



Cartography M.Sc.

Master thesis

Assessing cyclist safety using infrastructure parameters from OpenStreetMap: The case of Leipzig, Marseille and Edinburgh

Camilo Cardona Torres



2022

Master Thesis

Assessing cyclist safety using infrastructure parameters from OpenStreetMap: The case of Leipzig, Marseille and Edinburgh

submitted by **Camilo Cardona Torres**
born on 26.01.1992 in Cali, Colombia

submitted for the academic degree of
Master of Science (M.Sc.)

Submission on 10/10/2022

Supervisors Mathias Gröbe, M.Sc.
Technical University of Dresden

Dipl.-Ing. Albine Chanove
Fraunhofer Institute for Transportation and
Infrastructure Systems

Reviewer Dr. Barend Köbben
University of Twente

Statement of Authorship

Herewith I declare that I am the sole author of the submitted Master's thesis entitled:

"Assessing cyclist safety using infrastructure parameters from OpenStreetMap: The case of Leipzig, Marseille and Edinburgh"

I have fully referenced the ideas and work of others, whether published or unpublished. Literal or analogous citations are clearly marked as such.

Dresden, 10/10/2022

Camilo Cardona Torres

Abstract

Road crashes produce more than one million deaths every year around the world, and they are considered as the eighth cause of death for people of all ages. In addition to loss of life, traffic crashes carry more consequences with high impact on the society, including reduction of quality of life, medical costs, property damage and administrative costs. Cyclists are one of the most vulnerable road users, and there has been an increment on cycling victims in European roads in the last years.

Previous research on road safety has shown the impact of road infrastructure on crash risk and severity, which makes essential to consider it in road safety analyses. Although official traffic authorities have infrastructure inventories, this information is rather difficult to get, and usually it is not suitable for road safety assessments. However, due to the increasing use and contribution of Volunteered Geographic Information (VGI), crowdsourced geographic data is being created around the world, and particularly infrastructure-related data is freely available in OpenStreetMap (OSM).

Motivated by the challenge of reducing cyclist victims from traffic crashes, and the availability of data through VGI, the main objective of this research was assessing cyclist safety considering infrastructure parameters from OSM. Therefore, an inventory of intersections at the city level was obtained, focusing on road infrastructure parameters and cyclist victims. In terms of infrastructure, the intersections were classified by type (number of arms), presence of cycling infrastructure and traffic signals. On the other hand, the cyclist victims were analysed in three different categories: all victims (regardless severity), slightly injured victims, and seriously injured and killed victims.

A methodology to cluster the nodes and ways conforming the traffic network from OSM was presented, taking into account the road hierarchy. This process led to identify the intersections as single points from the clustered nodes, and to count the number of streets arriving to each intersection from the clustered ways. Later, the geolocated victims were matched and assigned to the intersections, leading to conduct a spatial and statistical analysis of cycling victims, including the identification of hotspots. The proposed methodology was implemented in Leipzig (Germany), Marseille (France) and Edinburgh (Great Britain), which were found to be comparable in terms of area and population density, and taking into consideration the crash data availability.

Based on the results by typology, 3-arm intersections presented the highest distribution of intersections with cyclist victims, with at least 50% in each of the three cities. However, comparing the rate of intersections with victims among all intersections of the same type, intersections with five or more arms, and roundabouts had the highest percentages in each city. When studying the different infrastructure parameters, in most cases the results suggested that intersections with traffic signals were safer for cyclists. Similarly, this pattern was also obtained when exploring the different categories associated with the severity level of the victims. Regarding the hotspots analysis, it was found that typically the hotspots were located in non-residential streets, particularly when analysing the most critical severity level (seriously injured and killed cyclists).

Since OSM is a promising data source for replicable road safety assessments, more researchers are encouraged to include it in their analysis worldwide, being aware of data correctness and availability. Additional data from other VGI sources is also advised to complement future analyses.

Acknowledgement

First of all, I would like to thank my supervisors Mathias Gröbe, M.Sc. from Technische Universität Dresden, and Dipl.-Ing. Albine Chanove from Fraunhofer Institute, for their academic and scientific guidance throughout this thesis. This work would not have been possible without their valuable time and professional advice, particularly in data collection, processing and analysis, as well as, in academic writing. Special thanks to my reviewer Dr. Barend Köbben, whose comments and feedbacks were key during this process; and to my classmate and friend Enock Seth Nyamador, who was extremely helpful providing his valuable expertise in OpenStreetMap and setting up the software required for executing this research.

Since this thesis is the final step of the Cartography M.Sc. program, I want to express my deepest gratefulness to all the persons who guided me along this beautiful two-year journey. I must mention especially my parents, Clara and Roberto; and my girlfriend, Marta. They gave me support, motivation and love to successfully finish my studies. I also want to say thousands of thanks to my closest friends in Munich, Vienna and Cairo, who always welcomed me in their houses and hearts. In particular, I want to thank Gonzalo and Estefania, who took care of me in the very beginning, and gave me a new home being far away from my original one. Additionally, my sincere gratitude to the family Köhler, the Deutschlandstipendium, and the DAAD for their sponsorship during my studies.

A special thanks also must go to the Carto family, who was an essential component during the whole process, personally and academically. To my professors, who taught me new cartographic and geospatial data technics. To my classmates, who allowed me to explore new cultures, and made possible to accomplish our academic tasks, having fun, trying international food, and traveling around Europe. And finally, thanks to the coordinators, especially to Juliane Cron, M.Sc. and Dr.-Ing. Eva Hauthal, for their continuous support throughout the master's program.

Contents

Statement of Authorship.....	i
Abstract	ii
Acknowledgement	iii
List of figures	vii
List of tables.....	viii
List of formulas.....	viii
Glossary	ix
1 Introduction	1
1.1 Motivation and problem statement.....	1
1.2 Research identification	2
1.2.1 Research objectives	2
1.2.2 Research questions.....	2
1.3 Document structure.....	3
2 Background	4
2.1 Literature review on road infrastructure and cyclist safety	4
2.1.1 Study cases without using Volunteered Geographic Information	4
2.1.2 Study cases using Volunteered Geographic Information	5
2.1.2.1 Applications exclusively to cycling infrastructure.....	5
2.1.2.2 Applications to cycling infrastructure and safety	5
2.2 Data	6
2.2.1 OpenStreetMap data	6
2.2.2 Crash data	7
3 Methods	8
3.1 Definition of parameters	9
3.1.1 Infrastructure parameters	9
3.1.2 Crash data parameters.....	9
3.2 Processing of OpenStreetMap data.....	10
3.2.1 Import of data from OpenStreetMap	10
3.2.2 Identification of intersections by nodes clustering	11
3.2.2.1 First DBSCAN	11

3.2.2.2	Second DBSCAN	12
3.2.2.3	Third DBSCAN	12
3.2.3	Definition of influence zone for intersections	12
3.2.4	Classification of intersections considering infrastructure parameters.....	13
3.2.4.1	Number of arms by clustering of ways	13
3.2.4.2	Roundabouts, cycling infrastructure and traffic signals	14
3.3	Processing of crash data	14
3.4	Analysis	14
3.4.1	Statistical analysis.....	15
3.4.2	Spatial analysis	15
4	Implementation	17
4.1	Study area.....	17
4.1.1	Other variables	19
4.2	OpenStreetMap data.....	19
4.2.1	Identification of intersections	20
4.2.2	Influence zones.....	21
4.2.3	Classification of intersections	21
4.2.3.1	Number of arms.....	21
4.2.3.2	Accuracy evaluation	23
4.3	Crash data	24
5	Results.....	25
5.1	General comparison between cities	25
5.2	Count of intersections.....	27
5.2.1	Intersections by typology based on number of arms	27
5.2.2	Intersections by category based on infrastructure attributes	27
5.3	Intersections with cyclist victims	29
5.3.1	Cyclist victims in intersections by type	29
5.3.2	Cyclist victims in intersections by infrastructure attributes	30
5.4	Intersections with slightly injured cyclists	32
5.4.1	Slightly injured cyclists in intersections by type	33
5.4.2	Slightly injured cyclists in intersections by infrastructure attributes	34
5.5	Intersections with seriously injured and killed cyclists	36

5.5.1	Seriously injured and killed cyclists in intersections by type	36
5.5.2	Seriously injured and killed cyclists in intersections by infrastructure attributes ...	37
5.6	Hotspots analysis	40
5.6.1	Leipzig	41
5.6.2	Marseille	42
5.6.3	Edinburgh	43
6	Discussion	44
6.1	Data	44
6.1.1	Definition of variables to select and compare cities	44
6.1.2	Definition of infrastructure parameters	45
6.1.3	Definition of cyclist safety indicators	46
6.2	Analysis	46
6.2.1	Identification and classification of intersections	46
6.2.2	Spatial match of cyclist victims with intersections	47
6.2.3	Spatial and statistical analysis	48
7	Conclusion	49
	References	50
	Appendix	53
	Appendix A: Metadata and details of available crash databases	53
	Appendix B: Results tables	56
	Appendix C: Hotspots for intersections with all cyclist victims	59

List of figures

Figure 3.1 Workflow of methodology applied in this research	8
Figure 3.2 Parameters for definition of influence zones	13
Figure 4.3 Population in potential cities for study area	17
Figure 4.4 Area, and population density in potential cities for study area.....	18
Figure 4.5 Localisation of chosen cities for the study area	18
Figure 4.6 Output of nodes and ways using pgRouting (intersection Gerberstraße - Tröndlinring in Leipzig).....	20
Figure 4.7 Clusters after first DBSCAN for identifying intersections (example in Leipzig)	20
Figure 4.8 Output of intersections after clustering (example in Leipzig).....	21
Figure 4.9 Output of influence zones and simplification of ways (example in Leipzig)	22
Figure 4.10 Identification of number of arms by intersecting influence zones and ways simplification.....	22
Figure 4.11 Output of intersection's classification by arms counting	23
Figure 4.12 Reported traffic crash victims with and without coordinates in original database (2015-2017).....	24
Figure 4.13 Reported cyclist victims with and without coordinates in original database (2015-2017)	24
Figure 5.14 Victims per city by vehicle type, between 2015 and 2017	26
Figure 5.15 Percentage of intersections by type and city	27
Figure 5.16 Percentage of intersections by city and category based on infrastructure attributes.	28
Figure 5.17 Percentage of intersections with cyclist victims by type and city	29
Figure 5.18 Percentage of intersections with and without cyclist victims by type	30
Figure 5.19 Percentage of intersections with cyclist victims by city and infrastructure attributes.	31
Figure 5.20 Percentage of intersections with and without cyclist victims by infrastructure attributes	32
Figure 5.21 Percentage of intersections with slightly injured cyclists by type and city	33
Figure 5.22 Percentage of intersections with and without slightly injured cyclists by type	33
Figure 5.23 Percentage of intersections with slightly injured cyclists by city and infrastructure attributes	34
Figure 5.24 Percentage of intersections with and without slightly injured cyclists by infrastructure attributes	35
Figure 5.25 Percentage of intersections with seriously injured and killed cyclists by type and city	36
Figure 5.26 Percentage of intersections with and without seriously injured and killed cyclists by type.....	37
Figure 5.27 Percentage of intersections with seriously injured and killed cyclists by infrastructure attributes	38
Figure 5.28 Percentage of intersections with and without seriously injured and killed cyclists by infrastructure attributes.....	39
Figure 5.29 Hotspots for cyclist victims by injury severity in Leipzig city centre	41
Figure 5.30 Hotspots for cyclist victims by injury severity in Marseille city centre.....	42
Figure 5.31 Hotspots for cyclist victims by injury severity in Edinburgh city centre	43
Figure A.32 Hotspots for all cyclist victims in Leipzig	59

Figure A.33 Hotspots for cyclist victims by injury severity in Leipzig	60
Figure A.34 Hotspots for all cyclist victims in Marseille	61
Figure A.35 Hotspots for cyclist victims by injury severity in Marseille	62
Figure A.36 Hotspots for all cyclist victims in Edinburgh	63
Figure A.37 Hotspots for cyclist victims by injury severity in Edinburgh	64

List of tables

Table 3.1 Importing parameters of osm2pgsql	10
Table 3.2 Importing parameters of pgRouting.....	10
Table 3.3 Values associated with the road hierarchy of the ways.....	11
Table 3.4 Parameters for first DBSCAN of nodes	11
Table 3.5 Parameters for second DBSCAN of nodes.....	12
Table 3.6 Parameters for third DBSCAN of nodes.....	12
Table 3.7 Values of z-score and p-value to reject the null hypothesis	16
Table 4.8 Sources of population and modal split data	19
Table 4.9 Tags applied to the imported OSM data	19
Table 4.10 Accuracy of intersection's categorisation by arm counting	23
Table 5.11 Comparison of selected cities through road infrastructure and safety variables	25
Table A.12 Metadata of crash databases in Saxony (Germany), France and Great Britain	53
Table A.13 Information at the crash level	54
Table A.14 Information at the participant/victim level	55
Table A.15 Cities per territory ordered by population density	56
Table A.16 Count of intersections by type and city	56
Table A.17 Count of intersections by city and category based on infrastructure attributes	56
Table A.18 Count of intersections with cyclist victims by type and city.....	57
Table A.19 Count of intersections with cyclist victims by city and infrastructure attributes	57
Table A.20 Count of intersections with slightly injured cyclists by type and city	57
Table A.21 Count of intersections with slightly injured cyclists by city and infrastructure attributes	58
Table A.22 Count of intersections with seriously injured and killed cyclists by type and city	58
Table A.23 Count of intersections with seriously injured and killed cyclists by city and infrastructure attributes	58

List of formulas

Formula 3.1 Getis-Ort G_i^* statistic.....	15
Formula 3.2 Sample mean	15
Formula 3.3 Sample standard deviation.....	15

Glossary

Term	Definition
Crash	It occurs when a motorised or non-motorised vehicle collides with another vehicle, a person or an object, resulting in injury, death and property damage. Also known as traffic accident.
Database	Organized collection of data stored and accessed electronically. The most common model includes rows (features) and columns (fields) in a series of tables.
Feature	Entity of a vector dataset, which contains a unique identifier and optional attribute records. It can be a point, line or polygon feature.
Feature layer	GIS data layer containing vector data (features) being connected to visualisation.
Field	Attribute of the features represented as a column in the tabular structure of a feature layer.
Geographic Information System (GIS)	It is a system that creates, manages, analyses and maps all types of spatial data.
Intersection	A place or point where two or more streets meet. Also known as junction.
Intersection arms	Elements of the traffic network, representing the independent streets which meet in the same intersection.
OpenStreetMap (OSM)	It is a map of the world, created by volunteers and free to use under an open license.
Volunteered Geographic Information (VGI)	It refers to the creation, assemblage and dissemination of geographic data provided voluntarily by individuals. It is also a form of user-generated content.
Node	It refers to a point feature in OSM. Using pgRouting, the concept of node is directly associated with the starting and ending point of a traffic network's edge (line, link or way).
Way	It refers to linear features in OSM. Using pgRouting, a way strictly refers to a traffic network's edge (line or link) associated with two distinct nodes.
Traffic network	It is a system of interconnecting lines and points (called edges and nodes) that represent a system of streets or roads for a given area.

1 Introduction

1.1 Motivation and problem statement

Deaths and injuries resulting from traffic crashes are a serious problem around the globe, and they cost millions to the society. In addition to loss of life or reduction of quality of life, traffic crashes carry more consequences associated with medical costs, property damage and administrative costs (European Transport Safety Council, 2007). In 2016 the number of casualties on the roads reached 1.35 million worldwide, which led to identify crashes as the eighth cause of death for people of all ages and the primary cause of death for children and young adults between 5 and 29 years old (World Health Organization, 2018).

Reviewing the data in Europe, 23,800 deaths were reported during 2016, in which 8% were cyclist fatalities. In 2019 the deaths number sank to 22,700, implying a 5% reduction in the total of people killed on European roads compared to the year 2016. However, cyclist fatalities presented an increase of 2% in the same amount of time (European Commission, 2020). This means that in Europe cycling became deadlier, which suggests taking serious actions on this regard, especially taking into account the benefits of cycling.

Cycling is an attractive and sustainable transportation mode, that brings benefits to the society including easing of road congestion, reduction of air and noise pollution, longer and healthier lives, cycling tourism and market. All in one, its annual benefits in Europe have been estimated between 150 and 155 billion euros (European Cyclists' Federation, 2018). In addition to that, it has been demonstrated that improvements in road safety and better cycling infrastructure encourage people to cycle more (Buehler & Pucher, 2012; Hong et al., 2020). Hence, it is key to improve safety conditions for cyclists.

In order to make cycling safer, it is necessary to carry out a cycling safety assessment. This assessment allows to understand the current situation, and helps to identify problems associated with the crashes considering the analysis of the existing data. Moreover, it is important to include road infrastructure parameters in the analysis, since their repercussion in injury and crash risk has been demonstrated by several authors (Reynolds et al., 2009; Saad et al., 2019; Wang & Akar, 2018). Although official traffic authorities gather information regarding road infrastructure, they typically do it at the local level for specific operational and maintenance needs, and this information is not suitable for safety assessments and it is rather difficult to get (Collins & Graham, 2019). Therefore, it is necessary to explore other open sources that bring useful information to enhance road safety assessments based on infrastructure parameters. For this task, OpenStreetMap (OSM) plays a significant role as a Volunteered Geographic Information (VGI) tool, since the crowdsourced data is freely accessible and updated, and its processing can be replicable not only locally, but also at a higher level around the world (Jokar Arsanjani et al., 2015).

Motivated by the challenge of traffic crashes, the potential of cycling, and the open access to crowdsourced data, the main objective of this research was assessing cyclist safety including road infrastructure parameters from OSM. Due to crash data availability and homogeneity, the analysis was focused on cities in Germany, France and Great Britain. In this way, the cyclist safety assessment led to identify crash patterns, both spatially and statistically, associated with the infrastructure characteristics at road intersections (also known as junctions). This will encourage local authorities to improve safety on their city's streets, and thereby to start decreasing the number of reported crashes and cyclist victims.

1.2 Research identification

1.2.1 Research objectives

The main objective of this research is:

Assessing cyclist safety considering official traffic crash data and road infrastructure parameters using OpenStreetMap data in Germany, France and Great Britain.

Thus, the following sub-objectives should be met:

Preparing data to make it comparable:

- Defining variables associated with road infrastructure and road safety that lead to establish comparisons between cities of the three countries.
- Defining infrastructure characteristics to classify traffic network intersections considering OSM data.
- Defining cyclist safety indicators based on the official traffic crash and OSM data.

Analysis of intersections and crashes:

- Building an algorithm which identifies and classifies all the intersections of the traffic network from OSM data at a city level.
- Matching cyclist crashes with intersections spatially.
- Carrying out a spatial and statistical analysis based on cyclist safety indicators by considering the parameters obtained from traffic crash and OSM data.

1.2.2 Research questions

According to the research objectives the main research question is:

How can official traffic crash data and road infrastructure parameters gathered from OpenStreetMap be used to assess cyclist safety in Germany, France and Great Britain?

To fulfil the sub-objectives, this thesis should answer the following research questions:

Preparing data to make it comparable:

- What variables associated with road infrastructure and road safety may be considered to establish comparisons between cities of the three countries?
- What parameters should be considered to classify traffic network intersections based on OSM data?
- What cyclist safety indicators might be used considering the traffic crash and OSM data?

Analysis of intersections and crashes:

- How is it possible to identify and classify all the intersections of a city's traffic network from OSM data?
- What kind of spatial relationships exist between cyclist crashes and the network intersections?
- What results can be obtained after carrying out a spatial and statistical analysis based on cyclist safety indicators by considering the traffic crash data and the parameters obtained from OSM data?

1.3 Document structure

This thesis includes seven chapters in total. The current chapter is the introduction of the thesis and presents the motivation, as well as the research objectives and questions. The Chapter 2 shows a background regarding the relevant literature review, and about the data. Later, the Chapter 3 explains the methodology applied to answer the research questions. In Chapter 4 the methodology is implemented in three European cities: Leipzig, Marseille and Edinburgh. Results and discussion about the implementation are presented in Chapters 5 and 6, highlighting the use of infrastructure parameters in cycling safety assessments. Finally, the Chapter 7 concludes the research and discusses future recommendations.

2 Background

This chapter includes two main sections. First, a revision of relevant academic projects related with this research's scope, including Volunteered Geographic Information (VGI), road infrastructure and road safety focused on cyclists. The second section includes a general description of the crash data and OpenStreetMap (OSM) data, which were used to fulfil the research objectives.

2.1 Literature review on road infrastructure and cyclist safety

Several authors have assessed road infrastructure and cycling safety around the world. However, it is possible to distinguish two main currents of studies: one without using VGI, and the other one using VGI.

2.1.1 Study cases without using Volunteered Geographic Information

Being aware about the influence of safety concerns in cycling, Wang & Akar (2018) built a model to understand cyclist's safety perception in Ohio (United States) based on different intersection features, controlling for sociodemographic variables and bicycling experience. Among the variables related with the intersection infrastructure, the authors chose types of intersection (based on number of arms, including roundabouts), presence of traffic signal controls, width of curb lanes, number of through auto lanes, traffic volume, presence of cycle tracks and sidewalks.

Wang & Akar (2018) found that intersections of five or more arms without traffic signals are negatively associated with cyclist's safety perception. Their results also suggested that cyclists would feel safer riding through roundabouts compared with traditional intersections, supported by the fact that roundabouts reduce the number of potential conflict points. In terms of cycling infrastructure, the presence of cycle tracks increased the safety perception as well.

Similarly, Shen et al. (2020) analysed and compared the influences of different intersection features on cyclist crashes in the United Kingdom, including cyclist and intersection characteristics, as well as environmental conditions. Particularly, the authors explored three types of intersections (roundabouts, 3-arm and 4-arm intersections), and found nine variables with significant impact on cyclist injury severity at those intersections: male cyclists, age, speed limit, traffic control strategies at intersections, urban junctions, overtaking inside the intersections, the collision point at the back of the bicycle, the collision point at the right of the bike, and the secondary collision happened on the roadside.

Reynolds et al. (2009) provided a list of studies that investigated relationships between safety for cyclists and road infrastructure, in Europe and North America. From the intersection-related analysis, the authors collected information mainly about roundabouts and bicycle crossings in Europe. Findings of the roundabout studies showed higher risk for cyclist after installation of roundabouts with multiple traffic lanes or with on-road bike lanes, whereas there was a lower risk at roundabouts with segregated cycle infrastructure. These results allow to confirm the importance of good-quality cycling infrastructure to increase road safety at intersections.

Bearing in mind that the authors reviewed several studies in different cities and countries, they also mentioned the significant variety in infrastructure design that can exist from one city to another. The reason for this may be the differences in urban form, cycling infrastructure, cycling rates, and even the culture of cycling.

2.1.2 Study cases using Volunteered Geographic Information

2.1.2.1 Applications exclusively to cycling infrastructure

Encouraged by the availability of crowdsourced data Ferster et al. (2020) and Hong et al. (2020) decided to explore the cycling infrastructure. The former evaluated the ability to use OSM for identifying and inventorying bicycling infrastructure, by comparing it to open data provided by city governments. On the other hand, the latter examined the impact of cycling infrastructure in the number of cycling trips, by considering Strava (VGI tool) cycle counts, and infrastructure data provided by official authorities. Although neither of the publications considered any aspect related with road safety in their research's scope, they highlighted the usability of crowdsourced data in the field of cycling.

Being aware about tagging and labelling issues in OSM, Ferster et al. (2020) found relatively good concordance in the overall length of bicycle infrastructure between OSM and official data. Nevertheless, concordance at infrastructure categories was low for cycle tracks and local street bikeways, both of which being less common types of bicycle infrastructure in the study area. The authors also suggest that OSM is a promising source for real time spatial data on bicycle infrastructure, and brought attention to the fact that, in some cases, OSM was more detailed and timelier than the open data provided by local authorities. In addition, they motivated practitioners to consider OSM data for multicity studies, being mindful of potential inconsistencies in attribution and local definitions.

2.1.2.2 Applications to cycling infrastructure and safety

Other authors also included the use of crowdsourced data in their research focused on road infrastructure and safety. That is the case of Bruzzone and Broccoli (2021), who obtained the length of the Italian traffic network from OSM, and used it as reference to compare different indicators. Therefore, they found the different ratios of road crashes, deaths, injuries and vehicles by the road length. As regards their results, it was found that a maximum exposure to the risk of crashes and number of vehicles involved, for motorways and urban roads, happens commonly in main cities. In their conclusions, the authors also mentioned the need to expand the statistical information with the supply of traffic flows (vehicles/km) on the national road network. However, this information is rather difficult to get from local authorities and is not available in OSM.

Similarly, Saad et al. (2019) carried out a study in Florida (United States) estimating safety performance functions for cycling crashes at intersections, by extracting volumes from Strava, and road geometric characteristics from official transportation authorities. Based on their adjusted model, it was concluded that traffic volume, bicycle volume, intersection size, signal control type, number of intersection arms, bike lanes, sidewalk width, median width, and speed limit are the significant factors that affect bicycle crashes at the intersections of this study area.

Also, Collins and Graham (2019) assessed cyclist safety by extracting road infrastructure data from OSM focused on specific collision hotspots in London. From official authorities they got the crash database and also the bus lane database. Whereas, from OSM they analysed variables such as junction density, signalized junctions, roundabouts, simple junctions (without traffic signals), road hierarchy, road lanes, speed limits and cycle routes. Based on their results, multilane roads and bus lanes affected cycle collision counts, 20-mph speed limits had less collisions than 30-mph, and junction density was found to obtain the highest impact on collision density. Additionally, they found that one-way roads had the largest effect on reducing collision risk along with the provision of junctions without traffic signals, which infers that other junction types, such as roundabouts and signalized junctions, present higher risk.

2.2 Data

This research was based on two main datasets, one related with the infrastructure parameters from OSM, and one related with the crash data.

2.2.1 OpenStreetMap data

OpenStreetMap (OSM) is a project that creates and distributes free geographic crowdsourced data for everybody. In order to model the physical world, OSM has three basic components (OpenStreetMap Wiki contributors, 2022):

- Nodes: A node represents a specific point on the earth's surface defined by its latitude and longitude. Each node comprises at least an id number and a pair of coordinates. Nodes can be used to define standalone point features.
- Ways: A way is an ordered list of between 2 and 2,000 nodes that define a polyline. Ways are used to represent linear features such as rivers and roads.
- Relations: A relation is a multi-purpose data structure that documents a relationship between two or more data elements (nodes, ways, and/or other relations)

These components have attributes, which are represented by tags. A tag consists of two free format text fields: a 'key' and a 'value'. For example, a way element representing a residential street can have the tag (*key=value*) *name=king street* and *highway=residential*.

Having free access to geographic crowdsourced data through OSM, allows to get information related with the road infrastructure, as well as the data corresponding to the administrative division of different places around the world. In order to get and work with the OSM data, the process started by downloading the .osm file of each territory through the Overpass API (Olbricht, 2019). This file is updated on a regular basis and contains all OSM data for the desired region, including map elements (nodes, ways and relations) with their specific attributes (tags). Then, the data from the .osm file was imported into relational databases in PostgreSQL, using *osm2pgsql* (Burgess & Pavlenko, 2020) and *pgRouting* (Kumar et al., 2022).

2.2.2 Crash data

Information about road crashes resulting in death or injury is collected from the police officers in most European countries. Moreover, the final reports are complemented with additional information from the hospitals, the National Statistical office, and witnesses (Adminaite et al., 2018). Yet, each country has its own reporting system, which makes challenging to merge the data and make it comparable.

However, Chanove (2021) developed a method to harmonise crash databases between different countries, and thanks to her work there is a ready-to-use database with crashes and victims for the federal state of Saxony in Germany, as well as the whole territory of France and Great Britain. Hence, for the purpose of this research that harmonised database was used, with data from 2015 to 2017. Additionally, the database included attributes associate with unique crash index, year, longitude, latitude, severity, junction type, road class and country of the crash, as well as victims' age and type of vehicle. For more details about the metadata and the available variables, see Appendix A.

In terms of injury severity, each country has its own definition to distinguish seriously injured from slightly injured. Nevertheless, for Germany, France and Great Britain, the definition is similar. In these countries a serious injured is consider when the person was hospitalised for at least 24 hours. In addition to the hospitalisations for at least one day, in the Great Britain the following are also considered serious victims whether or not they are detained in hospital: fractures, concussion, internal injuries, crushing, burns (excluding friction burns), severe cuts and lacerations, and severe general shock (Jost et al., 2022).

3 Methods

In order to carry out the cyclist safety assessment between different cities, taking into account not only traffic crash data, but also infrastructure data, the methodology applied for this research followed four main sections: definition of parameter, processing of OSM data, processing of crash data, and analysis. During the first section, infrastructure and crash data parameters were defined, including intersection type, cycling infrastructure, traffic signals, and victims by severity level. In the second section, the OSM data was imported and processed to identify the intersections, define their influence zone, and also to classify them according to the infrastructure parameters. Later, the crash data was processed, and the cyclist victims were matched with the intersections. Finally, the analysis was executed including a statistical component, with the distribution of intersections considering the different parameters; and also including a spatial component with hotspots identification. Before presenting a deeper description of each of these sections, in Figure 3.1 it is possible to see the workflow of this methodology.

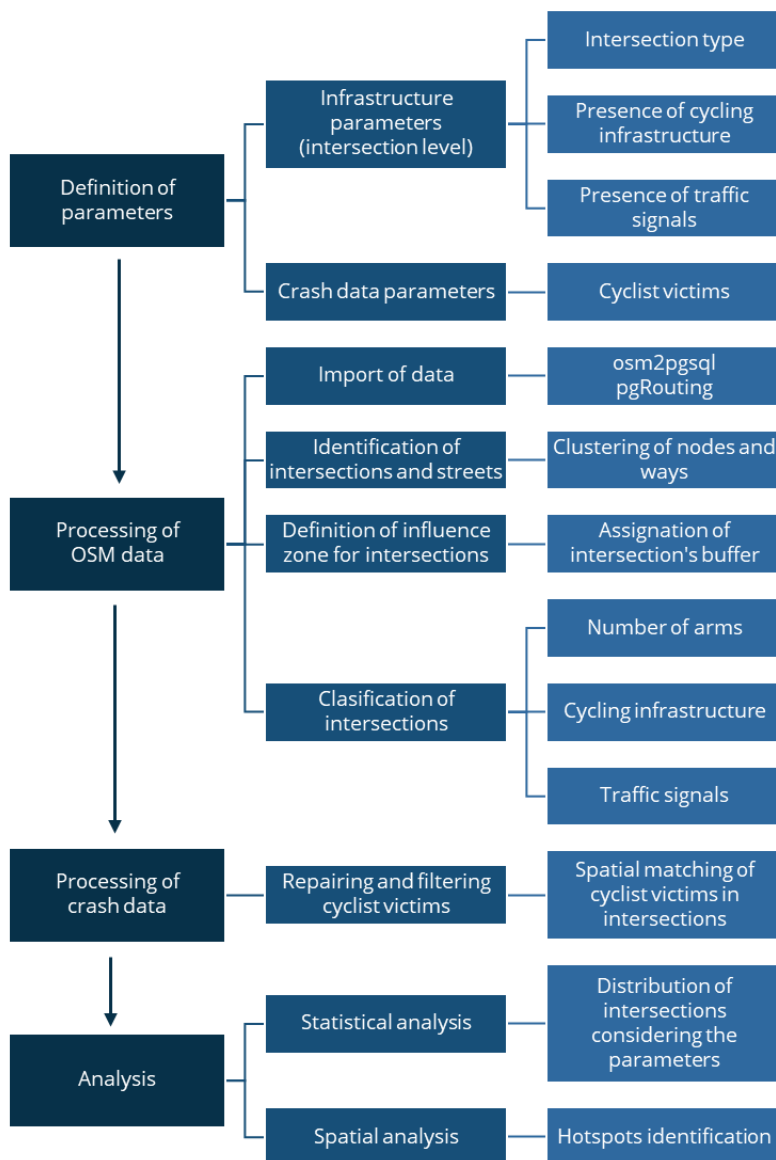


Figure 3.1 Workflow of methodology applied in this research

3.1 Definition of parameters

3.1.1 Infrastructure parameters

Several researchers have showed that road intersections (also known as junctions) are dangerous areas for cyclist due to the crossing traffic streams and the high level of interactions with the motorised traffic (Saad et al., 2019; Shen et al., 2020; Wang & Akar, 2018). Thereby, the current study was carried out at the intersection level for the analysis of cyclist victims.

Taking into account the infrastructure-related information available from the OSM data, three different parameters were chosen to classify the junctions:

- Intersection type (based on number of arms):
 - 3 arms
 - 4 arms
 - 5 or more arms
 - Roundabouts
- Presence of cycling infrastructure in the intersection:
 - Yes
 - No
- Presence of traffic signals in the intersection:
 - Yes
 - No

Although applied in different methodologies, these three parameters were also used in other cycling safety studies (Collins & Graham, 2019; Saad et al., 2019; Shen et al., 2020; Wang & Akar, 2018).

3.1.2 Crash data parameters

For the specific purposes of this research, only cyclist victims were studied, and based on the victim's severity, three groups were considered for the analysis:

- All cyclist victims (regardless severity)
- Slightly injured cyclists
- Seriously injured and killed cyclists

The first group produced results at a general level about the victims in the study area. Nevertheless, disaggregating the severity level allowed to identify specific findings and patterns that might help cities prioritize their actions in order to save lives and reduce the amount of traffic crash victims.

Even though killed victims are regularly considered in a different category (Shen et al., 2020), in some cases the reported numbers were found very low in order to reach representativity and consistency in the resulting analyses. Thus, for the purpose of this research it was necessary to merge seriously injured and killed cyclists in one category.

3.2 Processing of OpenStreetMap data

After defining the parameters, the data processing started with the import of nodes and ways from OSM. Then, these elements were simplified through clustering operations that allowed to group several nodes into single intersections, and several ways into single streets. In addition to this, once the intersections were obtained, it was necessary to define an influence zone so that a spatial matching with different attributes, as well as with the crash data could be executed. These steps are explained in detail as follows.

3.2.1 Import of data from OpenStreetMap

Using the `osm2pgsql` tool it was possible to get city boundaries, road infrastructure and traffic signals from OSM data, as relational databases in PostgreSQL. For achieving this, a `.lua` file was configured defining a function to process the OSM elements from the `.osm` file, defining the geometry types supported by PostgreSQL, and including the tags of interest (see Table 3.1).

Data	OSM's element	Geometry type	Tag's keys
City boundaries	relations	multipolygon	<i>admin_level, name</i>
Road infrastructure	ways	linestring	<i>name, highway, lanes, junction, bicycle, cycleway, cycleway:right, cycleway:left, cycleway:both</i>
Traffic signals	nodes	points	<i>highway</i>

Table 3.1 Importing parameters of `osm2pgsql`

Due to the tag's keys, the specific elements associated with the cycling infrastructure were identified from the traffic network. Consequently, this led to calculate the length of the cycling infrastructure, as well as the length of the traffic network.

Using `pgRouting`, the nodes and ways conforming the traffic network were imported. Similar to the process with `osm2pgsql`, a special file had to be configured to define importing parameters. In `pgRouting` this is a `.xml` file, where tag's names and values are defined (see Table 3.2).

Data	Tag's name	Tag's value	Id (pgRouting)
Nodes and ways from traffic network	<i>highway</i>	<i>motorway</i>	<i>101</i>
		<i>trunk</i>	<i>104</i>
		<i>primary</i>	<i>106</i>
		<i>secondary</i>	<i>108</i>
		<i>tertiary</i>	<i>109</i>
		<i>residential</i>	<i>110</i>

Table 3.2 Importing parameters of `pgRouting`

Note that elements associated to the link roads of the network (e.g., *highway=primary_link*) were not included, since they are not part of the main traffic network based on this research's scope. In addition to this, nodes with only two intersecting ways were removed, since they cannot be classified as road intersections (more than two arms).

3.2.2 Identification of intersections by nodes clustering

The original nodes were grouped using the density-based spatial clustering of applications with noise, also known as DBSCAN (Ester et al., 1996). However, due to the different road hierarchies, as well as diversity, complexity, and interaction of the nodes in the traffic network, it was not possible to apply a simple parameter for the clustering. Hence, the DBSCAN was executed three times using different criteria in order to get the final points at the intersections, as explained in the following subsections.

3.2.2.1 First DBSCAN

Since pgRouting allows to retrieve every way and node from the traffic network, the ways intersecting at the nodes were identified. First, based on the road hierarchy of the ways, a rank was given as a simple integer (see Table 3.3). Then, at each node a road score was calculated based on the sum of values according to the ways that were intersecting it. For example, a node intersected by two primary and one residential ways would have a road score of nine ($4+4+1$).

Road hierarchy	Value
Motorway	6
Trunk	5
Primary	4
Secondary	3
Tertiary	2
Residential	1

Table 3.3 Values associated with the road hierarchy of the ways

This also allowed to assign the corresponding minor and major road hierarchy per node, with which it was feasible to identify the pgRouting nodes in three categories:

- Nodes of non-residential ways
- Nodes of residential and non-residential ways
- Nodes of only residential ways

According to these categories, the first DBSCAN was applied. For this research's purposes, the higher the road classification associated with the node was, then the higher the search radius of a neighbourhood. The parameters used are shown in Table 3.4

Category of the nodes	Radius (m)	Min. Points
Nodes of non-residential ways	50	2
Nodes of residential and non-residential ways	35	2
Nodes of only residential ways	20	2

Table 3.4 Parameters for first DBSCAN of nodes

Later, the centroid of each cluster was found, by weighting the road score of each node. After that, the centroids were joined with the nodes that were not clustered.

3.2.2.2 Second DBSCAN

Since three different categories for the first DBSCAN were considered, at some intersections there was more than one cluster point. Therefore, a second DBSCAN was executed without discrimination of categories and applying the following cluster's parameters:

Category of the nodes	Radius (m)	Min. Points
Nodes resulting from first DBSCAN	20	2

Table 3.5 Parameters for second DBSCAN of nodes

For the second DBSCAN, the centroid of each cluster was found, by weighting the values associated with the radius of the first DBSCAN. Then, the maximum value of the radius between the clustered nodes to each centroid was assigned, allowing to keep identifying the different categories of the nodes (see Table 3.4). Later, the non-clustered points were combined with the clustered ones, to obtain the output of the second DBSCAN.

3.2.2.3 Third DBSCAN

When checking the clustered nodes, it was found that clusters with maximum category different to "nodes of only residential ways" (equivalent to 20 m of radius from the first DBSCAN) had not been clustered in the second DBSCAN. This finding was associated with the fact that clusters belonging to "nodes of non-residential ways" and "nodes of residential and non-residential ways" had mostly a distance greater than 20 meters between them. Thus, a third DBSCAN was applied.

This time, the output from the second DBSCAN was taken considering only categories not associated with "nodes of only residential ways", and a greater radius was applied (see Table 3.6).

Category of the nodes	Radius (m)	Min. Points
Nodes resulting from second DBSCAN not associated with "nodes of only residential ways"	30	2

Table 3.6 Parameters for third DBSCAN of nodes

Then, and similarly to the previous DBSCANS, in the third one the centroid of each cluster was calculated by weighting the values associated with the maximum radius of the second DBSCAN. After running the three DBSCANS, the gathered clusters were combined and this led to get one point per intersection, as desired.

3.2.3 Definition of influence zone for intersections

Since the junctions are points, it was necessary to create an area around them, in order to make it possible to assign other elements to them, like the crash victims and the infrastructure parameters. That is why, a circular buffer surrounding each intersection was defined as the influence zone. The radius of each buffer was assigned, based on the type of ways and number of nodes conforming the junction, and the maximum distance between the junction and its furthest node. Thereby, the specific chosen parameters, as shown in Figure 3.2 led to produce influence zones adaptable to the diverse conditions of the traffic network.

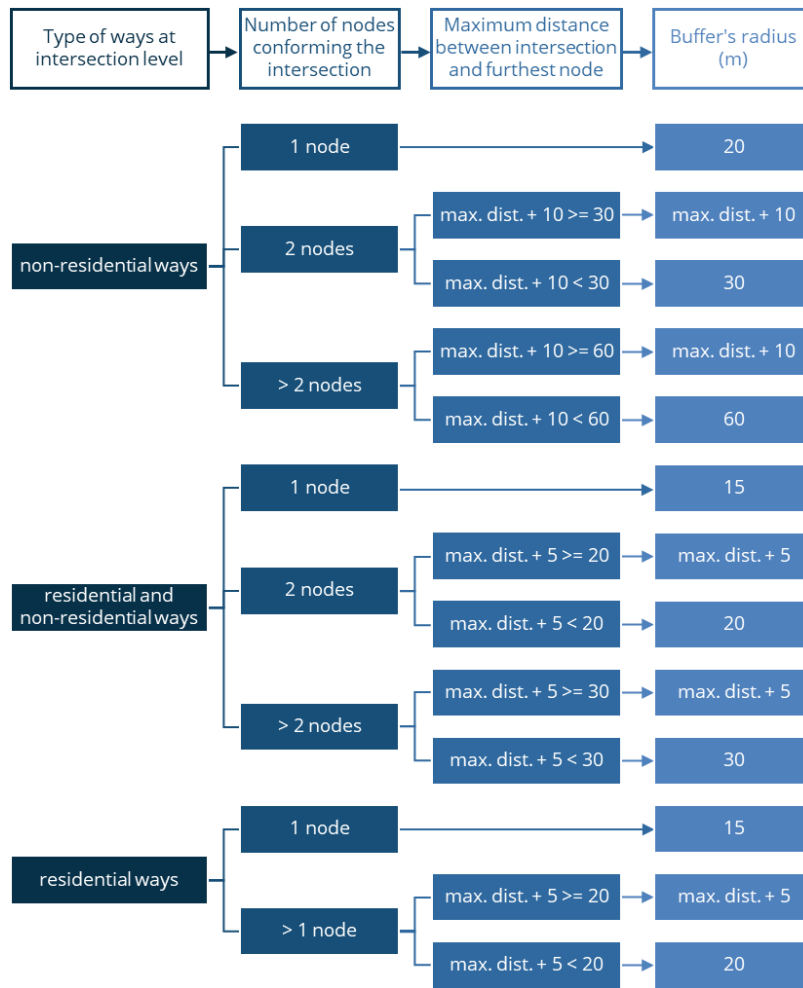


Figure 3.2 Parameters for definition of influence zones

In order to get to this approach, each node was provided with the different identifiers and radiuses from the clustering process, including each DBSCAN run and the final intersection. Hence, it was feasible to know to what junction each node belonged, and to calculate the distance between each intersection and its further corresponding node, which would be key for creating the influence zones. In this case, the type of ways was studied based on the minimum and maximum road hierarchy of the nodes shaping each junction.

3.2.4 Classification of intersections considering infrastructure parameters

3.2.4.1 Number of arms by clustering of ways

In order to count the number of arms per junction, the ways from OSM were clustered and simplified. This step was executed by running a negative buffer of previously aggregated buffers of the ways, taking into account the road hierarchy. The process was run as follows:

- Buffer of ways: 25 meters for non-residential ways and 10 meters for residential ways
- Aggregation through simple cluster of previous buffer elements within one meter distance
- Negative buffer of aggregated elements: -24.9 meters for non-residential ways and -9.9 meters for residential ways

The output of this process was a group of polygons of different width (minimum 0.1 meters) representing the simplification of the ways. This approach allowed to identify different streets, and hence, it was possible to count the number of them crossing at each intersection.

To find the number of arms, the streets from the output of the clustered ways were intersected with the exterior ring (also perimeter) of the influence zone per junction. At this step, the connectivity between the nodes and the ways was considered, and therefore, the actual streets that were containing the specific nodes of the respective intersections were counted.

3.2.4.2 Roundabouts, cycling infrastructure and traffic signals

In addition to the arms counting, further attributes related to the intersections were explored. Thereby, this research included presence of roundabouts, cycling infrastructure and traffic signals for each junction using its influence zone. For achieving this, OSM data was imported with `osm2pgsql` (see 3.2.1 Import of data from OpenStreetMap), and the corresponding tags to associate the attributes were identified.

Later, these elements were associated with the output of intersections (clustered nodes). In roundabouts, the intersection points located within the ways tagged as roundabout were identified. Whereas for cycling infrastructure and traffic signals, the elements were intersected with the buffers of the junctions (zone of influence).

3.3 Processing of crash data

Crash databases from different cities use to contain similar attributes and fields, however the notation between them is not consistent. Thus, the databases must be checked and repaired before executing a data analysis with them. Focusing on this research's scope, it was strictly necessary to fix the attributes of victim's vehicle, severity and coordinates.

Once the unified database reached consistency between the different cities, the victims were georeferenced based on the reported coordinates. Later, a spatial matching was executed in order to identify the victims located within the influence zone of each intersection.

3.4 Analysis

After the data processing, all parameters were assigned to the intersections. This made possible to characterise each junction with the following parameters:

- Intersection type (number of arms)
- Presence of cycling infrastructure
- Presence of traffic lights
- Number of cyclist victims
- Number of slightly injured cyclists
- Number of seriously injured and killed cyclists

Based on this information, the analysis was executed from an statistical perspective, including counts and proportions of the gathered data; and from a spatial perspective, including the identification of hotspots for the different victims categorisation.

3.4.1 Statistical analysis

From a descriptive approach, this analysis included the following steps:

- Count of all intersections
- Count of intersections with cyclist victims
- Count of intersections with slightly injured cyclists
- Count of intersections with seriously injured and killed cyclists

Moreover, for each of these counts the different distributions were calculated, considering the intersection typology and the other infrastructure parameters (cycling infrastructure and traffic signals).

3.4.2 Spatial analysis

Using local spatial autocorrelation analysis methods, it is possible to identify where the crashes gather, and hence, statistically significant hotspots (Cheng et al., 2019). Therefore, in this research spatial patterns of the intersections with cyclist victims were assessed with the Getis-Ord G_i^* statistic (Getis & Ord, 1995), which is a technique for local spatial autocorrelation, and the equation is as follows:

$$G_i^* = \frac{\sum_{j=1}^n w_{i,j} x_j - \bar{X} \sum_{j=1}^n w_{i,j}}{S \sqrt{\frac{n \sum_{j=1}^n w_{i,j}^2 - (\sum_{j=1}^n w_{i,j})^2}{n-1}}}$$

Formula 3.1 Getis-Ord G_i^* statistic

where x_j is the attribute value for feature j , $w_{i,j}$ is the spatial weight between feature i and j , n is equal to the total number of features. Additionally, \bar{X} is the usual sample mean, and S the standard deviation:

$$\bar{X} = \frac{\sum_{j=1}^n x_j}{n}$$

Formula 3.2 Sample mean

$$S = \sqrt{\frac{\sum_{j=1}^n x_j^2}{n} - \bar{X}^2}$$

Formula 3.3 Sample standard deviation

These formulas were already included in QGIS, through the Hotspot Analysis Plugin (Oxoli et al., 2017). By using this plugin, z-scores and p-values are obtained under the Complete Spatial Randomness hypothesis (null hypothesis) of the G_i^* statistic. The z-score is a standard deviation, the p-value is a probability, and both tell whether the null hypothesis can be rejected or not. When the null hypothesis is rejected, there is a statistically significant spatial pattern, that lead to identify hotspots.

The values of z-score and p-value for which the null hypothesis is rejected, are presented in Table 3.7.

z-core	p-value	Confidence level
< -1.65 or > 1.65	< 0.10	90%
< -1.96 or > 1.96	< 0.05	95%
< -2.58 or > 2.58	< 0.01	99%

Table 3.7 Values of z-score and p-value to reject the null hypothesis

When the null hypothesis is rejected under positive values of z-score, then that is considered as a hotspot. Hotspots represent atypical high-value location surrounded by other high-value location as well; whereas not significant points represent location in which local values are likely random distributed.

For the purpose of this research, the three groups of victim's severity were studied (see subsection 3.1.2 Crash data parameters). Thereby, it was possible to find hotspots of intersections with all cyclist victims, with only slightly injured cyclists, and with seriously injured and killed cyclists. Moreover, the hotspots were chosen for a confidence level of 99% (p-value < 0.01 and z-value > 2.58), and filtering only intersections with cyclist victims. Additionally, in order to compute the local Getis-Ord G_i^* statistic a fixed distance band of 200 meters was used, guaranteeing a coverage of at least two urban streets per intersection in most cases.

4 Implementation

4.1 Study area

Taking into account the crash data availability, this research's methodology was implemented after choosing one city from Germany, France and Great Britain. Since the data was only available for the federal state of Saxony, in Germany, the selection of the city was made consequently in that area. For the other two countries, it was possible to consider all their cities for the definition of the study area.

After exploring several socio-demographic and transportation-related parameters, the variables of population and area were picked to prioritize the selection. These variables, which are extensively available and used in other related studies (Branion-Calles et al., 2020; Klanjčić et al., 2022; Santacreu, 2018), allowed to reach comparability in terms of population density between cities in Saxony, France and Great Britain.

To find the trio of cities, the process started by filtering cities between 500,000 and one million inhabitants in Saxony, France and Great Britain (see Figure 4.3).

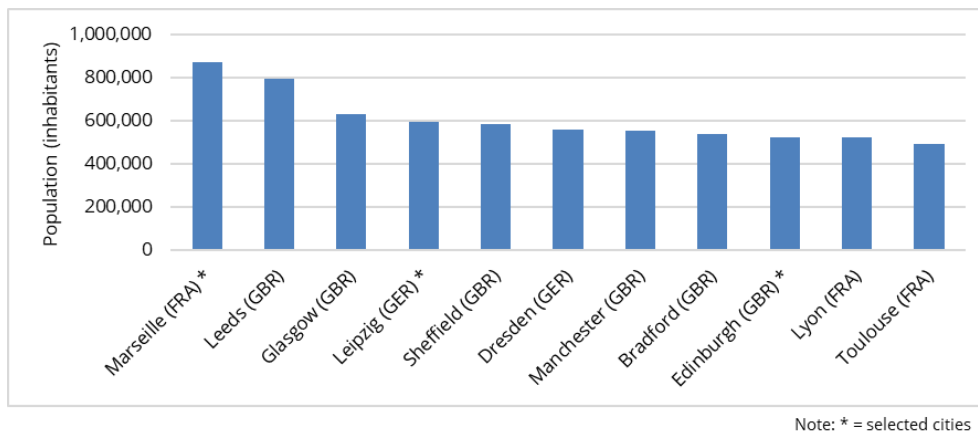


Figure 4.3 Population in potential cities for study area

Later, the combination of three cities from different countries with the smallest difference in area and population density was found (see Figure 4.4). This led to choose the cities of Leipzig in Saxony, Marseille in France, and Edinburgh in Great Britain.

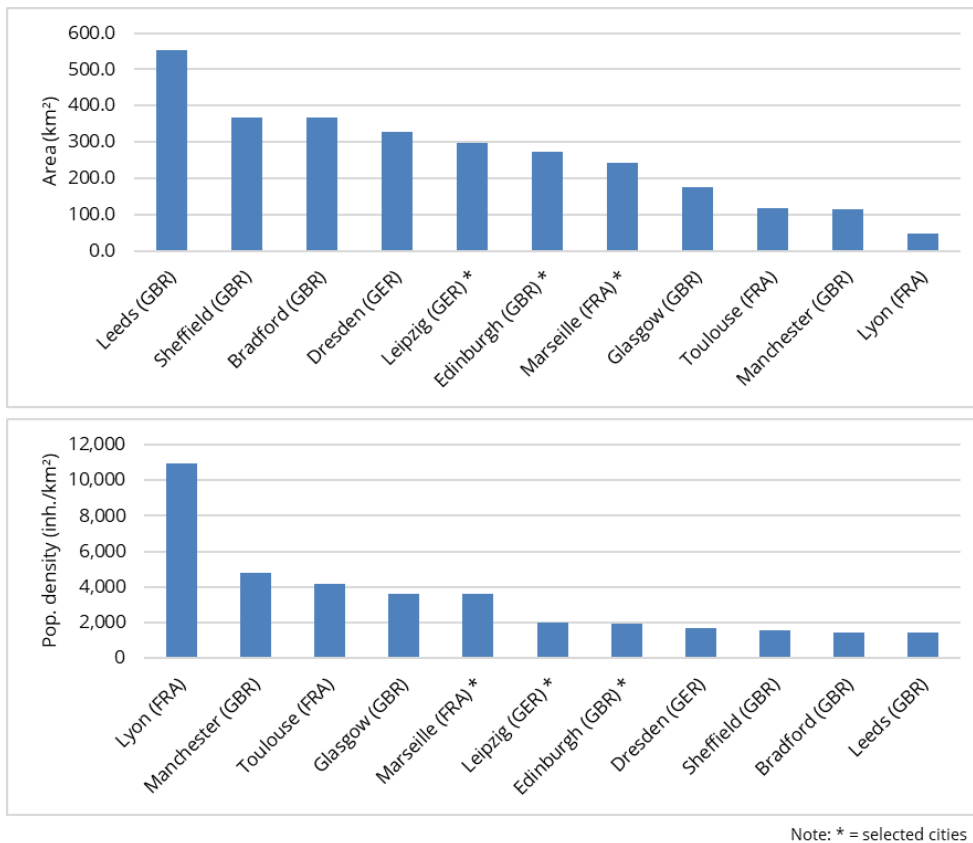


Figure 4.4 Area, and population density in potential cities for study area

The localisation of Leipzig, Marseille and Edinburgh is shown in Figure 4.5.

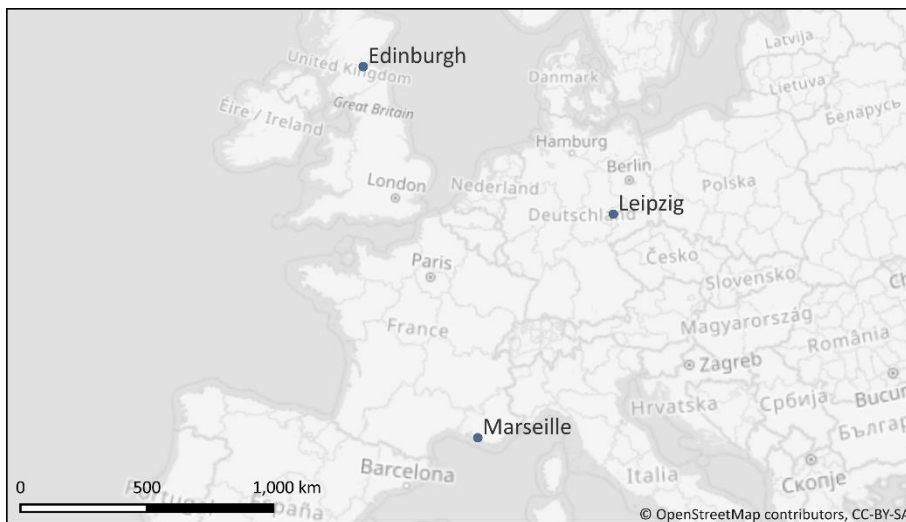


Figure 4.5 Localisation of chosen cities for the study area

4.1.1 Other variables

In addition to crash and OSM data, variables related with socio-demographics and mobility patterns were collected from different sources.

Data regarding population and modal split was collected from different local sources (Table 4.8). In this case, population datasets for Germany, France and Great Britain had information for 2019, since it was the common and most updated year for the three countries. Regarding modal split, the most updated information was gathered for the three chosen cities. However, in this case, it was not possible to establish a common year of the data for them.

Variable	Year	Country/City	Source
Population	2019	Germany	Federal and state statistical offices of Germany (2022)
		France	French National Institute of Statistics and Economic Studies (2021)
		Great Britain	British Office for National Statistics (2021)
Modal split	2018	Leipzig	Mobility Office of Leipzig (2020)
	2017	Marseille	General direction for sustainable city and expansion (2020)
	2019	Edinburgh	Transport and Environment Committee (2021)

Table 4.8 Sources of population and modal split data

4.2 OpenStreetMap data

During the process of getting data from OSM to PostgreSQL, it was required to define the parameters of the imported elements, according to the values available for each city of the study area, and focusing on the specifications required for this research. This included elements associated with city boundaries, cycling infrastructure, road infrastructure, roundabouts and traffic signals, as presented in Table 4.9.

Imported elements	Key	Values
City boundaries	<i>admin_level</i>	6 (for Leipzig and Edinburgh), 8 (for Marseille)
Cycling infrastructure	<i>bicycle</i>	customers, designated, destination, mtb, official, permissive, private, use_sidepath, yes
	<i>cycleway</i>	advisory, crossing, designated, lane, lane:right, left, opposite, opposite_lane, opposite_share_busway, opposite_track, segregated, separate, share_busway, shared, shared_lane, sidepath, sideride, sidewalk, soft_lane, track, track; lane, yes
	<i>cycleway:right</i>	lane, opposite, opposite_lane, separate, share_busway, shared_lane, soft_lane, track
	<i>cycleway:left</i>	advisory, buffered_lane, lane, opposite, opposite_lane, opposite_track, separate, separate, share_busway, shared_lane, soft_lane, track, use_sidepath, yes
	<i>cycleway:both</i>	lane, separate, separate, share_busway, shared_lane, track, yes
Road infrastructure	<i>highway</i>	motorway, trunk, primary, secondary, tertiary, residential
Roundabouts	<i>junction</i>	roundabout
Traffic signals	<i>highway</i>	traffic_signals

Table 4.9 Tags applied to the imported OSM data

4.2.1 Identification of intersections

After importing the data, the layers of each city were merged, joined by location with the city boundary, and clipped. This process allowed to have the features associated within each city.

In the Figure 4.6 a particular example is shown, including the imported OSM elements in the surroundings of the intersection Gerberstraße - Tröndlinring in Leipzig. As it can be noted from the figure, OSM data includes every single way element according to each carriageway and each change of directionality of the streets (turning lanes), which leads to present concentration of ways and nodes in some intersections. In this case there is one intersection (Gerberstraße - Tröndlinring) with 16 nodes and nine external ways coming to them. This is one junction, and based on this research's scope, the aim was to obtain one point representing it instead of 16. Additionally, aiming at simplifying the ways, instead of nine, it should be four streets coming to the intersection (in this case, one from each cardinal point). This was achieved after applying the clustering and several spatial functions, which were presented in a previous chapter (see 3 Methods).

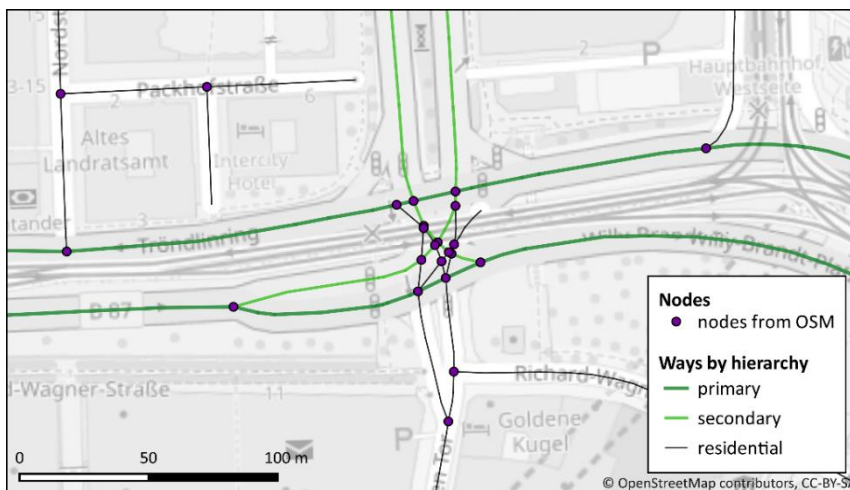


Figure 4.6 Output of nodes and ways using pgRouting (intersection Gerberstraße - Tröndlinring in Leipzig)

Following this example, in Figure 4.7 it is possible to see the output of the clustered nodes after applying the first DBSCAN.

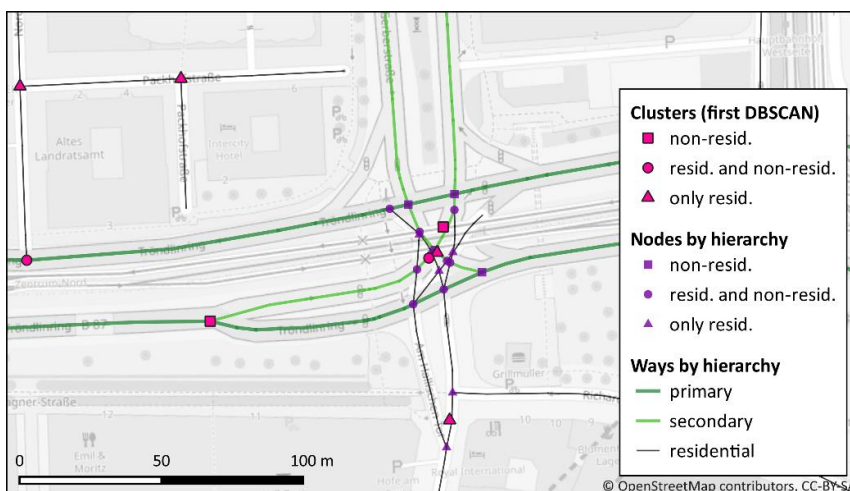


Figure 4.7 Clusters after first DBSCAN for identifying intersections (example in Leipzig)

After running the three DBSCANs and combining the clusters gathered, it was feasible to get one point per intersection, as desired (see Figure 4.8).

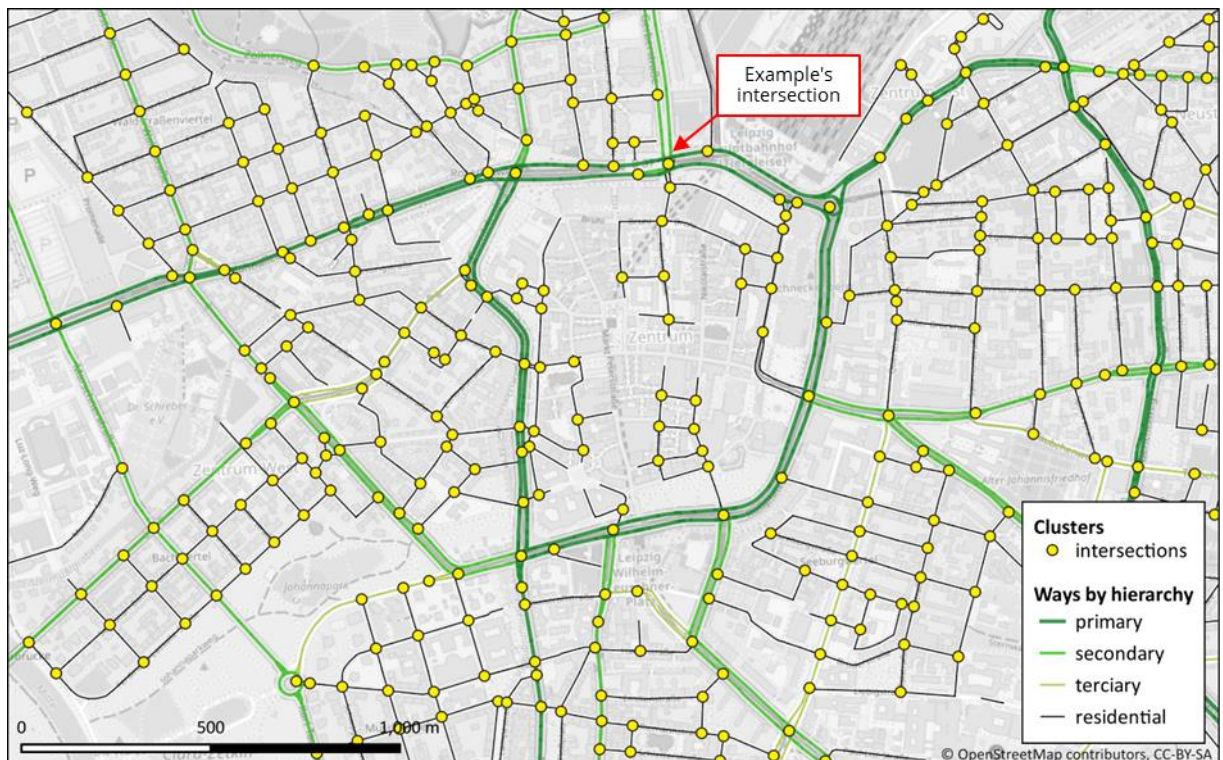


Figure 4.8 Output of intersections after clustering (example in Leipzig)

4.2.2 Influence zones

Next, using the specific parameters showed in the methodology (see Figure 3.2,) the influence zones for the intersections were defined (see Figure 4.9).

4.2.3 Classification of intersections

4.2.3.1 Number of arms

Each way element was aggregated and clustered in order to get simplified streets. The final output was a group of polygons of different width (minimum 0.1 meters) representing the simplification of the ways (see Figure 4.9). Later, the influence zone at each intersection was intersected with the simplified ways, and this led to find the number of arms per intersection.



Figure 4.9 Output of influence zones and simplification of ways (example in Leipzig)

Intersections of one node with residential and non-residential ways had a buffer's radius of 15 meters of the influence zone, which was convenient for most cases. However, in a few intersections the residential way was connecting non-residential ways with bigger geometry, and the intersection's buffer was smaller than the street's buffer (see Figure 4.10 at intersections on the primary road). Thus, the counting was reporting two arms, when there were actually three of them.



Figure 4.10 Identification of number of arms by intersecting influence zones and ways simplification

After solving this issue, the final output of the arms counting was obtained as showed in Figure 4.11.



Figure 4.11 Output of intersection's classification by arms counting

4.2.3.2 Accuracy evaluation

In order to check the accuracy of this methodology in the study area, 30 intersections per type were randomly chosen in Leipzig, Marseille and Edinburgh. Later, for each of them the number of arms was manually found using the traffic network from OSM, and then compared with the value obtained from the applied methodology. As expected, the highest accuracy value was got for the simple intersections (three arms) with 97%, and the lowest one for the complex intersections (five or more arms) with 87%. The overall accuracy of the applied methodology remained high with 93% (see Table 4.10).

		Expected results				Total	Accuracy
Obtained results	Typology	3 arms	4 arms	5 or more arms	roundabout		
	3 arms	29	1	0	0	30	97%
	4 arms	2	28	0	0	30	93%
	5 or more arms	2	2	26	0	30	87%
	roundabout	1	1	0	28	30	93%
Overall accuracy							93%

Table 4.10 Accuracy of intersection's categorisation by arm counting

4.3 Crash data

The crash data was explored using PostgreSQL, and imported to QGIS as a feature layer based on the coordinates provided. This allowed to have a visual overview of the crash events in the different cities. During this step, it was found that the coordinate values for crashes in France did not have decimal separator due to data codification of French local authorities. Nevertheless, this issue was fixed, and the position of the crash events could be correctly visualised.

Once the database was corrected including the data from Saxony, France and Great Britain, in total 653,657 reported victims were obtained between 2015 and 2017. However, not all of them were referenced, due to lack of information when collecting the crash reports. For Saxony and Great Britain, 100% of the reported victims had coordinates. But, in the French data, the reported victims with coordinates rose to 69% (see Figure 4.12).

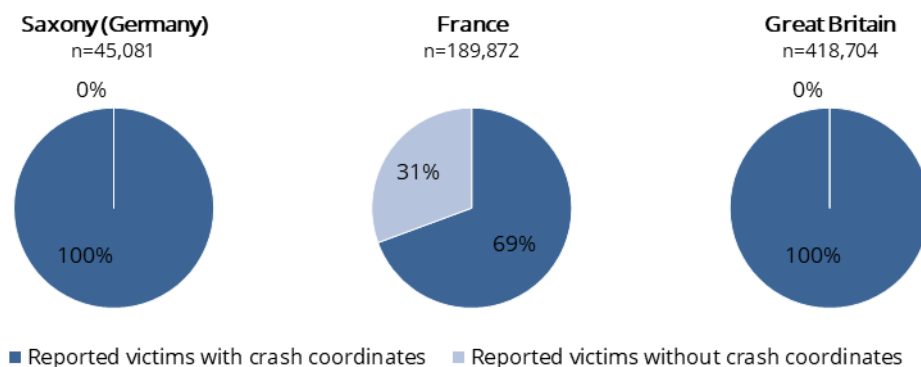


Figure 4.12 Reported traffic crash victims with and without coordinates in original database (2015-2017)

Looking at the distribution of victims who were riding a bicycle, Saxony reported 12,225 and Great Britain 55,558, all of them with the respective coordinates. France reported 13,086, from which the 74% included coordinates in the database (see Figure 4.13).

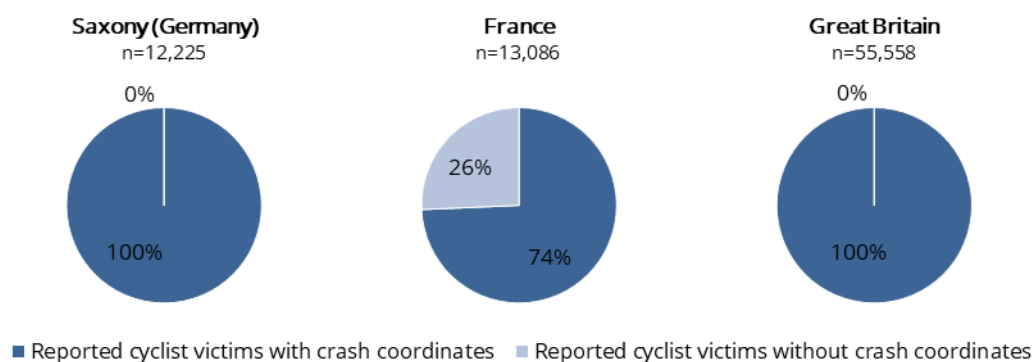


Figure 4.13 Reported cyclist victims with and without coordinates in original database (2015-2017)

5 Results

After implementing the methodology in Leipzig, Marseille and Edinburgh, in the following sections the obtained results are presented. For additional information about the results tables, it is suggested to see Appendix B.

5.1 General comparison between cities

Once the three cities of the study area were identified, more specific information was collected looking forward to having a wider characterization of them. The different variables are presented in the next table:

Variables	Leipzig	Marseille	Edinburgh	Mean	Std. Deviation
General					
Population (2019)	593,145	870,731	524,930	662,935	183,160
Area (km ²)	297.9	242.1	273.1	271.1	27.9
Population density (inh./km ²)	1,991.2	3,596.0	1,921.8	2,503.0	947.2
Minimum elevation (m)	97	0	0	32	56
Maximum elevation (m)	178	652	251	360	255
Infrastructure and cycling					
Length of cycling infrastructure (km)	967.1	241.5	521.5	576.7	365.9
Length of traffic network (km)	1,572.2	1,538.7	1,513.5	1,541.4	29.5
Cycling inf. by area (km/km ²)	3.2	1.0	1.9	2.1	1.1
Traffic network by area (km/km ²)	5.3	6.4	5.5	5.7	0.6
Cycling modal split (%)	18.7	1.0	4.0	7.9	9.5
Crash victims (yearly average)					
Victims	2,305.3	2,293.0	978.0	1,858.8	762.8
Cyclist victims	950.3	39.3	212.3	400.7	483.8
Seriously injured and killed cyclists	145.0	12.7	35.7	64.4	70.7
Total victims / 100,000 inh.	388.7	263.3	186.3	279.4	102.1
Cyclist victims / 100,000 inh.	160.2	4.5	40.4	68.4	81.5
Seriously injured and killed cyclists / 100,000 inh.	24.4	1.5	6.8	10.9	12.0

Note: Highlighted values are the highest per row

Table 5.11 Comparison of selected cities through road infrastructure and safety variables

As already stated in a previous section (see 4.1 Study area), the trio of cities was chosen based on area and population density. Yet, it was possible to find differences between each city in other aspects, as shown in Table 5.11.

In terms of geography, Leipzig is an inland city, mostly flat and with an elevation difference of less than 100 m. On the other hand, Marseille and Edinburgh are coastal and hilly cities. In Marseille, the maximum elevation is 652 m, and in Edinburgh is 251 m. In both cases, these heights are found on the outskirts of the urban area.

Looking at the cycling patterns and infrastructure, Leipzig had the longest cycling network, and also the biggest cycling modal split, even though the traffic network's length was rather similar in the three cities. This pattern shows that having more facilities for cyclist, encourage more people to cycle.

For road victims, the highest values were found in Leipzig for each of the variables. Now, comparing Marseille and Edinburgh, more victims were reported in Marseille, however, there were more cyclist victims in the British city.

Intersecting the points from the crash database with the polygons of the study area, in total there were 6,916 victims in Leipzig, 6,879 in Marseille, and 2,934 in Edinburgh, between 2015 and 2017. The distribution of victims by mode of transportation is shown in Figure 5.14, where it can be seen different patterns between the cities. In Leipzig the highest proportion of crash victims was made by cyclists (41%), whereas in Marseille it was the motorcyclists (43%), and in Edinburgh the car drivers and passengers (41%). Since this research was focused on cyclist victims, only them were considered for further analyses. In this case, for Leipzig the proportion of cyclist victims was 41%, which is equivalent to 2,851; in Marseille, the proportion of cyclist victims was 2%, equivalent to 118; and in Edinburgh was 22%, with 637 victims who were riding a bike.

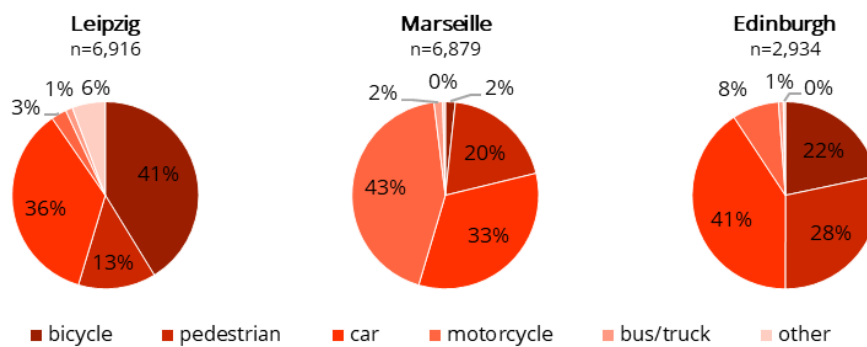


Figure 5.14 Victims per city by vehicle type, between 2015 and 2017

5.2 Count of intersections

Once the nodes from the OSM traffic network were clustered, a total of 16,162 intersections within the boundaries of the three cities were identified. In this section, different distributions are presented focusing on the assessed infrastructure parameters: number of arms, presence of cycling infrastructure, and presence of traffic signals.

5.2.1 Intersections by typology based on number of arms

The distribution showing the different typologies per city is presented in the Figure 5.15

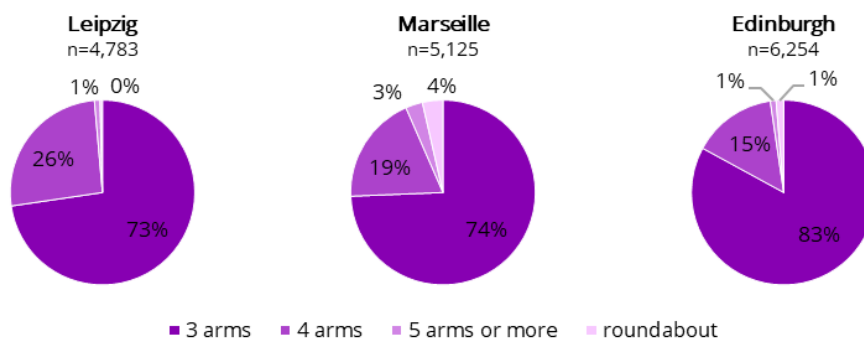


Figure 5.15 Percentage of intersections by type and city

According to this results, 3-arm intersections were the majority in each city, with the highest rate found in Edinburgh (83%). Moreover, it can be noted that in each city more than the 90% of the junctions had three or four arms. For roundabouts and intersections with five or more arms, the highest proportion comparing the other cities was found in Marseille (4% for roundabouts, and 3% for junctions with five or more arms).

5.2.2 Intersections by category based on infrastructure attributes

The 16,162 intersections were also categorised according to the infrastructure parameters, including number of arms (intersection type), presence of cycling infrastructure and presence of traffic signals. In the Figure 5.16 it is possible to visualise the distribution of the infrastructure characteristics based on cycling infrastructure and traffic signal, per intersection type and city.

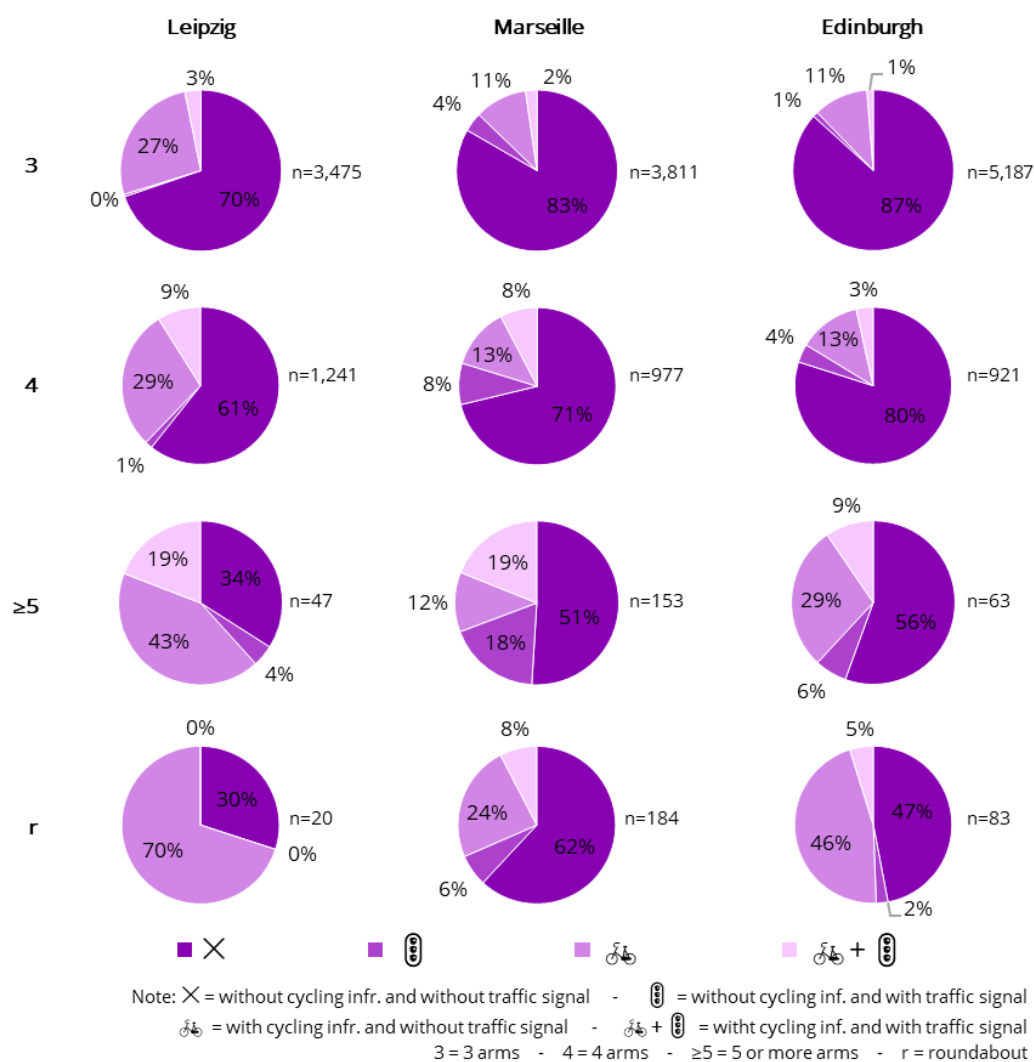


Figure 5.16 Percentage of intersections by city and category based on infrastructure attributes

According to these results, and looking at the distribution of the categories by city and number of arms, in Marseille and Edinburgh the biggest proportion of junctions was the one without cycling infrastructure and without traffic signals, for each case (regardless the number of arms). In these cities, specially the 3-arm and 4-arm junctions had the highest rates of intersections without cycling infrastructure and without traffic signals with at least 61%.

For Leipzig, when having 3-arm and 4-arm junctions, the highest proportion belonged to intersections without cycling infrastructure and without traffic signals. However, for roundabouts or intersections with five or more arms in the German city, the highest proportion was found at intersections with cycling infrastructure but without traffic signals.

Junctions with both, cycling infrastructure and traffic signals, were rather few in each city. For roundabouts, 3-arm and 4-arm intersections the proportion of junctions having cycling infrastructure and traffic signals was not bigger than 9%. Regarding, the highest values for intersections with cycling infrastructure and traffic signals, these were found for intersections with five or more arms in Leipzig and Marseille, where 19% of the intersections fulfilled both characteristics.

5.3 Intersections with cyclist victims

Once the victims were matched within the intersections, a total of 1,150 intersections with cyclist victims were found in the study area. In this section, different distributions of intersections with victims are presented based on the assessed infrastructure parameters.

5.3.1 Cyclist victims in intersections by type

From a total of 1,150 intersections with cyclist victims in the study area, Leipzig presented the highest amount (837). The results obtained per city are presented in the Figure 5.17.

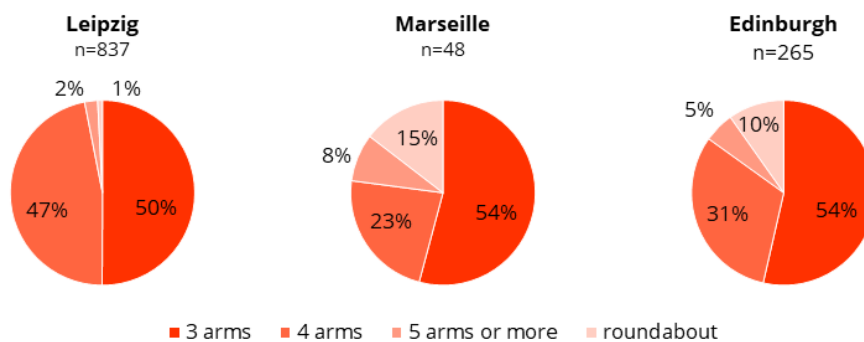


Figure 5.17 Percentage of intersections with cyclist victims by type and city

Based on these results, it was found that the highest distribution of junctions with cyclist victims belonged to 3-arm intersections, with at least 50% in each city. In other words, from the total amount of intersections with victims, more than half were intersections with three arms, and this was a common pattern in each city. Nevertheless, the distribution of the other types had more differences between the cities. For instance, in roundabouts and intersections with five or more arms, the proportion reached 3% in Leipzig, whereas in Marseille it was the 23%, and in Edinburgh the 15%.

From a different perspective, it was also possible to look at the distribution of junctions with cyclist victims among all the intersections sharing the same typology (see Figure 5.18). Hence, this representation, considered not only junctions with cyclist victims, but also without them, which allowed to visualise the level of incidence of intersections with victims according to its typology.

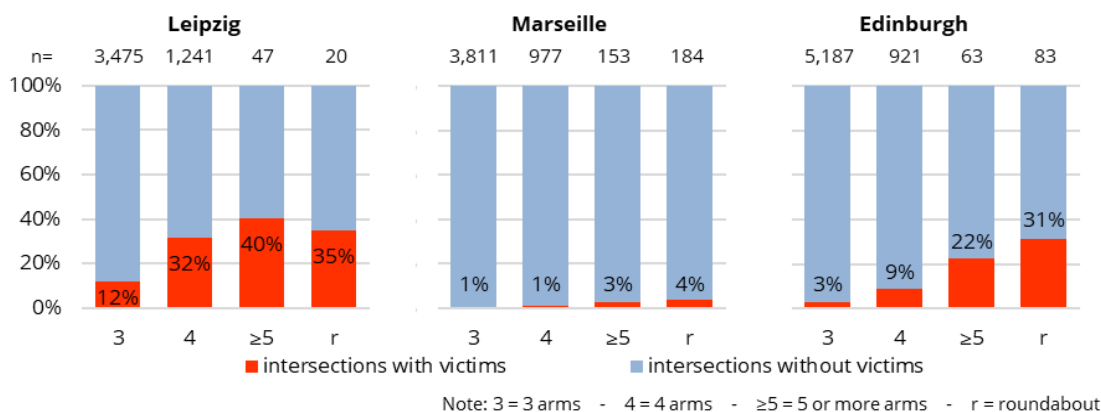


Figure 5.18 Percentage of intersections with and without cyclist victims by type

In this case the results showed that the highest rates were presented in junctions categorized as five or more arms, and roundabouts. For instance, from all the intersections with five or more arms in Leipzig, 40% had cyclist victims; in Marseille, 3%; and in Edinburgh, 22%. From all the roundabouts in Leipzig, 35% had cyclist victims; in Marseille, 4%; and in Edinburgh, 31%.

As stated previously, in the three cities, 3-arm junctions were having the highest distribution among intersections with cyclist victims (Figure 5.17). However, when focusing on the distribution of intersections of cyclist victims among the total number of intersections of the same type, this is also including the intersections without victims (Figure 5.18), it was found the 3-arm intersections had the lowest rates in the three cities. As a matter of fact, from all the junctions with three arms in Leipzig, 12% had cyclist victims; in Marseille it was the 1%; and in Edinburgh it was the 3%.

5.3.2 Cyclist victims in intersections by infrastructure attributes

The resulting 1,150 intersections with cyclist victims were also categorised taking into account the different infrastructure attributes.

In the Figure 5.19 it is possible to visualise the distribution of intersections with victims per intersection type and city, according to its characteristics based on cycling infrastructure and traffic signals.

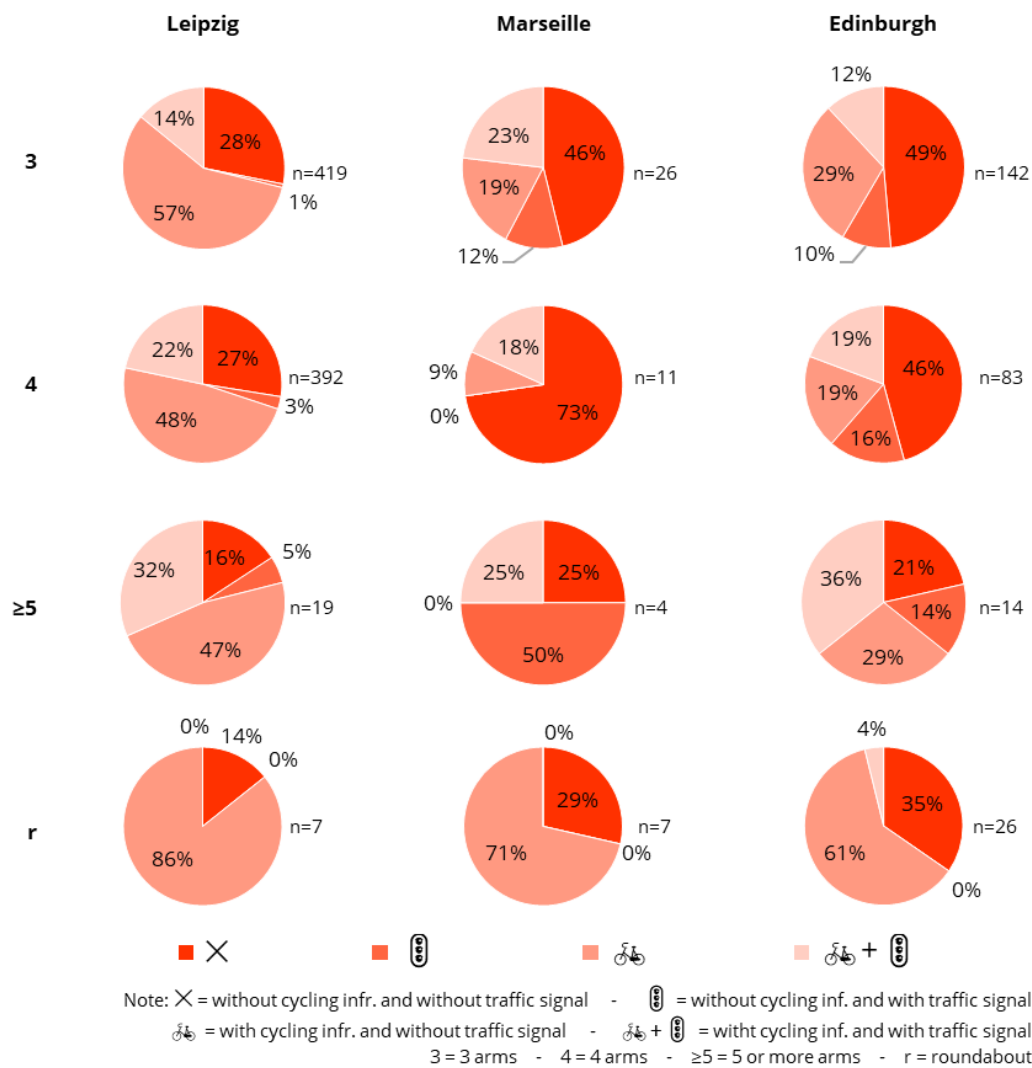


Figure 5.19 Percentage of intersections with cyclist victims by city and infrastructure attributes

According to these results the distribution of junctions with cyclist victims followed a special pattern per city. In Leipzig the highest values of intersections with victims, were the ones with cycling infrastructure and no traffic signals, regardless the number of arms. In Marseille and Edinburgh the highest values were found at junctions with no cycling infrastructure and no traffic signals, when having three and four arms; and at intersections with cycling infrastructure and no traffic signals, when having roundabouts.

When comparing the intersections with cyclist victims among all the intersections of the same type, it was possible to identify the distribution of victims considering the presence of cycling infrastructure and presence of traffic signals, by number of arms (see Figure 5.20). In this case, the highest percentage per city of 3-arm and 4-arm junctions with victims, were the ones with cycling infrastructure and traffic signal.

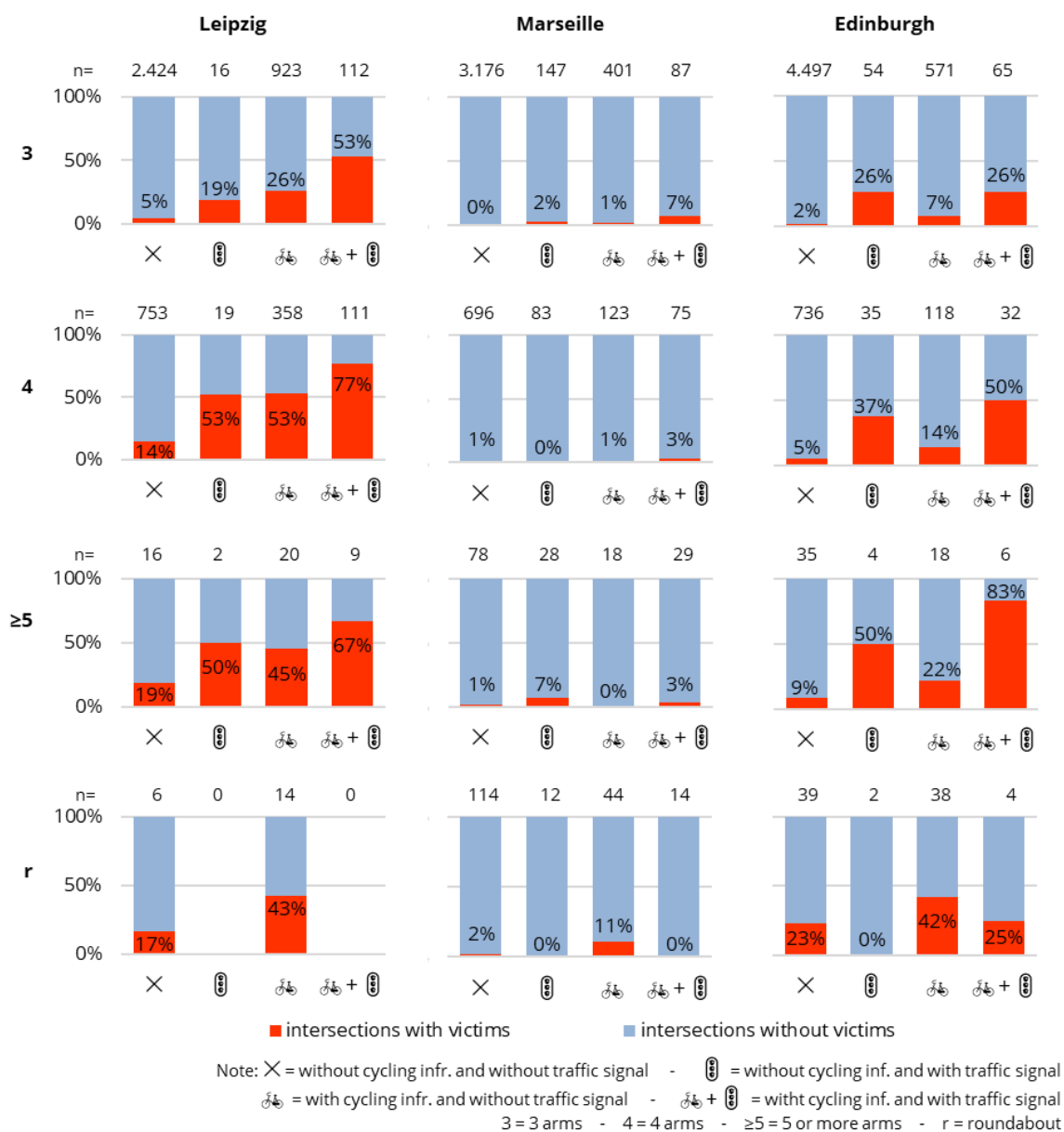


Figure 5.20 Percentage of intersections with and without cyclist victims by infrastructure attributes

Looking at the intersections with five or more arms, the trend was similar between Leipzig and Edinburgh, with the highest distribution in junctions with both infrastructure parameters (67% for Leipzig and 83% for Edinburgh). Whereas for Marseille, the highest distribution in this intersection type corresponded to intersections without cycling infrastructure but with traffic signal. For roundabouts in the three cities, the biggest proportions were found when having cycling infrastructure but no traffic signal.

5.4 Intersections with slightly injured cyclists

Based on the severity level, the results were explored at the intersection level, this time filtering the cyclist victims who had slight injuries. This led to a total of 1,027 intersections with slightly

injured cyclists in the study area. In this section, different distributions of intersections with slightly injured cyclists are presented focusing on the assessed infrastructure parameters.

5.4.1 Slightly injured cyclists in intersections by type

From a total of 1,027 intersections with slightly injured cyclists in the study area, Leipzig had the highest amount (756). The results obtained per city are presented in the Figure 5.21

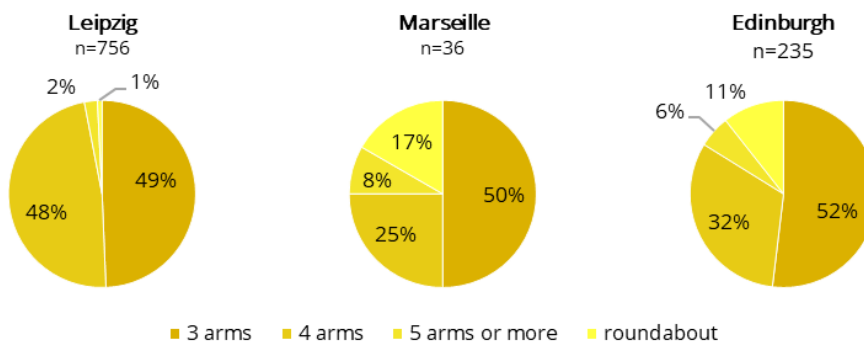


Figure 5.21 Percentage of intersections with slightly injured cyclists by type and city

Based on these results, it was found that the highest distribution of junctions with slightly injured cyclists belonged to 3-arm intersections, with at least 49% in each city. In other words, from the total amount of intersections with slight victims, almost half of them had three arms, and this was a common pattern in each city. Nevertheless, the distribution of the other types has more differences between the cities. For instance, in roundabouts and junctions with five or more arms, the proportion reached 3% in Leipzig, whereas in Marseille it was the 23%, and in Edinburgh the 17%.

Regarding the distribution of intersections with and without slightly injured cyclists, results can be seen in Figure 5.22.

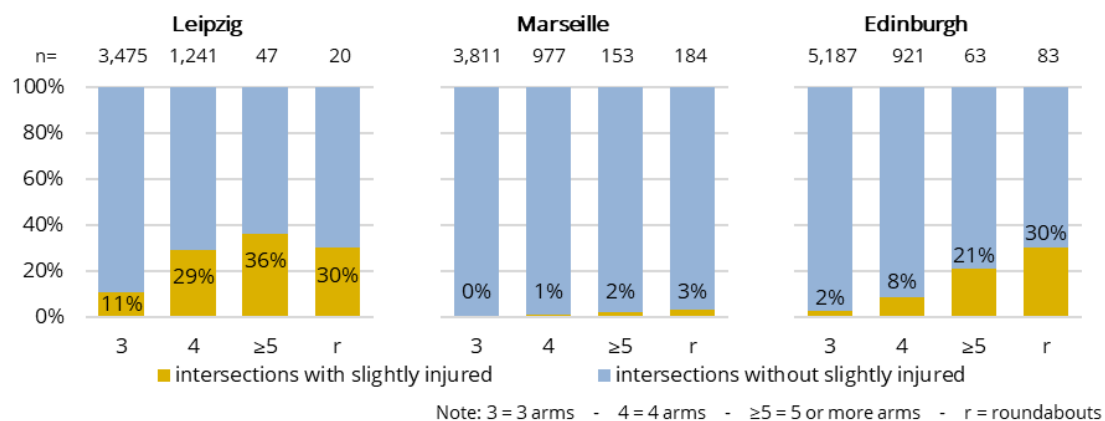


Figure 5.22 Percentage of intersections with and without slightly injured cyclists by type

According to these results, for the three cities, the highest rates of junctions with slight injured were found in those with five or more arms, and roundabouts. In Leipzig, the highest rate was presented in intersections with five or more arms (36%), whereas in the other two cities it was found in roundabouts (3% in Marseille and 30% in Edinburgh).

5.4.2 Slightly injured cyclists in intersections by infrastructure attributes

This time, the 1,027 intersections with slightly injured cyclists were chosen to run the analysis considering the different infrastructure attributes. In the Figure 5.23 it is possible to visualise the distribution of intersections with slight injured per typology and city, according to its characteristics based on cycling infrastructure and traffic signal.

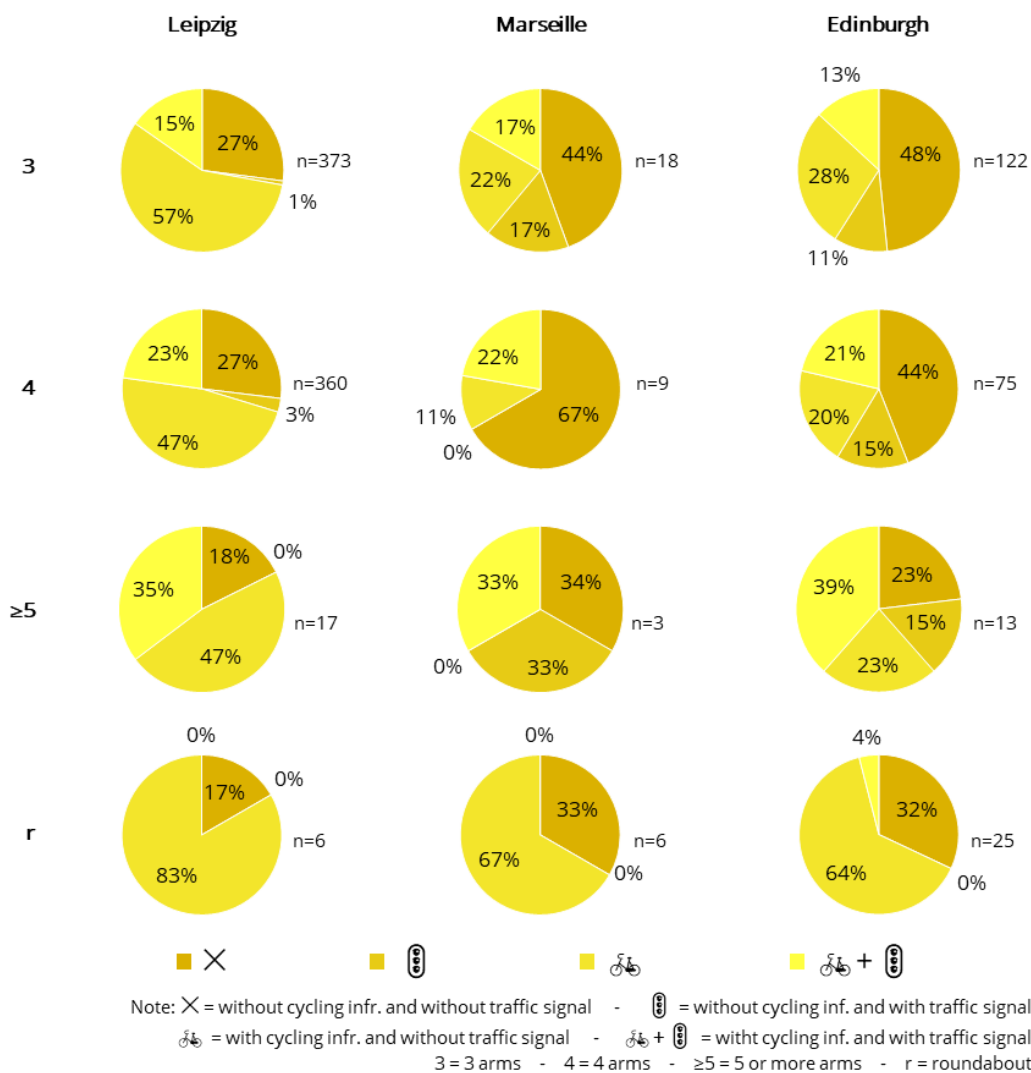


Figure 5.23 Percentage of intersections with slightly injured cyclists by city and infrastructure attributes

Based on these results, it was found that the distribution of intersections with slightly injured cyclists follow a special pattern per city. In Leipzig the highest values of intersections with slight injured, were the ones with cycling infrastructure and no traffic signals, regardless the number of arms. In Marseille the highest distributions were found at intersections with no cycling infrastructure and no traffic signals, no matter the number of arms except roundabouts. In Edinburgh the biggest proportions were at intersections with no cycling infrastructure and no traffic signals, when having three and four arms.

For intersections of five or more arms, in Leipzig the highest distribution of intersections with slight injured had cycling infrastructure but no traffic signals (47%). In Marseille the distribution for this typology was equally distributed for junctions with cycling infrastructure and traffic signals, junctions with neither of those parameters, and junctions with only traffic signal. In Edinburgh, the 39% of intersections with five or more arms and slight injured, had cycling infrastructure and traffic signal, and this was the highest distribution in this city.

Looking at the roundabouts, the trend remained similar for the three cities, and most of the intersections with slight injured were the ones with cycling infrastructure and without traffic signals. In Leipzig the proportion was 83%, in Marseille 67%, and in Edinburgh 64%.

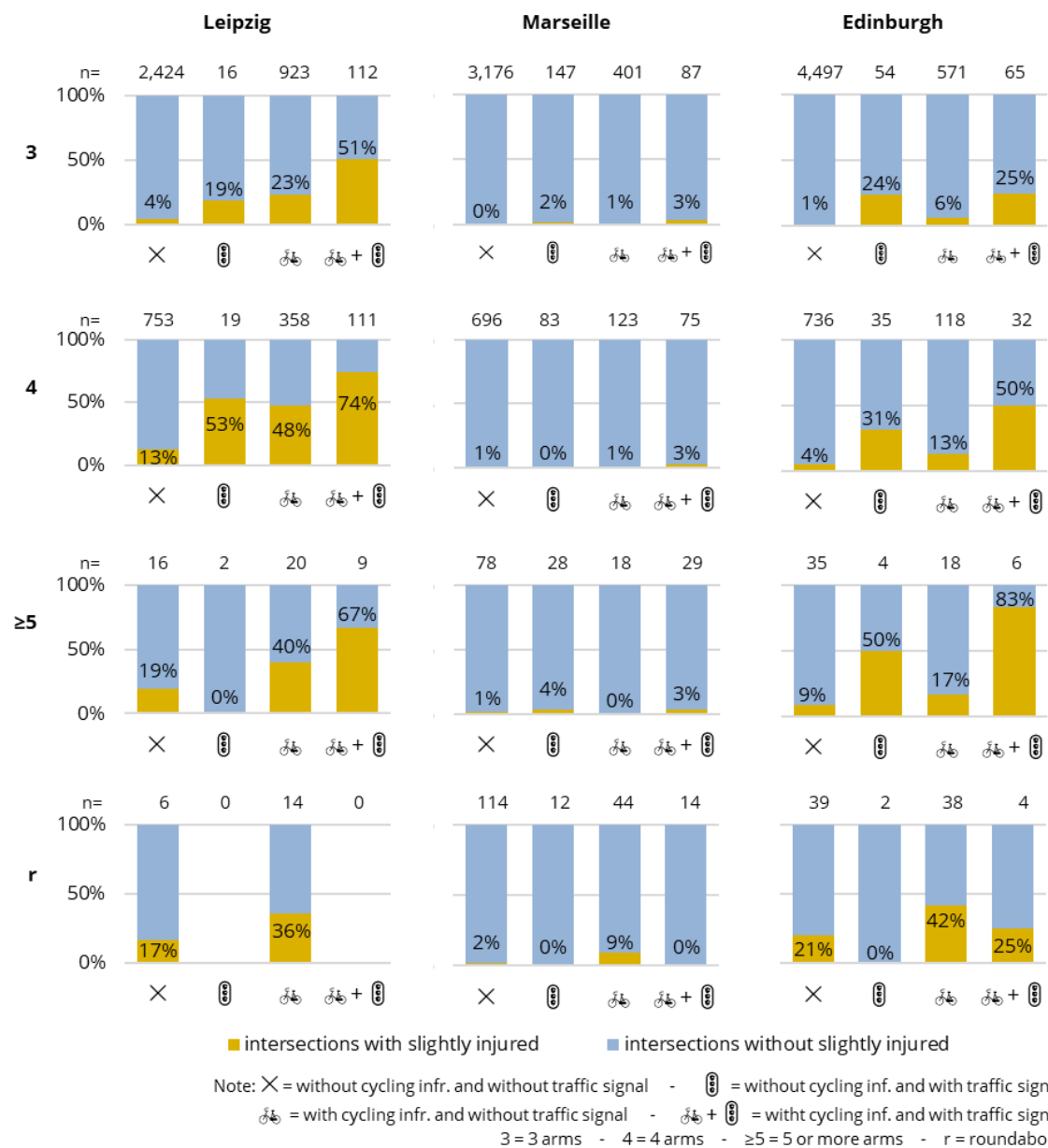


Figure 5.24 Percentage of intersections with and without slightly injured cyclists by infrastructure attributes

When comparing the intersections with slightly injured cyclists among all the intersections of the same type, it was possible to identify their distribution considering the presence of cycling infrastructure and presence of traffic signals, by number of arms (see Figure 5.24).

In this case, the highest percentage per city of 3-arm and 4-arm junctions with victims, were the ones with cycling infrastructure and traffic signal. In Leipzig, 51% of 3-arm intersections with both infrastructure parameters presented slightly injured cyclist; in Marseille, 3%; and in Edinburgh, 25%. When looking at the 4-arm intersections with the same infrastructure parameters the values were 74% for Leipzig, 3% for Marseille, and 50% for Edinburgh.

Looking at the junctions with five or more arms, Leipzig and Edinburgh had the highest distribution in those with both infrastructure parameters (67% for Leipzig and 83% for Edinburgh). Whereas for Marseille, the highest distribution in this intersection type corresponded to intersections without cycling infrastructure but with traffic signal (4%).

For roundabouts in the three cities, the biggest proportions were found when having cycling infrastructure but no traffic signal. In Leipzig the distribution reached 36%, in Marseille 9%, and in Edinburgh 42%.

5.5 Intersections with seriously injured and killed cyclists

Based on the severity level, the results were explored at the intersection level, this time filtering the cyclist victims who had serious injuries and who were killed. This led to a total of 274 intersections with this type of victims in the study area. In this section, different distributions of intersections with seriously injured and killed cyclists are presented based on the assessed infrastructure parameters.

5.5.1 Seriously injured and killed cyclists in intersections by type

From a total of 274 intersections with seriously injured and killed cyclists in the study area, Leipzig was the city with the highest amount (210). The results obtained per city are presented in the Figure 5.25.

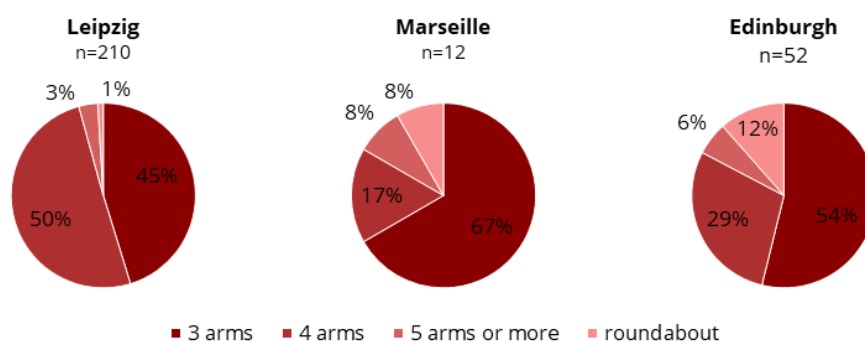


Figure 5.25 Percentage of intersections with seriously injured and killed cyclists by type and city

As regards these results, in Marseille and Edinburgh it was found that the highest distribution of intersections with seriously injured and killed cyclists had three arms, with at least 54%. However,

in Leipzig the biggest proportion of junctions with these victims had four arms, with 50%. In other words, from the total amount of intersections with seriously injured and killed cyclists, more than half were intersections with three arms in Marseille and Edinburgh, whereas in Leipzig more than half were intersections with four arms. In roundabouts and junctions with five or more arms the proportion reached 4% in Leipzig, whereas in Marseille it was the 16%, and in Edinburgh the 18%.

From a different perspective, it was also possible to look at the distribution of intersections with seriously injured and killed cyclists among all the intersections of the same type (see Figure 5.26). In this case the results are not similar between cities, but in each of them the highest rates are presented in junctions classified as five or more arms, and roundabouts. For instance, from all the intersections with five or more arms in Leipzig, 15% had seriously injured and killed cyclists; in Marseille, 1%; and in Edinburgh, 5%. From all the roundabouts in Leipzig, 10% had seriously injured and killed cyclists; in Marseille, 1%; and in Edinburgh, 7%.

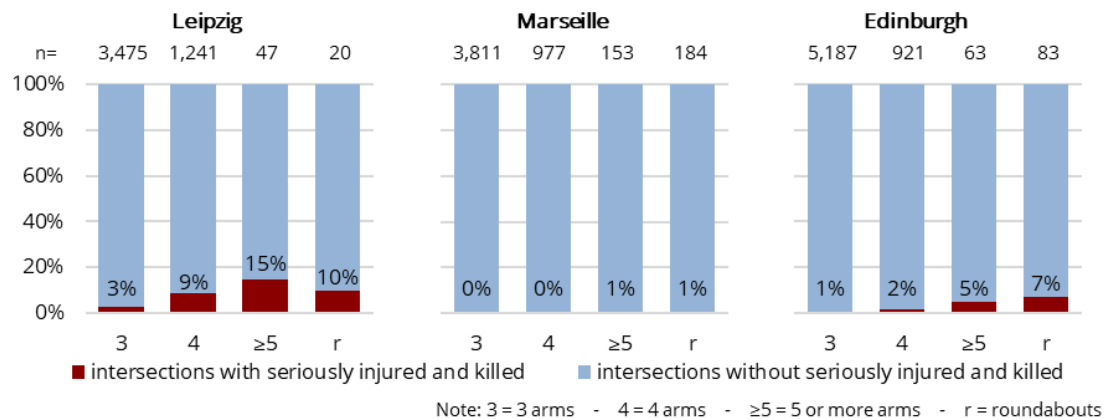


Figure 5.26 Percentage of intersections with and without seriously injured and killed cyclists by type

5.5.2 Seriously injured and killed cyclists in intersections by infrastructure attributes

This time, the 274 serious injured and dead victims were extracted for the analysis considering the different infrastructure attributes. In Figure 5.27 it is possible to visualise the distribution of intersections with these specific victims per typology and city, according to its characteristics based on cycling infrastructure and traffic signal.

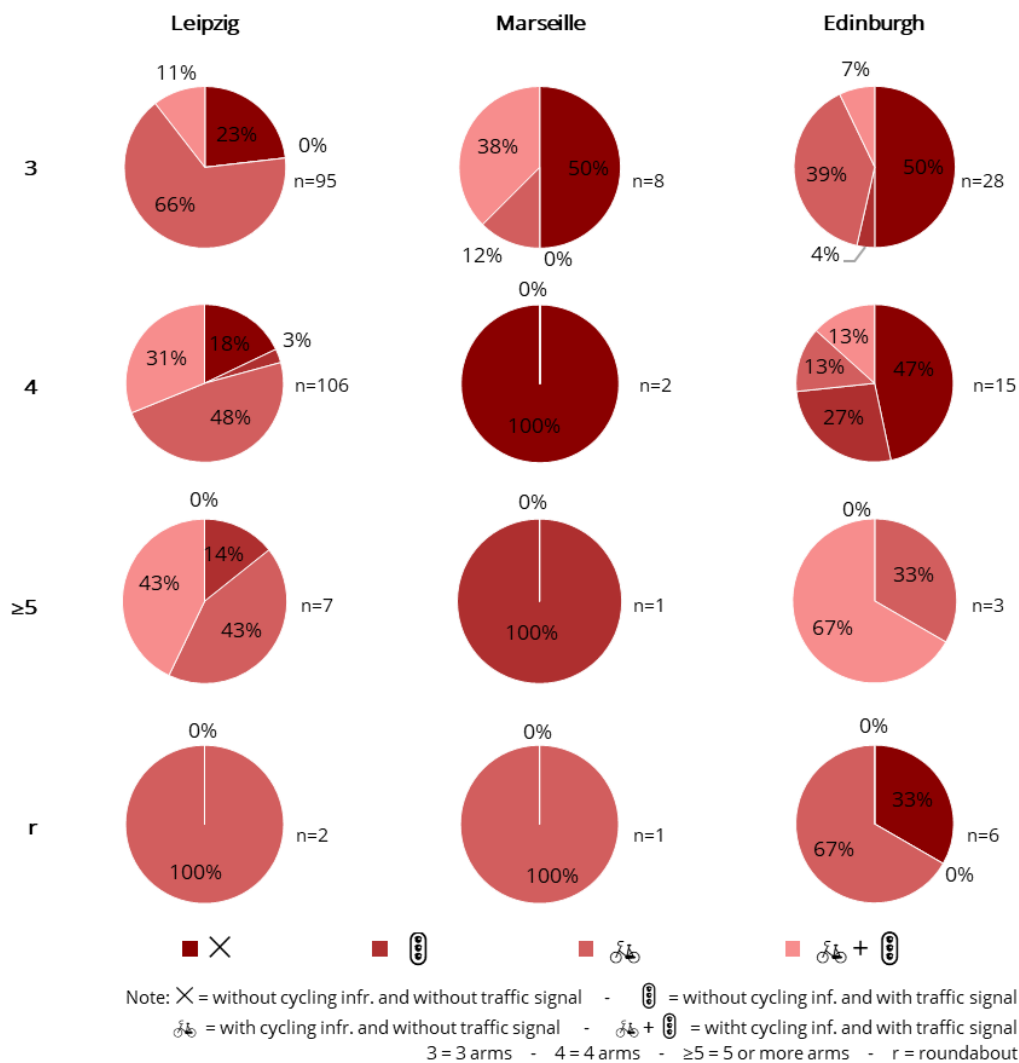


Figure 5.27 Percentage of intersections with seriously injured and killed cyclists by infrastructure attributes

Looking at this, it was found that the distribution of intersections with seriously injured and killed cyclists follow a special pattern in each city. In Marseille and Edinburgh the highest values were found at intersections with no cycling infrastructure and no traffic signals, when having three and four arms; while in Leipzig this case was found within intersections with cycling infrastructure but without traffic signals, when having three and four arms. When having five or more arms, the largest distribution of junctions with seriously injured and killed cyclists belonged to those with either cycling infrastructure or traffic signal in Leipzig (43%), to the ones without cycling infrastructure but with traffic signal in Marseille (100%), and to those with cycling infrastructure and traffic signal in Edinburgh (67%). For roundabouts the largest distribution of intersections with seriously injured and killed cyclists in each city, belonged to junctions with cycling infrastructure, but without traffic signals.

In this case, when comparing the intersections with seriously injured and killed cyclists among all the intersections of the same type, it was possible to identify their distribution considering the presence of cycling infrastructure and presence of traffic signals, by number of arms (see Figure 5.28).

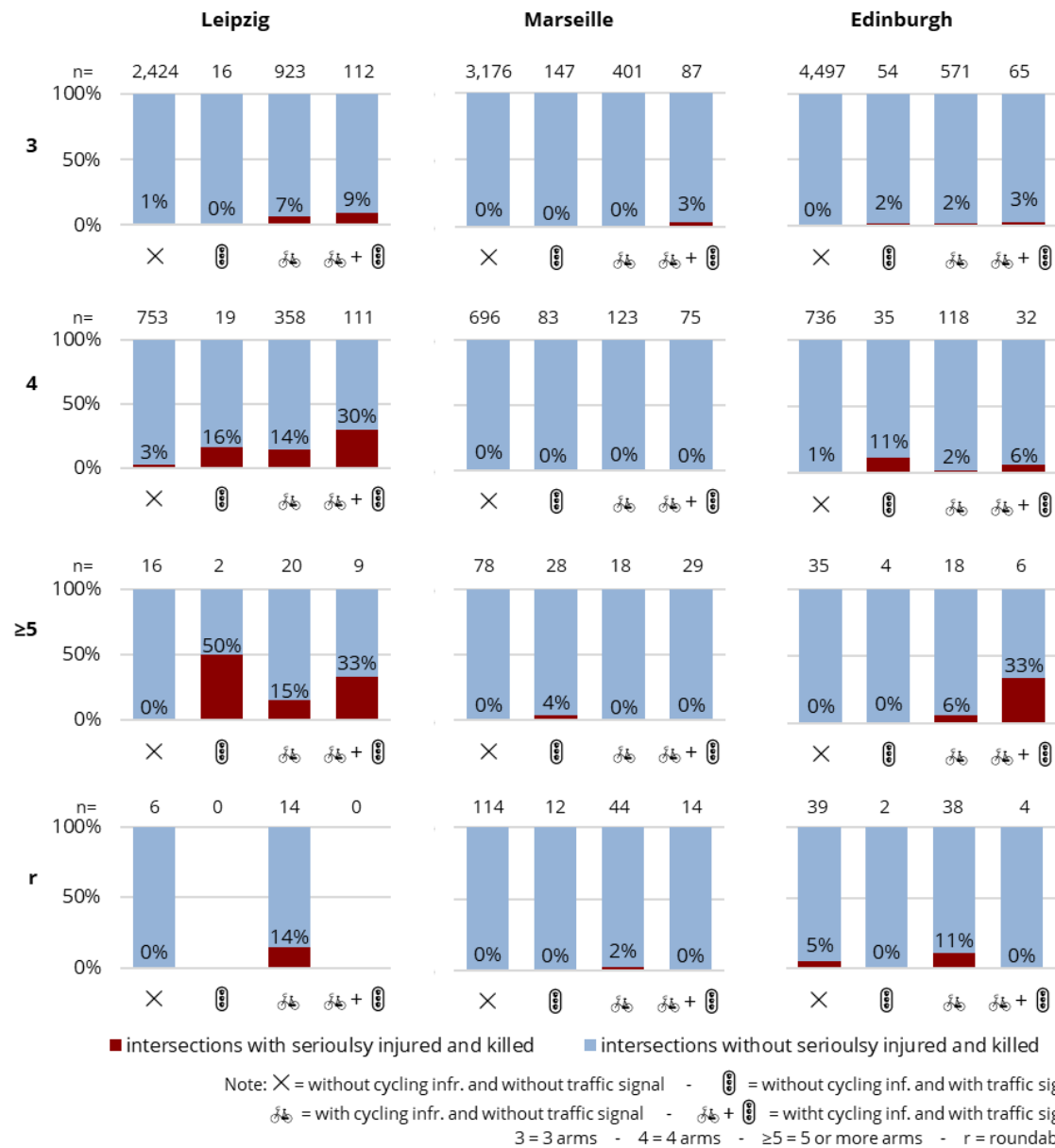


Figure 5.28 Percentage of intersections with and without seriously injured and killed cyclists by infrastructure attributes

Here, the highest percentage per city of intersections with three arms and severe and dead victims, were the ones with cycling infrastructure and traffic signal. In Leipzig the distribution reached 9%, and in Marseille, as well as in Edinburgh 3%.

Looking at the intersections with four arms, in Leipzig the highest distribution was found for intersections with cycling infrastructure and traffic signal (with 30%). Whereas in Edinburgh the highest value was for intersections without cycling infrastructure but with traffic signal (with 11%). For Marseille no distribution was obtained here, since there were neither severe nor dead victims reported.

Within junctions of five or more arms, in Leipzig and Marseille the highest distribution was found for those without cycling infrastructure but with traffic signal (50% for Leipzig and 4% for Marseille). Whereas for Edinburgh, the highest distribution in this intersection type corresponded to the ones with cycling infrastructure and with traffic signal (33%).

For roundabouts in each of the three cities, the biggest proportions were found when having cycling infrastructure but no traffic signal (14% in Leipzig, 2% in Marseille, and 11% in Edinburgh).

5.6 Hotspots analysis

The results of the hotspots analysis were produced at the city level, taking into account the identified intersections and their number of cyclist victims. In this case, the results were plot in two type of maps per city. First, a map showing the hotspots of intersections with all cyclist victims (regardless severity). And second, a map showing the hotspots of intersections with cyclist victims differentiating slightly injured from seriously injured and killed. In both cases the confidence level for determining the hotspots was set to 99%.

When exploring these two types of maps, it was possible to have a general approach of all cyclist victims, and also a more detailed approach, where intersections with seriously injured and killed victims were easily identified. This differentiation was also useful to distinguish intersections, which were hotspots for both categories of severity (slightly injured, and seriously injured and killed).

Since the maps at the city level covered a large extension difficult to represent in this thesis, it was decided to scale the maps, and focused on a smaller area per city. Therefore, in the following subsections, the maps of hotspots in the city centre in Leipzig, Marseille and Edinburgh are shown, considering the two categories of severity (slightly injured cyclists, and seriously injured and killed cyclists). Nevertheless, in Appendix C it is possible to see the original maps at the city level, showing the hotspots in each city, and also including the hotspots with all cyclist victims.

5.6.1 Leipzig

In Leipzig city centre there was a concentration of hotspots located in the primary and secondary roads, and this was the case for slightly injured, as well as seriously injured and killed (see Figure 5.29). Focusing on only hotspots of intersections with slightly injured, there were also several of them in residential streets, specially within the inner-city ring, and out of it in the north-western side. Regarding the hotspots of intersections with seriously injured and killed cyclists, they were found only in non-residential streets, especially along Johannisplatz, Ranstädter Steinweg, and Martin-Luther-Ring. Additionally, according to the results, it is possible to note that several hotspots with seriously injured and killed cyclists in the centre, were also hotspots with slightly injured.

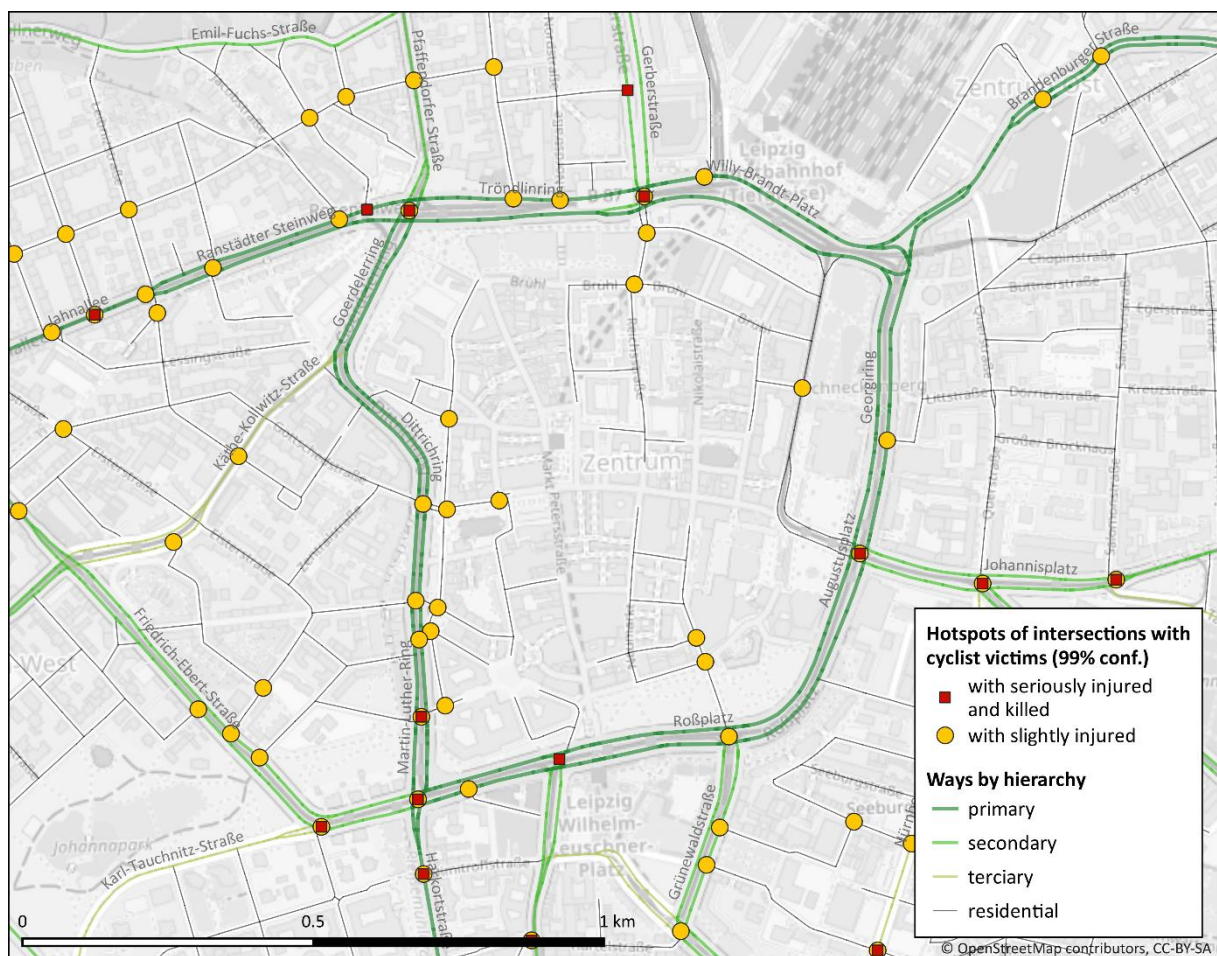


Figure 5.29 Hotspots for cyclist victims by injury severity in Leipzig city centre

5.6.2 Marseille

The Figure 5.30 presents the hotspots in Marseille city centre, where it was not possible to identify a general trend of hotspots. There were found four hotspots of intersections with slightly injured, from which two of them were in the residential street Rue Colbert, and the other two in Quai du Port (primary road), and in Boulevard de Montricher (tertiary road). In this case, only one hotspot for intersections with seriously injured and killed cyclist was found, and it was located at the intersection of Cours Lieutaud (primary road) and La Canebière (tertiary and residential road).

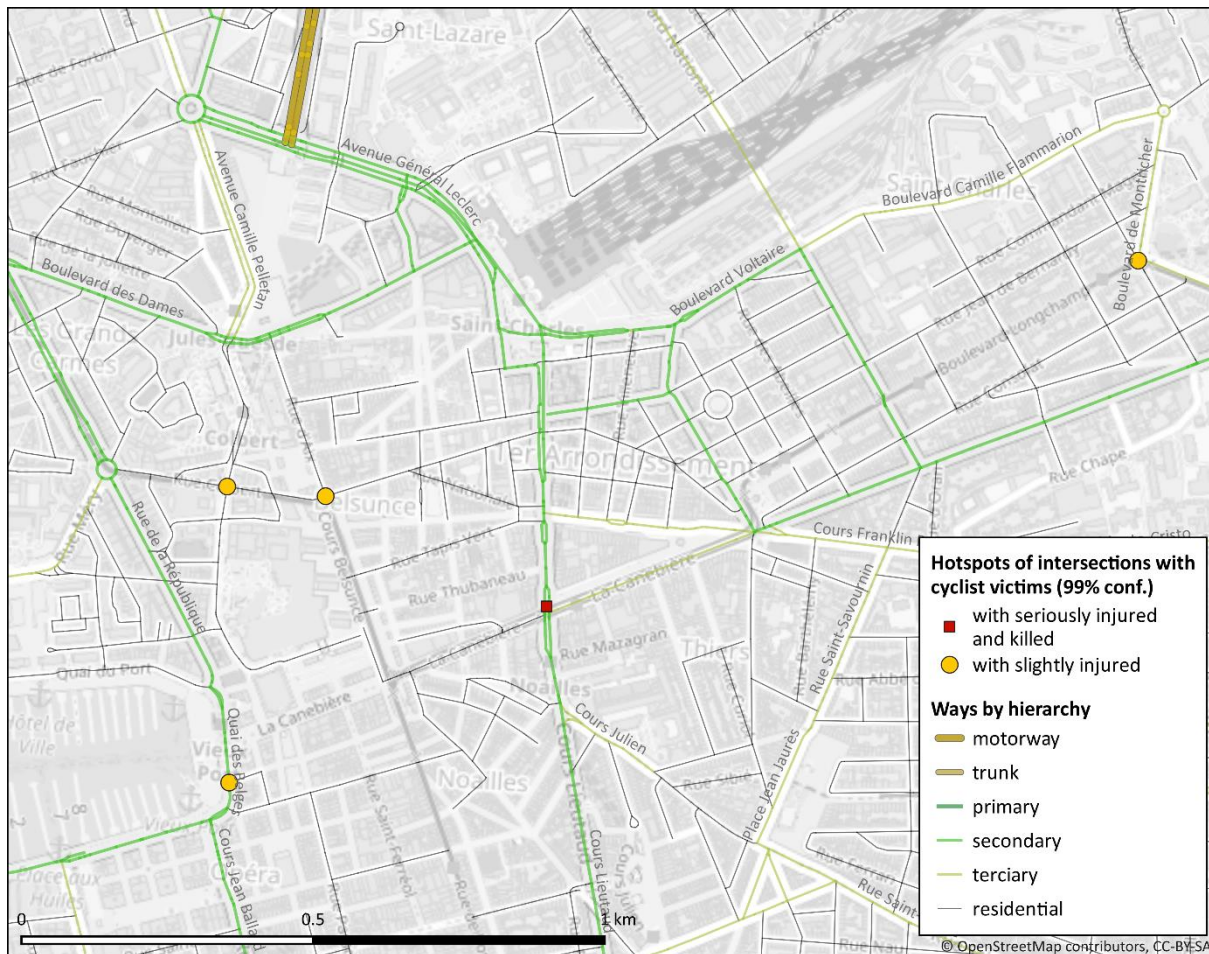


Figure 5.30 Hotspots for cyclist victims by injury severity in Marseille city centre

5.6.3 Edinburgh

The hotspots found in Edinburgh city centre were rather dispersed, with a small concentration in the south-western part of it. In this case, most of the hotspots were located in non-residential streets, and this is the case for slightly injured, as well as seriously injured and killed. However, one hotspot for the most critical severity level was found in a residential street (Tarvit Street), which is connecting two primary roads (Home Street and Brougham Street). Similarly as in Leipzig, some intersections were identified as hotspots of both categories for injury severity.

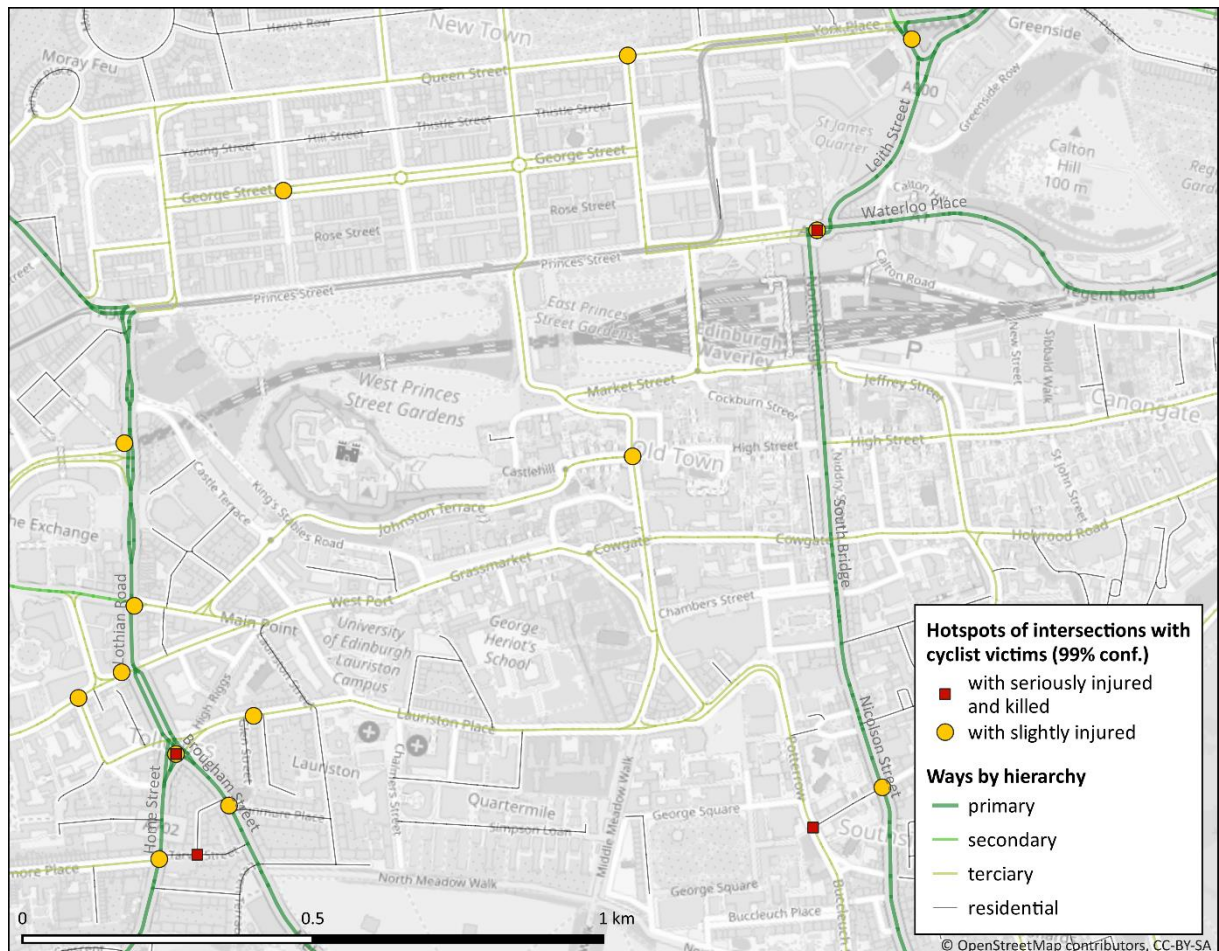


Figure 5.31 Hotspots for cyclist victims by injury severity in Edinburgh city centre

6 Discussion

The proposed methodology and the conducted analysis allowed to assess cyclist safety considering several parameters related with road infrastructure and crash data. Additionally, it was possible to obtain results in different European cities, and to establish comparisons between them. Thereby, the main objective of this research was achieved, and in the following sections a discussion about the steps to get to it are presented.

6.1 Data

Data availability and accuracy played an important role during this research, particularly for the OSM and the road crash data. This data was key on the definition of the study area and the parameters associated with road infrastructure and cyclist safety.

On the one hand, OSM brings a lot of useful and updated information available for everyone, thanks to the contribution of its volunteers. However, the activity of the contributors might change from place to place, in terms of quality and quantity (Ferster et al., 2020; Jokar Arsanjani et al., 2015). In fact, when processing the data for this research, it was found that some tags had either wrong or unreadable values, and in some other cases the tags were incomplete.

On the other hand, the crash data highly depends on the way it is collected, which is a task of the police. Nevertheless, several data registration errors, underreporting or lack of information might occur, since the police don't attend all collisions, and the communication with the hospitals and other possible information providers (e.g., witnesses) doesn't always happen (Adminaite et al., 2018). For the purpose of this research, the quality of the crash data was not measured, but, there might be an underreporting issue in Marseille, since only 2% of the reported victims were cyclists during 2015 and 2017, compared to 41% in Leipzig and 22% in Edinburgh. This situation restricted most of the analyses in Marseille, where the amount of cyclist victims did not allow to come up with clear results.

6.1.1 Definition of variables to select and compare cities

Regarding the city selection in Germany, France and Great Britain, crash data availability also played a big role. In this case, the information was gathered at the country level in France and Great Britain, but only for the state of Saxony in Germany. This situation forced to narrow the city selection, since only cities from Saxony were considered for the analysis in Germany.

At the beginning of this research, it was tried to choose similar cities based on several parameters, especially those associated with road infrastructure and safety. However, it was encountered that each city had different characteristics in terms of sociodemographic, transport, and road crash, which made difficult to directly reach comparability between them. After exploring several combinations, it was decided to drop from the city selection the parameters associated with transport and road crashes, which were not always available and presented different collection methodologies. Therefore, the chosen approach included the variables of population, area and population density, which were simple and replicable variables.

Even though this common approach was set to compare the cities, in average, French cities are denser than in Saxony and Great Britain. As a matter of fact, although the population density of Marseille is around 80% higher than the one of Leipzig and Edinburg, it is the most similar to them out of French cities.

Beyond the limitations to reach comparability between the cities, it was key to understand their differences regarding the parameters of interest, including cycling infrastructure and safety. In this case, after choosing the trio of cities, a general comparison was presented, including several variables related with population, area, elevation, cycling and road infrastructure, modal split and crash victims.

Since no cities are the same, it will be a difficult task to find similar criteria within a group of them, particularly if they are from different countries. Hence, it is recommended to carry out further research in this topic, including more parameters and new methods that were not explored in this research, in order to reach comparability between several cities.

6.1.2 Definition of infrastructure parameters

The definition of parameters associated with road infrastructure was linked to the data accessible through OSM, and to previous research conducted in cyclist safety assessments (Collins & Graham, 2019; Saad et al., 2019; Shen et al., 2020; Wang & Akar, 2018). This led to focus on three main parameters: intersection type, presence of cycling infrastructure, and presence of traffic signals.

Taking into account the chosen parameters, the cyclist assessment was executed based on specific infrastructure attributes of the intersections. In this way, the analysis led to characterize the intersections where the cyclist victims occurred, in terms of number of arms, as well as presence of cycling infrastructure and traffic signals.

Even though, some authors have also studied additional variables, such as traffic volume, street width, road lanes, median width and speed limit, these elements imply a higher level of detail of the information, and they are rather difficult to get from OSM. The contributors may tag the way elements with some of these variables in OSM, but the information is mostly insufficient to make analysis at the city level. Thus, further collaboration is required, so that more detailed information can be available in the traffic network from OSM.

Considering the limitations of OSM data, for further research it is also suggested to use complementary data sources (e.g., official authorities, Strava, Google Maps API). In this way, using OSM and more data sources, it will be possible to get more infrastructure-related parameters, and thereby, this process will enrich the cycling safety assessments.

6.1.3 Definition of cyclist safety indicators

In this research, the injury severity was treated as the main parameter associated with crash data. Therefore, in addition to the general assessment including all cyclist victims (regardless severity), a differentiation was made between slightly injured, and seriously injured and killed. Through this approach, it was possible to distinguish where the more severe victims occurred, which are the costliest to the society (including medical costs, property damage and administrative costs). At the same time, this differentiation is key for decision makers to establish a prioritisation of counter-measures, which help reducing traffic crashes victims, specially seriously injured and killed.

In terms of injury severity from the crash database, there were three original categories: slightly injured, seriously injured and killed. However, in this research it was decided to merge as one category the seriously injured and killed, since the amount of data in each category was rather insufficient to make conclusions after the analysis. Moreover, taking into account that in Great Britain some victims who were not hospitalised more than 24 hours, are also considered seriously injured, further work is suggested in this regard so that the severity levels share the same criteria between cities, and can be directly comparable.

In addition to the severity injury, there were more specific variables that were not studied in this analysis. Looking forward to complementing the executed analysis, further research might include more specific variables associated with the victims and the traffic crash, such as victim's age and gender, as well as weather, risk manoeuvre and involved vehicles.

6.2 Analysis

6.2.1 Identification and classification of intersections

The traffic network obtained from OSM was built by nodes and ways, representing each carriageway, including different elements when there is a physical channelization of turning lanes. This representation generated concentration of ways and nodes in some intersections, that for the purpose of this research was inconvenient. Thus, the clustering process had to be executed, and this allowed to have one node per intersection, instead of several. In addition, thanks to the simplification of the way elements, the general number of intersection arms could be found, instead of several carriageways and turning lanes.

When clustering the nodes, it was challenging to find the right combination of parameters, that led to identify the junctions. Nevertheless, after running the DBSCAN in different steps and considering the road hierarchy, the output led to satisfactory results. Further research might include exploring the use of different clustering methods, such as single linkage clustering.

During the process where the ways were clustered and simplified, instead of applying a negative buffer over the aggregated elements, a function to get the approximate medial axis (also known as skeleton) was unsuccessfully applied. Due to the complexity of the traffic network at the city level and the road hierarchy classification, the output of the skeleton function included some random residual lines located at the ends of the elements, which were not part of the original traffic network, and were disrupting the count of arms per intersection.

Since the simplification of the ways was finally executed with the negative buffer approach, the resulting output was polygonal instead of lineal. This approach allowed to identify different streets, and hence, it was feasible to count the number of them crossing at each intersection. However, in some junctions the resulting polygonal geometry was bigger than the generated influence zone, which prevented to differentiate between arms, and led to wrong classification of the intersections.

Regarding the definition of the influence zones, it was found that finding the correct parameters that match all the intersections is a hard task, because of the complexity and diversity of the traffic network imported from OSM. During the analysis of the data, for intersections of the same road hierarchy, the radius of the influence zone was sometimes small and sometimes big to accurately count the number of arms. Additionally, when selecting the radius, the possibility of overlapping between influence zones was ignored, and in the obtained results there were found some influence zones intersecting with each other. This led to assign the same elements (including cyclist victims and infrastructure parameters) to more than one intersection, when the overlapping existed.

Since significant variety in infrastructure design can exist from one city to another (Reynolds et al., 2009), it was challenging to calibrate the different variables in order to execute this research. In addition, this might lead to include some adjustments to the different parameters, for a replication of this methodology in different cities. Similarly, future research might include some strategies to overcome the already mentioned issues, especially the skeleton algorithm, the identification of arms per intersection, and the assignation of influence zone. Nevertheless, the obtained results through this research were satisfactory, and specific patterns associating cyclist victims' rates with infrastructure parameters were found.

6.2.2 Spatial match of cyclist victims with intersections

Since the crash database included the coordinates where the victims took place, this allowed to identify the location of the victims in the study area. For Saxony and Great Britain, 100% of the reported victims had coordinates. However, it was found that crashes reported in France did not include coordinates in 31% of the cases, which made impossible to locate the position of the victims. Thus, for the results obtained in the French city of Marseille an underestimation in the spatial match of cyclist victims was expected, whereas in Leipzig and Edinburgh this was not the case.

Due to the definition of the influence zones, the cyclist victims were matched and assigned to the intersections. This approach allowed to execute the cyclist safety assessment focused on the intersections with victims, leading to identify their patterns related with the road infrastructure and the location of hotspots.

The spatial match was entirely dependent of the coverage of the influence zones, and this brought some limitations. In some cases, the influence zones were too big, and they were overlapping with each other, meaning that the same victim could be assigned to different intersections. In addition, when having big influence zones, there was the uncertainty of matching victims that were not any more in the intersection itself, but rather in the link of the network (at the street level).

6.2.3 Spatial and statistical analysis

Beyond the previously described limitations, through this research it was possible to execute the cyclist assessment. This led to identify spatial and statistical patterns of the different parameters, including road infrastructure and cyclist safety.

Based on the results by typology, 3-arm junctions presented the highest distribution with cyclist victims, with at least 50% in each of the three cities. However, comparing the rate of intersections with victims among all intersections of the same type (including intersections without victims), the junctions with three arms had the lowest percentage, whereas those with five or more arms, and roundabouts had the highest percentages in each city. Since traffic networks were mainly integrated by 3-arm intersections, it is not surprising that they had the lowest rates of intersections with victims among all intersections of the same type. In addition, this result implies that intersections with five or more arms, and roundabouts should be further studied, since their rates with cyclist victims were the highest.

When studying the different infrastructure parameters in terms of cyclist victims, there was a common pattern in the chosen cities, regardless number of arms per intersection. In most cases the highest proportions of intersection with victims, were the ones with cycling infrastructure and without traffic signal, followed by intersections without any of these two infrastructure facilities. This suggests that intersections with traffic signals were safer for cyclists, coinciding with findings from another research in the United Kingdom (Shen et al., 2020) and in the United States (Saad et al., 2019; Wang & Akar, 2018).

The fact that presence of cycling ways is associated with high rates of victims, invites to explore in a deeper way the conditions of the existing cycling infrastructure, looking forward to improving the safety conditions in those places. However, this may also be associated with the cyclist volume in these intersections, which could be significantly higher than in intersections without cycling ways, and therefore the possibility of having a crash with victims is higher as well.

When analysing the victims by severity level, the pattern was similar to the previous discussion, meaning that most of the intersections with all cyclist victims, slightly injured victims, and serious injured including killed cyclist, had better safety performance when having traffic signals. Even though there were similarities in the analysis, the distinction of victims by severity level help to understand the overall situation of cyclist victims, and to prioritise actions that allow to reduce the crash indicators. Especially severe injured and killed cyclists stand out, since, in a one-to-one basis, they cost more and have a worse negative impact to the society than the slightly injured.

Regarding the hotspots analysis, each of the cities had several locations with statistically significant concentration of intersection with cyclist victims. Nevertheless, it was found that in most cases the hotspots were found in non-residential streets, particularly when studying the most critical severity level (seriously injured and killed cyclists). This information must be considered by decision makers in cities, and therefore improve the cyclist safety conditions in the particular junctions, where the hotspots were identified.

7 Conclusion

There are several works and academic papers regarding cyclist safety analyses. However, none of them has approached completely the way it was intended in this research. Here, the approach was focused on assessing cyclist safety, not only with the official crash data, but also including infrastructure parameters of the traffic network gathered from OpenStreetMap.

Infrastructure-related data is rather difficult to get from the local authorities, and in some cases it is obsolete. That is why, OSM is a promising source for updated and freely accessible geodata, including different infrastructure parameters, that were useful during this research. Even though there are some issues concerning availability and accuracy of OSM data, it was feasible to implement the proposed methodology in three European cities: Leipzig, Marseille and Edinburgh. Thus, statistical and spatial patterns were found for each city, which will be useful for decision makers and mobility planners, looking forward to improving road safety and reducing victims on the streets.

After implementing the data analysis, the results showed general patterns of intersections with cyclist victims, considering their severity level and also different infrastructure variables shaping the intersections: number of arms, presence of cycling infrastructure and presence of traffic signals. These results were obtained for the chosen cities, and it was also possible to compare the resulting data between them. Hence, it was found that in most cases the highest proportions of intersection with cyclist victims, are the ones with cycling infrastructure and without traffic signal, followed by intersections without any of these two infrastructure facilities. This suggests that intersections with traffic signals were safer for cyclists, and this applied for the different categories adopted for cyclist victims: all victims, slightly injured, and seriously injured including killed cyclists.

This thesis allowed to identify different aspects, that should be improved in the future. More work should be done to improve data availability and quality, including OSM and crash data, allowing to replicate this kind of analysis in different places around the world. Additionally, for future research it is recommended to explore better algorithms to simplify nodes and ways from OSM, and get the number of arms per intersection, more accurately and also automatically. In order to boost the scope of this thesis, further research is suggested taking into account complementary variables, such as traffic volumes and victims from other transportation modes.

References

- Adminaite, D., Jost, G., Stipdonk, H., & Ward, H. (2018). *An overview of road death data collection in the EU, PIN Flash Report 35*. <https://etsc.eu/wp-content/uploads/PIN-FLASH-35-final.pdf>
- Branion-Calles, M., Götschi, T., Nelson, T., Anaya-Boig, E., Avila-Palencia, I., Castro, A., Cole-Hunter, T., de Nazelle, A., Dons, E., Gaupp-Berghausen, M., Gerike, R., Int Panis, L., Kahlmeier, S., Nieuwenhuijsen, M., Rojas-Rueda, D., & Winters, M. (2020). Cyclist crash rates and risk factors in a prospective cohort in seven European cities. *Accident Analysis and Prevention*, 141. <https://doi.org/10.1016/j.aap.2020.105540>
- British Office for National Statistics. (2021). *Estimates of the population for the United Kingdom*. [Dataset from 2019]. <https://www.ons.gov.uk/peoplepopulationandcommunity/populationandmigration/populationestimates/datasets/populationestimatesforukenglandandwalesscotlandandnorthernireland>
- Bruzzzone, S., & Broccoli, M. (2021). *Use of Open Street Map for accident investigation on the road and motorways networks*.
- Buehler, R., & Pucher, J. (2012). Cycling to work in 90 large American cities: New evidence on the role of bike paths and lanes. *Transportation*, 39(2), 409–432. <https://doi.org/10.1007/s11116-011-9355-8>
- Burgess, J., & Pavlenko, A. (2020). *osm2pgsql manual*. <https://osm2pgsql.org/doc/manual.html>
- Chanove, A. (2021). A method to harmonize accident databases between different countries. *9th International Expert Symposium on Accident Research ESAR*.
- Cheng, Z., Zu, Z., & Lu, J. (2019). Traffic crash evolution characteristic analysis and spatiotemporal hotspot identification of urban road intersections. *Sustainability (Switzerland)*, 11(1). <https://doi.org/10.3390/su11010160>
- Collins, D. J., & Graham, D. J. (2019). Use of Open Data to Assess Cyclist Safety in London. *Transportation Research Record*, 2673(4), 27–35. <https://doi.org/10.1177/0361198119837221>
- Ester, M., Kriegel, H.-P., Sander, J., & Xu, X. (1996). A Density-Based Algorithm for Discovering Clusters in Large Spatial Databases with Noise. *Proceedings of the Second International Conference on Knowledge Discovery and Data Mining*, 226–231.
- European Commission. (2020). *Annual statistical report on road safety in the European Union*. https://ec.europa.eu/transport/road_safety/system/files/2021-07/asr2020.pdf
- European Cyclists' Federation. (2018). *The benefits of cycling: Unlocking their potential for Europe*. <https://ecf.com/sites/ecf.com/files/TheBenefitsOfCycling2018.pdf>

- European Transport Safety Council. (2007). *Social and Economic Consequences of Traffic Injury in Europe*. <https://etsc.eu/wp-content/uploads/Social-and-economic-consequences-of-road-traffic-injury-in-Europe.pdf>
- Federal and state statistical offices of Germany. (2022). *Population in Germany*. [Dataset from 2019]. <https://www.regionalstatistik.de/genesis//online?operation=table&code=12411-01-01-5&bypass=true&levelindex=1&levelid=1654532064331#abreadcrumb>
- Ferster, C., Fischer, J., Manaugh, K., Nelson, T., & Winters, M. (2020). Using OpenStreetMap to inventory bicycle infrastructure: A comparison with open data from cities. *International Journal of Sustainable Transportation*, 14(1), 64–73. <https://doi.org/10.1080/15568318.2018.1519746>
- French National Institute of Statistics and Economic Studies. (2021). *Legal population in France*. [Dataset from 2019]. <https://www.insee.fr/fr/statistiques/6011070?sommaire=6011075>
- General direction for sustainable city and expansion. (2020). *Notice of the City of Marseille on the urban transport plan of the Aix-Marseille-Provence Metropolis*. https://www.marseille.fr/sites/default/files/contenu/se-deplacer/extraits-des-deliberations_conseil-municipal_05-octobre-2021.pdf
- Getis, A., & Ord, J. K. (1995). Local Spatial Autocorrelation Statistics: Distributional Issues and an Application. *Geographical Analysis*, 27(4), 286–306.
- Hong, J., McArthur, D. P., & Livingston, M. (2020). The evaluation of large cycling infrastructure investments in Glasgow using crowdsourced cycle data. *Transportation*, 47(6), 2859–2872. <https://doi.org/10.1007/s11116-019-09988-4>
- Jokar Arsanjani, J., Zipf, A., Mooney, P., & Helbich, M. (2015). An Introduction to OpenStreetMap in Geographic Information Science: Experiences, Research, and Applications. In *OpenStreetMap in GIScience, Lecture Notes in Geoinformation and Cartography* (pp. 1–15). https://doi.org/10.1007/978-3-319-14280-7_1
- Jost, G., Carson, J., & Meinerio, M. (2022). *Ranking EU Progress on Road Safety, 16th Road Safety Performance Index Report*. https://etsc.eu/wp-content/uploads/16-PIN-annual-report_FINAL_WEB_1506_2.pdf
- Klanjčič, M., Gauvin, L., Tizzoni, M., & Szell, M. (2022). Identifying urban features for vulnerable road user safety in Europe. *EPJ Data Science*, 11(1). <https://doi.org/10.1140/epjds/s13688-022-00339-5>
- Kumar, A., Benavent, C., Kastl, D., Shinde, R., Obe, R., Joshi, S., & Vergara, V. (2022). *pgRouting manual*. <https://docs.pgrouting.org/latest/en/index.html>
- Mobility Office of Leipzig. (2020). *Traffic survey "Mobility in cities - SrV 2018, City of Leipzig"*. <https://static.leipzig.de/fileadmin/mediendatenbank/leipzig->

- de/Stadt/02.6_Deiz6_Stadtentwicklung_Bau/66_Verkehrs_und_Tiefbauamt/Verkehrskonzepte/Verkehrsz%C3%A4hlungen/SrV_2018_Kennziffern_Internet.pdf
- Olbricht, R. (2019). *Overpass API User's Manual*. <https://dev.overpass-api.de/overpass-doc/en/index.html>
- OpenStreetMap Wiki contributors. (2022). *OpenStreetMap*. OpenStreetMap Wiki. https://wiki.openstreetmap.org/w/index.php?title=Main_Page&oldid=2301410
- Oxoli, D., Prestifilippo, G., Bertocchi, D., & Zurbaràn, M. (2017). Enabling spatial autocorrelation mapping in QGIS: The Hotspot Analysis Plugin. *Geoingegneria Ambientale e Mineraria - GEAM*, 151(2), 45–50.
- Reynolds, C. C. O., Harris, M. A., Teschke, K., Crompton, P. A., & Winters, M. (2009). The impact of transportation infrastructure on bicycling injuries and crashes: A review of the literature. *Environmental Health: A Global Access Science Source*, 8(1). <https://doi.org/10.1186/1476-069X-8-47>
- Saad, M., Abdel-Aty, M., Lee, J., & Cai, Q. (2019). Bicycle Safety Analysis at Intersections from Crowdsourced Data. *Transportation Research Record*, 2673(4), 1–14. <https://doi.org/10.1177/0361198119836764>
- Santacreu, A. (2018). *Safer City Streets: Global Benchmarking for Urban Road Safety*. International Transport Forum Working Document, OECD Publishing. <https://www.itf-oecd.org/safer-city-streets-global-benchmarking-urban-road-safety>
- Shen, J., Wang, T., Zheng, C., & Yu, M. (2020). Determinants of Bicyclist Injury Severity Resulting from Crashes at Roundabouts, Crossroads, and T-Junctions. *Journal of Advanced Transportation*, 2020. <https://doi.org/10.1155/2020/6513128>
- Transport and Environment Committee. (2021). *City Mobility Plan – Mode Share Targets*. <https://democracy.edinburgh.gov.uk/documents/s40132/7.4%20-%20City%20Mobility%20Plan%20-%20Mode%20Share%20Target.pdf>
- Wang, K., & Akar, G. (2018). Street Intersection Characteristics and Their Impacts on Perceived Bicycling Safety. *Transportation Research Record*, 2672(46), 41–54. <https://doi.org/10.1177/0361198118801349>
- World Health Organization. (2018). *Global status report on road safety*. <https://www.who.int/publications/i/item/9789241565684>

Appendix

Appendix A: Metadata and details of available crash databases

Territory	Saxony (Germany)	France	Great Britain
Country code	GER	FRA	GBR
Years in study	2015 - 2017	2015 - 2017	2015 - 2017
Tables	3	4	4
Structure	Accident, participant, passenger	Accident, place, vehicle, person	Accident, vehicle, casualty, contributory factors
Variables (characteristics)	71 (342)	65 (317)	70 (275)
Victims of the years in study (annual avg.)	45,081 (15,027)	189,872 (63,291)	418,704 (139,568)
Crashes of the years in study (annual avg.)	82,951 (27,650)	332,529 (110,843)	709,009 (236,336)
Condition of acquisition	At least one victim, one vehicle, on the public road		
Severity definition for serious injured	Hospitalised for at least 24 hours.	Hospitalised for at least 24 hours. Police records. People injured are asked to go to the police to report details	Hospitalised for at least 24 hours or any of the following injuries whether or not they are detained in hospital: fractures, concussion, internal injuries, crushing, burns (excluding friction burns), severe cuts and lacerations, severe general shock. S
Source	Police records, Federal Ministry of Justice and Consumer Protection	Police records. French Road Safety Observatory, Transport Ministry	Police records, Department for Transport

Table A.12 Metadata of crash databases in Saxony (Germany), France and Great Britain

Territory	Saxony (Germany)	France	Great Britain
Weather	x	rain, snow, hail, fog, blinding sun, wind	rain, snow, hail, fog, all with or without high wind
Date, time	✓	✓	✓
Location type	Urban, rural	Urban, rural	Urban, rural
Location of the accident	X,Y + milestone + address	X,Y + milestone + address	X,Y
Light conditions	Daylight	Daylight + street lighting	Daylight + street lighting
Collision type	Frontal, side collision, from behind, stopped vehicle. Obstacle, no collision, pedestrian collision	2-vehicles: frontal, from behind, from side. 3 vehicles and more: series, multiple. Pedestrian and animal collision	Frontal, side collision, from behind, stopped vehicle. Obstacle, no collision, pedestrian collision
Junction type	Intersection shape, property access	Intersection shape, roundabout, crossing level	Intersection shape, roundabout, crossing level, property access
Junction special	x		
Accident position	x	On the road, sidewalk, bicycle lane, emergency lane, shoulder	Other variables on the position of vehicles on the road
Accident severity	Accident resulting of slight, severe injuries or death	x	Accident resulting of slight, severe injuries or death
Accident factor	List of 10 accident causes	x	x
Cost estimations	At the accident level and per vehicle	x	x
Additional information	x	Next to a school	x

Table A.13 Information at the crash level

Territory	Saxony (Germany)	France	Great Britain
Participant type	Vehicle type, bicycle, pedestrian, passenger	Driver, passenger, pedestrian, pedestrian on wheels	Driver, passenger, pedestrian
Participant reference	✓	✓	✓
Position in vehicle	x	✓	For <u>car</u> passenger only
Gender	✓	✓	✓
Birth	✓	✓	✓
Special participant information	x	Profession, postcode, nationality	Driver's & injured postcode
Injury severity	Unarmed, slight, severe, dead	Unarmed, slight, severe, dead	Unarmed, slight, severe, dead
Presumed responsible	✓	✓	x
Participant's address	x	x	x
Alcohol test, drug test	Alcohol & drug test are mandatory for drivers and pedestrians	Alcohol & drug test are mandatory for drivers and pedestrians	Alcohol & drug test are not always mandatory
Driver license information	License date & type	License date & type	x
Personal safety equipment	x	Main equipment: belt, helmet, kid restraint. Second equipment: reflective jacket, airbag, gloves	For bicycle: helmet worn. Other: seatbelt use
Accident factor due to the participant	List of 65 causes	Up to 3 possible factors: fatigue, drugs, infirmity, disturbed attention, alcohol, medication, phone, priority error, manoeuvre error, speed	Driver: error, distraction, impairment, behaviour. Vision affected. Pedestrian error.
Trip	x	✓	✓
Additional information	x	Pedestrian manoeuvre, pedestrian location	Pedestrian manoeuvre, pedestrian location and direction

Table A.14 Information at the participant/victim level

Appendix B: Results tables

Country/Territory	City	Population	Area (km ²)	Pop. Density
France	Lyon	522,969	47.9	10,924.8
Great Britain	Manchester	552,858	115.6	4,780.5
France	Toulouse	493,465	118.3	4,171.3
Great Britain	Glasgow	633,120	174.6	3,625.7
France	Marseille	870,731	242.1	3,596.0
Saxony (Germany)	Leipzig	593,145	297.9	1,991.2
Great Britain	Edinburgh	524,930	273.1	1,921.8
Saxony (Germany)	Dresden	556,780	328.5	1,695.0
Great Britain	Sheffield	584,853	367.9	1,589.6
Great Britain	Bradford	539,776	366.4	1,473.1
Great Britain	Leeds	793,139	551.7	1,437.6

Table A.15 Cities per territory ordered by population density

Intersection type	Leipzig	Marseille	Edinburgh	Total
3 arms	3,475	3,811	5,187	12,473
4 arms	1,241	977	921	3,139
5 or more arms	47	153	63	263
roundabout	20	184	83	287
Total	4,783	5,125	6,254	16,162

Table A.16 Count of intersections by type and city

Intersection type	Cycling inf.	Traf. signal	Leipzig	Marseille	Edinburgh	Total
3 arms	no	no	2,424	3,176	4,497	10,097
		yes	16	147	54	217
	yes	no	923	401	571	1,895
		yes	112	87	65	264
4 arms	no	no	753	696	736	2,185
		yes	19	83	35	137
	yes	no	358	123	118	599
		yes	111	75	32	218
5 or more arms	no	no	16	78	35	129
		yes	2	28	4	34
	yes	no	20	18	18	56
		yes	9	29	6	44
roundabout	no	no	6	114	39	159
		yes	0	12	2	14
	yes	no	14	44	38	96
		yes	0	14	4	18
Total			4,783	5,125	6,254	16,162

Table A.17 Count of intersections by city and category based on infrastructure attributes

Intersection type	Leipzig	Marseille	Edinburgh	Total
3 arms	419	26	142	587
4 arms	392	11	83	486
5 or more arms	19	4	14	37
roundabout	7	7	26	40
Total	837	48	265	1,150

Table A.18 Count of intersections with cyclist victims by type and city

Intersection type	Cycling inf.	Traf. signal	Leipzig	Marseille	Edinburgh	Total
3 arms	no	no	118	12	69	199
		yes	3	3	14	20
	yes	no	239	5	42	286
4 arms	no	yes	59	6	17	82
		no	108	8	38	154
	yes	yes	10	0	13	23
5 or more arms	no	no	189	1	16	206
		yes	85	2	16	103
	yes	no	3	1	3	7
roundabout	no	yes	1	2	2	5
		no	9	0	4	13
	yes	yes	6	1	5	12
Total	no	no	1	2	9	12
		yes	0	0	0	0
	yes	no	6	5	16	27
		yes	0	0	1	1
Total			837	48	265	1,150

Table A.19 Count of intersections with cyclist victims by city and infrastructure attributes

Intersection type	Leipzig	Marseille	Edinburgh	Total
3 arms	373	18	122	513
4 arms	360	9	75	444
5 or more arms	17	3	13	33
roundabout	6	6	25	37
Total	756	36	235	1,027

Table A.20 Count of intersections with slightly injured cyclists by type and city

Intersection type	Cycling inf.	Traf. signal	Leipzig	Marseille	Edinburgh	Total
3 arms	no	no	101	8	59	168
		yes	3	3	13	19
	yes	no	212	4	34	250
		yes	57	3	16	76
4 arms	no	no	97	6	33	136
		yes	10	0	11	21
	yes	no	171	1	15	187
		yes	82	2	16	100
5 or more arms	no	no	3	1	3	7
		yes	0	1	2	3
	yes	no	8	0	3	11
		yes	6	1	5	12
roundabout	no	no	1	2	8	11
		yes	0	0	0	0
	yes	no	5	4	16	25
		yes	0	0	1	1
Total			756	36	235	1,027

Table A.21 Count of intersections with slightly injured cyclists by city and infrastructure attributes

Intersection type	Leipzig	Marseille	Edinburgh	Total
3 arms	95	8	28	131
4 arms	106	2	15	123
5 or more arms	7	1	3	11
roundabout	2	1	6	9
Total	210	12	52	274

Table A.22 Count of intersections with seriously injured and killed cyclists by type and city

Intersection type	Cycling inf.	Traf. signal	Leipzig	Marseille	Edinburgh	Total
3 arms	no	no	22	4	14	40
		yes	0	0	1	1
	yes	no	63	1	11	75
		yes	10	3	2	15
4 arms	no	no	19	2	7	28
		yes	3	0	4	7
	yes	no	51	0	2	53
		yes	33	0	2	35
5 or more arms	no	no	0	0	0	0
		yes	1	1	0	2
	yes	no	3	0	1	4
		yes	3	0	2	5
roundabout	no	no	0	0	2	2
		yes	0	0	0	0
	yes	no	2	1	4	7
		yes	0	0	0	0
Total			210	12	52	274

Table A.23 Count of intersections with seriously injured and killed cyclists by city and infrastructure attributes

Appendix C: Hotspots for intersections with all cyclist victims

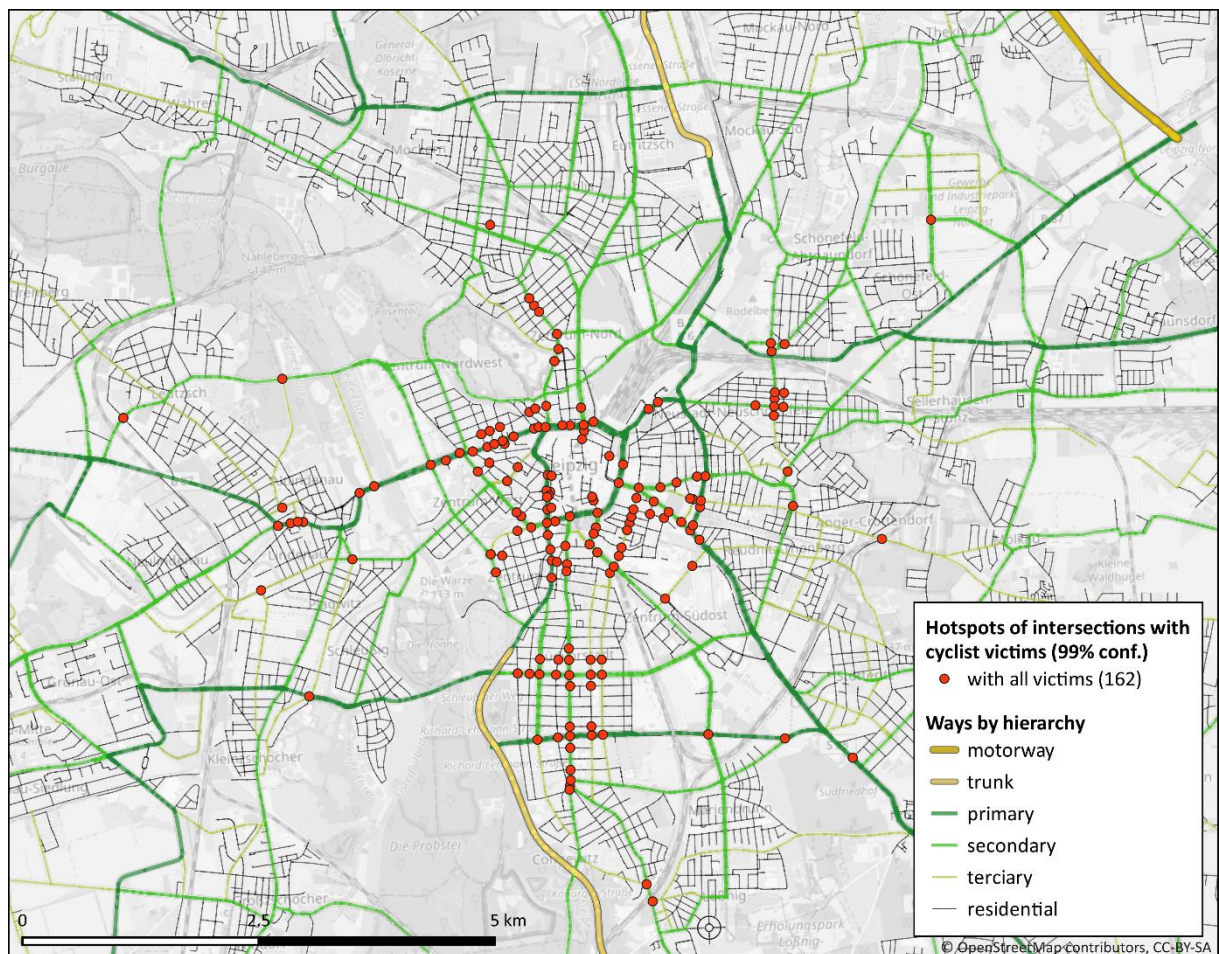


Figure A.32 Hotspots for all cyclist victims in Leipzig

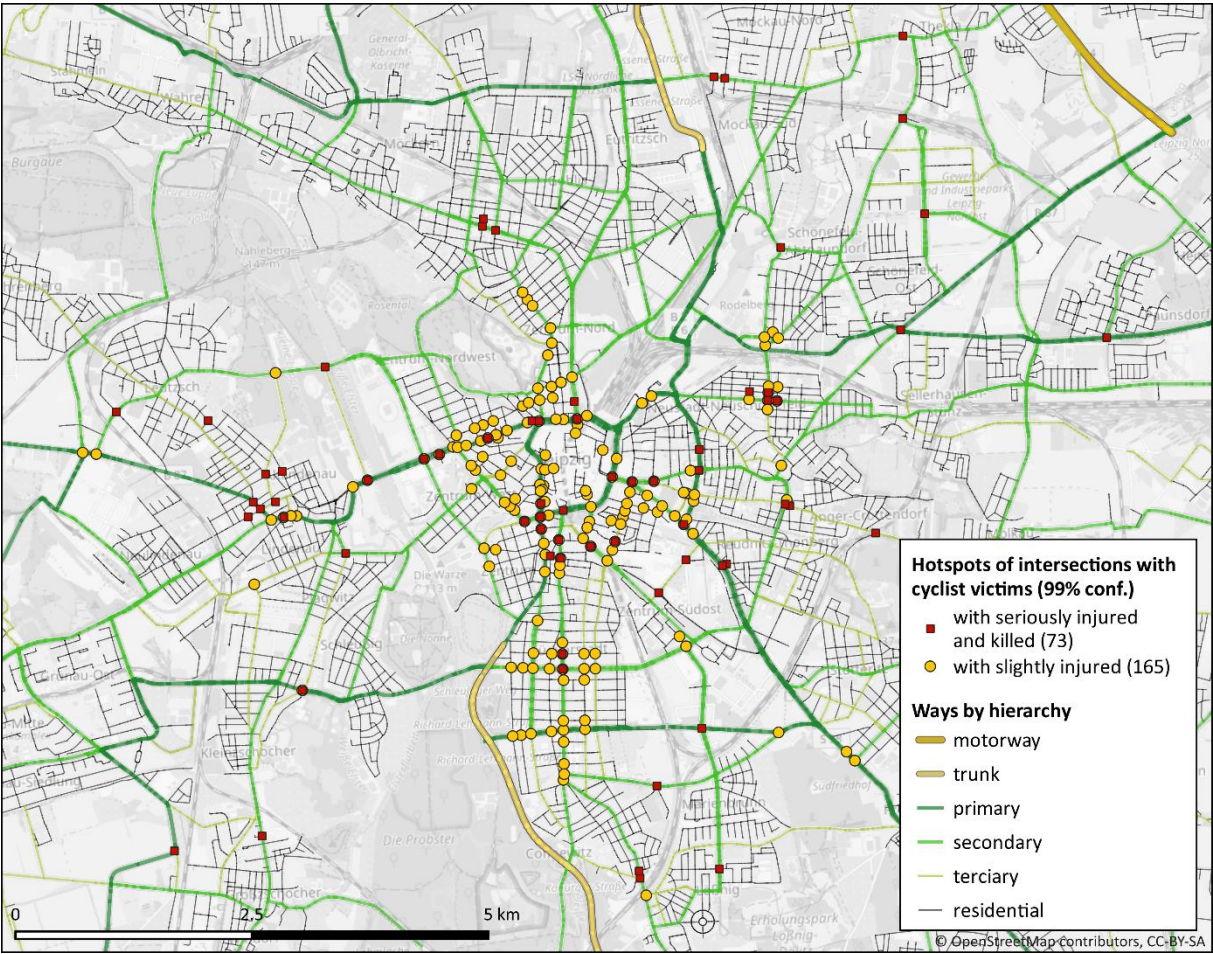


Figure A.33 Hotspots for cyclist victims by injury severity in Leipzig

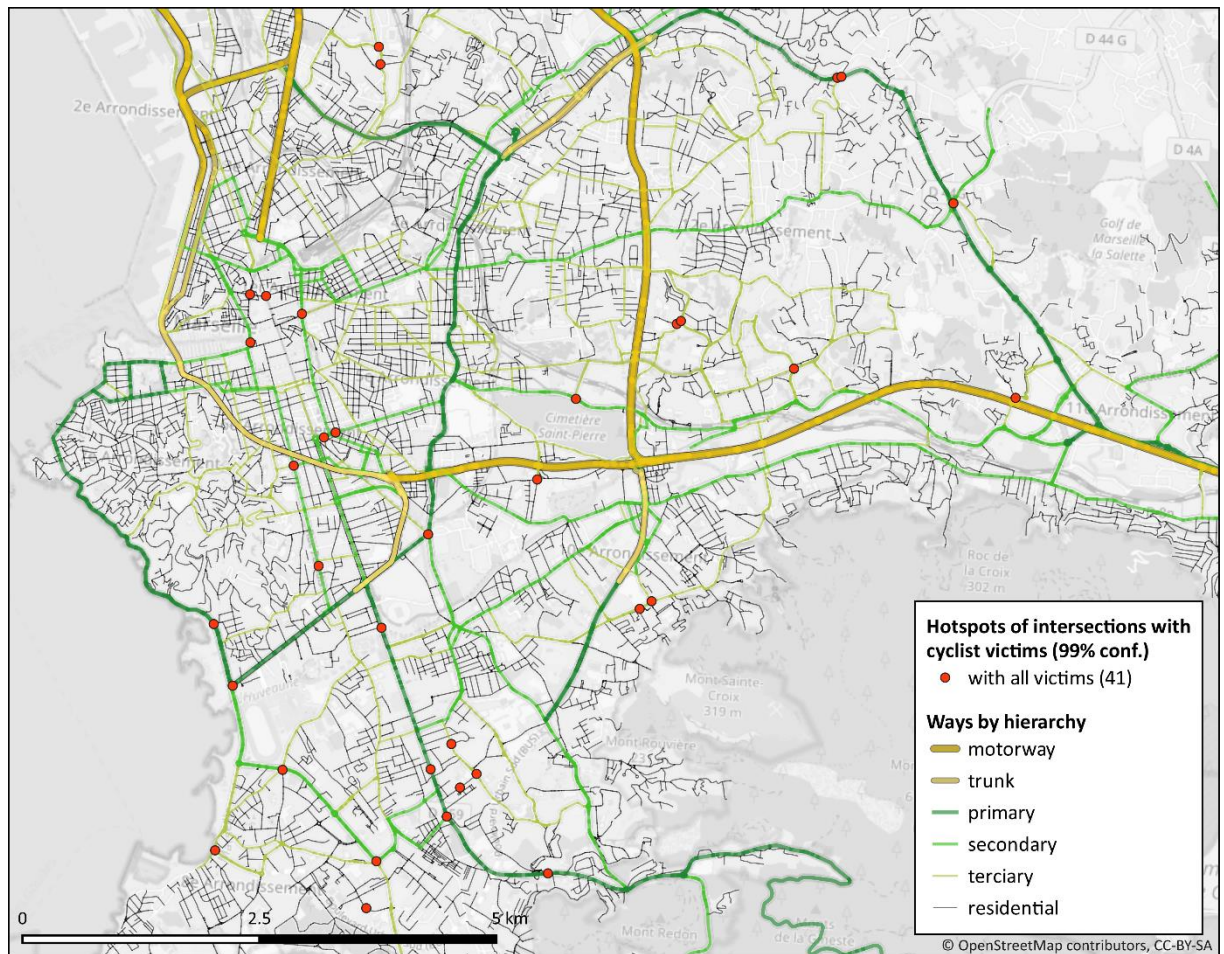


Figure A.34 Hotspots for all cyclist victims in Marseille

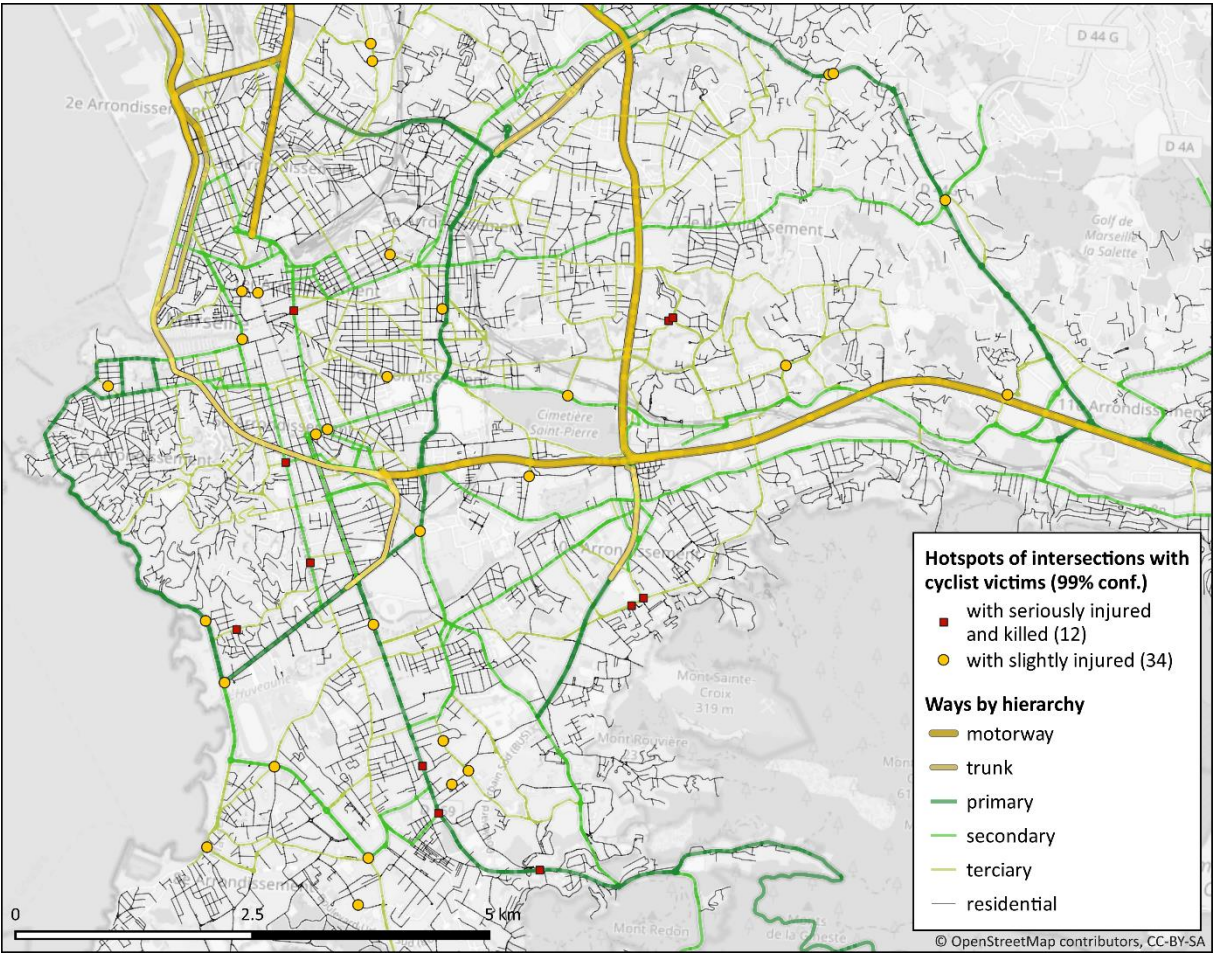


Figure A.35 Hotspots for cyclist victims by injury severity in Marseille

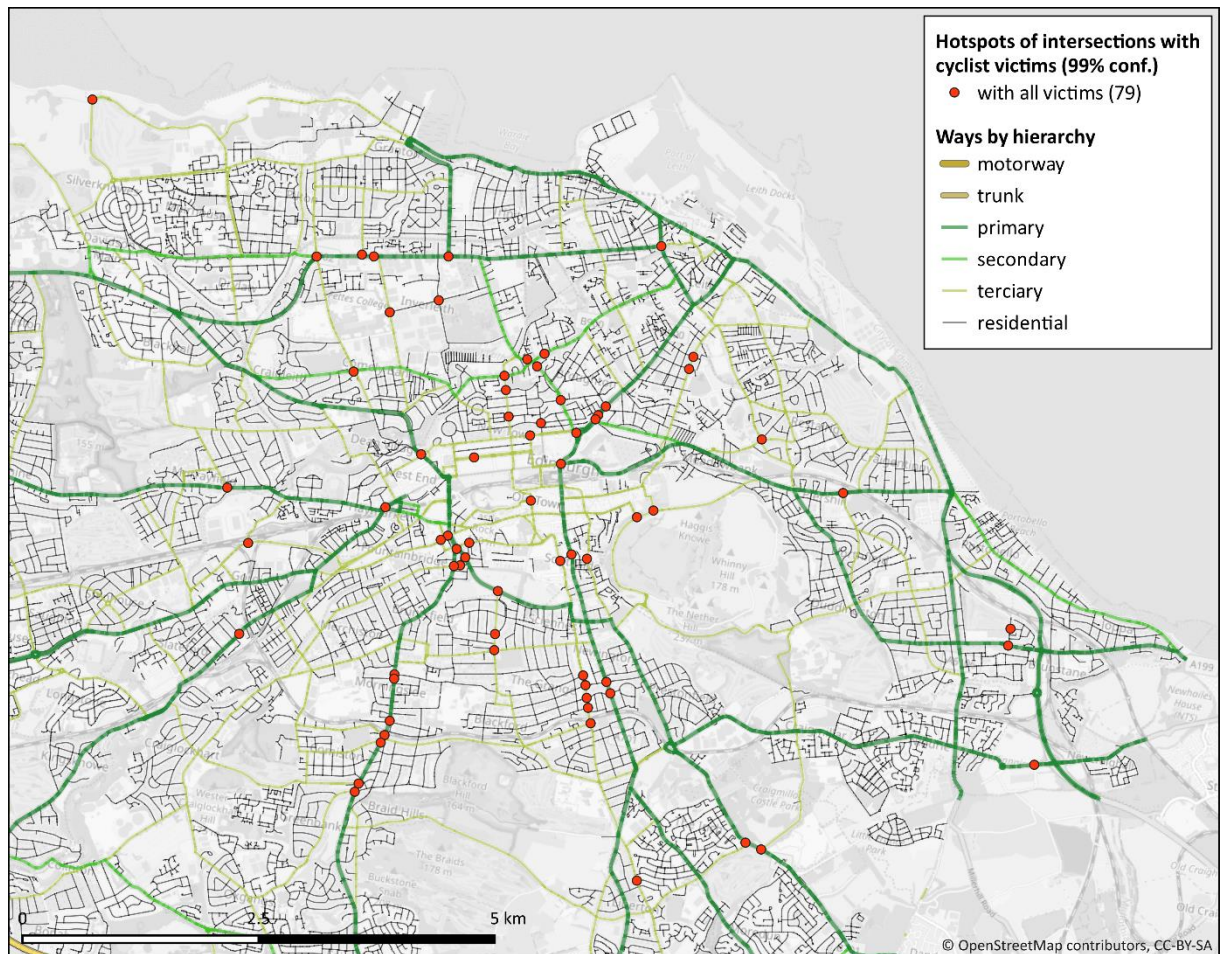


Figure A.36 Hotspots for all cyclist victims in Edinburgh

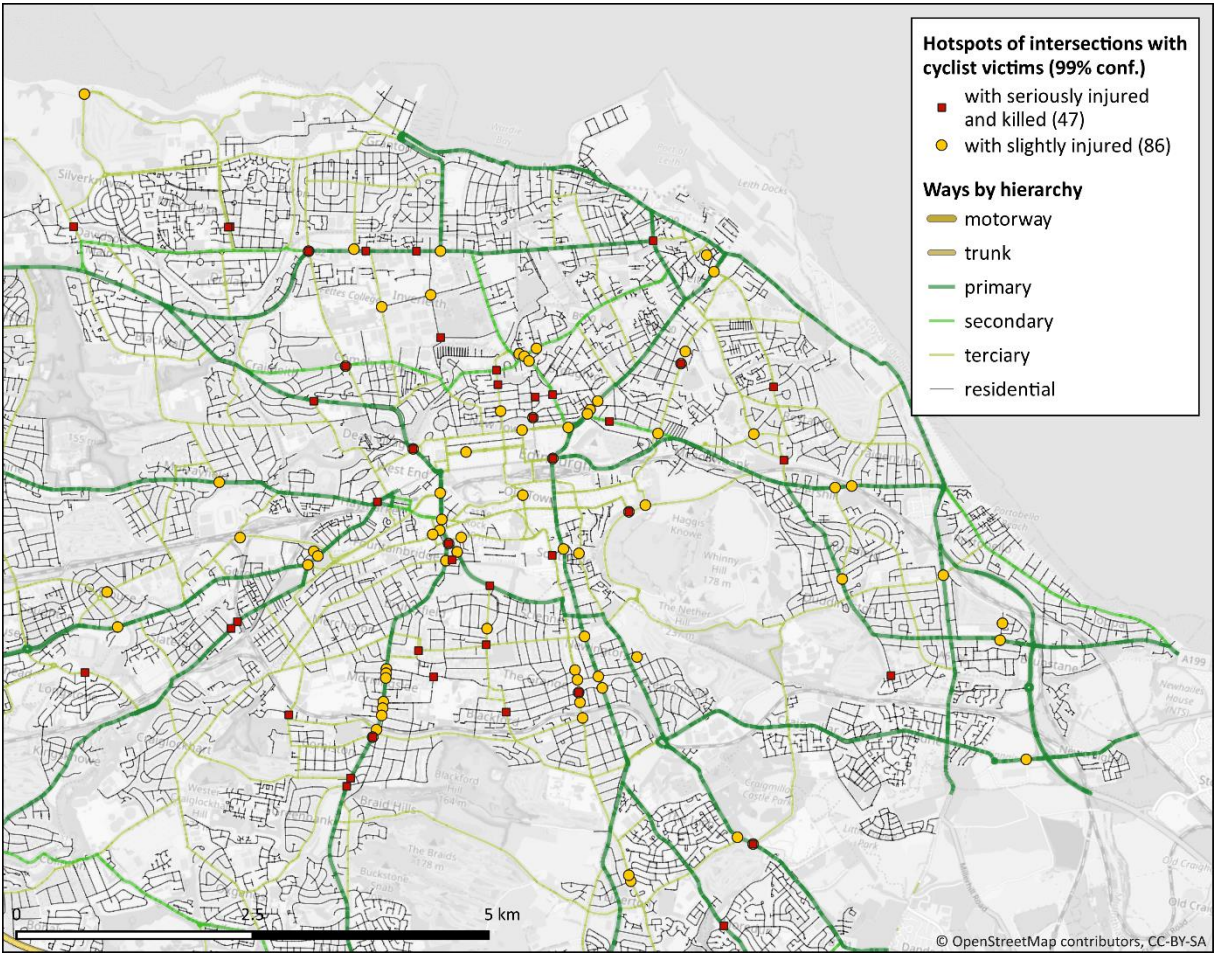


Figure A.37 Hotspots for cyclist victims by injury severity in Edinburgh