to occur, a sound knowledge of the individual network variables and their interactions in this area is essential, especially in preventing bicycle accidents (United Nations, 2021). It is critical to understand the relationship between accidents and the factors that cause them and to know which variables need to be addressed more forcefully and to what degree. Although attempts have been made to create safer roads for all users, there still needs to be more comprehensive strategies to minimize accidents between cars and bicycles. As shown below, there are research gaps and uncertainties when analyzing road safety due to the growing popularity of bicycles as a preferred mode of transportation. As more people choose to ride bicycles, the proportion of bicycle traffic compared to car traffic is gradually increasing (Federal Office of Public Health FOPH, 2023). This presents a crucial challenge for urban planners and policymakers to address, as there is limited space available to accommodate every mode of transportation safely and effectively in the network. Additionally, cycling education and intelligent transport systems can be implemented to enhance safety and efficiency. With car traffic slowly decreasing and limited space for expanding and improving bicycle infrastructure in Zurich, the risk of accidents between cars and bicycles has been only minimally reduced, and exposure remains an ongoing issue. The COVID-19 pandemic has presented an opportunity to study Zurich’s predicted 2025 traffic composition and its potential impact on road safety when the exposure changes. This unique situation allows for crucial insights to be drawn for the future.

## 2.2 Literature review

Numerous studies have focused on the spatial and temporal interplay between car traffic density, bicycle accidents, and the network infrastructure. Still, the COVID-19 pandemic’s onset has added a fresh dimension to these relationships. This section examines critical studies in this area, emphasizing the evolving dynamics in the wake of the pandemic, particularly in Zurich.

### 2.2.1 Time-Series and Exposure

**Exposure** Among the multiple factors identified as influential in the road safety literature, traffic density has emerged as crucial in determining the likelihood and severity of bicycle accidents. The term “*Exposure*” is used in ecology and traffic safety - traffic density, traffic volume, and distance from congested roads are among the metrics used to measure exposure in traffic safety (Pratt et al., 2014). The term is used to quantify the exposure of cyclists to car traffic on the road. This relationship’s complexities unfold as research shows increased bicycle accidents corresponding to heavy traffic. As cyclists increase, the risk of accidents grows, particularly in areas with high car traffic volumes (Miranda-Moreno, Strauss, and Morency, 2011). Due to the increased exposure, more accidents occur, but it remains open to what extent the traffic of bicycles and cars alone influences the number of accidents. Contrarily, some research has supported the “safety-in-numbers” phenomenon, suggesting that an individual cyclist’s risk diminishes as bicycle use surges (Jacobsen, 2015). When there are

more cyclists on the road, the risk for each cyclist decreases, known as the “safety in numbers” phenomenon, especially in areas with high traffic volumes. This suggests their susceptibility to accidents decreases as more cyclists share the road with vehicles. However, this also highlights the intricate relationship between increased cycling, traffic density, and accident rates, warranting further exploration. According to a study conducted by Li et al. in 2011, as the number of cars on the road increases, the likelihood of cyclists making sudden lane changes is reduced. However, this also heightens the risk of accidents for cyclists due to the increased stress caused by the higher traffic density (Li, Shi, and D. Chen, 2011). These relationships could exhibit similar trends in Zurich’s context, characterized by mixed traffic where bicycles and cars often share the roadway. Further bolstering this correlation, a study by Wang and Nihan, 2004 indicated a strong relationship between car traffic density and bicycle accidents. Increased traffic volume generally escalated the interactions between vehicles and bicycles, thereby elevating collision risk. A consequent reduction in car traffic could decrease accident numbers, enhance environmental quality, and foster a healthier, more active population (Meschik, 2012). However, according to a study conducted in China, bicycle traffic cannot be blamed solely on traffic. According to that study, the severity of accidents resulting from high traffic density is usually due to the cyclist’s failure rather than the car driver’s (Kim et al., 2007). Thus, even in the literature, there is a clear trend that the density of cars on the roads contributes significantly to the number of bicycle accidents. However, the issue remains inconclusive due to discrepancies and ambiguities in the existing research. There are still uncertainties regarding how exposure affects safety and which factors in road traffic play a significant role in achieving “Vision Zero” within a specific timeframe. However, few studies have investigated and assessed a thorough analysis of the temporal shifts and connections between car and bicycle accidents, utilizing sophisticated methodologies (J. Lee, Abdel-Aty, and Jiang, 2015).

**Accident Time Series** The COVID-19 outbreak significantly reshaped these dynamics by triggering sweeping changes in mobility patterns worldwide. On 28 February 2020, the Federal Office of Public Health launched the campaign "How to protect ourselves" with hygiene recommendations to protect against the new coronavirus, which marked the start of the pandemic in Switzerland. A study by the Swiss Federal Office of Public Health revealed a substantial drop in mobility during the pandemic’s early phase in the Spring of 2020, with a discernible shift towards active transportation modes such as walking and cycling (Federal Office of Public Health FOPH, 2023). These shifts were also observed in several other nations, reflecting a global trend (Transportation Bureau of Statistics, 2023; Kellermann et al., 2022). Consequently, the reduced car traffic combined with an increased number of cyclists has been associated with a significant decrease in accidents involving cyclists (Monfort, Cicchino, and Patton, 2021; Ebrahim Shaik and Ahmed, 2022; Barnes et al., 2020). In some regions in Italy, the decrease in road traffic accidents was up to 70% compared to before the pandemic, indicating the massive influence the pandemic had on road safety (Valent, 2022). Even though the existing body of research provides a robust base, a need for more detailed research remains, particularly considering the varied trends that emerged during the pandemic. Although some studies reported an uptick in bicycle usage, a simultaneous increase in private car usage has also been noted (Francke, 2022).

Analyzed accident data in Zurich showed that a drastic increase in bicycle accidents accompanied the outbreak of the pandemic before accidents began to decline, as in other countries (Tiefbauamt Stadt Zürich, 2021). The number of bicycle traffic accidents significantly increased in 2020, and even during the lockdown period, there were 50 more accidents (or 9.19 %) compared to 2019. It is concerning that e-bicycle accidents with a speed limit of 45 km/h have increased substantially more than those with 25 km/h. Further, over half of all severe injuries occurred while cycling, and four traffic fatalities in 2020 were cyclists (Brucks, 2020). Moreover, although the rise in bicycle use during the pandemic has been documented in numerous studies, it is noteworthy that the incidence of bicycle accidents

has also risen during the pandemic, particularly during leisure time (Unfallversicherung (UVG), 2021). The paradoxical rise in bicycle usage and accidents during the pandemic underscores the need for a more nuanced understanding of the factors contributing to bicycle accidents. This finding was also recorded in a study of accidents involving minors, where an increase in accidents during the pandemic was recorded, in contrast to the findings in Italy (Failing et al., 2023). A similar picture was found in Zurich, where the number of bicycle accidents fell only briefly after the pandemic’s start before rising to an all-time high. In many places, however, as in Zurich, it has been shown that the first wave had a much more pronounced negative impact on accidents and traffic volume among cyclists than the second wave of the pandemic. Similar results were obtained in a study by Failing et al. (2023), where cyclists in the USA recorded 48 % more accidents during the pandemic than in the previous reference period (Failing et al., 2023).

The impact of these shifts in Zurich’s unique context has yet to be explored in the form of time-series analyses. So far, only a fraction of the analyses in the literature cover this. Understanding how the “safety in numbers” phenomenon works under Zurich’s mixed traffic conditions is critical to how planning for individual networks should proceed. In addition, a deeper understanding of how pandemic-related changes have affected the dynamics between traffic density, bicycle use, and accident rates in Zurich is needed.

When analyzing changes over time in variables that relate to safety, a significant amount of uncertainty remains in the accident data. This uncertainty is due to accidents that were not registered or were registered incorrectly mainly minor accidents. For example, 2 % of all countries worldwide do not collect any bicycle accident data, and 50 % of countries, including Switzerland, only use police data when analyzing accident trends. Likewise, the number of unreported cases increases with the severity of the accident and the fact that it was a self-inflicted accident (Wegman, Zhang, and Dijkstra, 2012). To guarantee accurate and impartial analysis, adopting a consistent method for documenting accidents is crucial. Reliable data is necessary for accurate analysis. A study from Germany demonstrates that reliable data is only sometimes available, highlighting the ambiguity in accident statistics (Harkort, Walker, and Lakes, 2023). In Switzerland, much is known about each accident based on the data collected in the *Unfallaufnahmeprotokoll* (UAP), which reports many accident-related variables (Section 3.3). However, much remains unclear about the temporal context within which accidents happen based on the data collected. Likewise, it has been shown that with increased residential and traffic density, the number of accidents increases, as does the accident severity, which is typically highly dependent on the network’s topology (Harkort, Walker, and Lakes, 2023).

**Approaches** These two contradictory trends indicate that it is necessary to examine the specific factors that influence different regions and time periods more closely. Existing research has often employed numerical comparisons without a comprehensive statistical analysis, suggesting the need for more rigorous investigations considering various variables and their interrelationships (Katrakazas et al., 2020). While numerous studies have assessed the absolute difference in accident numbers, more have yet to delve into statistical correlations and variations in accident figures. Nevertheless, some research has juxtaposed accident trends with traffic data, socio-demographic factors, and environmental elements, often employing time series methodologies such as the generalized additive model (GAM) and autoregressive integrated moving average (ARIMA) (Quddus, 2008; Getahun, 2021; Jaber, Balint Csonka, and Juhasz, 2022). Specific investigations have also explored the association between accident likelihood, driving distance, and weather conditions within time series, aiming to elucidate the underlying patterns of change attributable to pandemic influences (Boucher and Turcotte, 2020; Pazdan, 2020). In the post-pandemic era, regression models have become more commonplace for studying the temporal interrelation between diverse variables, such as traffic volume and accident occurrences (Yang et al., 2021). Negative binomial and Poisson regressions have been mainly instrumental in examining

these temporal correlations. More recently, alongside generalized additive models, ARIMA models have been incorporated to study the relationship of variables within road traffic scenarios (Quddus, 2008). An evaluation of various methodologies utilized in past research on traffic accident-related time series revealed the prominence of the generalized additive model (GAM). Its usefulness in modeling seasonality has led to its widespread application in several studies (Boucher and Turcotte, 2020; Xia et al., 2023; W.-K. Lee et al., 2014; P. Chen and Shen, 2016).

Traffic density and the frequency and severity of bicycle and car accidents are intertwined. The COVID-19 pandemic introduced a significant shift in these dynamics, with a global trend towards active transportation and reduced car traffic, leading to a decline in accidents involving cyclists in many regions. Nevertheless, an initial increase in bicycle accidents was observed in areas such as Zurich. Despite the extensive body of research, there remains a need for a deeper investigation, given the contradicting trends observed during the pandemic. Current studies have primarily focused on numerical comparisons, suggesting the necessity for comprehensive statistical analyses that consider multiple variables and their interactions. The generalized additive model (GAM) has been identified as a standout methodology due to its effectiveness in modeling seasonality.

### 2.2.2 Infrastructure and Severity

The construction of roads and the overall transportation network plays a crucial role in determining the safety of those using them and the likelihood of accidents occurring. As a result, the United Nations and the World Health Organization have identified the network infrastructure as a vital element in their plan to create a safe accident prevention system by 2030 (United Nations, 2021). Improving road safety can be achieved in two ways: either by constructing new infrastructure or by making better use of the existing infrastructure. However, current road infrastructure is primarily designed to accommodate motorized traffic and is not adequately adapted to cater to the needs of cyclists (Wegman, Zhang, and Dijkstra, 2012). For example, a 2008 study by Pucher and Buehler, 2017 first examined global infrastructure and outlined direct proposals in the abovementioned approaches: creating safe infrastructure for bicycles, bicycle lanes, and bicycle paths and reducing car traffic through congestion pricing and car-free zones. Regarding the infrastructure analysis, a well-developed and maintained bicycle infrastructure is essential for safety. However, it has also been shown that the definition of a well-developed infrastructure is only vaguely addressed in the literature, and the data situation here is still precarious (Vanparijs et al., 2015). So far, it remains undisputed in the literature that road infrastructure significantly influences safety. The only remaining question is how much infrastructure contributes to the safety of cyclists and how the infrastructure behaves in terms of road safety when infrastructure elements occur in combination.

Spatial studies of accident data are of great value to effectively adapt infrastructure to given conditions. G. Vandenbulcke and colleagues have already shown in 2009 that there are massive spatial differences in accident data related to infrastructure and that effective planning of this infrastructure is only possible if the spatial approaches are known. Thus, with the necessary knowledge about the effect of individual variables, the infrastructure can be spatially adjusted such that it is most effective (Vandenbulcke, I. Thomas, Geus, et al., 2009). One crucial aspect in which these variables manifest themselves is at complex intersections.

**Risk Factors** *Complex intersections* A study from Belgium shows how infrastructure affects the risk of accidents. For example, tram tracks, bridges without bicycle facilities, and complex intersections were statistically found to have an increased risk of bicycle accidents (Vandenbulcke, I. Thomas, and Int Panis, 2014). Similar results were shown in a different study from Tokyo, marking complex intersections

as especially risky (Wang and Nihan, 2004). Some geometric design factors were tested concerning intersections and junctions, such as a median, parking access, and the number of intersection cuts. However, it was found that the effects are not statistically significant and are not the determining factor for increased accident chances (Miranda-Moreno, Strauss, and Morency, 2011). The complexity index shows the most significant effect on the chances of accidents, which indicates that the more complex a traffic intersection is, the higher the chances of accidents. For example, as shown in other studies, it is very likely that with higher traffic density and narrower roads, the number of crashes increases and correlates with each other (Guo, Osama, and Sayed, 2018). For example, areas and zones with higher densities of complicated traffic infrastructure and higher-than-average traffic density will likely experience more crashes. A study by P. Chen and Shen, 2016 showed a positive correlation between negative environmental factors such as weather, the time of day, and network topologies such as traffic circles and intersections. Finally, a study from Japan found that the most common cause of bicycle accidents is collisions with cars at intersections (Stone and Broughton, 2003; Koike, Morimoto, and Kitazawa, 2003).

*Traffic areas* Parking spaces and high-density pedestrian areas are expected to influence the accident risk negatively, leading to increased levels (Vandenbulcke, I. Thomas, and Int Panis, 2014). In contrast, another study suggested that traffic calming measures such as pedestrian zones, driving bans, and 30 km/h zones significantly impact safety positively (Kim et al., 2007).

*Bicycle lanes/paths* In the case of cycling infrastructure, the study by Vandenbulcke, I. Thomas, and Int Panis, 2014 showed two different results in terms of safety. Separate bicycle lanes and bicycle paths were found to be a positive factor in reducing the number of accidents. However, separate bicycle lanes/paths in combination with complex intersections, right of way, and traffic circles were shown to have an increased chance of accidents. Another literature review on the influence of transport infrastructure has shown that infrastructure specifically adapted to cyclists, such as separate bicycle lanes/paths, street lighting, and inclination, significantly influences the chances of accidents negatively (Reynolds et al., 2009). Wegman and his team have devised two methods to decrease the likelihood of accidents involving bicycle paths. The first is to have a secure separation of bicycles and cars with dedicated bicycle lanes/paths—the safest option. The second is to reduce the speed of cars if bicycle lanes/paths are not feasible, so the speed difference between the two is minimal (Wegman, Zhang, and Dijkstra, 2012).

*Speed* It has been evident since the early 1950s that higher speeds drastically increase the risk of accidents for all road users (Smeed, 1949; Boyer and Dionne, 1987). Furthermore, the involvement of cyclists massively increases the severity of accidents at higher speeds. Since the early 2000s, road safety in connection with speed has increasingly become the focus of research. Thus, road safety can be noticeably increased by reducing the speed (Stone and Broughton, 2003; Koike, Morimoto, and Kitazawa, 2003). The accident severity and chances strongly depend on the speed and the difference in speed between the cars and the bicycles (Janstrup et al., 2023; Kim et al., 2007; P. Chen and Shen, 2016). For example, the chance of being killed as a cyclist in an accident is 16 times higher at a speed of more than 80 km/h (Kim et al., 2007). Similarly, adequate street lighting positively impacts accident rates in terms of lighting for bicycles and car infrastructure, making visibility at higher speeds better (P. Chen and Shen, 2016).

*Severity* According to Cheng and Peng's study in 2016, older cyclists are more prone to experiencing severe injuries in accidents. Furthermore, accidents involving large vehicles such as buses, trams, or trucks are more likely to cause severe injuries to cyclists (P. Chen and Shen, 2016). Very little research has been done in the area of infrastructure-related accident severity. In the past, analysis has focused more on the effect of helmets or only a single type of infrastructure per study, for instance the traffic circle in Wegman, Zhang, and Dijkstra, 2012. When modeling the location and severity of accidents,

it is crucial to consider land use surrounding accident locations, such as residential or industrial areas. This can significantly impact the severity of accidents. Furthermore, cyclists' perception of accident risks can also play a role in determining accident severity (Janstrup et al., 2023). Likewise, the average severity of an accident is lower on separate bicycle lanes/paths than when the bicycles are on the road. Additionally, the severity of accidents is also lower when bicycles can avoid complex intersections or have a separate track to traverse the intersection (B. Thomas and DeRobertis, 2013). Additionally, a study conducted in China has identified three categories of variables that can influence the severity of accidents. The first category includes cyclist age, district, day, and encountering vehicle type. The second includes road type, and the third includes street lighting and road surface (Yang et al., 2021).

**Approaches** Numerous theoretical frameworks and models have been leveraged to understand the spatial correlation between bicycle accidents, car traffic, and road infrastructure. For instance, applying Geographical Information Systems (GIS) in spatial analyses has become a critical tool for researchers, helping identify hotspots and the spatial distribution of accidents (Ziari and Khabiri, 2005). Statistical models such as regression analysis and machine learning (ML) methods such as Random Forests (RF) have also been employed to determine the influence of various factors on accident occurrence and severity, with the ability to predict accidents with an accuracy above 70 % (Santos, Dias, and Amado, 2022). In addition, kernel density estimation and cluster significance evaluations are suitable for modeling accident black spots where only a minimal amount of data is required (Bíl, Andrášik, and Janoška, 2013). Likewise, the Bayesian model, which uses pseudo-absence points to model binary factors, has proven to be an excellent methodological approach. Thus, pseudo-absence points are a good way to validate the accidents spatially and to train the model using the risk factors (Vandenbulcke, I. Thomas, and Int Panis, 2014). Janstrup et al., 2023 propose using a combination of neural networks and natural language processing to analyze perceived safety and accident severity to predict the severity of accidents on the network.

The current body of literature offers a solid basis for comprehending the connection between bicycle accidents and road infrastructure in terms of spatial and statistical correlation. However, further research is still needed to concentrate on Zurich's distinct context and the combination of attributes, as demonstrated by (Vandenbulcke, I. Thomas, and Int Panis, 2014). In addition, a thorough study is needed to explore the impact of infrastructure elements on the severity of accidents, as it is not possible to generalize the impact of individual elements, as there are significant local differences (Wegman, Zhang, and Dijkstra, 2012).

## 2.3 Research Gaps

The above review of the existing literature in the road safety domain reveals essential research gaps. Notably, three prominent gaps are:

1. Prior studies have investigated the association between roadway infrastructure and accidents involving bicyclists and motorized traffic. However, these studies tend to focus on singular aspects, leaving a comprehensive examination of a broad spectrum of infrastructure elements wanting.
2. The temporal correlation between motorized traffic density and bicycle accidents, particularly in specific urban contexts such as this thesis, has been largely unexplored. Most existing research fails to provide a nuanced, temporal understanding over a continuous period.

3. Research is also scant on how the various variables of road infrastructure contribute to the spatial correlation and severity of bicycle accidents in urban settings. Prevailing studies have typically focused on the origin/type of accidents and the parties involved, neglecting the potential impact of diverse infrastructure elements, especially in a combined approach.

Consequently, this thesis addresses these gaps, providing a comprehensive and nuanced understanding of road safety dynamics.

## 2.4 Research Questions

The three research gaps and uncertainties mentioned above give rise to the following two research questions and hypotheses, which are to be considered within the framework of this master thesis:

**RQ1: What is the temporal correlation between car traffic density and bicycle accidents in Zurich, and does this relationship change with the COVID-19 pandemic?**

**Research Hypothesis:** In the city of Zurich, a positive temporal correlation exists between car traffic density and bicycle accidents (De Rome et al., 2014; Bauman et al., 2008; Haworth, Heesch, and Schramm, 2018), which decreases during the COVID-19 pandemic due to reduced car traffic density and increased availability and usage of bicycle infrastructure (Kellermann et al., 2022; Lovelace et al., 2020). Literature suggests that increased car traffic density is associated with higher rates of bicycle accidents (De Rome et al., 2014; Bauman et al., 2008; Haworth, Heesch, and Schramm, 2018), as drivers may not be aware of or prioritize the safety of bicyclists. Additionally, studies have shown that the COVID-19 pandemic has decreased car traffic density and increased the use of bicycles as people try to avoid public transportation and prioritize physical distancing (Kellermann et al., 2022). Therefore, it is likely that the relationship between car traffic density and bicycle accidents in Zurich has changed during the pandemic, with a reduced positive correlation due to decreased car traffic density and increased usage of bicycle infrastructure.

**RQ2: How do bicycle accidents, car traffic, and the road infrastructure in Zurich correlate spatially, and how does the latter influence accident severity?**

**Research Hypothesis:** This thesis predicts a correlation between bicycle accidents and the network infrastructure in Zurich, contingent upon the proximity to different types of road infrastructure, such as intersections, traffic circles, and bicycle lanes/paths. The hypothesis postulates that the distance to these infrastructural elements and the local variations in speed and road width are crucial determinants of accident frequency and severity (Pucher and Buehler, 2008; Wegman, Zhang, and Dijkstra, 2012). In particular, it is anticipated that areas with high car traffic volumes, especially at junctions and traffic circles, could see elevated bicycle accident rates. Furthermore, the complexity of the infrastructure at a given location or the coexistence of multiple infrastructural types is expected to influence the rate and severity of accidents, considering additional factors such as the time of day, weather conditions, and the involved vehicles (Pucher and Buehler, 2008; Wegman, Zhang, and Dijkstra, 2012). However, a contrasting situation is predicted for areas with dedicated bicycle lanes/paths. The correlation between bicycle accidents and car traffic is expected to weaken. Reducing the distance to bicycle infrastructure, i.e., the presence of bicycle lanes and paths, can increase safety and reduce the rate and severity of accidents. (B. Thomas and DeRobertis, 2013; Garrard, Rose, and Lo, 2008). This hypothesis supports

previous studies linking traffic volume, infrastructure, and accident incidence and severity (Pucher and Buehler, 2008; Bíl, Andrášik, and Janoška, 2013). Nevertheless, the unique characteristics of Zurich—marked by its comprehensive cycling network and distinctive traffic policies—may lead to nuanced or even unique outcomes, thereby necessitating this analysis (Wegman, Zhang, and Dijkstra, 2012).

# Chapter 3

## Data

### 3.1 Traffic volume data

The traffic data in Zurich are collected from various road traffic counting stations throughout the city (Figure 3.1). These stations continuously measure the traffic volume of motorized vehicles (97 stations) and bicycles (22 stations) using induction loops in the asphalt. Each measuring point consists of several counting stations. If any counting station experiences temporary failures, the missing values are replaced with imputed values based on data from the same day and time over the past five years. The construction department oversees these counting stations and is responsible for data processing. All the data collected is available to the public <sup>4</sup> and is current and accurate as of May 2023.

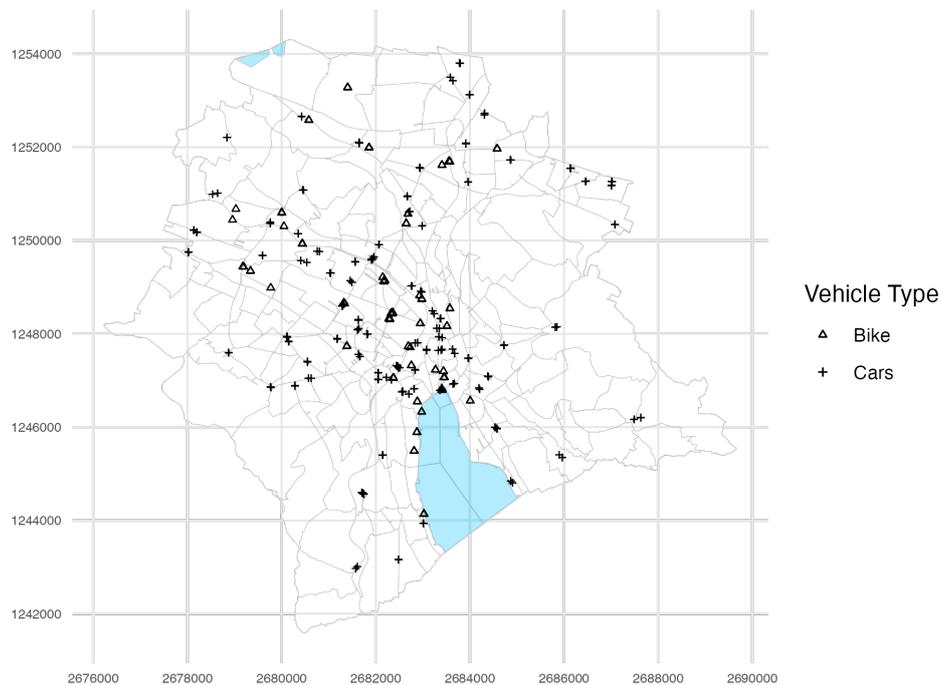


Figure 3.1: Locations of the traffic counting stations on Zurich’s motorized transport and bicycle traffic. Counting points for bicycle traffic (triangles) and locations for motorized transport (crosses). Lakes are colored in light blue for orientation. Status of the locations: May 2023.

<sup>4</sup><https://data.stadt-zuerich.ch>

### 3.1.1 Motorized vehicle count

To accurately track the flow of motorized traffic in Zurich, a comprehensive system is in place to measure traffic at 97 different counting stations, as depicted in Figure 3.1. This data is then gathered and aggregated into hourly intervals, providing a detailed record of the number of vehicles passing through during specific periods spanning from 2012 to 2021. The motorized vehicle count, which includes all motorized vehicles legally allowed to drive on the road, is referred to as road traffic in this paper.

### 3.1.2 Bicycle counts

In Zurich, cyclists are counted at 22 stations every 15 minutes, as shown in Figure 3.1. This data is used to determine the total number of cyclists during each period between 2012 and 2021. Measurements are conducted on location to guarantee precision, and any uncertainties or errors are considered. Additionally, a correction factor of +0.841 % is applied to the counts at each station and time interval to account for any device-related uncertainties.

### 3.1.3 Complete traffic model (GVM-ZH)

This study measured bicycle exposure to traffic using the GVM-ZH 2018 as a proxy variable. The GVM-ZH is a traffic forecast model that provides short-term and long-term planning data for 2018. For this research, the variable “traffic density” was chosen from the various model states that differ in settlement, structural development, and traffic supply in motorized transport. Traffic density is calculated based on the daily traffic of an average working day and the two peak hours in the morning and evening. It is given numerically for the number of vehicles for each road segment. The National Passenger Traffic Model of the Federal Government (ARE) was utilized to analyze traffic relations to more distant destinations and transit traffic. However, only the spatial coverage of the city of Zurich was utilized in this thesis.

## 3.2 Network and transport-related features

Within this thesis, data on network characteristics and the transport-related features of the road network were used. Thus, network properties consist of geodata, which appear as an independent spatial entity and can be individually geolocated. These topological geodata comprise data that characterize and classify the existing road and bicycle network. Tables 3.1 and 3.2 list all the relevant variables for network and transport-related features. Each attribute listed in the “Details” column briefly explains its purpose or characteristics. The variables comprise unprocessed raw data obtained from Open Data Zurich<sup>5</sup> and OpenStreetMap (OSM)<sup>6</sup>, along with data from DAV that is not publicly accessible and must be requested for use. All the data included within this section is current and accurate as of May 2023.

**OpenStreetMap** To complement the open data and DAV geodata, simple features were obtained from the OpenStreetMap (OSM) database using the `osmdata`<sup>7</sup> package in R. OSM data was downloaded

---

<sup>5</sup><https://data.stadt-zuerich.ch>

<sup>6</sup><https://www.osm.org>

<sup>7</sup><https://cran.r-project.org/web/packages/osmdata/vignettes/osmdata.html>

for the simple features: stop signs, right of way, and red lights. The minimum bounding rectangle of the Zurich city boundary was used to obtain the data.

Table 3.1: Network features on the transport network for the City of Zurich. Source: OSM, DAV, Open Data Zurich.

Attribute	Details	Count/Length	Variable Name
<b>Bicycle network</b>	Designated bicycles paths/lanes	1118.7 km	<i>BicycleNetwork</i>
<b>Bus lanes</b>	Designated bus lanes	1693.670 km	<i>BusLane</i>
<b>Road network</b>	Designated lanes for motorized vehicles	2566.4 km	<i>RoadNetwork</i>
<b>Tram track</b>	Tracks for trams including depots	346.2 km	<i>TramTrack</i>
<b>Road width</b>	Width of the road [m]	7 classes	<i>RoadWidth</i>
<b>GVM Kt. ZH</b>	Complete traffic model of the Canton of Zurich [count]	(-)	<i>Exposure</i>

The count/length of the geodata includes all network features and refers to the unprocessed version of the variables. If multiple geometries existed for the same network feature, they have been collapsed to a single geometry. It is important to note that the length of the variables does not refer to the actual length but rather the measurements of the parameters queried by the geodata. The geodata pertains to Zurich’s geographical area and does not distinguish between cantonal and municipal jurisdiction. A map of all complex intersections in Zurich is shown in the left-hand panel of Figure 3.2, while the right-hand panel displays the road network accessible to both motorized vehicles and bicycles.

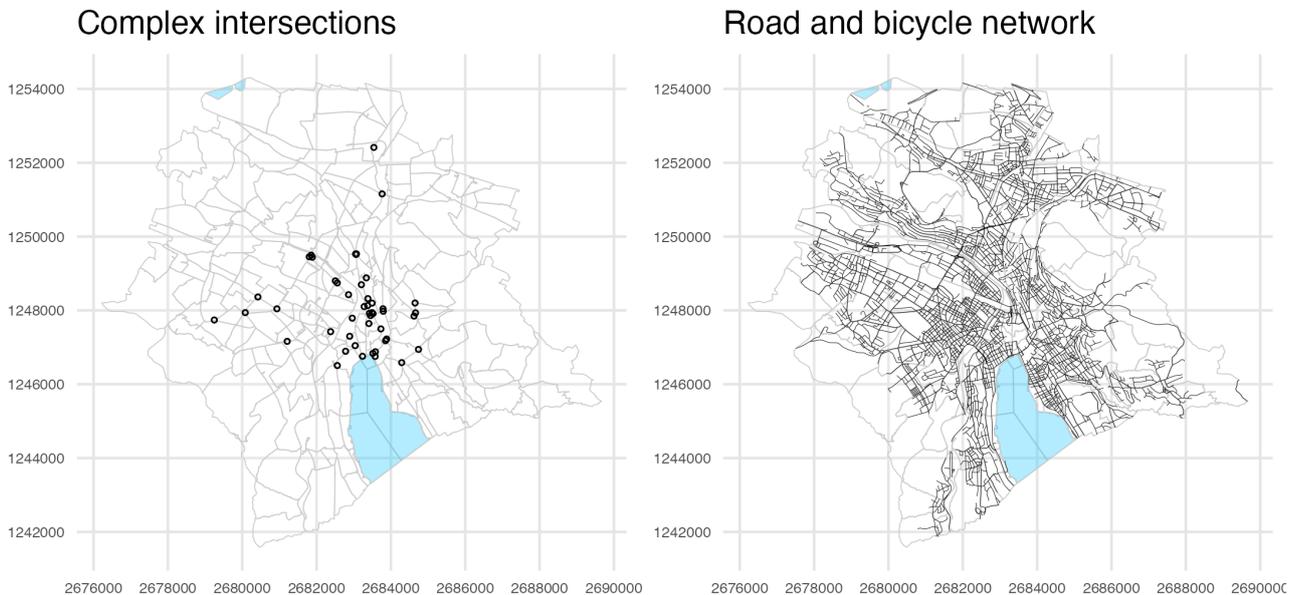


Figure 3.2: Complex intersections on the left. The right panel shows the road network of Zurich. Lakes are colored in light blue.

The locations of complex intersections are not publicly accessible, as they are gathered from Open Data

data sets. Please refer to Section 4.2. The network features utilized in this thesis are the following (Table 3.2).

Table 3.2: Transport-related features on the transport network for the City of Zurich.

Attribute	Details	Count	Variable Name
<b>Traffic areas</b>	Pedestrian area, T20 area, T30 area, Mixed speed limit* area (MSLA)	292	<i>TrafficArea</i>
<b>Signalized speed limit</b>	Indicated speed limit [km/h]	8 classes	<i>SpeedLimit</i>
<b>Right-of-way</b>	Priority given to a particular vehicle or road user	89/2613	<i>ROW</i>
<b>Red lights</b>	Traffic signals indicating to stop	1305	<i>RedLight</i>
<b>Stop signs</b>	Signs requiring drivers to come to a complete stop	91	<i>StopSign</i>
<b>Pedestrian crossing</b>	Designated areas for pedestrians to cross the road	2841	<i>PedestrianCrossing</i>
<b>Junction</b>	Intersection between a main road and a side road	849	<i>Junction</i>
<b>Complex intersections</b>	Complex intersections & traffic circles	47	<i>ComplexIntersection</i>
<b>Public parking spaces</b>	Parking spaces for MPT	46282	<i>PPS</i>

\*Note: In “*mixed speed limit areas*”, no particular speed limit has been defined in spatial planning, unlike in the T20 and T30 areas, where a general speed limit of 20 or 30 km/h applies, respectively. In these mixed speed limit areas, a speed limit of 50 km/h generally applies, while roads with a speed limit of 30 km/h may also be found. Accidents may thus occur on roads with both, a 50 km/h or a 30 km/h speed limit in these mixed areas.

### 3.3 Bicycle accident data

The bicycle accident data set provides information on all 5656 bicycle accidents reported between 2012 and 2021 involving at least one or more bicycles, including e-bikes. The city police maintain this data set based on the accident recording protocol (UAP) of the Federal roads Office (ASTRA), provided by the DAV. Note that an open version of this data set with accident coordinates and a few attributes is available through Open Data Zurich.

This data set concerns accidents within Zurich’s city limits and provides information about all such incidents. It is essential to note that only police-registered accidents are included in this data set, that is, those that do not result in a police report are not recorded. The data were anonymized and did not allow tracing back to the individuals involved. The data set comprises 52 parameters, but this report focuses on a reduced set of 24 parameters listed in Table 3.3. These parameters have been chosen based on the existing literature and studies relevant to this thesis’ topic. Regarding age and vehicle type, the accident-causing vehicle and age are referred to as Vehicle Type 1 and Age 1, respectively. In contrast, Vehicle types and Ages 2-4 are considered the affected traffic participant and their respective ages.

Table 3.3: Summary of Accident variables and corresponding information.

Category	Parameters	Description
<b>General information</b>	Date	Date of the; accident
	Year	Year of the accident
	Time	Time of the accident
	Accident type Group	Group of the accident type
	Accident type	Type of the accident
	Dead/ LV and HV	Number of fatally, lightly, heavily wounded
	Root cause	Root cause of the accident
<b>Infrastructure</b>	Point geometry	Coordinates of the accident location
	road Classification/ Outside City parameters	
	Regulation of Priority road type	Regulation of Priority Gravel, Paved
	Maximum speed road lighting	Maximum signalized speed limit Availability of road lighting
<b>Circumstances</b>	Weather conditions	Weather type (Rain, Sun, Snow, Hail, Cloudy)
	Light conditions	Natural light conditions (Day, Night, Dawn)
	road conditions	Dry, Wet, Icy, Damp, Snowy
	Traffic volume	Light, Normal, Strong, Congestion
<b>Involvement</b>	Bicycle count	Number of bicycles involved in the accident
	Vehicle type [1:4]	Type of vehicle involved in the accident
	Age [1:4]	Age of the involved
<b>Date</b>	Month	Month of the accident
	Week	Week of the year of the accident
	COVID-19	Binary before or during the pandemic

The classification of the motorized vehicles involved (vehicle type) is as follows:

E-bike, truck, bus, bicycle, tram, no vehicle (if the accident was self-caused, no other vehicle was involved), and all other vehicles like motorbikes and cars are MIV. The term *motorized vehicle* refers to all vehicles, while *MIV* is explicitly used for cars/ and motorbikes. The registered accidents are distributed over the entire city area, with apparent gaps in the forest areas. The spatial distribution is shown in Figure 3.3:



Figure 3.3: All locations of bicycle accidents within the Zurich city limits between 2012 and 2021. Lakes are colored in light blue. Status of the locations: March 2023

## COVID-19 Pandemic in Switzerland

In Switzerland, the inception of the COVID-19 pandemic can be traced back to late February 2020. The data used to evaluate the influences and impacts of the pandemic on road safety are as shown in Table 3.4, as follows:

Table 3.4: Timeline of the COVID-19 pandemic in Switzerland.

<b>Event</b>	<b>Date</b>
First confirmed COVID-19 case	<i>25th Feb 2020</i>
Start of the pandemic according to the BAG	<i>28th Feb 2020</i>
Federal government explains the exceptional situation	<i>16th March 2020</i>
First Wave	<i>March 2020</i>
Second Wave	<i>October 2020</i>

# Chapter 4

## Methodology

This chapter explains the software used, the preparation and cleaning of the data, and the statistical analysis methods employed. The statistical analysis is divided into two parts, each focusing on one of the two research questions. The approach for analysis was based on the most promising variables and approached from a prior exploratory phase. Figure 4.1 illustrates the simplified methodological flow that includes all the stages, from raw data to data preprocessing and final statistical analysis. For a detailed breakdown of the modeling and prediction work steps, refer to Figures 4.2 and 4.6.

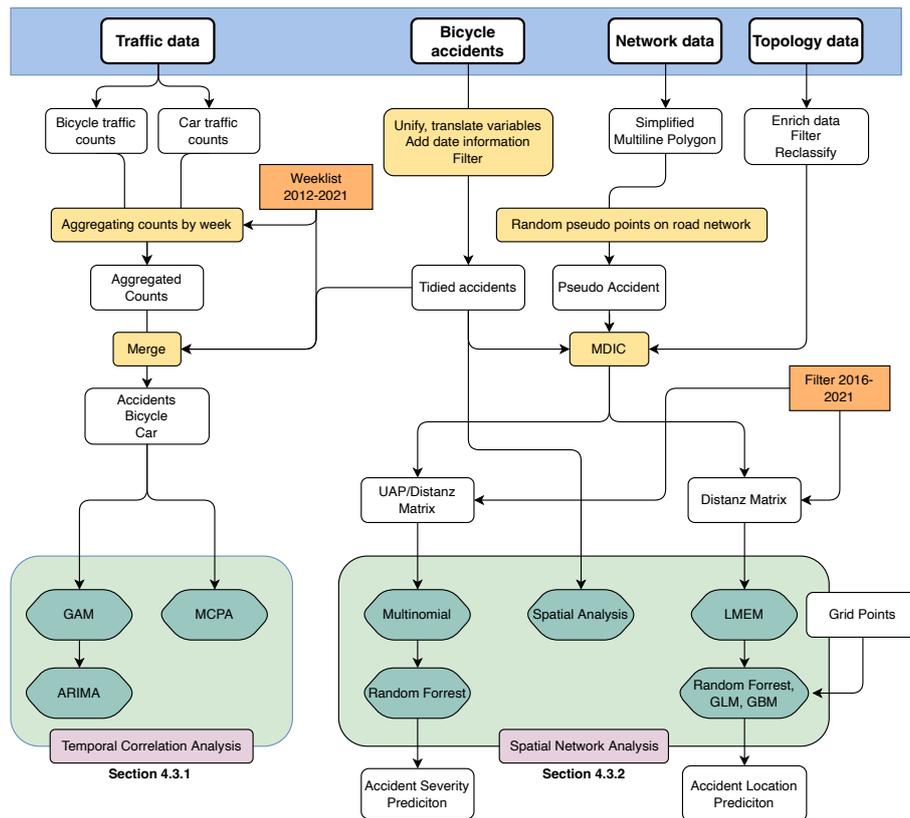


Figure 4.1: Flowchart of the methodology of this thesis, running from the raw data to the models and analysis. Data sets are indicated in white, work steps in yellow, parameters in orange, and models in dark green. The statistical models used are structured according to the two work packages and highlighted in light green. Abbreviations: GAM (generalized additive model), MCPA (multiple change point analysis), LMEM (linear mixed effect model), GBM (generalized boosted regression models), MDIC (minimum distance intersection calculations).

## 4.1 Technical Setup

R 4.3.0 software and RStudio 2023.06.0 were used for data processing and analysis - *used R packages are listed in the Appendix*. For visualizing spatial data, creating specific point data, and conducting essential geospatial analysis, QGIS (3.30.0-'s-Hertogenbosch) was utilized. In addition, ArcGIS Pro 3.1.1 was used to provide more advanced spatial analytical capabilities and render layout maps for clear visual representation. All software was run on a MacBook Pro using macOS 13.4.1, with 16GB RAM and a 3.2 GHz 8-core CPU.

## 4.2 Data Preparation

All relevant data were tailored to the Zurich urban area for processing and preparation to ensure the geospatial data were appropriately formatted for network and transport-related features. Aiming at enduring consistency, all naming conventions were standardized. All spatial data was reprojected and georeferenced utilizing the Swiss coordinate system LV95, EPSG:2056. Furthermore, uniform numeric values were assigned to all the data for a more straightforward analysis. All 5,656 accidents in the dataset were used to model the *time series* accurately. However, the *spatial analysis* was conducted solely on accidents between 2016 and 2021 to predict accident locations and severity, resulting in a subset of 3,411 accidents. This choice of study years ensured that the network and topological properties associated with the accidents were current and relevant. Including older accidents could introduce inaccuracies into the results, as they may be linked to a road section or infrastructure that no longer exists in its original form<sup>8</sup>. The dataset includes 2202 slightly injured, 479 seriously injured, and 13 fatal bicycle accidents.

In the following, only those work steps that have changed the public data in its raw form by changing the format or data structure, including the values due to recalculations, are explained.

**Right of way (ROW)** Two data sets were used to create the right of way data: 1) All documented right of ways from OSM 2) Using the network geodata, all intersections in the road network were located where two roads with a speed of 30 km/h meet, and one road is counterclockwise to the right of the other. The angle of the intersection, on the other hand, did not play a role.

**Junction** For the Junction, all primary and secondary roads where a road with a speed limit of 50 km/h meets a road with a speed limit of 30 km/h were geo-referenced.

**Complex intersections (CI)** To generate the complex intersections, all intersections between roads, bicycle lanes/paths, and tram tracks were computed and intersected with a 10 m buffer from the nearest crosswalk, thus creating an intersection point. Intersections were identified where at least three such points were present within a 50-meter radius.

**Network** The road network of motorized vehicles and bicycles was linked due to the different data availability of the city of Zurich and the canton of Zurich. As a result, all redundant and inaccurate roads were removed or simplified from this complete network. Likewise, network sections for motorized vehicles that were not publicly accessible were omitted.

---

<sup>8</sup>Based on personal communication with Markus Deublein, BFU

**Traffic and accident data** Motorized vehicle and bicycle traffic data were aggregated based on hourly aggregations from the raw data, using a weekly list that aggregated the first day of each week between 2012 and 2021 and combined into a dataset of 520 readings. The traffic data was then centered and scaled so it was standardized. Similarly, the accidents were aggregated, centered, and scaled based on the same weekly list as the traffic data every week, see Section 4.1. Additionally, the complete accident data set created two new data sets. The first one contains 2122 accidents caused by self-infliction (SIA), while the second data set includes the accidents induced by external causes (ECA), totaling 3534.

**Public parking spaces** The available public parking spaces were filtered based on the width of the road. A buffer was placed around each road with a  $width = (roadwidth/2)$  and intersected with the parking spaces. This resulted in two subsets: one with all the parking spaces on or at the edge of the road and another, binary dataset indicating whether there were any parking spaces on a particular road.

### 4.2.1 Exploratory Analysis

The initial exploratory phase of this study involved extensive data visualization and accumulating preliminary insights that could guide the subsequent analysis. Simple visualizations and spatial representations were implemented, as well as comprehensive regression models using all available variables in a second stage. This process enabled the identification of apparent trends and patterns, providing a foundational understanding for the subsequent research. Using R and QGIS, charts and maps were developed to quantify the accidents' spatial dimensions and evaluate the distribution of accident-related variables. This stage of the exploratory phase focused exclusively on the accident data, without considering any additional variables related to the network and transport-related features.

Following this, the network and transport-related features data were incorporated and modeled to ascertain the relevance of different variables. Variables that did not contribute significantly to the model in the first phase were eliminated to refine and optimize the model's performance (James et al., 2021). Thus, this initial exploratory analysis set the stage for a more detailed and targeted examination of the data in the work packages of the subsequent statistical analysis.

## 4.3 Statistical Analysis

### 4.3.1 Temporal Correlation Analysis

A Seasonal Mann-Kendall test was conducted to analyze the trend in accident and traffic data and the relationship between bicycle traffic and accidents. Pearson's product-moment correlation coefficient was used to estimate the correlation between bicycle accidents and motorized vehicle traffic and supplemented with a lag analysis, in which the lowest root mean square error (RMSE) of the lags from a linear model was used to model the time-lagged correlation between the accidents and the motorized vehicle traffic.

#### 4.3.1.1 Multiple change point analysis

A change point detection method was utilized to identify potential shifts in the data, focusing on variations in the mean or variance over time (Rebecca Killick, 2022). This analysis encompassed two detection techniques (*PELT* and *SegNeigh*) and two forms of penalty: Schwarz information criterion

(*SIC*) and Bayesian information criterion (*BIC*). The PELT (Pruned Exact Linear Time) detects changes in mean and variance by considering all possible change points and pruning sub-optimal solutions, allowing for exact inference in linear time complexity (Killick, Fearnhead, and Eckley, 2012). The BinSeg (Binary Segmentation) method applies a binary partitioning strategy to detect changes in mean and variance, dividing the data into segments and testing for changes within those segments, offering a more computationally efficient but approximate solution (Scott and Knott, 1974). These change points were then implemented on specific data sets encompassing motorized vehicles, bicycles, and traffic accident records. Another round of change-point detection was conducted using the optimal number of change points from the previous function. The BinSeg method and a BIC penalty were utilized to identify the change points in the time series of accidents, motorized vehicle traffic, and bicycle traffic.

#### 4.3.1.2 Generalized additive models (GAM)

To model the number of accidents in the traffic data using a GAM, the data sets with the aggregated accidents and traffic counts, based on the weekly data, were first merged. In statistical analysis, GAMs are an extension of the well-established GLMs that enable us to account for non-linear relationships between predictors and the response variable. This serves as an enhancement to GLMs, overcoming the potential limitation of strict linearity imposed in GLMs. A GAM can be implemented with the following equation:

$$g(E(Y_i)) = \beta_0 + \sum_{j=1}^p f_j(x_{i,j}) \quad (4.1)$$

In this equation,  $g(\cdot)$  denotes the chosen link function, which connects the expected value of the response variable  $E(Y_i)$  to the linear predictor. On the other hand,  $f_j(\cdot)$  represents smooth functions of the predictors that play a crucial role in capturing the non-linear relationships in the model. These smooth functions are estimated using techniques such as spline or kernel smoothing.

The analysis focused on a particular model derived from the Equation 4.1, examining the influences of traffic numbers and a seasonal relationship. The impact of the COVID-19 pandemic was introduced through the parameter ‘‘Period’’, which was binary and reflected the periods before and during the pandemic as in Table 3.4 . The model (Equation 4.2) incorporated smoothed weekly fluctuations, enabling seasonality effects:

$$\begin{aligned} \text{GAM Seasonal} = & \text{gam}(\text{Accidents} \sim s(\text{MotorizedVehicle}) + s(\text{Bicycle}) \\ & + \text{MotorizedVehicle} + \text{Bicycle} + s(\text{Week}) + \text{Period}) \end{aligned} \quad (4.2)$$

The model utilized the Poisson family and spline smooth functions to accommodate non-binary responses and non-linear patterns. The terms  $s(\text{MotorizedVehicle})$  and  $s(\text{Bicycle})$  applied spline smooth functions to the traffic volumes of *MotorizedVehicle* and *Bicycle*, enabling the model to capture non-linear patterns. Linear terms ‘*MotorizedVehicle*’ and ‘*Bicycle*’ were also included, indicating a mixed effect of linearity and non-linearity on bicycle accidents.

The model integrated temporal patterns using the term  $s(\text{Week}, bs = \text{’’ps,’’ } k = 52)$ , where *ps* denotes a periodic spline, and  $k = 52$  signifies the inclusion of 52 knots for the weeks in a year. The estimation of smoothing parameters was conducted using the Residual Maximum Likelihood (REML) approach, thereby reducing the risk of overfitting.

Following the GAM modeling, an ARIMA model was constructed based on the residuals from the GAM model, thereby addressing any temporal patterns not captured by the GAM model (Jaber,

Balint Csonka, and Juhasz, 2022). This step incorporated an Augmented Dickey-Fuller test to check the stationarity of the data. The ARIMA model parameters were optimized using the *AIC* (Corrected Akaike Information Criterion), providing a robust time series and accident data analysis method. Despite their robustness, these methods still have limitations. For instance, the requirement for stationarity and assumptions of normality in ARIMA models can constrain their use in specific scenarios (Petrică, Stancu, and Tindeche, 2016) - the smoothing parameters in GAMs can be sensitive to the choice of estimation method. They might miss non-linear trends (Dr. S. Jackson, 2023).

### 4.3.2 Spatial Network Analysis

#### 4.3.2.1 Accident Infrastructure Analysis

The combined road network was used to model bicycle accident severity. This consists of the entire road network, filtered by roads with a speed limit between 0 km/h and 80 km/h, and the entire bicycle network, with all bicycle paths and lanes. This was because, in both cases, neither bicycles nor motorized vehicles might be present. Therefore, evaluating the interaction between the two vehicle types would not be possible. Figure 4.2 shows the methodological steps schematically and is labeled according to the sections in which the methods are implemented:

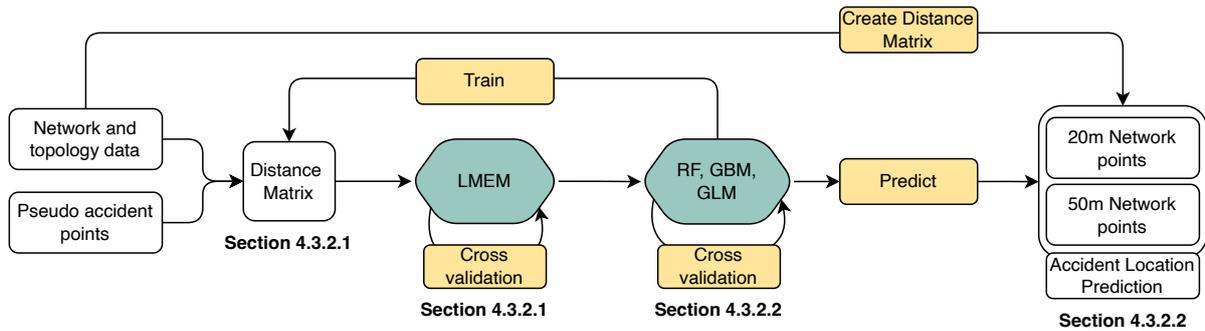


Figure 4.2: Workflow for training and predicting the *location* of bicycle accidents. Data sets are indicated in white, work steps in yellow, and models in dark green.

**Generation of Pseudo-Absence Points (PAP)** Ten thousand pseudo-absence points were generated to analyze traffic accidents on the road network. These points were selected randomly as locations where no accidents had occurred. They were also chosen to be at least 5 meters away from existing accident sites and 10 meters from each other (Figure 4.3). The generation of pseudo-absence points provides a strategy for modeling the absence of events, in this case, traffic accidents. To create the pseudo-absence points, 10'000 accident points were generated along the combined road network using the “*Random Points on Lines*” function in QGIS.

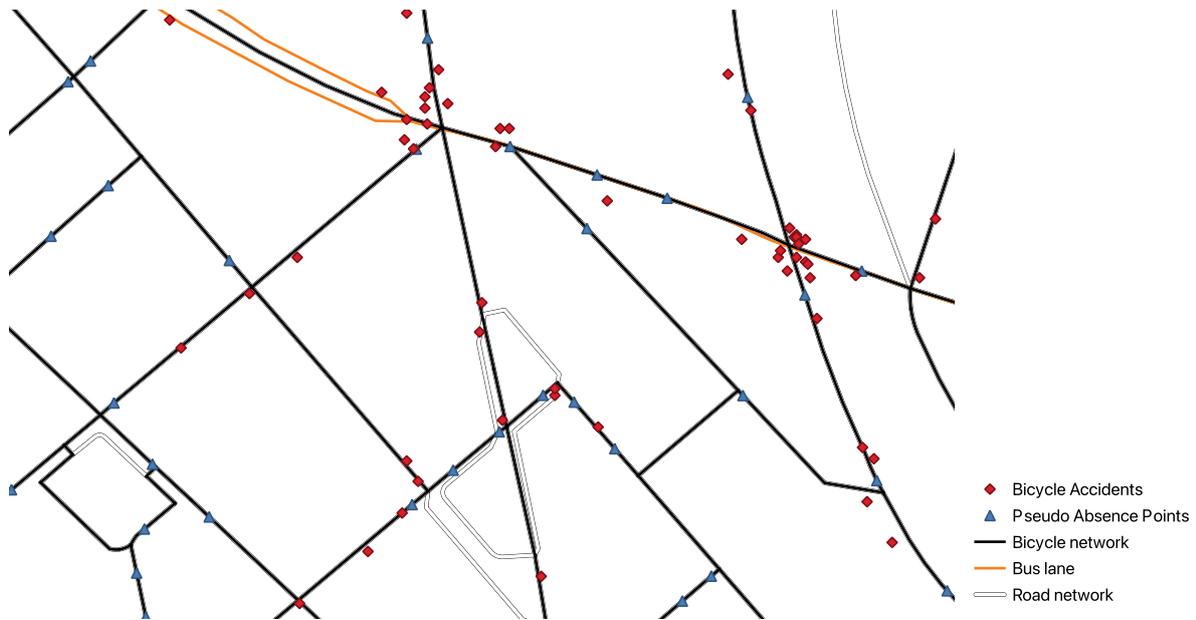


Figure 4.3: Bicycle accidents (*red*) and PAPs for randomly distributed accident locations (*blue*), the road network (*orange*), as well as the tram lines and bicycle paths.

The number of pseudo-absence points was based on a previous study by Barbet-Massin et al., 2012. The focus in generating these points was to ensure a high spatial resolution and proportional representation while mitigating spatial autocorrelation.

**Distance/Feature Matrix** The minimum distance intersection calculations (MDIC) were determined using two approaches visualized in Figure 4.4:

1. The first approach (1) calculates the smallest distance [m] between an accident (red) and a network or transport-related feature. The attribute of this next feature is then assigned to the accident - in this case, the speed limit of 50 km/h.
2. The second approach (2) is based on the same principle. Still, it calculates only the distance [m] to the nearest network feature from an accident (red) and assigns the smallest distance between the accident and this feature to the accident.

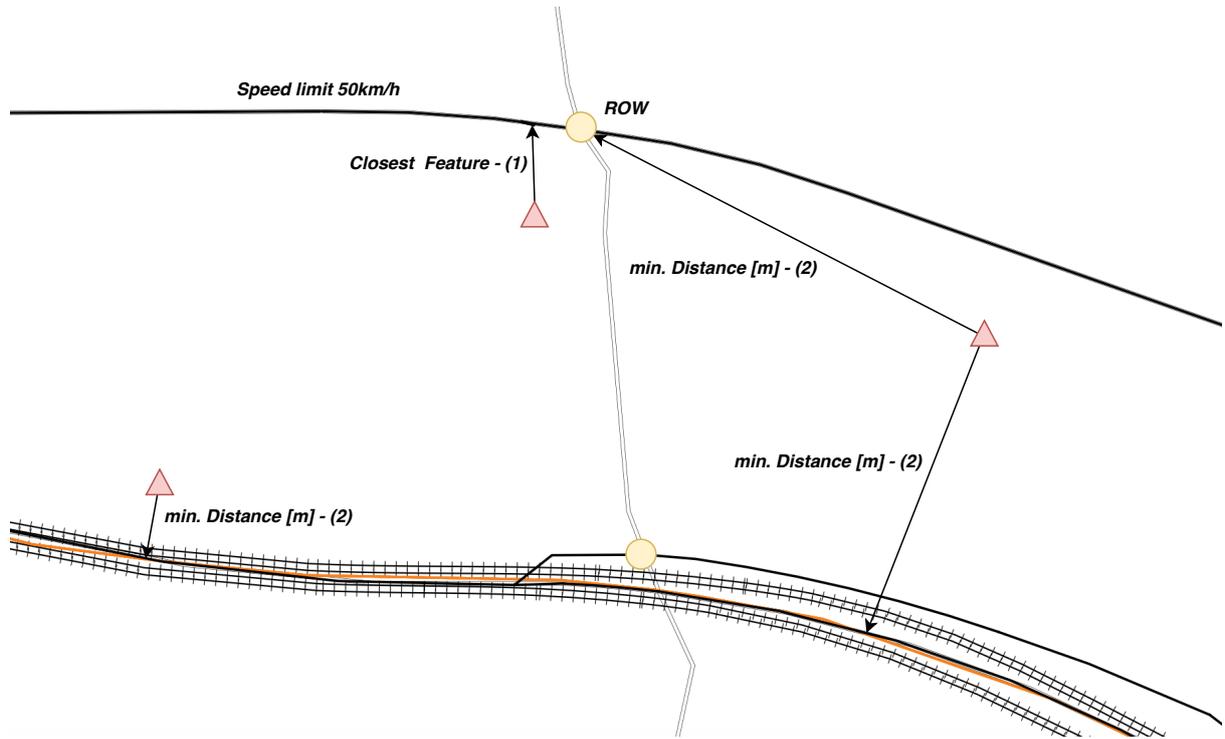


Figure 4.4: MDIC calculation schemata. Accidents (red triangle), ROW (right-of-way, yellow circles). The numbers in brackets correspond to the two approaches above respectively

**Linear Mixed Effect Models** To probe the spatial relationship between bicycle accidents across different road infrastructures in Zurich, a Linear Mixed Effects Model (LMEM) was employed. This approach, advantageous for accommodating both fixed effects (variables of interest) and random effects (additional variabilities such as geographical regions or periods), models a binary outcome (accident occurrence vs. Pseudo absence Points [PAP]) against various explanatory variables. These variables include proximity to infrastructural elements, road characteristics, traffic exposure, and a time indicator representing the COVID-19 pandemic.

The LMEM equation is as follows:

$$Y = X\beta + Z\gamma + \epsilon \quad (4.3)$$

$Y$  denotes the observed outcome,  $X$  and  $Z$  are known as design matrices,  $\beta$  and  $\gamma$  represent fixed and random effects vectors, respectively, and  $\epsilon$  indicates residual errors. Fixed effects include proximity to various road infrastructures, traffic speed and volume, road width, and the COVID-19 pandemic period. Random effects include district-level variations and speeds nested within road width.

Three distinct models were created for forecasting accidents using Equation 4.3. These models were constructed to provide a comprehensive overview of the accident patterns and delve into specific types of accidents. The first model, the “Comprehensive Model”, used Equation 4.4 and was created using the complete dataset. This model was developed using continuous scaled predictors to improve interpretability and ensure effective convergence.

Subsequently, two subset models were constructed. The first subset model, “Self-Inflicted Accidents Model (SIAM)”, focused on self-inflicted incidents. In contrast, the second subset model, the “External Causes Accidents Model (ECAM)”, was centered around all other types of accidents. These subset models also employed continuous scaled predictors for consistency and to facilitate interpretability. Using filtered data sets for these models allowed for a more nuanced analysis of accident causes.

Each model was designed to examine how the network infrastructure influences the cause of the accident. To assess and compare how well these models performed, the Mean Squared Error (MSE) was calculated. Finally, the models were checked for the assumptions of linearity, normality, and homogeneity of variance - this was done using a normal Q-Q plot and fitted vs. residual plots.

$$\begin{aligned}
 \text{LMEM} = \text{lmer}(\text{Accident} \sim & \text{PedestrianCrossing} + \text{StopSign} \\
 & + \text{TramTracks} + \text{ComplexIntersections} \\
 & + \text{RoadWidth} + \text{Exposure} + \text{BusLane} \\
 & + \text{TrafficAreas} + \text{ROW} + \text{COVID} \\
 & + (1|\text{Neighborhood}) + (\text{RoadWidth}|\text{SpeedLimit}) \\
 & + (\text{SpeedLimit}|\text{Exposure}) + \text{PPS} + \text{BicycleNetwork} \\
 & + \text{SpeedLimit}, \\
 \text{data} = & \text{DistancesMatrix})
 \end{aligned} \tag{4.4}$$

**Cross-Validation** The Comprehensive Model, SIAM, and ECAM prediction accuracy were evaluated by 10-fold cross-validation using a 90:10 split for training and test data. This split was performed seven times to ensure accurate representation. The mixed-effects linear regression model was then fitted to the training data using various predictor variables and random effects. The "bobyqa" optimization method was used for model fitting, with restricted maximum likelihood (REML) estimation (Bates et al., 2023). Predictions were made for the test data and compared to actual scores, with the mean square error calculated to quantify prediction error.

#### 4.3.2.2 Accident Location Prediction

Data from actual accidents and pseudo-absence underwent further predictive modeling procedures to predict bicycle accidents. This modeling involved three distinct methods: Gradient Boosting Machine (GBM), Random Forest (RF), and Generalized Linear Model (GLM), with the workflow depicted in Figure 4.2. Variables from the previous mixed-effect model (Equation 4.4) were reused for this analysis task, but this time, the interest was in predicting accident *locations*.

Two point grids were prepared within a 50 m road network buffer to generate the prediction points, adopting a 20x20 m and a 50x50 m grid, respectively. This creates a 50m buffer, which according to the black spot management standards of the BFU <sup>9</sup> forms the inner-city distance in which infrastructure elements and accidents can be interlinked. A buffer was therefore placed around each road with a *width* = 25 m, in which consequently the two grid sizes were computed using the "*st\_make\_grid*" function in R (Figure 4.5). Thus creating a 20 m and 50 m *network point grid*.

---

<sup>9</sup><https://www.bfu.ch/de/ratgeber/infrastruktur-sicherheitsinstrumente-issi>

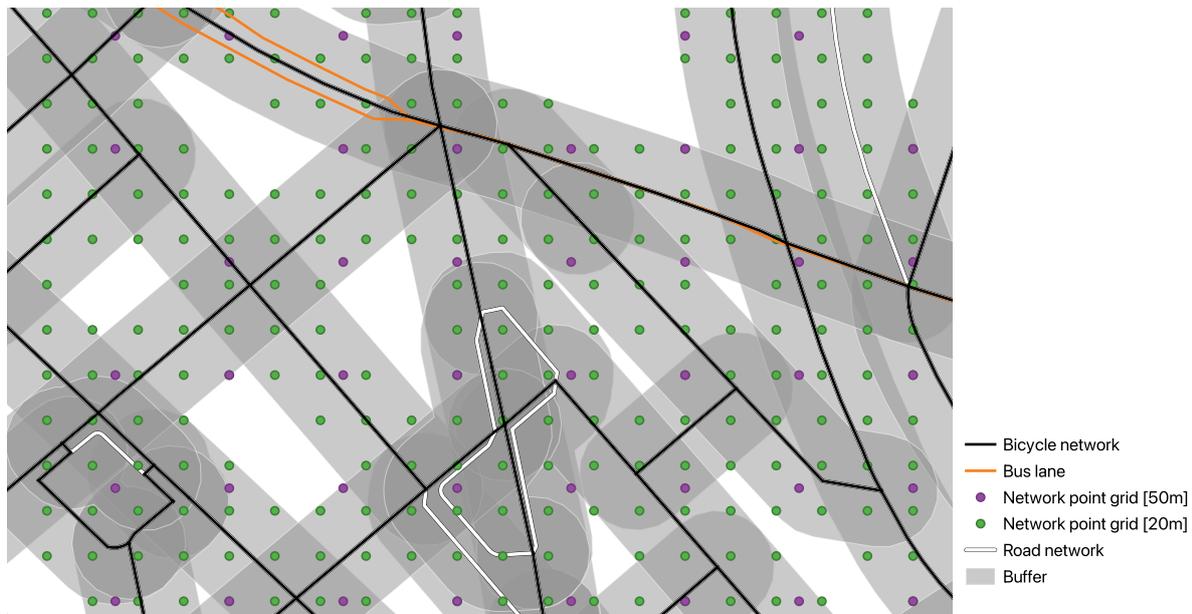


Figure 4.5: Prediction points on the road network for bicycle accidents in a 20x20 m (green points) and 50x50 m (purple points) network points grid. All grid points are clipped to a 50 m buffer around the street axes.

The three models were applied for each network point grid [20m,50m]. For the GBM, a Bernoulli distribution with 200 trees and ten cross-validation folds was used. The interaction depth was 9 (Greenwell et al., 2022). For the hyper parametrization of the RF model, see section 4.3.2.3. The GLM uses a binomial “logit” as a link function and a “glm. fit” as a method <sup>10</sup>.

On the other hand, GLMs are implemented with a binomial distribution and a logit link function to fit the binary response variable appropriately. The predicted accident probabilities derived from each model are then appended to the corresponding grid-based and network-based data sets, serving as valuable outputs for subsequent spatial analysis.

To evaluate the effectiveness of these three models, 10-fold cross-validation was used, splitting the data into 90 % training data and 10 % test data. The three models were trained and cross-validated using bicycle accidents and PAs, and the probability was predicted utilizing these models (GBM, RF, and GLM). These models facilitate the identification of the likelihood of an accident occurring at any specific spot within the network by leveraging significant interactions gleaned from actual accidents. The 350 most likely network grid points were selected and compared with the accidents for the three classification models for the 20 m and 50 m grid. This quantity represents approximately 10% of the 3411 accident points used in the analysis. This was done using a Mantel- test that finds the Mantel statistic as a matrix correlation between two dissimilarity matrices. Additionally, the mean distance between the actual and fitted accidents was computed. For visual clarity, only the 100 most probable points are consequently plotted.

#### 4.3.2.3 Accident Severity Analysis

**Infrastructure Effect** The steps for modeling and predicting accident severity (light, heavy, and fatal) are depicted in Figure 4.6. These steps are outlined in detail below.

<sup>10</sup><https://www.rdocumentation.org/packages/stats/versions/3.6.2/topics/glm>

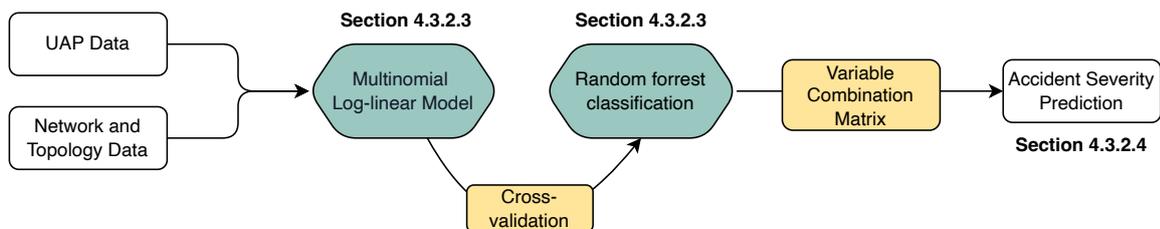


Figure 4.6: Workflow for modeling and predicting the *severity* of bicycle accidents. Data sets are indicated in white, work steps in yellow, and models in dark green.

For the modeling of accident severity, the most severe degree of accident was used as the severity for each accident. In other words, if an accident involves one severely injured person and one slightly injured person, the severity is set to severe. These were first reclassified if the parameters were not stored as factors for modeling and classifying the UAP and network/transport-related features variables. The exposure values involved categorizing values into seven classes using the Jenks method. This resulted in the following traffic count classes (Table 4.1).

Table 4.1: Upper and lower bounds for the reclassified traffic exposure. Values represent the number of motorized vehicles.

Class	Lower Bound	Upper Bound
1	0.00	3300.00
2	3300.00	8100.00
3	8100.00	13100.00
4	13100.00	18900.00
5	18900.00	27800.00
6	27800.00	40600.00
7	40600.00	63800.00

For the distances calculated between the accidents and the infrastructure objects, the intervals were chosen as shown in Table 4.2.

Table 4.2: Breaks and labels for reclassified distance-to-infrastructure categories

Category	Min. Distance [m]	Max. Distance [m]
Very Close	0	10
Close	10	50
Close-proximity	50	100
Distant	100	500
Far	500	$\infty$

Additionally, the persons involved in an accident were categorized into distinct age classes. There are five age classes, as outlined in Table 4.3.

Table 4.3: Breaks and labels for age categories

Category	Min. Age	Max. Age
Children	0	14
Teenager	15	23
Young Adult	24	35
Adult	36	65
Elderly	66	99

Public parking spaces were categorized into “Parking spaces on the road” and “Parking Spaces beside the road”. This was done following the steps outlined in Section 4.2. Likewise, the values of traffic zones were categorized into “Pedestrian area”, “T20 area”, “T30 area”, and “MSLA”.

For both methods, the multinomial and the random forest models, the severities were first balanced using the SMOTE algorithm <sup>11</sup>. 2500 was chosen as perc.over and 750 as perc.under with  $k = 10$ .

This ensured that all degrees of severity were equally present in the model and were not modelled and predicted in an unbalanced way. The equation implemented in R using the SMOTE method is the same as equation .

**Multinomial Log-linear Model (Mulinom)** To gain a clearer understanding of an accident’s severity, a multinomial log-linear model was implemented. This method is typically employed when the dependent variable has more than two classes and is a factor (Brian Ripley and William Venables, 2016) . The approach involves predicting the likelihood of an observation falling into a specific category based on a set of predictors  $X = (X_1, X_2, \dots, X_p)$ . The model develops  $(K - 1)$  equations for a response variable  $Y$  with  $K$  categories.

The core equation is expressed as:

$$\log \left( \frac{P(Y = k|X)}{P(Y = K|X)} \right) = \beta_{k0} + \beta_{k1}X_1 + \beta_{k2}X_2 + \dots + \beta_{kp}X_p, \quad (4.5)$$

Coefficients  $\beta_{kj}$  are estimated using maximum likelihood estimation. Predicted probabilities are obtained using an inverse logit transformation.

Multinomial log-linear regression models allow for analyzing complex relationships between multiple predictors and categorical outcomes with more than two categories. However, it is essential to note that the independence of irrelevant alternatives (IIA) is assumed, and sufficient sample sizes for each outcome category are required to ensure unbiased and reliable estimates. In this particular case, multinomial logistic regression was utilized to examine accident severity, which was classified into three levels: “fatal”, “heavy”, and “light”. The model (Equation 4.6) was created using a set of explanatory variables obtained from the UAP and the distances and attributes to the nearest network infrastructure features, similar to those already implemented and found significant in Section 4.3.2.1.

<sup>11</sup><https://www.rdocumentation.org/packages/DMwR/versions/0.4.1/topics/SMOTE>

$$\begin{aligned}
\text{AccidentSeverity} = \text{multinom}(& \textit{Severity} \sim \textit{LightCondition} + \textit{RoadCondition} \\
& + \textit{VehicleType1} + \textit{VehicleType2} + \textit{StopSign} \\
& + \textit{ComplexIntersections} + \textit{RoadWidth} + \textit{TramTracks} \\
& + \textit{PedestrianCrossing} + \textit{TrafficAreas} \\
& + \textit{COVID} + \textit{Junction} + \textit{AccidentLocation} \\
& + \textit{Age1} + \textit{Age2} + \textit{Exposure} + \textit{BicycleCount} \\
& + \textit{SpeedLimit} + \textit{PPS} + \textit{BusLane} + \textit{TramTracks} \\
& + \textit{SpeedLimit} * \textit{RoadWidth} + \textit{Speed} * \textit{Exposure} \\
& + \textit{RoadWidth} * \textit{Exposure} + \textit{ROW} \\
& + \textit{PPS} + \textit{BicycleNetwork})
\end{aligned} \tag{4.6}$$

**Random Forest Classifications** The classification of accident severity was performed using a Random forest (RF) model using the same explanatory and response variables were used as in Equation 4.6 of the multinomial log-linear model.

In constructing the random forest model, several hyperparameters were selected to control the growth and performance of the trees within the forest. 1000 trees ( $\text{n tree} = 1000$ ) were grown in the ensemble to ensure a robust and stable prediction. The number of randomly selected predictors for each split was set to 15 ( $\text{m try} = 15$ ), guiding the diversity among the individual trees. To regulate the complexity of the trees, the minimum node size was fixed at 5 ( $\text{node size} = 5$ ). In contrast, the maximum number of terminal nodes was limited to 30 ( $\text{max nodes} = 30$ ), ensuring that the trees were not overfitting the training data. The model was configured to compute variable importance ( $\text{importance} = \text{TRUE}$ ) to identify significant predictors and interactions. Additionally, an interaction depth of 18 ( $\text{interaction.depth} = 18$ ) was applied to capture complex relationships within the predictors, and a 10-fold cross-validation ( $\text{cv.folds} = 10$ ) was employed to validate the model’s performance across different subsets of the training data (Breiman, 2001). These hyperparameters were chosen to enhance the model’s interpretability and generalization ability to unseen data. The model’s accuracy was then assessed using the Kappa statistic, which considers the possibility of random agreement between predicted and observed categories (Yang et al., 2021).

#### 4.3.2.4 Predictive Modelling of Accident Severity

A predictive model was created by analyzing factors contributing to varying accident severity levels. This was based on insights from the RF and multinomial models, respectively. This examination focused on the variables outlined in Section 4.3.2.3, observing their effects in mutual interactions. Due to computational limitations in R on the available hardware (Section 4.1), only the 10 most significant variables from the RF model were identified based on their effect size (Section 4.3.2.3). The variables used are listed in Table 4.4.

The resulting RF model used to predict the accident severity was the following:

Table 4.4: Top 10 important variables for parameter level combination matrix

Nr.	Variable
1	VehicleType2
2	VehicleType1
3	RoadWidth
4	ROW
5	Age2
6	Age1
7	TramTrack
8	SpeedLimit
9	PPS
10	BusLane

$$\begin{aligned}
 \text{PredictedAccidentSeverity} = \text{multinom}(& \textit{Severity} \sim + \textit{VehicleType1} + \textit{VehicleType2} \\
 & + \textit{RoadWidth} + \textit{ROW} + \textit{Age2} \\
 & + \textit{Age1} + \textit{TramTrack} + \textit{SpeedLimit} \\
 & + \textit{PPS} + \textit{BusLane}
 \end{aligned} \tag{4.7}$$

A matrix containing all possible parameter-level combinations was created to analyze the interplay of critical variables. This resulting matrix consisted of approximately 11 million variable combinations. The next step involved applying the RF model (Equation 4.7) to the matrix. For each severity level (light, heavy, and fatal), the variable combinations most likely to cause that particular severity level were predicted using the preceding RF model. This analytical process allowed for a deeper understanding of factors contributing to accident severity by highlighting the combinations of variables most likely to result in each of the three severity levels.

# Chapter 5

## Results

This thesis examines the relationship between motorized vehicles traffic density, bicycle accidents, and road infrastructure in Zurich, focusing first on the temporal evolution of bicycle accidents (Section 5.1) and second on the impact of accidents on the network infrastructure on the location and severity (Section 5.2). The chapter will present the detailed results of the analyses and statistical models introduced in Chapter 4.

### 5.1 Temporal Correlation Analysis

#### 5.1.1 Exploratory Analysis

Figure 5.1 shows that the trends of motorized vehicle and bicycle traffic data differ significantly. Motorized vehicle traffic remained higher than average at the start of 2012 but decreased significantly toward the summer of 2016, with a record high in the winter of the same year. Traffic volume peaked again in 2018 and has been decreasing steadily since then. Notably, there were some anomalies after the outbreak of the COVID-19 pandemic in 2020, as well as in the winter of 2015 and 2022. On the other hand, bicycle traffic has shown a consistent positive trend since 2012, with seasonal fluctuations becoming more evident and more cyclists in the summer. However, the positive trend has leveled off since 2018 and is now stable.

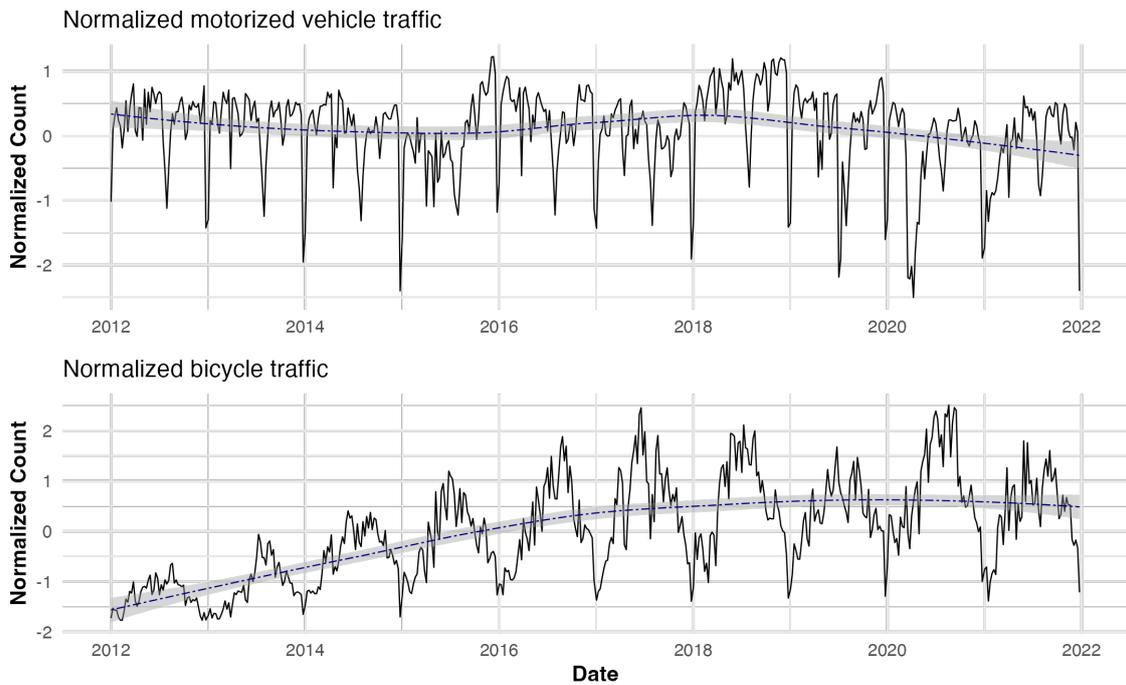


Figure 5.1: Time series of the normalized bicycle and motorized vehicle traffic counts, aggregated weekly between 2012 and 2021, in Zurich. Trend lines fitted with a LOESS function (Locally Estimated Scatterplot Smoothing, a method of local polynomial regression fitting), family used: “Gaussian” (Cleveland, Grosse, and Shyu, 1992).

In Figure 5.2, the time series data indicates a distinct trend in the number of accidents. From 2012 to 2018, the number steadily increased, with seasonal fluctuations. However, since the start of 2020, the trend has been decreasing and has continued to drop until 2022. It is worth noting that the summer of 2021 saw the highest number of bicycle accidents, with 2020 and 2019 following closely behind.

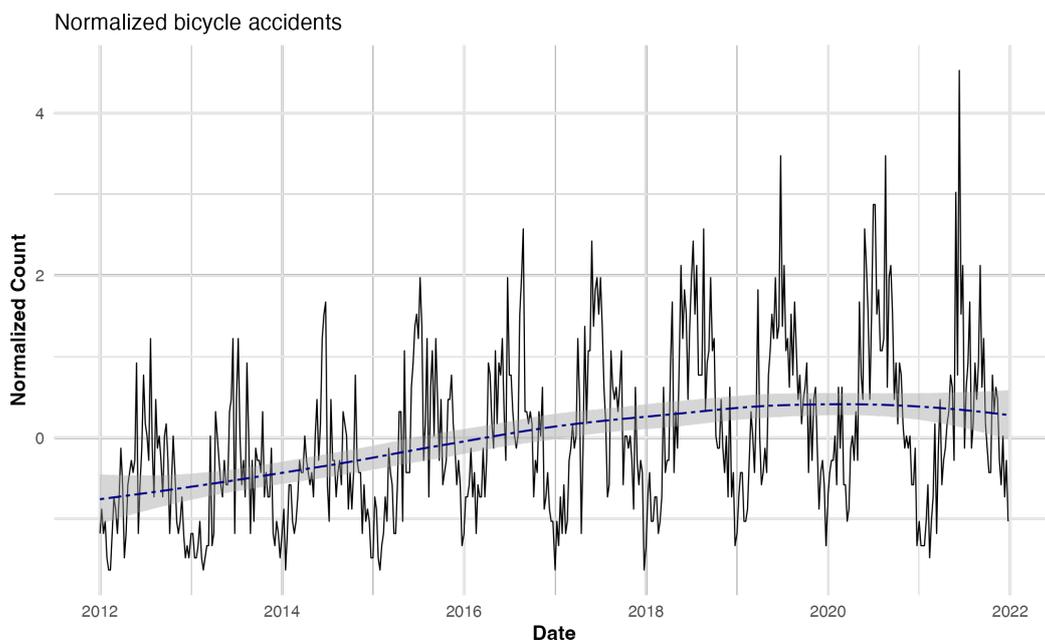


Figure 5.2: Time series of normalized counts of bicycle accidents aggregated weekly between 2012 and 2021 in Zurich. Trend lines fitted with a LOESS function (Locally Estimated Scatterplot Smoothing, a method of local polynomial regression fitting), family used: “Gaussian” (Cleveland, Grosse, and Shyu, 1992).

## 5.1.2 Statistical Analysis

**Trends** The Seasonal Mann-Kendall test has shown that there has been a strong positive trend in bicycle traffic ( $\tau = 0.644$ ,  $p < 2.22e-16$ ), indicating a significant increase over time. On the other hand, motorized vehicle traffic has shown a weak negative trend ( $\tau = -0.0923$ ,  $p = 0.0073809$ ), resulting in a slight decrease. Bicycle accidents have displayed a moderate positive trend ( $\tau = 0.421$ ,  $p < 2.22e-16$ ), reflecting a significant increase. The ratio of bicycle traffic to bicycle accidents has demonstrated an intermediate negative trend ( $\tau = -0.224$ ,  $p = 8.063e-11$ ), signifying a notable decline. These findings suggest that bicycle traffic and accidents have risen while motorized vehicle traffic has fallen. However, the ratio between bicycle traffic and bicycle accidents has been declining, indicating that the increase in bicycle accidents has outpaced the growth in bicycle traffic.

**Correlation** The data was normally distributed. Hence a Pearson’s product-moment correlation was implemented, showing a positive correlation between the “motorized vehicle” and “Accidents” variables. The correlation coefficient (*cor*) was estimated at 0.153, indicating a positive relationship between the two variables. The observed correlation is statistically significant, with a *t*-value of 3.5321, which resulted in a *p*-value of  $> 0.05$ . The 95 % confidence interval for the correlation coefficient is [0.068, 0.236], meaning the actual correlation falls within this range.

**Lag Analysis** Based on the GLM analysis results, it was found that none of the predictor variables, including the axis intercept, have a statistically significant effect on the “accidents” variable. The estimated coefficients for the intercept and the lagged values for “motorized vehicle” and “bicycle” were -0.3289, -0.4082, and 0.3922, respectively. However, their corresponding *p*-values of 0.535, 0.509, and 0.622 suggest that these coefficients are not significantly different from zero. In summary, lag analysis shows that the lagged values for motorized vehicles and bicycles have no significant influence on the accident count, and the traffic density in the past has no influence on the accident risk in the future.

### 5.1.2.1 Multiple Change Point Analysis

An analysis of change point dates using the SegNeigh method, with a BIC penalty and  $Q = 10$ , for bicycle accidents, bicycle traffic, and motorized vehicle traffic from 2012 to 2021 revealed several critical patterns. Notably, change points for bicycle traffic were observed predominantly in the initial years, specifically between 2012 and 2016, with seven distinct change points. However, after 2016, the number of change points dropped considerably, with only two occurrences noted in 2020.

In contrast, no change points were detected for motorized vehicle traffic before 2018. Post-2018, there was a marked surge, with eleven distinct change points identified from 2018 to 2021. All the change points for accidents coincided with motorized vehicle traffic, starting from 2018 to 2021 (Table 5.1). This synchronous pattern may suggest a potential correlation between motorized vehicle traffic and accident incidences.

The bicycle counts show two distinct change points before the COVID-19 pandemic: May 6, 2020, and September 16, 2020. These change points suggest some shift in bicycle usage during the pandemic. However, the motorized vehicle and accident groups show three similar change points post-pandemic: March 11, 2020, April 29, 2020, and December 16, 2020. The concurrence of these change points indicates that similar factors may have influenced motorized vehicle usage and accident rates during the pandemic. Another change point occurred in the motorized vehicle, and accident counts on February 19, 2021, possibly indicating another significant shift in the trend of these variables after the first year of the pandemic.

Table 5.1: Change point dates (YYYY:MM:DD) for bicycle accidents, bicycle traffic, and motorized vehicle traffic between 2012 and 2021. An 'X' indicates if a change point occurred on that date for the corresponding variable.

Date	Bicycle	Motorized vehicle	Accidents
2012-11-18	X		
2013-04-02	X		
2014-05-07	X		
2014-10-29	X		
2015-04-02	X		
2015-11-12	X		
2016-03-25	X		
2018-01-01		X	X
2018-10-15		X	X
2018-12-17		X	X
2019-01-01		X	X
2019-04-09		X	X
2020-03-11		X	X
2020-04-29		X	X
2020-05-06	X		
2020-09-16	X		
2020-12-16		X	X
2021-02-19		X	X

Figures 5.3 and 5.4 show the resulting change points for road and bicycle traffic and bicycle accidents<sup>12</sup>. The SegNeigh method showed five distinct periods from 2012 to 2015, then to 2018, and from the time before the pandemic to the second wave, but no changes at the start of the pandemic, though there were changes towards the summer of 2020. The Pelt method showed so many change points that the results were classified as implausible and disregarded for the results. There were many fluctuations in bicycle traffic between 2012 and 2016, which disappeared from 2016, showing a steady count in bicycle traffic until the beginning of the pandemic, when there was a marked increase. Beginning of 2021, bicycle traffic resumed at the pre-pandemic level. Motorized vehicle traffic remained relatively stable between 2012 and 2017, with only very little seasonality and fluctuations, with fluctuations starting only in 2017 with a notable increase beginning in 2018. Two more change points can be detected at the beginning of the first and second waves of the pandemic.

<sup>12</sup>The following dates are referenced here: start of pandemic (28 February 2020, start of first wave (March 2020), start of second wave (October 2020)).

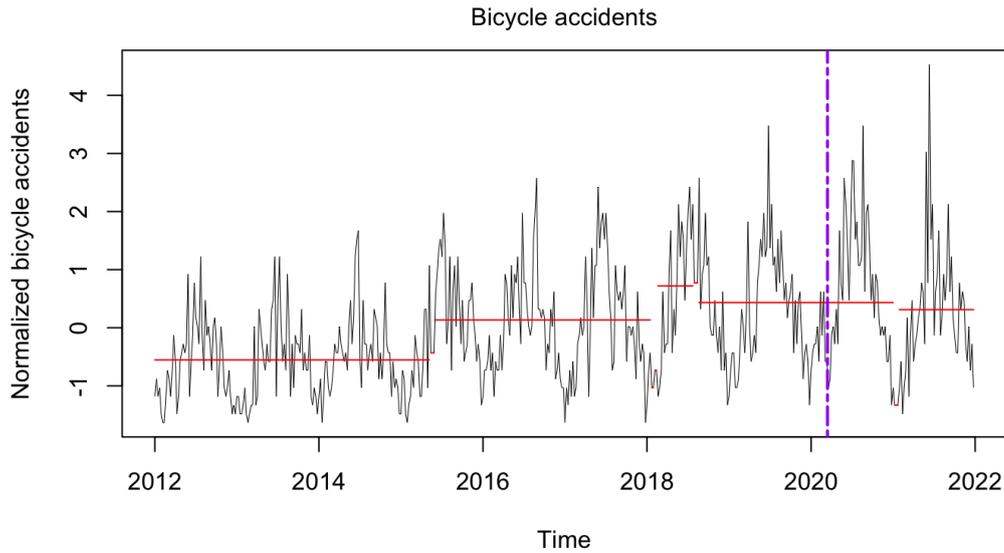


Figure 5.3: Change points for the accidents in Zurich detected with the SegNeigh method. The significant change points are indicated as horizontal red segments. Start pandemic Switzerland: 27th February 2020 (purple).

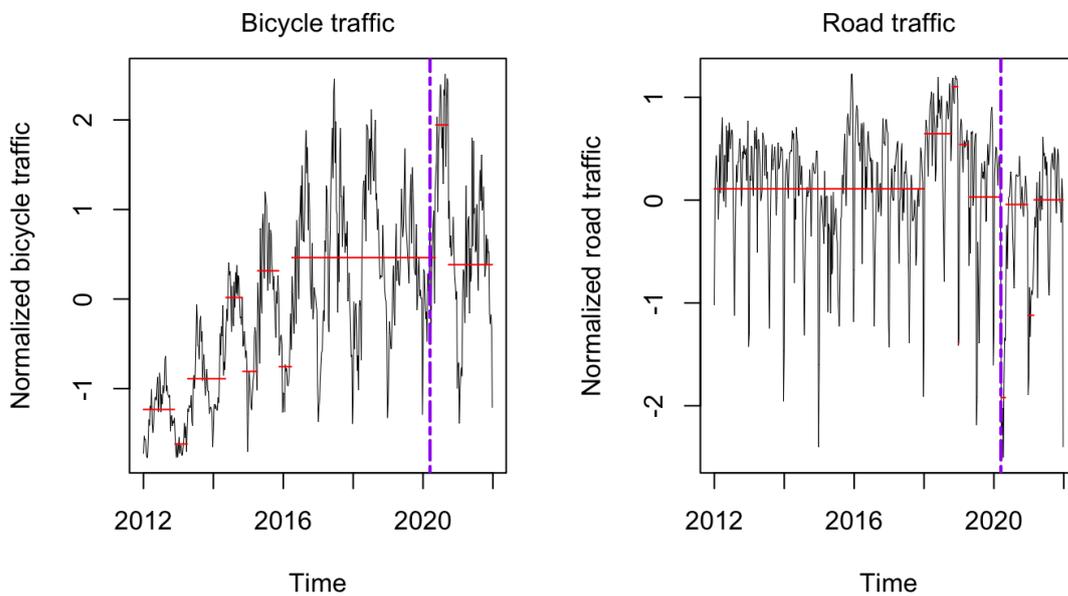


Figure 5.4: Change points for Zurich’s bicycle and motorized vehicle traffic with the SegNeigh method. The significant change points are indicated as horizontal red segments. Start pandemic Switzerland: 27th February 2020 (purple).

### 5.1.2.2 GAM

The results of the GAM (Table 5.2) with a Poisson family and log link function, applied to accident data, are presented below. The model includes both smooth and parametric components. The parametric components include bicycle traffic, road traffic, and before/after COVID predictors, while the smooth

components consist of smooth functions for bicycle traffic, road traffic, and weeks.

Regarding parametric coefficients, the presence of bicycles significantly influences the number of accidents ( $p < 0.001$ ). The effect of motorized vehicles on accident numbers is marginal and less statistically significant ( $p = 0.089$ ). The pre-pandemic period shows no difference from the post-pandemic period, meaning that there was no significant difference in accident rates before and after the start of the lockdown ( $p = 0.281$ ). The approximate significance of smooth terms suggests that non-linear effects associated with bicycles ( $p < 0.001$ ) and the week of the year ( $p < 2e-16$ ) have highly significant influences on the number of accidents. In contrast, the motorized vehicle presence's non-linear impact is insignificant ( $p = 0.301$ ). The week's smooth term had the most significant impact on the model, as indicated by the highest Chi-squared value, suggesting a robust seasonal effect. The model explained 72.7 % of the deviance, with an adjusted R-squared value of 0.715, signifying a relatively good fit to the data. This indicates that the number of accidents can be significantly explained by the presence of motorized vehicles, bicycles, and the specific week of the year, along with the non-linear effects of the presence of bicycles and the week of the year.

Table 5.2: Poisson GAM model estimates for accident prediction.

<i>Dependent variable:</i>				
Accidents				
	(Intercept)	Motorized vehicle traffic	Bicycle traffic	Period
Estimate	0.000e+00	5.952e-06	2.487e-04	-3.920e-02
Std. Error	0.000e+00	3.500e-06	5.202e-05	3.637e-02
z value	NaN	1.700	4.782	-1.078
Pr(> z )	NaN	0.089*	<2e-06***	0.281
Approximate significance of smooth terms				
	s(Motorized vehicle)	s(Bicycle)	s(Week)	
edf	1.608	5.072	6.077	
Ref.df	2.012	6.212	7.580	
Chi.sq	2.355	35.139	197.666	
p-value	0.301	<2e-16***	<2e-16***	
Observations	519			
R-squared (adj.)	0.715			
Deviance explained	72.7%			
Scale est.	1			

*Note:* Significance codes: \*  $p < 0.05$ , \*\*  $p < 0.01$ , and \*\*\*  $p < 0.001$ .

This is also reflected in Figure 5.5, where the modeled values accurately reflect the seasonality and the trend of the accident figures. There are deviations, especially around extreme values such as summer and winter. The model mainly overestimates high values and underestimates low values, respectively. The 72.7 % explained deviance is also reflected in this figure.

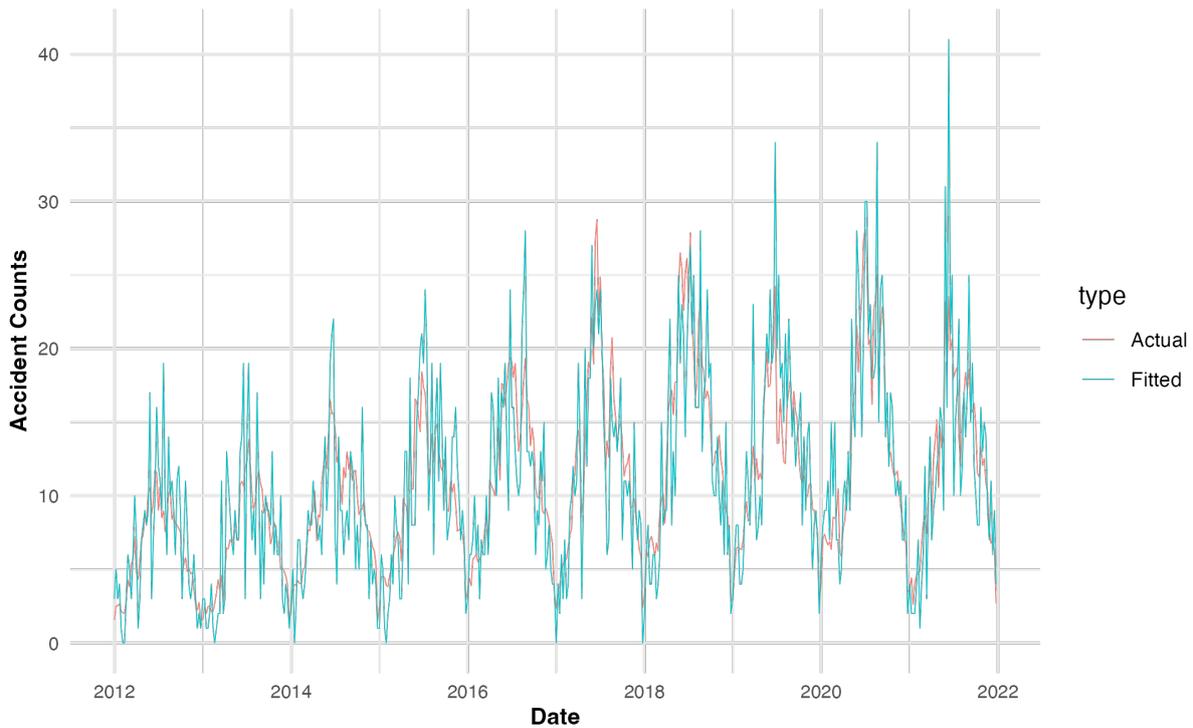


Figure 5.5: Modeled (red) versus actual accident values (blue) using a generalized additive model.

**ARIMA** The ARIMA(1,0,0) model on the GAM residuals successfully captures the correlation between motorized vehicles and bicycle traffic and bicycle accidents. The first-order autoregressive term (ar1) has a coefficient of 0.1189 (standard error = 0.0436), suggesting a mild positive autocorrelation in the residuals, indicative of an increased risk of bicycle accidents with increasing motorized vehicle traffic.

Despite a  $\sigma^2$  of 1.113, reflecting some unexplained variability in the relationship, the model's root mean square error (RMSE) is relatively low at 1.052818, demonstrating a decent model performance in predicting bicycle accidents. Additionally, the autocorrelation function at lag 1 (ACF1) near zero at -0.0002852135 supports that the model has effectively captured the underlying relationship between motorized vehicle traffic and bicycle accidents.

## 5.2 Spatial Network Analysis

### 5.2.1 Exploratory Analysis

#### 5.2.1.1 Average nearest neighbor analysis

The results of the analysis of accident locations and severities are explained below. Fundamental to this is that the average nearest neighbor analysis results show that the observed mean nearest neighbor distance is 30.39 m, while the expected mean nearest neighbor would be 72.75 m. This gives a nearest neighbor index (NNI) of 0.417802, indicating that the accident locations are significantly spatially clustered, with accident points closer together than expected from a random distribution (which would lead to an NNI around 1). Given the z-score of -83.76 ( $p$ -value = 0.001), there is less than a 1 % chance that this clustering pattern results from chance alone - indicating that the probability of observing such a low z-score by chance is improbable. These results suggest a significant spatial pattern in the accident

locations and that the observed clustering is not due to events, and there is an underlying pattern to the accidents in Zurich.

## 5.2.2 Statistical Analysis

### 5.2.2.1 Accident Infrastructure Analysis

In Figure 5.6, it can be observed that the density distributions of both pseudo and actual bicycle accidents are somewhat similar. These distributions are skewed towards the left, indicating that the highest density of accidents occurs within relatively short distances to infrastructure elements. Most of these accidents take place within a radius of 100 m around complex intersections and public parking spaces. Maximum density is around 25 m or less for bus lines, pedestrian crossings, red lights, and intersections. The distribution of right-of-way and stop signs is less clear, with the maximum density observed between 100 and 250 m. Motorized vehicle lines also show a precise distribution, with a maximum density of around 50 m.

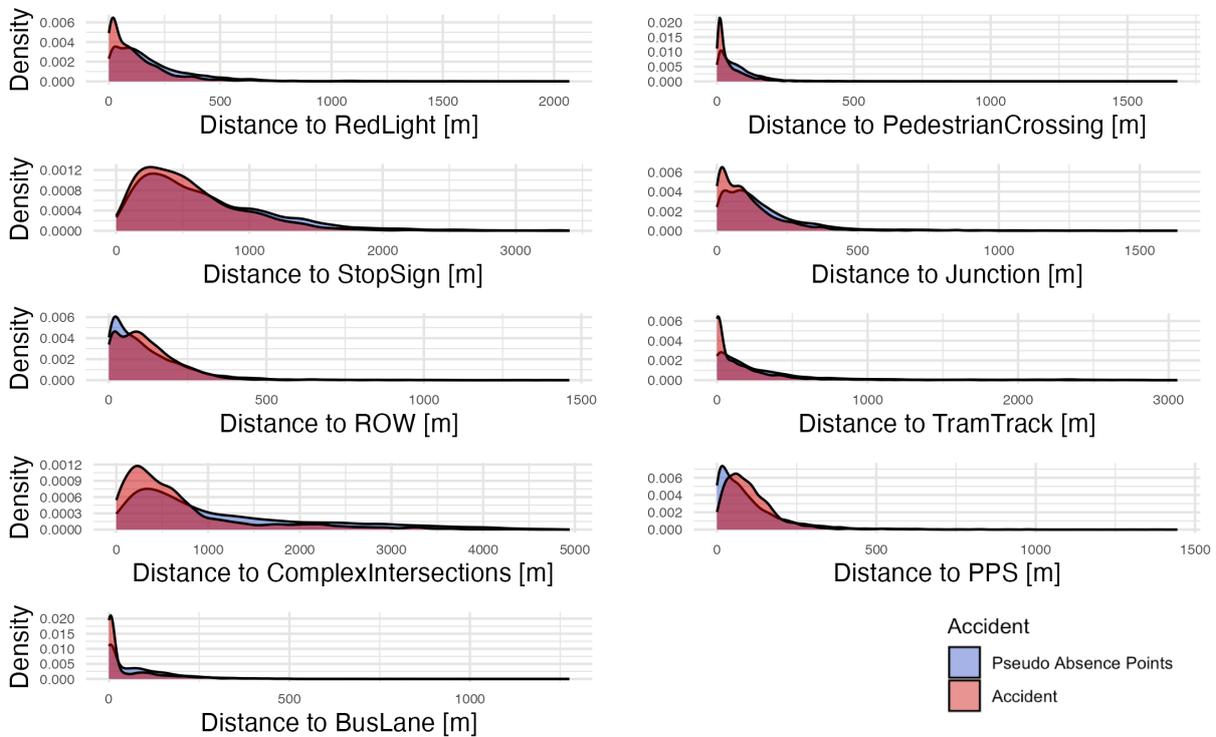


Figure 5.6: The density distribution of distances between accidents (red) and pseudo accidents (blue) concerning the nearest network or topology feature in meters.

Table 5.3 and Figure 5.7 show the distributions of the accidents from the analysis described above per network and topology feature. Most accidents (3307) occur at 30 and 50km/h speeds—also, 547 accidents and pedestrian zones. Regarding road width, most accidents (1290) occur on 10m roads and 1830 accidents on 4 and 6m roads. The most frequent accidents occur when the parking places are not on the road (3153).

Table 5.3: Bicycle accident counts by network and topology feature.

(a) Speed limit		(b) Road width		(c) Public parking spaces (PPS): <i>0 = off-road, 1 = on-road</i>	
Speed Limit [km/h]	Count	Road width [m]	Count	PPS	Count
0	547	1	42	0	3153
20	63	2	168	1	825
30	1238	3	253		
50	2069	4	930		
60	41	6	900		
80	16	8	395		
100	4	10	1290		

An equally skewed distribution applies to accidents by traffic volume (Figure 5.7). Most accidents occur with an annual traffic volume of less than 20,000 vehicles, with a strongly decreasing trend curve.

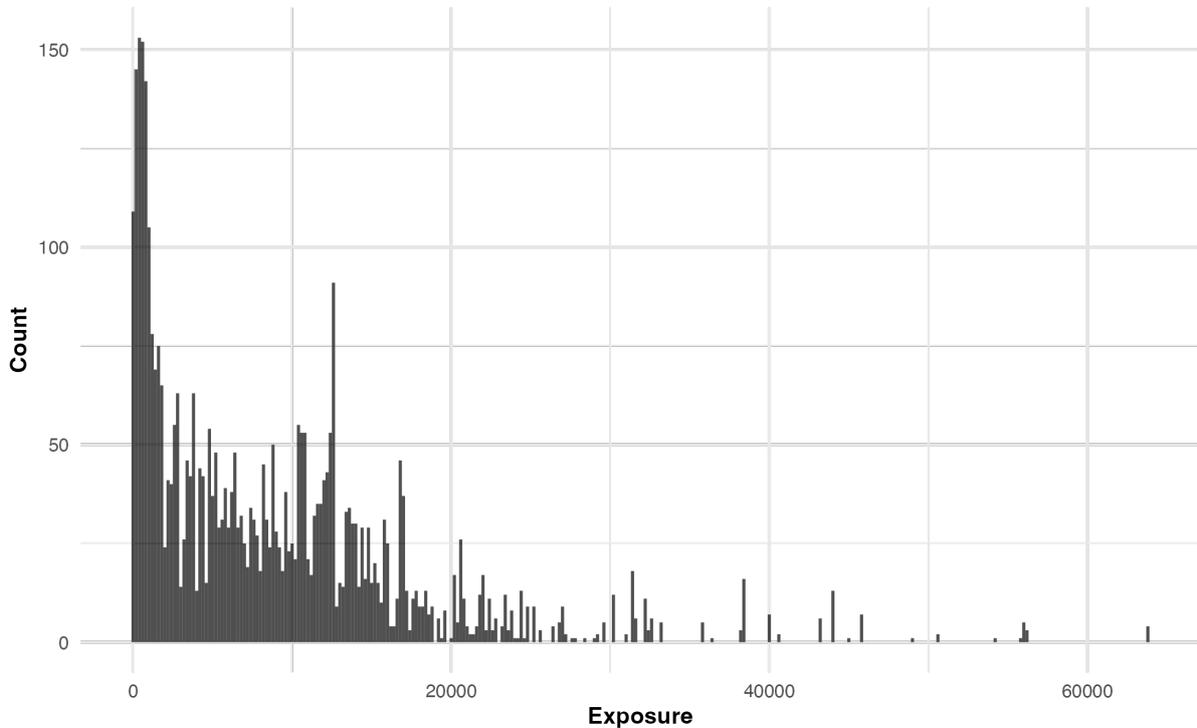


Figure 5.7: Number of bicycle accidents (dark grey) in thousands by motorized vehicle exposure, derived from GVM-ZH.

**Linear Mixed Effects Model Results** Table 5.4 lists the fixed and random effects of the linear mixed model concerning network infrastructure and topology. The most significant variables are the COVID pandemic ( $t$ -value: 14.887), BicycleNetwork type ( $t$ -value: -20.094), TrafficArea ( $t$ -value: 13.718), and ComplexIntersections ( $t$ -value: -7.116). Tiny to non-significant were SpeedLimit ( $t$ -value: 1.066), Exposure ( $t$ -value: 1.316), and most weakly, ROW ( $t$ -value: 0.237). The random effects model shows significant variation for different grouping factors. The Exposure had a variance of 0.00756 ( $sd = 0.08694$ ), while the SpeedLimit was 0.03312 ( $sd = 0.18199$ ). These values show the variability in accident risk across different exposure levels and speed limits. A substantial negative correlation existed between the speed limit and the intercept (-0.79). The neighborhood grouping factor showed a variance of 0.00593 ( $sd = 0.07699$ ), suggesting that accident risk differs across neigh-

borhoods. The variance was also observed for road width, while a perfect negative correlation between road width and the intercept indicated that wider roads are associated with lower accident risk. The residual variance was 0.15859, indicating unexplained variability in accident risk.

Table 5.4: Linear mixed model results.

Predictor	Fixed Effects		Random Effects	
	Estimate	t- value	Variance	Std.Dev.
(Intercept)	-0.04101	-1.196		
PedestrianCrossing	-0.45334	-4.464		
StopSign	0.11158	2.434		
TramTrack	-0.14170	-2.410		
ComplexIntersection	-0.34341	-7.116		
RoadWidth	0.13396	5.990		
Exposure	0.05643	1.316		
TrafficArea	0.40336	13.718		
ROW	0.01491	0.237		
COVID	0.12795	14.887		
PPS	-0.00180	-0.184		
BicycleNetwork	-0.25032	-20.094		
BusLane	-0.08248	-1.385		
SpeedLimit	0.03468	1.066		
Groups: exposure			0.00756	0.08694
SpeedLimit			0.03312	0.18199
Correlation: SpeedLimit, Intercept				-0.79
Groups: Neighborhood			0.00593	0.07699
Groups: SpeedLimit			0.00107	0.03274
road-width			0.00087	0.02952
Correlation: RoadWidth, Intercept				-1.00
Residual			0.15859	0.39824

The resulting most significant interaction terms from the linear mixed model are listed below:

1. ROW and PedestrianCrossing showed a strong interaction, with a coefficient of -0.0016.
2. The interaction between BusLane and PedestrianCrossing was substantial, showing a coefficient of -0.0023.
3. TramTrack presented a significant self-interaction, with a coefficient of 0.0035.
4. BusLane also demonstrated a strong self-interaction, with a coefficient of 0.0035.
5. The interaction between ComplexIntersection and the RoadWidth was notable, indicated by a coefficient of 0.000005.

The MSE of the linear mixed model fit from the LMEM (Table 5.4), after cross-validation with a 90:10 % split, is 0.0003.

**Variable Effect Size** The effect sizes of three LMEMs will be reported, including a comprehensive model and two distinctions for self-caused and externally caused accidents. The estimates demonstrate the change in the response variable (accidents) for a one-unit change in the predictor, and the relationship direction is indicated by the sign (+/-). For example, if the distance to the nearest pedestrian crossing (PedestrianCrossing) is increased by one unit, the probability of an accident would decrease by 0.45334 units. As accidents are binary (i.e., present or absent), this would suggest that the risk is highest within 2 m of a pedestrian crossing.

The results of the Comprehensive Model depicted in Figure 5.8 elucidate the significance of several infrastructure variables in their relationship with accident probability ( $p < 0.001$ ). The distance to the nearest pedestrian crossing (PedestrianCrossing) exhibits a significant inverse correlation with accident probability (-0.45). This suggests a decrease in the likelihood of accidents with increasing distance to a pedestrian crossing, assuming other factors are constant. Furthermore, the type of bicycle infrastructure (BicycleNetwork) profoundly impacts the accident probability, as suggested by an estimate of -0.25. This implies a reduced likelihood of accidents in areas with superior bicycle infrastructure, such as bicycle paths. The distance to complex intersections (ComplexIntersection) also displays a substantial inverse relationship with accident probability (-0.34), signifying a lower likelihood of an accident with increasing distance from a complex intersection (ComplexIntersection), provided other factors remain equal.

In contrast, traffic zones (Traffic area) positively correlate with accident probability (0.40), indicating an increased likelihood of accidents in these areas. Road width is another significant factor, with a positive correlation of 0.13, indicating a higher accident probability associated with wider roads. Interestingly, the variable representing the COVID-19 pandemic situation shows a significant positive association (0.13) with accident probability, implying increased accidents during the pandemic, considering other factors are controlled. Additionally, the effect of proximity to tram routes (TramTrack) is crucial, with an estimate of -0.14. Nevertheless, it is less significant than other predictors, signifying a decreased likelihood of accidents with an increased distance to the nearest tram route. The remaining variables in the model do not show significant relationships with accident probability.

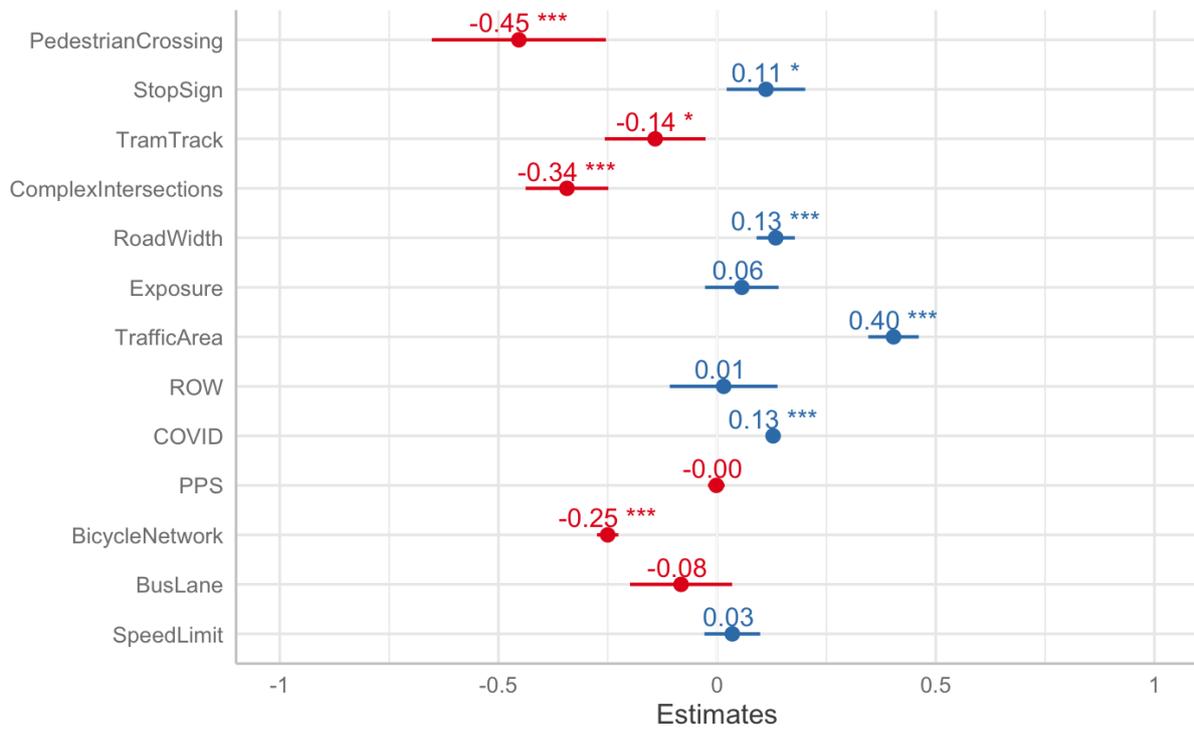


Figure 5.8: Estimates of the variable effects from the Comprehensive Model on the occurrence of a bicycle accident. Significance codes: \*  $p < 0.05$ , \*\*  $p < 0.01$ , and \*\*\*  $p < 0.001$ .

As for the Self-Inflicted Accidents (SIA) Model depicted in Figure 5.9, the type of bicycle infrastructure (BicycleNetwork) is the most impactful variable, with an estimate of -0.22. This implies that bicycle lanes/paths contribute to a substantial reduction in self-caused bicycle accidents. The traffic area (0.20) also significantly influences accident occurrence. Similarly, the distance to complex intersections (ComplexIntersection) presents a notable inverse relationship with accident probability (-0.15), indicating that increasing distance from such intersections decreases the likelihood of self-caused accidents. Significantly, the COVID-19 pandemic also demonstrated a meaningful positive relationship (0.09) with the occurrence of accidents. This suggests an increase in self-caused accidents during the pandemic. These four variables exhibit significance levels of  $p < 0.001$ . All the other variables are insignificant and can be neglected in the presence of the previously mentioned features.

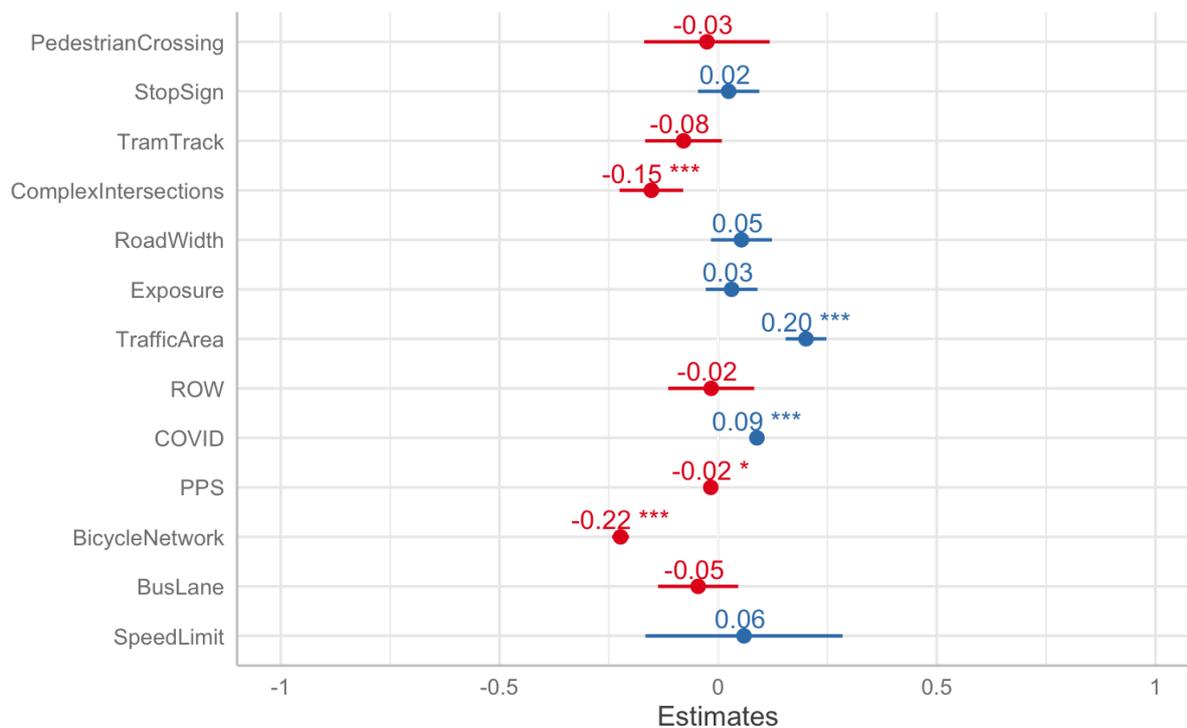


Figure 5.9: Estimates of the variable effects from the Self-Inflicted Accidents Model on a bicycle accident. Significance codes: \*  $p < 0.05$ , \*\*  $p < 0.01$ , and \*\*\*  $p < 0.001$ .

As presented in Figure 5.10, the dynamics of accidents caused by external factors align closely with the overall accident landscape. The most substantial effect size is the distance to the closest pedestrian crossing (PedestrianCrossing) (-0.51). This indicates that the incidence of externally-caused accidents decreases as the distance to a pedestrian crossing increases. This is followed closely by the distance to complex intersections (-0.27), denoting reduced accident probability with increasing distance from these intersections. The type of traffic zones also exerts a significant effect, presenting a positive estimate (0.36). This suggests that accident likelihood increases with higher speeds. Equally crucial is the presence and type of bicycle infrastructure, which exhibits a negative estimate (-0.21), reinforcing the interpretation that advanced bicycle infrastructure contributes to fewer accidents caused by external factors. Road width and the COVID-19 pandemic situation (COVID) also exhibit significant influences but relatively less intense effects, showing estimates of 0.10, respectively, suggesting that wider roads and the pandemic are associated with a slight increase in externally caused accidents. The above variables show high significance levels with  $p$ -values  $< 0.001$ . The remaining variables in the model reveal a less considerable impact on accident occurrence.

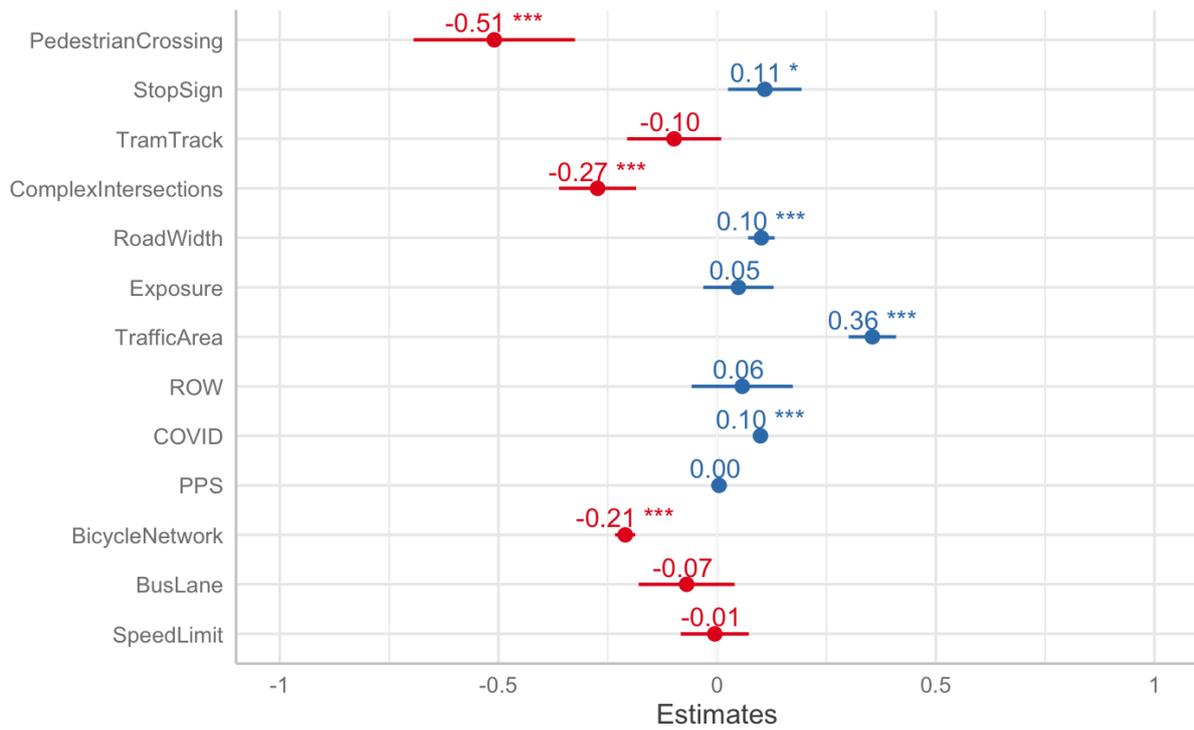


Figure 5.10: Estimates of the variable effects from the External Causes Accidents Model on a bicycle accident. Significance codes: \*  $p < 0.05$ , \*\*  $p < 0.01$ , and \*\*\*  $p < 0.001$ .

### 5.2.2.2 Accident Location Predictions

The gradient boosting machine (GBM) model for accident location prediction, as outlined in Table 5.5, yielded a substantial overall accuracy of 0.8123. The model's Kappa statistic was 0.4805, suggesting a reasonable level of agreement. With a sensitivity of 0.9325 and a detection rate of 0.6733, the GBM model demonstrated a high ability to identify actual accident locations correctly.

Table 5.5: Performance of the gradient boosting machine (GBM) model for accident location prediction.

		Prediction		Accuracy	0.8123
		0	1	Kappa	0.4805
Reference	0	843	174	Sensitivity	0.9325
	1	61	174	Detection Rate	0.6733

The random forest (RF) model results, detailed in Table 5.6, were similar to those for the GBM. The overall accuracy in predicting accident locations was slightly higher, at 0.8147. The Kappa statistic was 0.4984, indicating a moderately high level of agreement. The model also showed a high sensitivity of 0.9133 and a detection rate of 0.6645, underscoring its effectiveness in identifying accurate accident locations.

Table 5.6: Performance of the random forest (RF) model for accident location prediction.

		Prediction		<b>Accuracy</b>	0.8147
		0	1	<b>Kappa</b>	0.4984
Reference	0	832	832	<b>Sensitivity</b>	0.9133
	1	79	188	<b>Detection Rate</b>	0.6645

In contrast, both accuracy measures for the GLM model (Table 5.7) were lower than the corresponding values for the GBM and RF, respectively. The overall accuracy was only 0.7532, while the Kappa statistic was 0.2512, only about half the value of those achieved by the other two models, indicating a markedly less satisfactory level of agreement. However, the model exhibited a high sensitivity of 0.9265, and its detection rate of 0.6741 was slightly better than those of the other models.

Table 5.7: Performance of the GLM model for accident location prediction.

		Prediction		<b>Accuracy</b>	0.7532
		0	1	<b>Kappa</b>	0.2512
Reference	0	844	242	<b>Sensitivity</b>	0.9265
	1	67	99	<b>Detection Rate</b>	0.6741

As shown in Figure 5.11, there are significant differences in the spatial distribution of the modeled points. For example, the GLM forms only four critical prediction zones on the 20m network point grid. At the same time, the GBM and RF predict the accidents more widely distributedly over the urban area. The GBM is also more clustered than the RF.

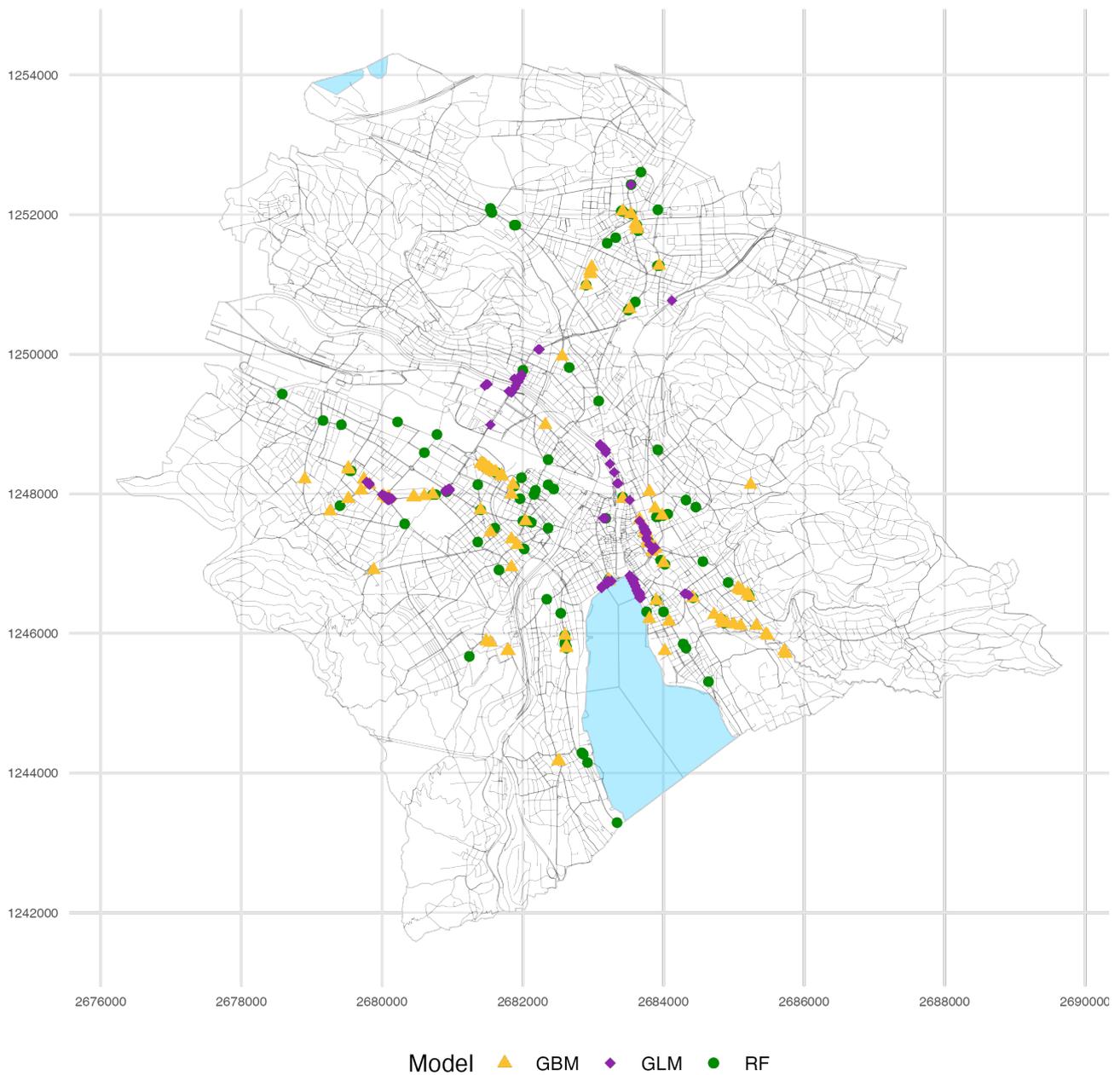


Figure 5.11: Predicted accident locations on the 20x20m network point grid using the three different models: RF (dark green), GBM (yellow), GLM (purple)

Figure 5.12 shows a similar pattern on the 50m network points grid. Thus, the GLM is very clustered, while the GBM and RM cover a larger prediction area. Similarly, all three models omit large areas where no accidents should occur, and the accident hotspots are more concentrated in the city center.

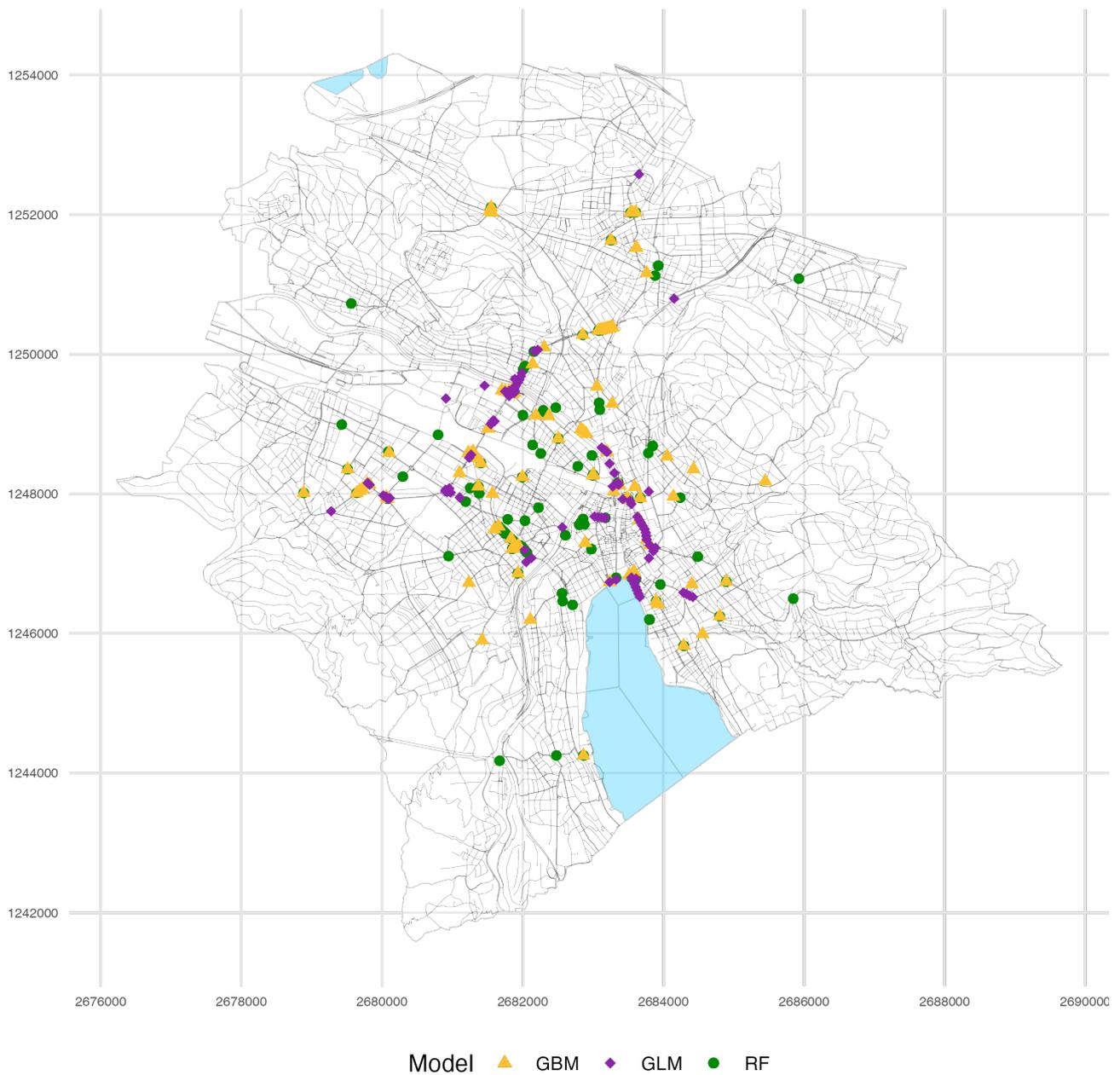


Figure 5.12: Predicted accident locations on the 20x20m network point grid using the three different models: RF (dark green), GBM (yellow), GLM (purple)

**Mean distance calculations** The average distance between a modeled accident point and the actual accidents on a 50 m network point grid is 2677.653 m, and on a 20 m grid, 2798.429 m, as depicted in Table 5.8.

Table 5.8: Mean Euclidean distances [m] between actual and predicted accident locations.

Network Grid Distance	Average Distance to Actual Accidents
50 Meters	2677.653 m
20 Meters	2798.429 m

**Mantel test** Spatial correlations between accident locations and point grids at 50 m and 20 m resolutions were assessed using the Mantel test. The 50 m grid showed a weak but significant correlation

with the accident locations ( $r = 0.0333, p = 0.019$ ), while the 20 m grid did not yield a significant correlation ( $r = 0.01205, p = 0.2$ ). These results suggest that the spatial pattern and resolution of the 50 m grid align with the accident locations, whereas the 20 m grid does not show a similar spatial correspondence.

**Location Prediction** The prediction of the accident points on the 20 m network point grid (Figure 5.13) reveals eight major point clusters, one in the area around *Bellevue* (A), around *Kunsthhaus/University District* (B), on the road from *Hubertus* towards *Altstetten* (C), along *Rosengartenstrasse* (D), around *Hardplatz* (E) and from *Central* towards *Beckenhof* (F) as well as in *Hottingen* (G) and *Oerlikon* (H). These predicted accident points are in spatial contrast to the actual bicycle accidents, distributed in the density map around the lake, the region around *Wiedikon Bahnhof* (I) and *Central*, and the *Langstrasse* (J).

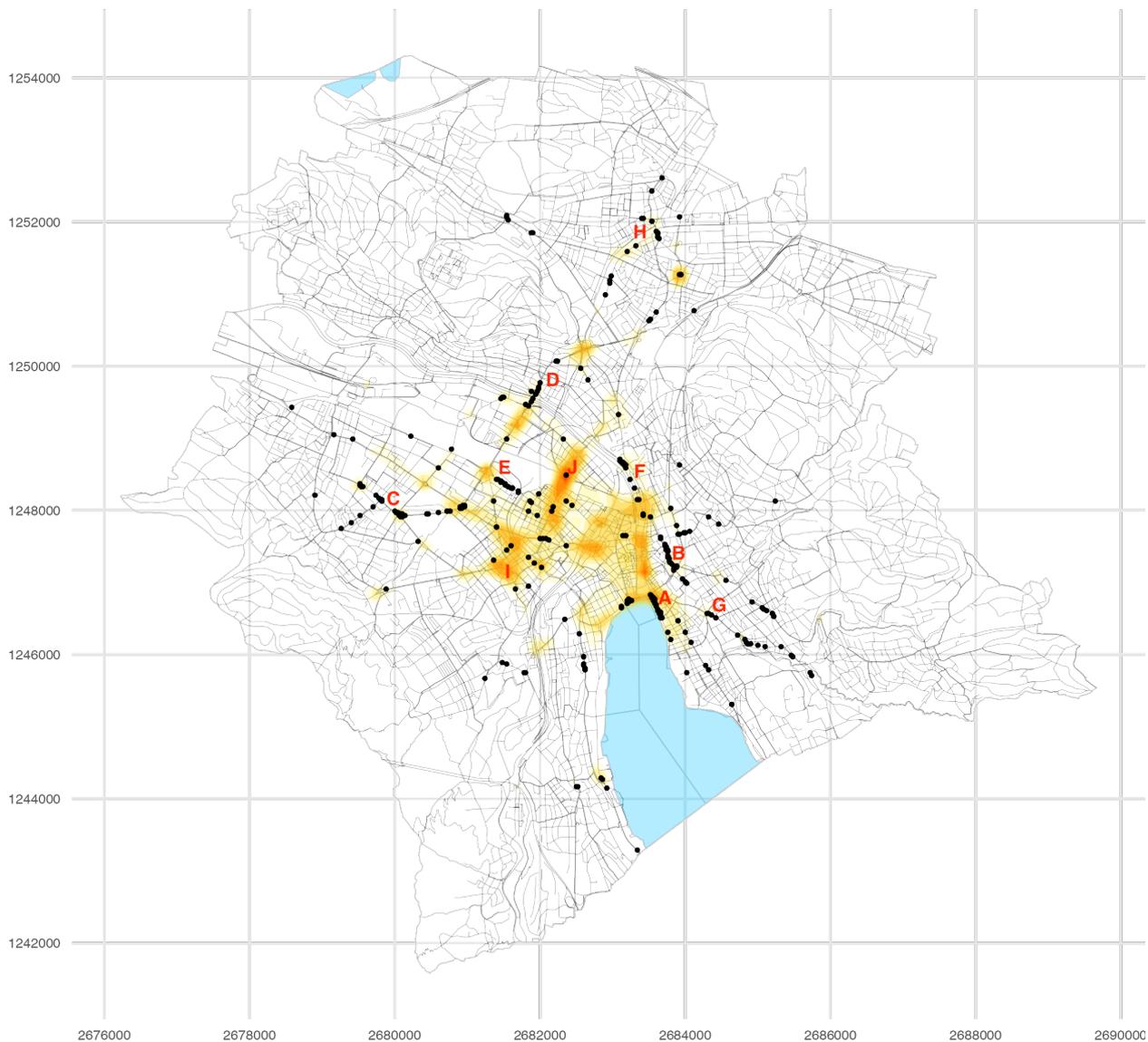


Figure 5.13: Predicted accident locations for bicycle accidents based on variables representing the road network infrastructure and topology (black dots). The point resolution on the network is 20 m. In the background, a density map of the actual reported accidents is shown (dark orange: high density, yellow: low density).

A closer look at the section in Figure 5.14 shows that the accident locations of the forecast and the actual points differ considerably. The predictions form 2 clusters, one around the *Bellevue* (A) and one in the *University District* (B). At the same time, the actual accidents are distributed more over the entire old town with a hotspot at the height of the *Helmhaus* (C).

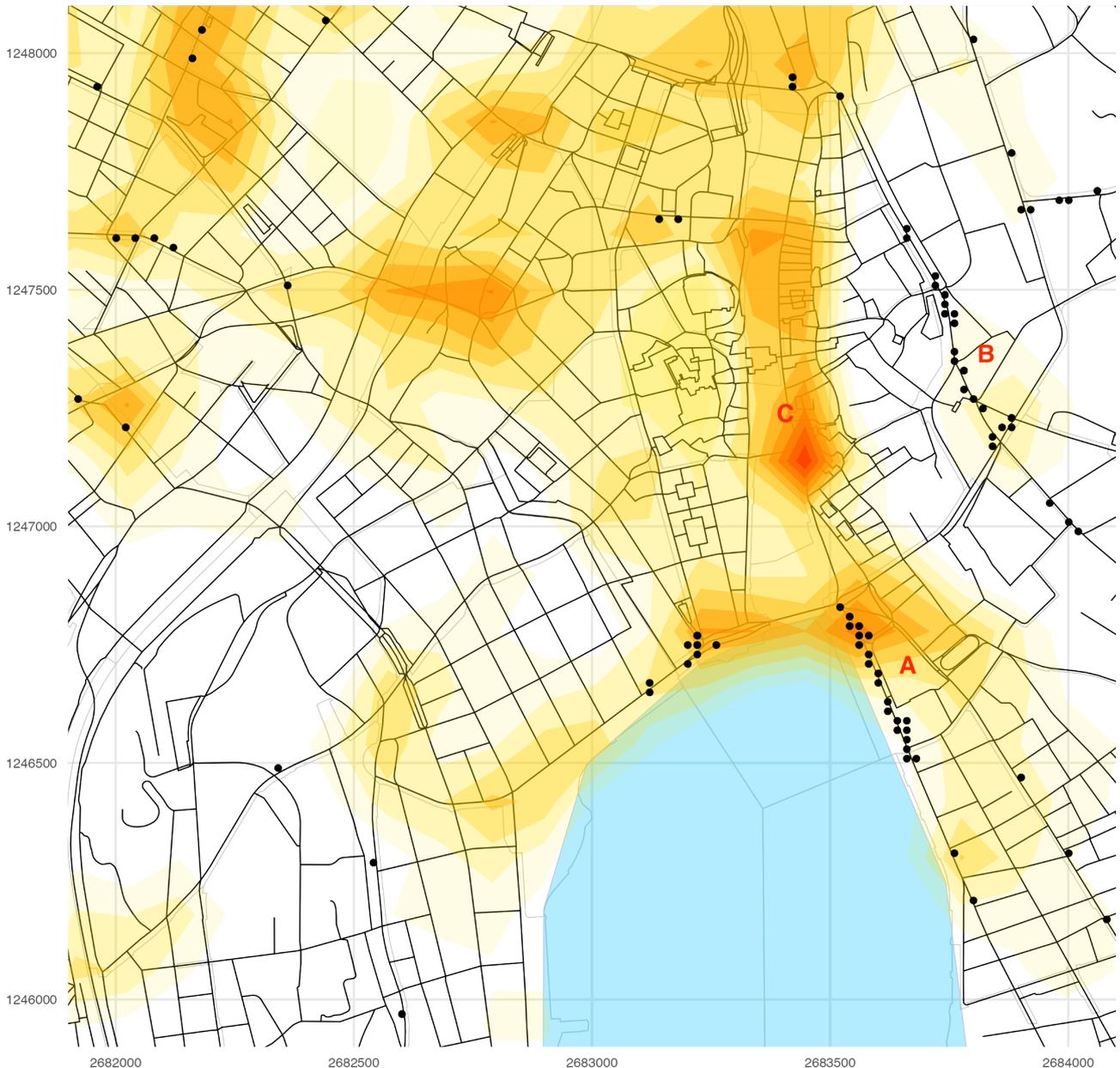


Figure 5.14: Predicted accident locations for bicycle accidents based on variables representing the road network infrastructure and topology (black dots). The point resolution on the network is 20 m. In the background, a density map of the actual reported accidents is shown (dark orange: high density, yellow: low density).

A similar prediction pattern can be seen for the predicted accident points on the 50 m grid points (Figure 5.15), which exhibit a high clustering at several locations in Zurich but are more concentrated than on the 20 m grid. The *Lake Basin* around *Bellevue* (A), the *University District* (B), *Rosengartenstrasse* (C), *Wiedikon Station* (D), *Albisriederplatz* (E), and *Hubertus* (F) are potential accident blackspots. The area around the *main station* (G) towards the *Central*, as well as the *Milchbuck* (H), also shows a clear accident black spot. Some other points are found in *Oerlikon* (I) and the *City Center*. The

predicted accident points, such as the along the *Lake Basin*, *Wiedikon Station*, and *Rosengartenstrasse*, are consistent with previous accidents. At the same time, the remaining cases show completely new accident locations that were not black spots in the reported data.

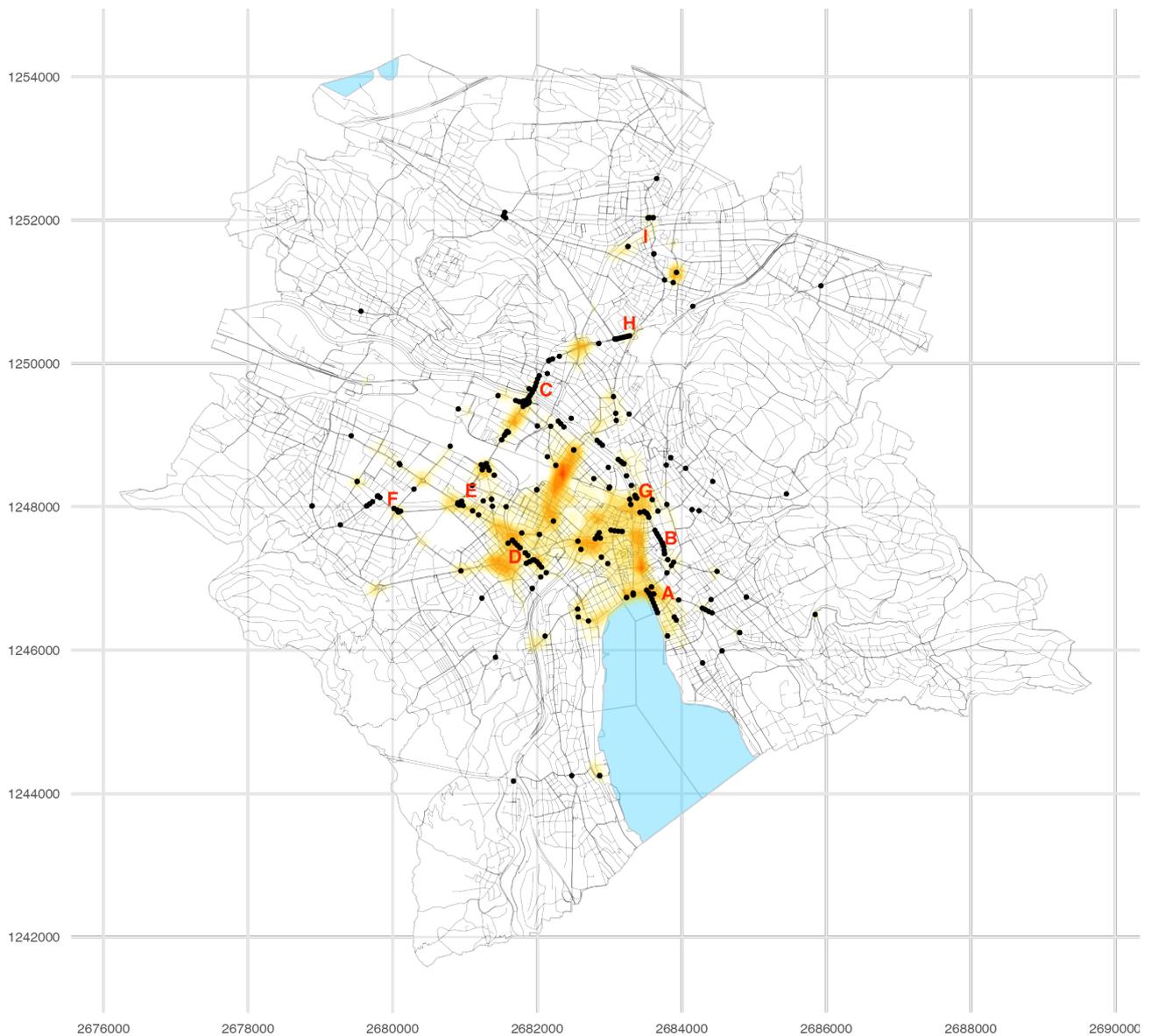


Figure 5.15: Predicted accident locations for bicycle accidents based on variables representing the road network infrastructure and topology (black dots). The point resolution on the network is 50 m. In the background, a density map of the actual reported accidents is shown (dark orange: high density, yellow: low density).

Figure 5.16 shows that the predicted accident points are very much concentrated at *Bellvue* (A) and the *University District* (B), while the actual accidents are very much distributed spatially with a hotspot at *Helmhaus* (C).

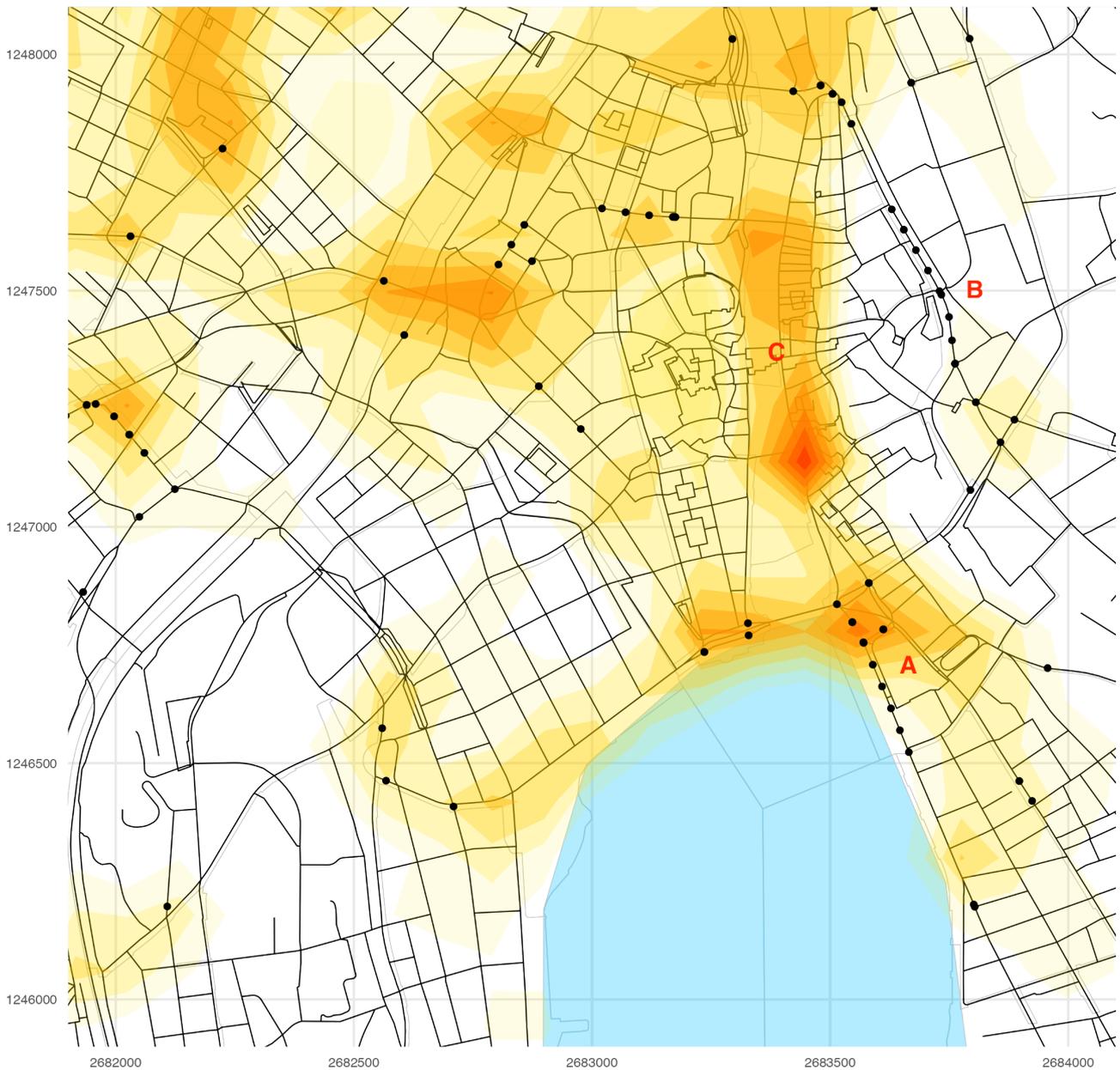


Figure 5.16: Predicted accident locations for bicycle accidents based on variables representing the road network infrastructure and topology (black dots). The point resolution on the network is 50 m. In the background, a density map of the actual reported accidents is shown (dark orange: high density, yellow: low density).

### 5.2.3 Accident Severity Analysis

While the preceding Section 5.2.2.1 presented the results of the prediction of the presence and locations of bicycle accidents in relation to road infrastructure and the road network, this section is concerned with analyzing and predicting *severity* of bicycle accidents based on the same infrastructure-related predictor variables.

### 5.2.3.1 Infrastructure Effect

**Multinomial Log-Linear Model (Multinom)** The multinomial effect diagrams show the different variables' influence on the probability of bicycle accidents, with the resulting severity as the response variable. For clarity and comprehensibility, the effects of the variables are shown only in the following diagrams, and the results of the multinomial model are not explained in tabular form. Thus, the effect graphs are taken from the output of the multinomial model and therefore reflect its results and significance.

The vehicles' influence shows considerable differences between the causing and affected vehicles. For the vehicles causing the accident (VehicleType1), the severe accident risk is highest for trucks followed by buses and lowest for buses (Fig. 5.17). The chance of a fatal accident is most significant by trams. The chance of a minor injury is highest when a bus causes an accident. Among the vehicles affected (VehicleType2), the accident severity is again exceptionally high for trams, but the chance of no injury is most elevated in combination with a MIV or truck. The truck is also found not responsible for a severe accident (Fig. 5.18).

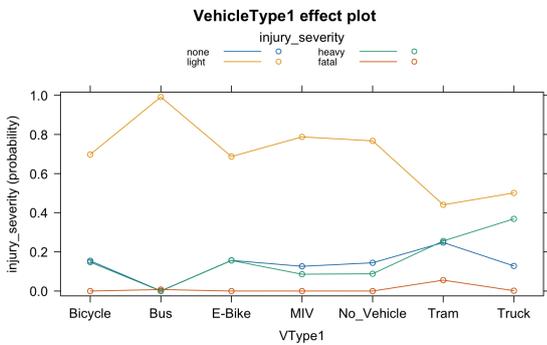


Figure 5.17: Effect probability per vehicle type causing the accident (VehicleType1), by severity.

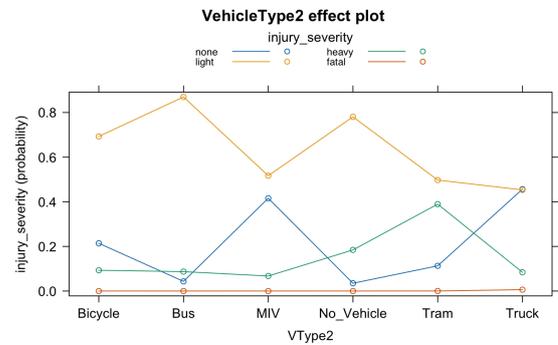


Figure 5.18: Effect probability per vehicle type affected by the accident (VehicleTYpe2), by severity.

For the right-of-way, the smaller the proximity to an infrastructure element, the more serious the accident, decreasing over distance. Surprisingly, the probability of a minor accident increases simultaneously (Fig. 5.19). Where for the road width no significant changes in the weight can be seen and the weights remain constant over all road widths (Fig. 5.20).

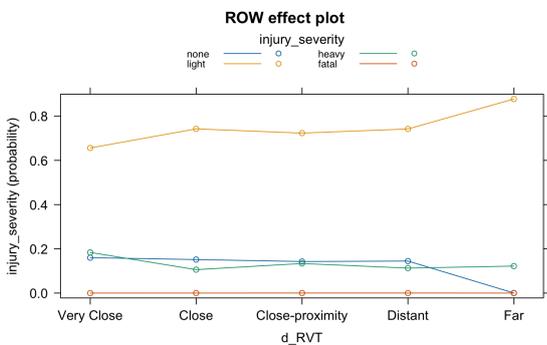


Figure 5.19: Effect probability per distance class from right-of-way (ROW).

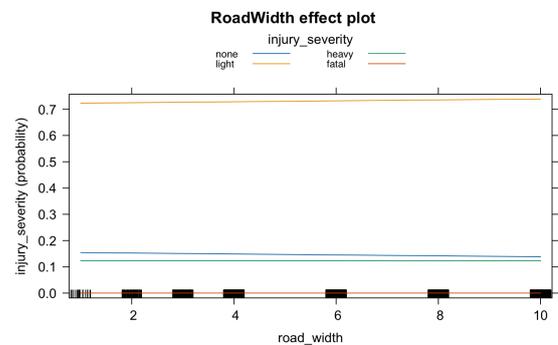


Figure 5.20: Effect probability per road width [m] (RoadWidth).

The age of those involved in accidents shows that the severity of accidents increases at older ages and is usually less severe for young people. For example, the risk of a serious accident is lowest for teenagers

in all age groups (Fig. 5.21). Similar patterns are also seen for the age of the individual causing the accident (Fig. 5.22).

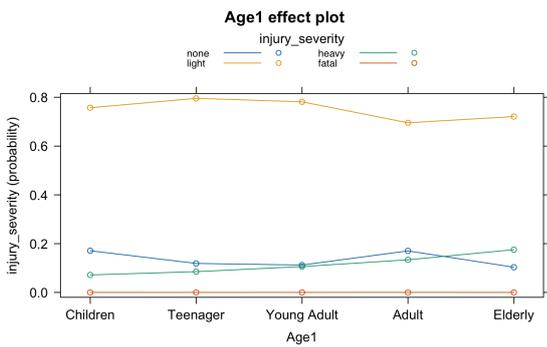


Figure 5.21: Effect probability per age group causing the accident (Age1), by severity.

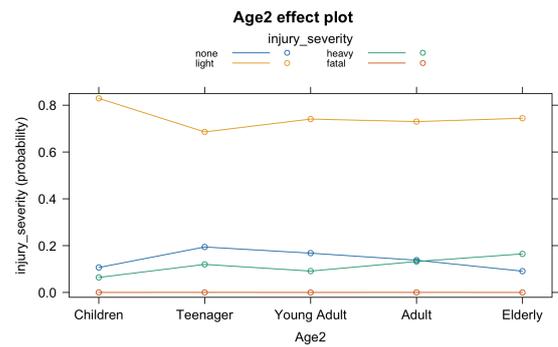


Figure 5.22: Effect probability per age group involved in the accident (Age2), by severity.

The bicycle lane/path infrastructure shows that the better the infrastructure is, the smaller the accident severity. Hence, accidents on separate bicycle lanes tend to result in minor injuries rather than on the road (Fig. 5.23). The distance to tram tracks shows no clear pattern in severity, with accident severity increasing slightly the further away the tram tracks are (Fig. 5.24).

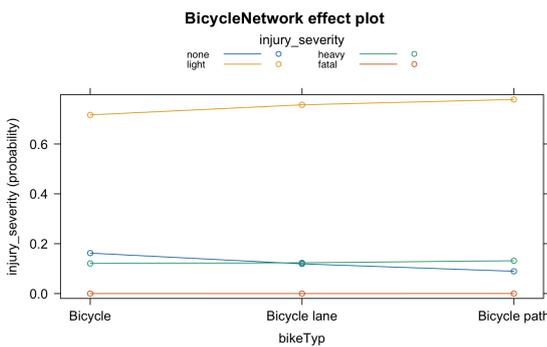


Figure 5.23: Effect probability per bicycle infrastructure types (BicycleNetwork)

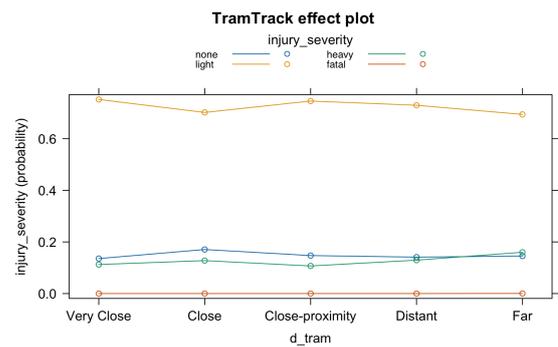


Figure 5.24: Effect probability per distance class from tram tracks (TramTrack).

For crosswalks, there is little difference in the resulting severity as a function of distance. Thus, here, the severity is similarly distributed for all distance classes, and the most common severity is minor accidents (Fig. 5.25). For bus lanes, on the other hand, the accident severity decreases steadily with increasing distance from the bus lane, and the accident most frequently ends with a minor accident (Fig. 5.26).

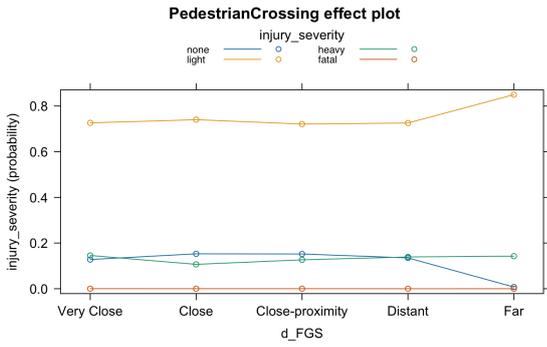


Figure 5.25: Effect probability per distance class from pedestrian crossings (Pedestrian-Crossing).

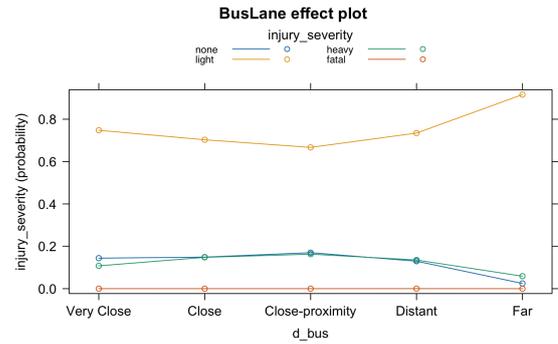


Figure 5.26: Effect probability per distance class from bus lanes (BusLane).

However, it is clear that the severity of accidents increases with increasing speed and that the lowest severity of accidents is most likely to occur at low speeds (20-30km/h) (Fig. 5.27). For junctions, the effect probability does not change much over distance. The only change is that the further the distance to a junction, the higher the chance of not being injured, but the fatal and heavy injuries still remain very small (Fig. 5.28).

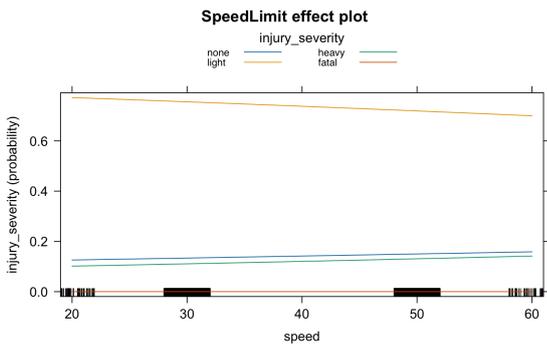


Figure 5.27: Effect probability per speed limit class (SpeedLimit).

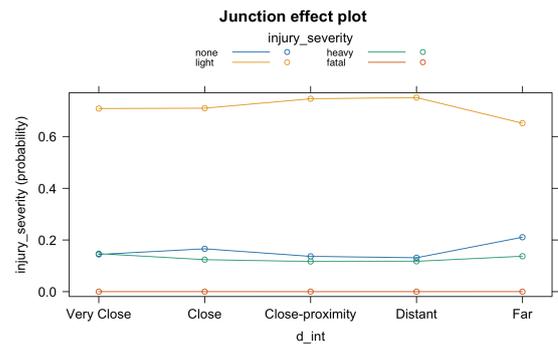


Figure 5.28: Effect probability per distance class from junctions (Junction).

**Random Forest Classification** Table 5.9 summarizes different performance measures for the random forest model used to predict the severity classes of bicycle accidents. The overall accuracy of this model was 0.6519, with a 95 % CI ranging from 0.5986 to 0.7026. This suggests a relatively good model performance. However, the more conservative Kappa statistic was only 0.1656, indicating minimal agreement beyond what would be expected by chance alone. Regarding category-specific predictions, the model correctly predicted 23 instances of no accidents and 198 lightly severe accidents. However, it failed to predict any instance of heavily severe and deadly accidents.

Table 5.9: Performance of random forest (RF) model for bicycle accident severity prediction.

Reference	Prediction				Sensitivity	Specificity
	None	Light	Heavy	Fatal		
None	27	15	5	0	0.38028	0.92537
Light	44	204	41	1	0.9273	0.2773
Heavy	0	0	0	0	0.0000	1.0000
Fatal	0	1	1	0	0.0000	0.99408
<b>Accuracy</b>	0.6814					
<b>Kappa</b>	0.2338					

As shown in Figure 5.29, the mean decrease accuracy for the relative importance of the predictor variables used in modeling accident severity yielded a value of 57.56 for the affected vehicle (VehicleType2), making it the most critical variable, followed by VehicleType1 (48.88), RoadWidth (23.91), ROW (19.60), Age 2(14.89), Age1(13.42), TramTrack (12.46) and the SpeedLimit(12.42). The least essential variable is the TrafficArea(0.38).

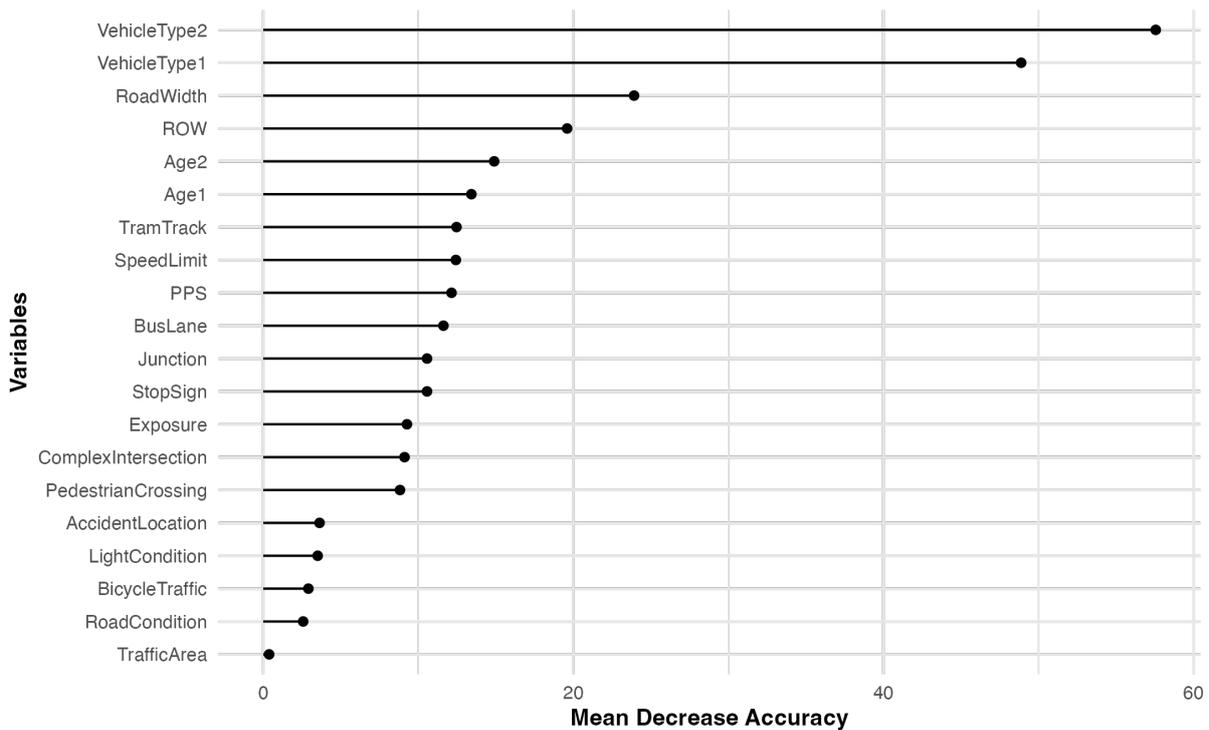


Figure 5.29: The mean decrease accuracy of a predictor variable on the prediction of injury severity classes of an accident.

### 5.2.3.2 Predictive Modelling of Accident Severity

The following tables list the ten most likely combinations of variables to result in fatal, heavy, and light crashes, respectively. Key factors such as vehicle type, age group, distance to various infrastructures, type of bicycle lane/path, speed, and accident likelihood are presented below using a series of abbreviations for clarity in Tables 5.10 to 5.12. These abbreviations are outlined below:

- VT1, VT2: Vehicle types (Tram: Tram, Bic: bicycle, NV: no second vehicle)
- A1, A2: Age groups (YA: young adult, Ad: adult, El: elderly, Tn: teenager)
- PPS: OffRoad (Parking space beside the road)
- tram, ROW, bus: distances to tram tracks, right-of-ways, and bus lanes (VCl: very close, Cl: close, Cp: close-proximity, Dst: distant, Far: far)
- RoadW: Road width [m]

The tables highlight the conditions most likely contributing to different accident severity levels.

Table 5.10 shows that collisions resulting in fatal accidents are predominantly between bicycles (Bic) and trams. Most of these interactions are consistent across most rows, especially concerning age groups. Adults (Ad) and Young Adults (YA) were frequently involved. Road widths of 3m and 8m appear regularly in the observations, suggesting narrow and relatively wider roads as significant contributors. The relative proximity of the tram, categorized as Very Close (VCl), emerged as a consistent pattern. The speed limit was uniformly recognized as 30 km/h, while all public parking spaces were beside the road (OffRoad). The "fatal" value consistently registers at 0.14, indicating a weak but stable relationship across the observed parameters. Therefore, based on the model, there is a 14% (or 0.14) probability of a fatal injury for the mentioned combination of variables.

The findings from table 5.11 on variable combinations leading to heavy accidents. A recurring road width of 6m was observed, and distances to trams were most frequently categorized as Distant (Dst). The age group, predominantly Adults (Ad), features in almost every row, with occasional appearances of Teenagers (Tn) and the Elderly (El). Speed limits vary, with 60 km/h being the most frequent, followed by 50 km/h and 20 km/h. The heavy accident value fluctuates slightly around a mean of 0.64, reflecting a moderate correlation with the listed variables, which means that based on the model, there is, therefore, a 64% (or 0.64) probability of a heavy injury for the mentioned combination of variables.

Table 5.12 shows relationships leading to light accidents. A more comprehensive array of age groups is present, with Adults (Ad) and Young Adults (YA) still dominant. However, including Teenagers (Tn) suggests a more varied demographic in these less severe incidents. The road width often clocks at 10m, with a few observations at 4m. Interestingly, there is a notable lack of trams in these accidents; most cases were self-inflicted bicycle (Bic) collisions. Another significant observation is the consistent value 1.00 for the "light" accident parameter, implying a strong correlation with the observed combinations. Therefore, based on the model, there is a 100% (or 1) probability of a light injury for the mentioned combination of variables.

Table 5.10: Ten most probable variable combinations leading to a fatal accident.

VT2	VT1	RoadW[m]	ROW	Age2	Age1	tram	speed	PPS	bus	prob.
Bic	Tram	3	Cl	Ad	YA	VCl	30	OffRoad	Far	0.14
Bic	Tram	8	Cl	Ad	YA	VCl	30	OffRoad	Far	0.14
Bic	Tram	3	VCl	Ad	YA	VCl	30	OffRoad	Far	0.14
Bic	Tram	8	VCl	Ad	YA	VCl	30	OffRoad	Far	0.14
Bic	Tram	3	Cl	YA	YA	VCl	30	OffRoad	Far	0.14
Bic	Tram	8	Cl	YA	YA	VCl	30	OffRoad	Far	0.14
Bic	Tram	8	VCl	YA	YA	VCl	30	OffRoad	Far	0.14
Bic	Tram	3	Cl	El	YA	VCl	30	OffRoad	Far	0.14
Bic	Tram	8	Cl	El	YA	VCl	30	OffRoad	Far	0.14
Bic	Tram	8	VCl	El	YA	VCl	30	OffRoad	Far	0.14

Table 5.11: Ten most probable variable combinations leading to a heavy accident.

VT2	VT1	RoadW[m]	ROW	Age2	Age1	tram	speed	PPS	bus	prob.
Tram	Bic	6	Dst	Tn	Ad	VCl	60	OffRoad	Far	0.65
Tram	Bic	6	Dst	Tn	Ad	VCl	50	OffRoad	Far	0.64
Tram	Bic	6	Dst	Ad	Ad	VCl	60	OffRoad	Far	0.64
Tram	Bic	6	Dst	Tn	Ad	VCl	20	OffRoad	Far	0.64
Tram	Bic	6	Dst	Ad	Ad	VCl	50	OffRoad	Far	0.64
Tram	Bic	6	Dst	Ad	Ad	VCl	20	OffRoad	Far	0.64
Tram	Bic	6	Dst	Tn	Ad	VCl	60	OffRoad	Cp	0.63
Tram	Bic	6	Dst	Tn	El	VCl	50	OffRoad	Far	0.63
Tram	Bic	6	Dst	Tn	El	VCl	60	OffRoad	Far	0.63
Tram	Bic	6	Dst	El	Ad	VCl	60	OffRoad	Far	0.63

Table 5.12: Ten most probable variable combinations leading to a light accident.

VT2	VT1	RoadW[m]	ROW	Age2	Age1	tram	speed	PPS	bus	prob.
Bic	Bic	10	Cp	Ad	YA	Far	50	OffRoad	Close	1.00
Bic	Bic	10	Cp	YA	YA	Far	50	OffRoad	Close	1.00
Bic	Bic	10	Cp	YA	Ad	Far	50	OffRoad	Close	1.00
Bic	Bic	10	Cp	Ad	Tn	Far	50	OffRoad	Close	1.00
Bic	Bic	10	Cp	YA	Tn	Far	50	OffRoad	Close	1.00
NV	Bic	4	Cl	Ad	YA	Dst	50	OffRoad	Close	1.00
NV	Bic	10	Cl	Ad	YA	Dst	50	OffRoad	Close	1.00
Bic	Bic	4	Dst	Ad	YA	Dst	50	OffRoad	Close	1.00
NV	Bic	10	Dst	Ad	YA	Dst	50	OffRoad	Close	1.00
Bic	Bic	10	Dst	Ad	YA	Dst	50	OffRoad	Close	1.00

# Chapter 6

## Discussion

### 6.1 Results

#### 6.1.1 Temporal Analysis

In the initial section of this discussion chapter, the thesis focuses on **Research Question 1**: “*What is the temporal correlation between car traffic density and bicycle accidents in Zurich, and how has this relationship been affected by the COVID-19 pandemic?*”

##### 6.1.1.1 Relationship Overview

A discussion of the time series will be conducted as a first step in assessing the impact of exposure on bicycle accidents and the impact of the pandemic on accident safety. None of the three normalized time series (Figure 5.1, 5.2) show a clear trend in one direction, in contrast to the trend analysis, which showed a strong positive trend for cycling, a moderate positive trend for bicycle accidents, and a weak negative trend for motorized vehicle traffic. However, the ratio between bicycle traffic and bicycle accidents has been declining, indicating that the increase in bicycle accidents has outpaced the growth in bicycle traffic. The thesis further shows in the Pearson correlation and the GAM the apparent positive relationship between motorized vehicle traffic and bicycle accidents and that traffic density negatively influences the number of accidents. This aligns with previous studies, which found a similar negative correlation (Li, Shi, and D. Chen, 2011; Wang and Nihan, 2004).

Furthermore, there is no time lag in the evolution of accidents concerning traffic volume. This is also in line with the consideration that the traffic volume on the previous day should not influence the probability of accidents on the following day. However, the MCPA study shows surprising patterns, with either motorized vehicle traffic and accidents or bicycle traffic always showing a turning point, but never together. Several change points showed up between 2012 and 2016 in bicycle traffic but were absent from 2016 onward. This leaves the question of what changed at that point to cause this stabilization—such as drastic changes in infrastructure or weather. On the other hand, the motorized vehicle traffic and accident groups show more similarities in their change points, especially from 2018 onwards, suggesting a possible common influence driving these changes, such as legislation or external environmental conditions. So, while it is surprising that there was no change point for motorized vehicles and accidents between 2012 and 2018, Figures 5.3 and 5.4 show that the pattern of accidents follows much more that of cycling than the one of motorized vehicle density. Nevertheless, there are similarities between motorized vehicle density and the number of accidents, which leads to the conclusion that motorized vehicle density indeed influences the number of accidents. This would be

consistent with previous studies that found a clear correlation between accident rates and motorized vehicle traffic (Miranda-Moreno, Strauss, and Morency, 2011; Li, Shi, and D. Chen, 2011).

These relationships are also supported by the GAM—showing the significant impact of cycling on accident rates, confirming the crucial role of cycling in influencing road safety. In contrast, while still playing a critical role, motorized vehicle traffic is relatively minor in influencing accident rates. Furthermore, the decisive significance of non-linear effects associated with bicycles and weeks of the year underscores the importance of these factors. The strong influence of the week suggests a pronounced seasonal variation that could be related to weather changes, holidays, or other seasonal behaviors. For example, the weekly change in the data is pronounced, which suggests that changes in the weather or holiday periods can cause the number of accidents to fall or rise markedly every week. The high explanatory power of the model, with a good 73 % of the variance, underlines its robustness. Focusing on the interplay between the presence of bicycles, the time of year, and the intertwined non-linear effects of these factors is of utmost importance for optimizing road safety measures.

The results of the ARIMA model, characterized by a slight positive autocorrelation and a low RMSE, suggest a nuanced correlation between motorized vehicle and bicycle traffic and bicycle accidents, with the model successfully reflecting an increased risk of bicycle accidents as motorized vehicle traffic increases. This supports the interplay between bicycle and motorized vehicle density and bicycle accidents, characterized by an underlying weekly seasonality trend. However, it should be determined how accident rates would develop independently of either traffic density, and a correlation between the three variables should be investigated separately.

#### 6.1.1.2 COVID-19 Impact

The GAM and the change point analysis highlight contradictory aspects regarding the impact of the pandemic on traffic safety. For example, the GAM analysis showed no significant difference in accident rates before and after the outbreak of the pandemic ( $p = 0.281$ ), which is in direct contrast to the counted accident rates of Zurich, where an increase was observed (Tiefbauamt Stadt Zürich, 2021). In Zurich, at least, there was the utterly opposite trend in the number of accidents as in all other studies around the world, where a drastic decrease in accidents took place during the COVID-19 protective measures (Kellermann et al., 2022; Monfort, Cicchino, and Patton, 2021). Many cities reported changes in travel behavior due to lockdowns, with the popularity of cycling increasing as people sought alternatives to public transport. However, this was not reflected in a measurable change in cycling accidents in Zurich. The same patterns can be seen in the time series and the change points. For example, there was a change point at the beginning of March for bicycle accidents and motorized vehicle traffic immediately after the start of the pandemic in Switzerland. This can also be seen in Figure 5.4. Bicycle traffic made a massive jump upward with the pandemic at the beginning of the first wave and then recovered towards the end of 2020, with the second wave back at the previous year's level.

Thus, there are significant differences in when a change point occurs based on temporal resolution. For example, Figure 5.4 and Table 5.1 do not directly show the exact data for the change points.

This discrepancy can be attributed to the temporal granularity of the study, which focuses on weekly intervals rather than absolute values or daily counts, which may hide variations such as different accident rates between pandemic waves. In addition, the analysis of change points in all groups after the pandemic outbreak identified new aspects that indicate possible behavioral changes in road use. However, these did not translate into a measurable difference in cycling accidents in Zurich. This complex relationship suggests that local conditions, policies, or specific behaviors have mitigated the expected increase in bicycle accidents typically associated with increased bicycle use during closures.

The analysis highlights the importance of a nuanced approach to understanding road safety trends during unprecedented global events such as the COVID-19 pandemic. Thus, the big question remains how to explain the changes in accident rates after the pandemic; there was clearly a notable change in accident rates, but it cannot be directly attributed to motorized vehicles and bicycle traffic. The question also arose that exposure to motorized vehicles is negligible within the pandemic. However, the increased number of cyclists on the roads with the lockdown has ultimately increased accident rates.

Thus, the first research question can be answered as follows: There is a positive correlation between accident rates and motorized vehicles and bicycle traffic, and increasing traffic numbers lead to an increase in accident rates. Similarly, contrary to expectations, the pandemic has led to a marked deterioration in road safety for bicyclists. It has caused the number of bicycle accidents in the first wave of the pandemic to soar due to the increased bicycle density.

## 6.1.2 Spatial Analysis

The following section of the discussion focuses on **Research Question 2**: “*How do bicycle accidents, car traffic, and the road infrastructure in Zurich correlate spatially, and how does the latter influence accident severity?*”

### 6.1.2.1 Spatial Patterns and Correlations

Regarding accident locations, the Nearest-Neighbor analysis shows that the accident data follow a clear pattern and are not simply randomly located in the network. This supports the assumption that accident locations are directly related to local conditions, such as infrastructure, and do not occur randomly. This is particularly evident in the mean nearest neighbor distance of 30 meters and a ratio of 0.417, a clear sign of accident clusters.

An equally clear pattern emerges here about the distance to the various infrastructure elements: accidents always occur in relative proximity to infrastructure. However, this is not particularly surprising, as the comparison with the pseudo-absence points shows. The distribution of the two data points is almost identical because, in Zurich, nearly every location is located in the small perimeter of an infrastructure. For this reason, a pure distance comparison is hardly suitable and does not allow any conclusions to be drawn.

The number of accidents per speed (Table 5.3) shows the expected pattern for most accidents with the highest speed limit. The road width is surprising because previous studies concluded that a large width would reduce the number of accidents (Pucher and Buehler, 2008; Wegman, Zhang, and Dijkstra, 2012). The public parking spaces (PPS) also show a robust clustering in situations where the parking space is not on the road but next to it. This raises the question of why the frequency is highest when the obstacle is not directly on the road. The distribution of accidents also shows a surprising picture: most accidents occur in the streets with the lowest traffic volume (Figure 5.7).

The linear mixed model results (as shown in Table 5.4) have several remarkable implications. Starting with the Comprehensive Model, the significant negative coefficients of PedestrianCrossing, TramTrack, ComplexIntersection, and BicycleNetwork suggest that these factors have a protective effect or correlate with decreased accident rates. For complex intersections, pedestrian crossings, and trams, the closer this infrastructure is, the higher the accident probability. Regarding the bicycle network, the study shows that separate bicycle paths and lanes substantially reduce the chance of accidents. This also aligns with other studies (Vandenbulcke, I. Thomas, and Int Panis, 2014). In contrast, StopSign, RoadWidth, TrafficArea, and COVID have positive coefficients, indicating potential risks or associations with increased accident rates. Thus, higher speed zones such as generally applicable 50 km/(h, wide

roads, and the pandemic lead to increased accidents. This is also directly in line with the study by Kim et al., who showed that pedestrian crossings and low speeds reduce accident rates (Kim et al., 2007). Of particular note is the significant  $t$ -value associated with TrafficArea and COVID, emphasizing their potentially essential roles in influencing bicycle accidents.

The interactions observed between variables further refine our understanding. The negative coefficient interactions between ROW and PedestrianCrossing and BusLane and PedestrianCrossing may indicate that the combined presence of these factors decreases the likelihood or severity of accidents even further than when considered in isolation. Similarly, the positive self-interaction of TramTrack and BusLane suggests that areas with higher densities or frequencies of these public transport vehicles might see exponential increases in accident risk. The notably small coefficient for the interaction between ComplexIntersection and RoadWidth implies that while this interaction exists, its influence might be minimal or subtle in practice.

Comparing the initial Comprehensive Model with the SIAM and ECAM models, key distinctions arise: SIAM's SpeedLimit random effect is more pronounced than ECAM's, implying that high speeds significantly negatively influence self-inflicted accidents. Both models show significant  $t$ -values for several predictors, but the magnitude and direction of influence vary. For instance, PedestrianCrossing and ComplexIntersection have a more significant negative impact on ECAM, highlighting their importance that pedestrian-involved accidents are much more frequent near crosswalks and pedestrian crossings. In contrast, BicycleNetwork is more negatively correlated in SIAM, suggesting its role in self-inflicted accidents, meaning that self-inflicted accidents occur more often on roads without bicycle infrastructure than on bicycle paths and lanes.

These differences underscore the complex relationship between predictors and accident types. The pronounced negative effect of certain factors in ECAM suggests that interventions targeting these could significantly reduce externally caused accidents. Similarly, SIAM's unique predictors provide insights into areas to focus on for reducing self-inflicted accidents. Recognizing these nuances is essential for effective safety interventions.

Comparing the gradient boosting machine (GBM), random forest (RF), and generalized linear model (GLM) for predicting accident locations, distinct differences emerge:

The GBM and RF models performed similarly, showcasing high overall accuracies, with RF slightly surpassing GBM (0.8147 vs. 0.8123). Their respective kappa statistics affirm a reasonable to moderate agreement, suggesting reliable predictions beyond chance. Conversely, the GLM model, while lagging in overall accuracy (0.7532) and displaying a lower Kappa, retains a high sensitivity and detection rate. This suggests a capability to accurately identify accident locations, albeit with potentially higher false positives. While GBM and RF offer more reliable outcomes, the GLM's strength lies in its detection capability, which stakeholders should weigh based on specific application needs.

These differences are again clearly shown in Figures 5.11 and 5.12, where the two models, GBM and RF, show a much larger dispersion of accident locations than the GLM. The summary of these three models is reflected in Figures 5.13, 5.14, 5.15 and 5.16, where it becomes clear that they can spatially classify the past accident hotspots well and assign similar trends and spatial hotspots. Finally, Why does the 50 m network point grid predict the accident locations more accurately than the finer resolution point grid? The average distance between the points of the 50 m-network points grid and the actual accidents is 250 m less than that of the 20 m grid. Thus, the average distance between the 50 m points and the actual accidents is 2677 m, and between the 20 m points is 2798 m. While infrastructure can provide a small measure of accuracy in modeling and predicting the location of bicycle accidents, more than infrastructure alone is needed to cover the entire picture. What is certain, however, is that the infrastructure has a definite influence on the distribution of accidents in space, and the hotspots can

be predicted with a relatively high degree of precision.

### 6.1.2.2 Influence on Accident Severity

The results drawn from the multinomial log-linear and random forest models provide detailed but cautious insights into the nexus between infrastructure variables and the severity of bicycle accidents. Through systematic analysis, the data illuminates several multifaceted relationships warranting further discussion as listed in Figures 5.17 - 5.28.

#### 1. Vehicular Dynamics and Accident Severity

- **Causing Vehicles (VehicleType1):** The disproportionate risk of trams and trucks, which may be related to their greater mass and potential blind spots, indicates that bicyclists are particularly vulnerable when roads must be shared with these much more extensive and heavier vehicles. This confirms the study findings of Chen et al. that vehicle size contributes significantly to crash severity (P. Chen and Shen, 2016). In contrast, buses are found to be comparatively safer vehicles. This counterintuitive result, especially given the size of buses, suggests other mitigating factors—perhaps slower speeds, routine routes, or advanced driver training.
- **Affected Vehicles (VehicleType2):** The recurrent theme of trams underscores a potential infrastructure inadequacy. Meanwhile, MIV vehicles correlate with less severe outcomes, requiring further contextual data about the nature and cause of such collisions.

#### 2. The Interplay of Proximity and Infrastructure

- **Road Width:** The road width does not show any significant changes in the severity of the accidents, which suggests that all casualties of the same type have the same seriousness regardless of the road width.
- **Right-of-Way (ROW):** The observed trend could indicate potential conflict points of mixed traffic flows. The simultaneous rise in minor accidents in close proximities suggests potential evasion maneuvers that avert severe outcomes.

#### 3. Demographic Aspects and Accident Outcomes

- Age emerges as a determinant of accident severity. The resilience of younger individuals—both physiological and perhaps cognitive—protects against severe injuries. Conversely, the increasing vulnerability of older individuals warrants tailored safety measures.

#### 4. Bicycle Infrastructure Nuances:

The results of this study corroborate the argument for robust bicycle infrastructure. Separate lanes ensure physical separation and delineate clear rights-of-way, substantially reducing conflict points, which would also directly cover the findings of (Janstrup et al., 2023).

#### 5. Tram Tracks Ambiguity:

The ambivalence of patterns related to tram infrastructure warrants further research. Factors like track conditions, cyclist behavior near tracks, or parallel vehicular movements might offer added clarity.

#### 6. Uniformity at Pedestrian Crossings:

The absence of pronounced variations across distances hints at possible design or regulatory adherence uniformity. However, the consistent prevalence of minor accidents suggests that while fatal outcomes might be averted, conflict points persist.

7. **Velocity and Impact Severity:** The correlation between speed and severity resonates with the principles of physics. Reducing speed limits, especially in areas with high bicycle traffic, emerges as an actionable recommendation.
8. **Public Transport:** Despite the tremendous benefits of public transport, buses are a source of safety due to their size and weight. Thus, the decreasing severity with distance indicates potential problems with bus lane demarcation or shared access.
9. **Junction Dynamics:** The protective factor of distance from junctions underscores the complexities at intersections: multiple vehicular paths, varied speeds, and pedestrian movements create a matrix of potential conflict.

Thus, the multinomial model results much better than the random forest model. Therefore, a kappa value of 0.23 and an overall accuracy of 0.68 provide a specific essential significance but also show that the model's overall performance was only moderate. This indicates that the accident's severity cannot be modeled with the infrastructure alone and that further parameters are needed. A collision between a motorized vehicle can be fatal or light and depends on the type of accident, the injuries, and other factors, not only on the vehicle itself. The same is true for self-inflicted accidents, which can be severe even without a significant infrastructure if the injury is serious but the cause is not.

Predicting the influence of road infrastructure on the severity of bicycle accidents offers multifaceted insights. One of the most notable infrastructure elements that have emerged is road width. As the results show, road widths of 3 m and 8 m seem to result in fatal accidents with marked frequency, while a width of 6 m is consistently found in severe accidents and 10 m in minor accidents. This suggests that narrow and relatively wide roads can contribute significantly to severe accidents - or the case numbers are too small for a clear picture. In particular, narrow streets can restrict the space for cyclists to maneuver safely, especially when trams or other large vehicles are on the road, leading to an increased risk of collision. Conversely, wider roads can lead to higher speeds or confusion without appropriate markings, increasing the risk of serious accidents. This is particularly surprising because, based on the literature, the assumption was that wide roads provide more safety (Pucher and Buehler, 2008; Wegman, Zhang, and Dijkstra, 2012).

Traffic elements such as trams also significantly influence the severity of accidents. The presence of trams often leads to fatal bicycle accidents. This could be because cyclists have a limited ability and experience to anticipate and avoid the movements of trams due to their size, relatively low frequency of occurrence, and fixed tracks. Since trams repeatedly appear in fatal accident scenarios, it is evident that in areas with active tram tracks, better infrastructure for cyclists or clearer warnings are needed to avoid such accidents. Likewise, a clear separation of bicycles and trams is more than advisable.

The results also show that the type of vehicles involved in the collision plays a decisive role. For example, collisions between bicycles and trams usually lead to fatal accidents. In contrast, light accidents are predominantly self-inflicted bicycle collisions, highlighting the difference in the severity of collisions between bicycles and heavier vehicles compared to solo accidents or accidents involving lighter vehicles.

In conclusion, the intricate web of infrastructure, vehicular dynamics, and demographics highlighted by the results underscores the multifactorial nature of bicycle accident severity. This is also underlined by the poor performance of the RF model, which clearly shows that important variables are missing from the modeling. As urban areas grapple with sustainable mobility challenges, these insights offer a roadmap for designing safer environments. Similarly, comparing the accident locations shows that the accident's severity follows similar patterns, even in different geographical areas. This is confirmed by the study of Yang et al., where age and the vehicles involved also had the most significant effect

on accident severity (Yang et al., 2021). This shows that studies in different countries can be used to inform measures to reduce the severity of accidents and increase road safety. Furthermore, anomalies and unexpected findings pave the way for more profound, focused research, emphasizing the need for a holistic, interdisciplinary approach to urban safety.

As such, the second research question can be answered as follows: There is a statistically significant correlation between traffic density and, above all, infrastructure. Individual infrastructure elements such as complex intersections, trams, wide roads, and high speeds can significantly increase the number of accidents. Concerning the severity of accidents, it is only partially possible to make a conclusive statement based on the infrastructure alone, as the age and type of vehicle influence the severity, and the infrastructure cannot conclusively explain the degree of severity.

## 6.2 Methodological Approach

### 6.2.1 Methodological Considerations: Temporal Analysis

Some limitations affecting the results and conclusions remain in the methodology used for the time series analysis. For example, while compiling the time series, the accident and traffic figures were aggregated weekly rather than daily, which led to a less detailed result. Daily counts would require including meteorological data, which significantly controls day-to-day fluctuations of bicycle traffic volume. The GAM assessment clearly showed these gaps in this thesis. Although accident rates can be inferred from traffic counts, using only two variables to make a final decision is particularly problematic. Therefore, using parameters such as weather, temperature, and additional information on the state of the infrastructure is essential to gain a holistic picture. Similarly, it is debatable whether approaches such as the change point analysis are the appropriate means to obtain a holistic view of how the pandemic has affected traffic safety and how traffic numbers have evolved. Still, the approach taken in this thesis builds a foundational core to explore trends and patterns in this area and to make initial assessments.

### 6.2.2 Model Evaluation: Spatial Analysis

The linear mixed effects model, which can capture fixed and random variations of predictor variables, offers nuanced insights into the relationship between road infrastructure and accidents. The reported Mean Squared Error (MSE) of the LMEM fit (0.0003) suggests an excellent fit of the model to the data. However, it is crucial to remember that the model is based on the available data only and lacks essential additional variables. Factors not captured in this dataset or not included in this model can still play a role in real-world scenarios. This also stresses the importance of continuous data collection and model refinement.

Using pseudo-absence points strengthens the model by addressing spatial bias. Still, it assumes that these points represent accident-free areas, which is only sometimes the practice case (due to potential under-reporting of accidents). Thus, a pseudo-absent point could also be an unreported accident point. On the other hand, using machine learning models such as GBM and RF to predict accident locations shows a multifaceted approach. While these models are highly adaptive and can deal with non-linearity, they are not immune to over-fitting, especially when dealing with complicated spatial data.

Using two network point grid sizes (20 m and 50 m) in the analysis is an innovative step that addresses the scaling issues associated with spatial analyses. Still, it also assumes that one of these specific grid sizes will be optimal for capturing the spatial patterns of accident occurrences rather than testing

a more comprehensive range of grid sizes, e.g., including a fine-grained grid corresponding to the estimated accuracy of 3-5 m of GPS devices used when reporting accident locations. Because of the increased computational cost of fine-grained grids, this option was not feasible within the scope of this thesis.

The multinomial model used for the severity analysis captures the discrete nature of accident outcomes. However, this model assumes that response categories are mutually exclusive and assumes the independence of irrelevant alternatives. This does not necessarily have to be the case, as in many cases, the infrastructures are connected. For example, a pedestrian crossing must have street lighting, which can lead to a correlation.

The random forest model used to predict the severity of bicycle accidents shows that no accidents and minor accidents are predicted accurately. Still, the model fails to predict severe and fatal accidents. This limitation raises concerns about the model's applicability in a real-world setting where predicting these more severe accidents is paramount. However, it should be noted that while the model captures the basic features of accident severity prediction, it cannot predict more accurate outcomes, indicating possible omissions or lack of specific essential data. The model may need more detailed data, a larger (and more balanced) data set, or a different algorithmic approach. Assumptions made during model development, such as the relationships between predictor variables or the distribution of response variables, may also need re-evaluation.

A certainly not negligible deficit in this thesis is the narrow range of variables. For example, this work focused on examining the impact of infrastructure. However, Spoerri et al., in a study in Switzerland, revealed the significant influence of social and demographic parameters, highlighting education and income as factors influencing road safety (Spoerri, Egger, and Elm, 2011). Thus, this thesis shows significant results about the influence of infrastructure and the change in the number of accidents related to motorized vehicle traffic. Still, there remain many variables that correlate with and influence infrastructure as well. For example, the infrastructure is also influenced by the financial resources of the municipalities and cities, the political situation, and the preferred means of transport used by the inhabitants. This is an excellent complement to this study and highlights the wide range of factors that influence accidents and weaknesses and missing approaches that need to be addressed in future research. In summary, while the models used in this study provide a robust framework for spatial analysis and offer valuable insights, it is crucial to view the results with an awareness of their inherent limitations, lack of data, and assumptions.

### 6.3 Data Availability

Data availability has been the biggest obstacle when analyzing accident figures. As Wegman and colleagues have described, not all countries collect accident data accurately, and if they do, accuracy often remains a problem (Wegman, Zhang, and Dijkstra, 2012). The situation is much different in Zurich, where the availability, accuracy, and breadth of data collected concerning an accident is substantial. For example, there is information in the accident recording protocol (UAP) on a rich set of parameters to be determined that relate to an accident, which allows a comprehensive analysis to be made. However, an essential element that the UAP does not provide is the direction of travel at the time of the accident—from this, it would be possible to investigate the gradient (i.e., steepness of a road) at which the accident took place.

Quite the opposite is true for data availability regarding the infrastructure on which the accidents occur. For example, although Zurich has a wide range of available geodata compared to other cities in Switzerland, many vital variables are still missing, or the existing data lack additional attributes. For

example, various studies have shown that the width of bicycle lanes and paths significantly influences cyclists' safety. However, this width has yet to be recorded, so recording it would be a considerable undertaking. About the missing data, there is also no data on curb height, which has been identified as a significant safety risk for bicyclists in various studies. One way to intensify data collection would be to incorporate machine/deep learning and image recognition algorithms. This way, network infrastructure elements could be recognized and extracted with image segmentation and recognition models from aerial images and 3D lidar point clouds. In addition, temporal and geospatial data on infrastructure needs to be improved, making it adequate for this type of work to examine the impact of changing infrastructure on the number of accidents over the past decade.

Not all accidents are registered by the police because they ended without injuries, or the involved parties could solve the problem bilaterally, and the police were not called. This creates a reporting bias, where minor accidents are reported far less frequently than significant accidents, leading to a bias in accident reporting. Thus, essential statistics must include data on other possible accident hotspots that must be made safer. Therefore, there are several sources of uncertainty in accident data, including errors in reporting, incomplete information, and missing data. These uncertainties can lead to inaccurate or biased analyses that significantly affect public safety initiatives and policy decisions.

**GVM-ZH** Regarding data availability, another elementary weakness in modeling temporal parameters related to bicycle accidents is the availability of traffic data. In Zurich, for example, there are only about 100 measuring points where traffic density is measured, and they are so spatially unevenly distributed that it is impossible to make robust inferences on the entire area from these measurements. The GMM-ZH addresses this problem, but the temporal resolution of 4 years makes it impossible to make up-to-date statements about the influence of traffic density and how traffic has changed during the COVID-19 pandemic. Thus, a higher temporal and spatial resolution would be essential for an accurate analysis, as is data on cycling in general.

## 6.4 Implications for Policy and Urban Design

The identified links between the presence of bicycles and motorized vehicles and accident rates, as well as the differential impact of the pandemic on road safety, have significant implications for urban planning and traffic management. This is directly related to previous studies highlighting the importance of improving road safety when the impact of planned infrastructure on accident risk is known (Vandenbulcke, I. Thomas, Geus, et al., 2009). Policymakers should consider the particular impact of bicycles on accident rates and the potential for unexpected outcomes during unprecedented events such as a pandemic. Only with integrative and interdisciplinary cooperation between road safety, spatial planning, and the geographical context of the infrastructure elements themselves will it be possible to implement a safe accident prevention system as planned by the United Nations by 2030 (United Nations, 2021). Implementing strategies to improve cyclist safety, optimizing traffic flow, and developing adaptive strategies to respond to sudden traffic behavior changes can improve city road safety. These findings argue for a dynamic and data-driven approach to transport policy that reflects the multifaceted nature of urban transport. Above all, this means for the government that with the modal split leveling off in the city by 2025, the changing traffic patterns also threaten to bring about a situation in which the number of accidents will continue to rise (Tiefbauamt Stadt Zürich, 2021). The expected value is that motorized traffic will decrease by 10 to 20 %, while bicycle traffic will account for  $1/10$  of the total traffic load. Preventive planning for this situation is needed to avoid such a scenario.

The spatial correlation between cycling accidents, motorized vehicle traffic, road infrastructure, and accident severity has significant implications for urban planning and traffic management in Zurich and

elsewhere. The data show that road infrastructures like roundabouts or high-speed roads are inherently more accident-prone for cyclists, primarily associated with high motorized vehicle traffic density - this is repeatedly found the same as in past studies and only further emphasizes the importance of solution-oriented urban planning Wegman, Zhang, and Dijkstra, 2012. This results in an urgent need for policymakers to rethink road design at these hotspots. Especially with the advent of *BIM* (Building information modeling) and in all areas involving road space, an integrative use of situational modeling is promising (Kubba, 2012). Especially in the increasing bicycle traffic in Zurich, installing bicycle lanes and creating divided zones where pedestrians, cyclists, and motor vehicles have equal rights could reduce the potential for conflict. Such approaches have already been proposed in past studies and are shown in this thesis to be supported by them (B. Thomas and DeRobertis, 2013). Knowledge and understanding of how infrastructures interact can be combined to separate vehicle corridors and reduce conflict points. The severity of crashes is also closely related to road infrastructure. Therefore, there is an urgent need to adapt infrastructure to avoid accidents and minimize their consequences, considering the *SERFOR*- principle.

# Chapter 7

## Conclusion

### 7.1 Contributions

**Analyzing Traffic Patterns and Bicycle Accidents** This master’s thesis delves into the intricate relationship between the density of motorized vehicles, the prevailing road infrastructure, and the occurrence of bicycle accidents in Zurich. The research illuminates the city’s traffic patterns through time series analysis and offers a historical perspective on how traffic dynamics have influenced bicycle accident rates.

**High-Resolution Traffic Data** Building on this foundational understanding, the study utilizes high-resolution traffic data to forecast fluctuations in traffic and accident figures more precisely. This offers insights into how these numbers have evolved concerning each other over the past decade.

**Leveraging Linear Mixed-Effects Models** One significant finding is the efficacy of linear mixed-effects models within the infrastructure context. These models elucidate accident locations influenced by traffic infrastructure and spotlight the intricate dynamics within traffic safety. They also highlight which infrastructure elements pose heightened risks to cyclists.

**Enhancing Predictive Accuracy with Advanced Models** Regression and machine learning models have been employed to augment further prediction accuracy concerning accident-prone zones influenced by road infrastructure. Their efficacy underscores their capacity to distill raw data into actionable insights, thereby aiding in safety measure prioritization. Such innovations pave the way for conceptualizing novel traffic designs, embedding state-of-the-art analytical methodologies within conventional road safety analyses, and ultimately catalyzing informed data-driven decisions.

**Geographical Approaches and Road Design Principles** Geographical methods applied to dissect crash severity have illuminated primary determinants such as vehicle type and age, relegating infrastructure to a contributory role. A critical recommendation from this study is adopting the *SER-FOR* principle—self-explaining roads (*SER*) coupled with forgiving roads (*FOR*). This design principle advocates infrastructures that induce instinctive user behavior adaptation (such as speed modulation) and ensures that mishaps do not escalate to severe accidents. In conclusion, this research offers a pioneering approach to quantify and address bicycle accidents, laying the foundation for innovative, future-centric network designs aligned with the *Vision Zero* ambition.

## 7.2 Main Findings

In the context of Zurich’s urban transport landscape, several pivotal trends and implications have been discerned. There has been a marked decline in motorized vehicle traffic, while there was, at the same time, an upswing in bicycle usage. Despite these shifts, bicycle-related accidents have maintained a high incidence rate, witnessing only a marginal downtrend. The COVID-19 pandemic has further exacerbated the challenges, casting a detrimental shadow over road safety, especially for bicyclists.

The thesis further reveals a strong correlation between the densities of motorized and non-motorized vehicles and the frequency of bicycle accidents. In addition, seasonal variations and environmental variables have significantly affected the traffic density, which in turn affects the accident history.

Delving into infrastructure, it is apparent that dedicated cycling infrastructure is a cornerstone for bolstering cyclist safety. The intricate interplay of the transportation network’s speed and complexity further dictates safety outcomes for cyclists. The importance of distinct pathways for cyclists, especially ensuring separation from tram and bus lines, emerges as a pivotal strategy for road safety enhancement. Adaptations in road widths, congruent with traffic demands and moderated speed limits, further amplify safety metrics. Thus, based on the road infrastructure, it is possible to predict the locations of accidents - which could allow the changing areas of accidents to be expected as the road infrastructure changes.

Lastly, the foundational role of Zurich’s road infrastructure in shaping accident probabilities and severities must be considered. Interestingly, accident severity correlates with the type and size of involved vehicles, right-of-ways as well as the road width, age demographics, and specific factors such as high-speed zones and trams. This comprehensive understanding forms the bedrock for future urban planning and road safety endeavors.

## 7.3 Limitations

Although this study provides valuable insights into the interaction of bicycle crashes, motorized vehicle traffic, and road infrastructure, it has limitations. The most severe and critical limitation is the availability and accuracy of the geospatial data (see 6.3). This leads to limitations in the accuracy of the model predictions and could be addressed in future work with a more diverse geospatial dataset—which would need to include all network-relevant infrastructure elements such as curbs, trees, signs, lights, and pavement markings. Therefore, integrating additional data, such as demographic, economic, and environmental data, is critical to improving the applicability of this study’s findings. Although limited by data constraints, this work provides a solid foundation for future research.

Moreover, the models developed did not incorporate other potentially influential factors such as weather conditions, human behavior and the cyclists’ perception of road safety, and demographic data, which, if included, could have enriched the understanding of bicycle accident dynamics. Road safety is inherently multifaceted and cannot be fully explained by infrastructure variables alone.

From a methodological perspective, while the regression and machine learning models employed provide advantages, they have inherent limitations. Their effectiveness hinges on the validity of initial assumptions, and deviations from these can weaken their predictive power. Furthermore, this study represents a snapshot of current conditions based on available data. Given the dynamic nature of traffic patterns and infrastructure development, the relationships and patterns identified in this study are likely to evolve and, hence, necessitate ongoing validation.

## 7.4 Outlook

This thesis underscores the interplay between speed, complexity, and infrastructure in cycling safety. There is an imperative to investigate how varied network infrastructure adjustments impact accident rates, which could guide policymakers. Extending this research to diverse urban environments would offer broader insights into the dynamics between motorized transport, bicycle accidents, and infrastructure. This expansion necessitates refining the existing machine-learning models by incorporating a broader spectrum of variables, ranging from meteorological conditions to cyclist behaviors and demographics. Moreover, unraveling the characteristics that define accident-prone locations is pivotal, focusing primarily on factors influencing accident severity. As we look forward, emphasizing subcategories within infrastructure elements, such as intersections or multi-lane configurations, will be crucial to discerning their distinct impacts on accident probabilities.

The intricate relationship between road use and accident frequencies, especially during the pandemic, presents numerous research opportunities. Analyzing specific pandemic periods or locales could yield detailed accident determinants. Considering weather or local regulations offers a holistic view of road safety. The pandemic's impact on traffic dynamics necessitates adaptive policy-making.

Unraveling the spatial ties between bicycle accidents and their influencing factors in Zurich prompts deeper exploration. Identifying high-risk timeframes can inform tailored traffic management. Including pedestrian dynamics will provide a broader perspective on urban mobility. Utilizing technologies such as *IoT* for real-time monitoring promises innovation. As cities globally lean towards sustainable mobility, comparing spatial analyses across different urban agglomerations will help to unearth global trends or pinpoint tailored interventions. The contextual nature of accidents emphasizes the need for a comprehensive geographical approach in future studies. Thus, more infrastructure data must be gathered and implemented into the methodological approaches put into operation in this thesis.

# Bibliography

- ASTRA (2016). “SERFOR Voranalyse Self Explaining and Forgiving Roads”. In: p. 96.
- (2022). *Unfallstatistik 2022*. Tech. rep.
- Barbet-Massin, Morgane et al. (Apr. 2012). “Selecting pseudo-absences for species distribution models: how, where and how many?” In: *Methods in Ecology and Evolution* 3.2, pp. 327–338. ISSN: 2041210X. DOI: 10.1111/j.2041-210X.2011.00172.x.
- Barnes, Stephen R et al. (2020). *The Effect of COVID-19 Lockdown on Mobility and Traffic Accidents: Evidence from Louisiana*. Tech. rep. 616. Essen.
- Bates, Author Douglas et al. (2023). *Package ‘lme4’*.
- Bauman, Adrian et al. (2008). “Cycling: Getting Australia moving - Barriers, facilitators and interventions to get more Australians physically active through cycling”. In: *31st Australasian Transport Research Forum, ATRF 2008* 2006, pp. 593–601.
- Bíl, Michal, Richard Andrášik, and Zbyněk Janoška (June 2013). “Identification of hazardous road locations of traffic accidents by means of kernel density estimation and cluster significance evaluation”. In: *Accident Analysis & Prevention* 55, pp. 265–273. ISSN: 00014575. DOI: 10.1016/j.aap.2013.03.003.
- Boucher, Jean-Philippe and Roxane Turcotte (Sept. 2020). “A Longitudinal Analysis of the Impact of Distance Driven on the Probability of Car Accidents”. In: *Risks* 8.3, p. 91. ISSN: 2227-9091. DOI: 10.3390/risks8030091.
- Boyer, Marcel and Georges Dionne (Oct. 1987). “The economics of road safety”. In: *Transportation Research Part B: Methodological* 21.5, pp. 413–431. ISSN: 01912615. DOI: 10.1016/0191-2615(87)90038-5.
- Breiman, Leo (2001). “Random Forests”. In: *Machine Learning* 45.1, pp. 5–32. ISSN: 1573-0565. DOI: 10.1023/A:1010933404324.
- Brian Ripley and William Venables (2016). “Package ‘nnet’”. In.
- Brucks, Wernher (2020). *Verkehrsunfallstatistik 2020*. Tech. rep. Stadt Zürich: Dienstabteilung Verkehr, p. 15.
- Bucsky, Péter (Nov. 2020). “Modal share changes due to COVID-19: The case of Budapest”. In: *Transportation Research Interdisciplinary Perspectives* 8, p. 100141. ISSN: 25901982. DOI: 10.1016/j.trip.2020.100141.
- Chen, Peng and Qing Shen (Jan. 2016). “Built environment effects on cyclist injury severity in automobile-involved bicycle crashes”. In: *Accident Analysis & Prevention* 86, pp. 239–246. ISSN: 00014575. DOI: 10.1016/j.aap.2015.11.002.
- Cleveland, W S, E Grosse, and W M Shyu (1992). “Local regression models”. In: *Statistical Models in S*. Ed. by J M Chambers and T J Hastie. Wadsworth & Brooks/Cole. Chap. 8.
- DAV (2021). *Verkehrsunfallstatistik 2021*. Tech. rep. Stadt Zürich.
- De Rome, Liz et al. (Jan. 2014). “Bicycle Crashes in Different Riding Environments in the Australian Capital Territory”. eng. In: *Traffic Injury Prevention* 15.1, pp. 81–88. ISSN: 1538-9588. DOI: 10.1080/15389588.2013.781591.
- Dr. S. Jackson (2023). *Machine learning*, p. 207. ISBN: 9781617294433.

- Ebrahim Shaik, Md. and Samsuddin Ahmed (Sept. 2022). “An overview of the impact of COVID-19 on road traffic safety and travel behavior”. In: *Transportation Engineering* 9, p. 100119. ISSN: 2666691X. DOI: 10.1016/j.treng.2022.100119.
- Failing, Gates R L et al. (2023). “The Impact of the COVID-19 Pandemic on Pediatric Bicycle Injury”. In: *International Journal of Environmental Research and Public Health* 20.8. ISSN: 1660-4601. DOI: 10.3390/ijerph20085515.
- Federal Office of Public Health FOPH (2023). *COVID-19 Switzerland*.
- Foley, J et al. (Aug. 2021). “Impact of a National Lockdown on Cycling Injuries.” eng. In: *Irish medical journal* 114.7, p. 412. ISSN: 0332-3102.
- Francke, Angela (2022). “Cycling during and after the COVID-19 pandemic”. eng. In: *Advances in Transport Policy and Planning*. Vol. 10, pp. 265–290. DOI: 10.1016/bs.atpp.2022.04.011.
- Garrard, Jan, Geoffrey Rose, and Sing Kai Lo (Jan. 2008). “Promoting transportation cycling for women: The role of bicycle infrastructure”. In: *Preventive Medicine* 46.1, pp. 55–59. ISSN: 00917435. DOI: 10.1016/j.ypmed.2007.07.010.
- Getahun, Kidane Alemtsega (Dec. 2021). “Time series modeling of road traffic accidents in Amhara Region”. In: *Journal of Big Data* 8.1, p. 102. ISSN: 2196-1115. DOI: 10.1186/s40537-021-00493-z.
- Greenwell, Brandon et al. (2022). *Package ‘gbm’*. Tech. rep.
- Guo, Yanyong, Ahmed Osama, and Tarek Sayed (Apr. 2018). “A cross-comparison of different techniques for modeling macro-level cyclist crashes”. In: *Accident Analysis & Prevention* 113, pp. 38–46. ISSN: 00014575. DOI: 10.1016/j.aap.2018.01.015.
- Harkort, Lasse, Byron Blake Walker, and Tobia Lakes (Mar. 2023). “Spatiotemporal Patterns of Cyclist Collisions in Germany: Variations in Frequency, Severity of Injury, and Type of Collision in 2019”. In: *Applied Spatial Analysis and Policy* 16.1, pp. 209–228. ISSN: 1874-463X. DOI: 10.1007/s12061-022-09476-w.
- Haworth, Narelle, Kristiann C Heesch, and Amy Schramm (Dec. 2018). “Drivers who don’t comply with a minimum passing distance rule when passing bicycle riders”. In: *Journal of Safety Research* 67, pp. 183–188. ISSN: 00224375. DOI: 10.1016/j.jsr.2018.10.008.
- Jaber, Ahmed, Balint Csonka, and Janos Juhasz (May 2022). “Long Term Time Series Prediction of Bike Sharing Trips: A Case Study of Budapest City”. In: *2022 Smart City Symposium Prague (SCSP)*. IEEE, pp. 1–5. ISBN: 978-1-6654-7923-3. DOI: 10.1109/SCSP54748.2022.9792540.
- Jaber, Ahmed, János Juhász, and Bálint Csonka (June 2021). “An Analysis of Factors Affecting the Severity of Cycling Crashes Using Binary Regression Model”. In: *Sustainability* 13.12, p. 6945. ISSN: 2071-1050. DOI: 10.3390/su13126945.
- Jacobsen, P L (2015). “Safety in numbers: more walkers and bicyclists, safer walking and bicycling”. In: *Injury Prevention* 21.4, pp. 271–275. ISSN: 1353-8047. DOI: 10.1136/ip.9.3.205rep.
- James, Gareth et al. (2021). *An Introduction to Statistical Learning*. Springer Texts in Statistics. New York, NY: Springer US. ISBN: 978-1-0716-1417-4. DOI: 10.1007/978-1-0716-1418-1.
- Janstrup, Kira Hyldekær et al. (Aug. 2023). “Predicting injury-severity for cyclist crashes using natural language processing and neural network modelling”. In: *Safety Science* 164, p. 106153. ISSN: 09257535. DOI: 10.1016/j.ssci.2023.106153.
- Kaplan, Sigal and Carlo Giacomo Prato (Oct. 2015). “A Spatial Analysis of Land Use and Network Effects on Frequency and Severity of Cyclist–Motorist Crashes in the Copenhagen Region”. In: *Traffic Injury Prevention* 16.7, pp. 724–731. ISSN: 1538-9588. DOI: 10.1080/15389588.2014.1003818.
- Katrakazas, Christos et al. (Sept. 2020). “A descriptive analysis of the effect of the COVID-19 pandemic on driving behavior and road safety”. In: *Transportation Research Interdisciplinary Perspectives* 7, p. 100186. ISSN: 25901982. DOI: 10.1016/j.trip.2020.100186.

- Kellermann, Robin et al. (Sept. 2022). “Mobility in pandemic times: Exploring changes and long-term effects of COVID-19 on urban mobility behavior”. In: *Transportation Research Interdisciplinary Perspectives* 15, p. 100668. ISSN: 25901982. DOI: 10.1016/j.trip.2022.100668.
- Killick, R, P Fearnhead, and I A Eckley (2012). “Optimal Detection of Changepoints With a Linear Computational Cost”. In: *Journal of the American Statistical Association* 107.500, pp. 1590–1598. DOI: 10.1080/01621459.2012.737745.
- Killick, Rebecca (2022). “Package ‘changept’ ”. In: pp. 1–28.
- Kim, Joon-Ki et al. (2007). “Bicyclist injury severities in bicycle–motor vehicle accidents”. In: *Accident Analysis & Prevention* 39.2, pp. 238–251. ISSN: 0001-4575. DOI: <https://doi.org/10.1016/j.aap.2006.07.002>.
- Koike, Hirotaka, Akinori Morimoto, and Atsushi Kitazawa (2003). “Unevenness of Intersection Pavement and Bicycle Safety”. In: *Transportation Research Record* 1846.1, pp. 56–61. DOI: 10.3141/1846-10.
- Kubba, Sam (2012). “Building Information Modeling”. In: *Handbook of Green Building Design and Construction*. Ed. by Sam B T - Handbook of Green Building Design Kubba and Construction. Boston: Elsevier, pp. 201–226. ISBN: 978-0-12-385128-4. DOI: 10.1016/B978-0-12-385128-4.00005-6.
- Kullgren, Anders et al. (June 2019). “The potential of vehicle and road infrastructure interventions in fatal bicyclist accidents on Swedish roads—What can in-depth studies tell us?” In: *Traffic Injury Prevention* 20.sup1, S7–S12. ISSN: 1538-9588. DOI: 10.1080/15389588.2019.1610171.
- Lee, Jaeyoung, Mohamed Abdel-Aty, and Ximiao Jiang (May 2015). “Multivariate crash modeling for motor vehicle and non-motorized modes at the macroscopic level”. In: *Accident Analysis & Prevention* 78, pp. 146–154. ISSN: 00014575. DOI: 10.1016/j.aap.2015.03.003.
- Lee, Won-Kyung et al. (July 2014). “A time series study on the effects of cold temperature on road traffic injuries in Seoul, Korea”. In: *Environmental Research* 132, pp. 290–296. ISSN: 00139351. DOI: 10.1016/j.envres.2014.04.019.
- Li, Maosheng, Feng Shi, and Dafei Chen (2011). “Analyze bicycle-car mixed flow by social force model for collision risk evaluation”. In: *Road Safety and Simulation*, pp. 1–22.
- Liu, Peipei and Stefanie Marker (Jan. 2020). “Evaluation of contributory factors’ effects on bicycle-car crash risk at signalized intersections”. In: *Journal of Transportation Safety & Security* 12.1, pp. 82–93. ISSN: 1943-9962. DOI: 10.1080/19439962.2019.1591555.
- Lord, Dominique, Abdelaziz Manar, and Anna Vizioli (Jan. 2005). “Modeling crash-flow-density and crash-flow-V/C ratio relationships for rural and urban freeway segments”. In: *Accident Analysis & Prevention* 37.1, pp. 185–199. ISSN: 00014575. DOI: 10.1016/j.aap.2004.07.003.
- Lovelace, Robin et al. (July 2020). “Methods to Prioritise Pop-up Active Transport Infrastructure”. In: *Findings*. ISSN: 2652-8800. DOI: 10.32866/001c.13421.
- Lusk, Anne C et al. (Apr. 2011). “Risk of injury for bicycling on cycle tracks versus in the street”. eng. In: *Injury Prevention* 17.2, pp. 131–135. ISSN: 1353-8047. DOI: 10.1136/ip.2010.028696. URL: <https://injuryprevention.bmj.com/lookup/doi/10.1136/ip.2010.028696>.
- Meschik, Michael (2012). “Reshaping City Traffic Towards Sustainability Why Transport Policy should Favor the Bicycle Instead of Car Traffic”. In: *Procedia - Social and Behavioral Sciences* 48, pp. 495–504. ISSN: 18770428. DOI: 10.1016/j.sbspro.2012.06.1028.
- Miranda-Moreno, Luis F., Jillian Strauss, and Patrick Morency (Jan. 2011). “Disaggregate Exposure Measures and Injury Frequency Models of Cyclist Safety at Signalized Intersections”. In: *Transportation Research Record: Journal of the Transportation Research Board* 2236.1, pp. 74–82. ISSN: 0361-1981. DOI: 10.3141/2236-09.
- Monfort, Samuel S., Jessica B. Cicchino, and David Patton (Dec. 2021). “Weekday bicycle traffic and crash rates during the COVID-19 pandemic”. In: *Journal of Transport & Health* 23.January, p. 101289. ISSN: 22141405. DOI: 10.1016/j.jth.2021.101289.

- Pazdan, Sylwia (Dec. 2020). “The impact of weather on bicycle risk exposure”. In: *Archives of Transport* 56.4, pp. 89–105. ISSN: 0866-9546. DOI: 10.5604/01.3001.0014.5629.
- Pedroso, Felipe E et al. (Dec. 2016). “Bicycle Use and Cyclist Safety Following Boston’s Bicycle Infrastructure Expansion, 2009–2012”. In: *American Journal of Public Health* 106.12, pp. 2171–2177. ISSN: 0090-0036. DOI: 10.2105/AJPH.2016.303454.
- Petrică, Andreea-Cristina, Stelian Stancu, and Alexandru Tindeche (2016). “Limitation of ARIMA models in financial and monetary economics”. In: *Theoretical and Applied Economics* XXIII.4, pp. 19–42.
- Prati, Gabriele, Luca Pietrantonio, and Federico Fraboni (Apr. 2017). “Using data mining techniques to predict the severity of bicycle crashes”. In: *Accident Analysis & Prevention* 101, pp. 44–54. ISSN: 00014575. DOI: 10.1016/j.aap.2017.01.008.
- Pratt, Gregory C et al. (May 2014). “Quantifying traffic exposure”. In: *Journal of Exposure Science & Environmental Epidemiology* 24.3, pp. 290–296. ISSN: 1559-0631. DOI: 10.1038/jes.2013.51.
- Pucher, John and Ralph Buehler (July 2008). “Making Cycling Irresistible: Lessons from The Netherlands, Denmark and Germany”. In: *Transport Reviews* 28.4, pp. 495–528. ISSN: 0144-1647. DOI: 10.1080/01441640701806612.
- (Nov. 2017). “Cycling towards a more sustainable transport future”. In: *Transport Reviews* 37.6, pp. 689–694. ISSN: 0144-1647. DOI: 10.1080/01441647.2017.1340234.
- Qiu, Lin and Wilfrid A Nixon (Jan. 2008). “Effects of Adverse Weather on Traffic Crashes”. In: *Transportation Research Record: Journal of the Transportation Research Board* 2055.1, pp. 139–146. ISSN: 0361-1981. DOI: 10.3141/2055-16.
- Quddus, Mohammed A (Sept. 2008). “Time series count data models: An empirical application to traffic accidents”. In: *Accident Analysis & Prevention* 40.5, pp. 1732–1741. ISSN: 00014575. DOI: 10.1016/j.aap.2008.06.011.
- Reynolds, Conor CO et al. (Dec. 2009). “The impact of transportation infrastructure on bicycling injuries and crashes: a review of the literature”. In: *Environmental Health* 8.1, p. 47. ISSN: 1476-069X. DOI: 10.1186/1476-069X-8-47.
- Santos, Kenny, João P Dias, and Conceição Amado (Feb. 2022). “A literature review of machine learning algorithms for crash injury severity prediction”. In: *Journal of Safety Research* 80, pp. 254–269. ISSN: 00224375. DOI: 10.1016/j.jsr.2021.12.007.
- Scott, A J and M Knott (1974). “A Cluster Analysis Method for Grouping Means in the Analysis of Variance”. In: *Biometrics* 30.3, pp. 507–512. ISSN: 0006341X, 15410420.
- Smeed, R J (1949). “Some Statistical Aspects of Road Safety Research”. In: *Journal of the Royal Statistical Society. Series A (General)* 112.1, pp. 1–34. ISSN: 00359238.
- Spoerri, Adrian, Matthias Egger, and Erik von Elm (Jan. 2011). “Mortality from road traffic accidents in Switzerland: Longitudinal and spatial analyses”. In: *Accident Analysis & Prevention* 43.1, pp. 40–48. ISSN: 00014575. DOI: 10.1016/j.aap.2010.06.009.
- Stone, Mervyn and Jeremy Broughton (2003). “Getting off your bike: cycling accidents in Great Britain in 1990–1999”. In: *Accident Analysis & Prevention* 35.4, pp. 549–556. ISSN: 0001-4575. DOI: [https://doi.org/10.1016/S0001-4575\(02\)00032-5](https://doi.org/10.1016/S0001-4575(02)00032-5).
- Teschke, Kay et al. (Dec. 2012). “Route Infrastructure and the Risk of Injuries to Bicyclists: A Case-Crossover Study”. eng. In: *American Journal of Public Health* 102.12, pp. 2336–2343. ISSN: 0090-0036. DOI: 10.2105/AJPH.2012.300762.
- Thomas, Beth and Michelle DeRobertis (Mar. 2013). “The safety of urban cycle tracks: A review of the literature”. In: *Accident Analysis & Prevention* 52, pp. 219–227. ISSN: 00014575. DOI: 10.1016/j.aap.2012.12.017.
- Tiefbauamt Stadt Zürich (2021). “Stadtverkehr 2025”. In: pp. 1–40.
- Transportation Bureau of Statistics (2023). *COVID-19 Related Transportation Statistics*. URL: <https://www.bts.gov/covid-19>.

- Unfallversicherung (UVG) (2021). “Unfallstatistik UVG 2021”. In: p. 67. URL: [www.unfallstatistik.ch](http://www.unfallstatistik.ch).
- United Nations (2021). *How safe roads feature in Global Plan for Decade of Action*. Tech. rep., p. 34.
- Valent, Francesca (May 2022). “Road traffic accidents in Italy during COVID-19”. In: *Traffic Injury Prevention* 23.4, pp. 193–197. ISSN: 1538-9588. DOI: [10.1080/15389588.2022.2047956](https://doi.org/10.1080/15389588.2022.2047956).
- Vandenbulcke, Grégory, Isabelle Thomas, Bas de Geus, et al. (2009). “Mapping bicycle use and the risk of accidents for commuters who cycle to work in Belgium”. In: *Transport Policy* 16.2, pp. 77–87. ISSN: 0967-070X. DOI: <https://doi.org/10.1016/j.tranpol.2009.03.004>.
- Vandenbulcke, Grégory, Isabelle Thomas, and Luc Int Panis (Jan. 2014). “Predicting cycling accident risk in Brussels: A spatial case-control approach”. In: *Accident Analysis & Prevention* 62, pp. 341–357. ISSN: 00014575. DOI: [10.1016/j.aap.2013.07.001](https://doi.org/10.1016/j.aap.2013.07.001).
- Vanparijs, Jef et al. (2015). “Exposure measurement in bicycle safety analysis: A review of the literature”. In: *Accident Analysis & Prevention* 84, pp. 9–19. ISSN: 0001-4575. DOI: <https://doi.org/10.1016/j.aap.2015.08.007>.
- Wang, Yinhai and Nancy L Nihan (May 2004). “Estimating the risk of collisions between bicycles and motor vehicles at signalized intersections”. In: *Accident Analysis & Prevention* 36.3, pp. 313–321. ISSN: 00014575. DOI: [10.1016/S0001-4575\(03\)00009-5](https://doi.org/10.1016/S0001-4575(03)00009-5).
- Wegman, Fred, Fan Zhang, and Atze Dijkstra (Jan. 2012). “How to make more cycling good for road safety?” In: *Accident Analysis & Prevention* 44.1, pp. 19–29. ISSN: 00014575. DOI: [10.1016/j.aap.2010.11.010](https://doi.org/10.1016/j.aap.2010.11.010).
- Xia, Jizhe et al. (Apr. 2023). “Impact of Human Mobility on COVID-19 Transmission According to Mobility Distance, Location, and Demographic Factors in the Greater Bay Area of China: Population-Based Study”. eng. In: *JMIR Public Health and Surveillance* 9, e39588. ISSN: 2369-2960. DOI: [10.2196/39588](https://doi.org/10.2196/39588).
- Yang, Zaili et al. (2021). “Risk analysis of bicycle accidents: A Bayesian approach”. In: *Reliability Engineering & System Safety* 209, p. 107460. ISSN: 0951-8320. DOI: <https://doi.org/10.1016/j.ress.2021.107460>.
- Ziari, Hasan and Mohammed M Khabiri (2005). “Applied Gis software for improving pedestrian & bicycle safety”. In: *Transport* 20.4, pp. 160–164. DOI: [10.1080/16484142.2005.9638014](https://doi.org/10.1080/16484142.2005.9638014).

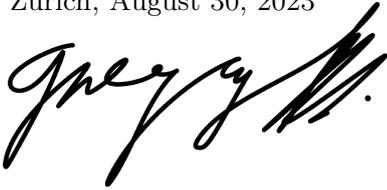
## Chapter 8

# Appendix

Package	Version	Description
lme4	1.1.33	Linear mixed-effects models
DHARMA	0.4.6	Residual diagnostics for hierarchical models
corrplot	0.92	Visualize a correlation matrix
car	3.1.2	Companion to Applied Regression
caret	6.0.94	Classification and Regression Training
randomForest	4.7.1.1	Random forests for classification and regression
MLmetrics	1.1.1	Machine learning evaluation metrics
gbm	2.1.8.1	Generalized Boosted Regression Models
glmnet	4.1.7	Lasso and elastic-net regularized regression
pROC	1.18.0	Display and analyze ROC curves
DMwR	0.4.1	Data Mining functions
lubridate	1.9.2	Work with date-times
feasts	0.3.1	Features and statistics for time series
TSstudio	0.1.6	Time series exploration tools
timetk	2.8.3	Toolkit for working with time series in R
BINCOR	0.2.0	Estimate bivariate binary correlations
lmtest	0.9.40	Testing linear regression models
strucchange	1.5.3	Testing for structural change
changepoint	2.2.4	Methods for changepoint detection
forecast	8.21	Forecasting functions for time series
zoo	1.8.12	S3 infrastructure for regular and irregular time series
tseries	0.10.53	Time series analysis and computational finance
cowplot	1.1.1	Streamlined plot theme and plot annotations
funtimes	9.1	Functions for time series analysis
Kendall	2.2.1	Kendall rank correlation and Mann-Kendall trend test
mgcv	1.8.42	Mixed GAM computation
bcp	4.0.3	Bayesian change point analysis
parallel	4.3.0	Enable parallel processing of functions
nnet	7.3.18	Feed-forward neural networks and multinomial logistic regression
MASS	7.3.58.4	Support functions and datasets for Venables and Ripley's MASS

**Personal declaration:** I hereby declare that the submitted thesis results from my own independent Work. All external sources are explicitly acknowledged in the thesis.

Zurich, August 30, 2023

A handwritten signature in black ink, appearing to read 'Gregory Biland', written in a cursive style.

Gregory Biland