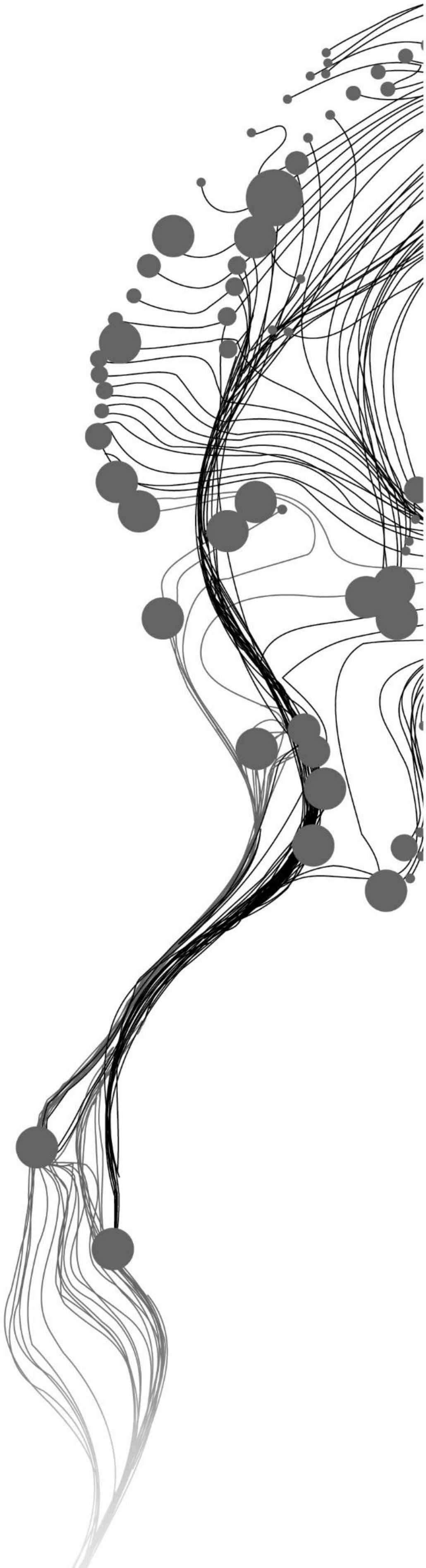
EXPLORING THE INFLUENCE OF THE ENVIRONMENT ON TRAFFIC ACCIDENTS IN ENSCHEDE, THE NETHERLANDS

LI XIONG
February, 2018

SUPERVISORS:
Dr. A.B. Grigolon
Ir. M.J.G. Brussel



EXPLORING THE INFLUENCE OF THE ENVIRONMENT ON TRAFFIC ACCIDENTS IN ENSCHEDE, THE NETHERLANDS

LI XIONG
Enschede, The Netherlands, [February, 2018]

Thesis submitted to the Faculty of Geo-Information Science and Earth Observation of the University of Twente in partial fulfilment of the requirements for the degree of Master of Science in Geo-information Science and Earth Observation.
Specialization: [Urban Planning and Management]

SUPERVISORS:
Dr. A.B. Grigolon
Ir. M.J.G. Brussel

THESIS ASSESSMENT BOARD:
Prof. dr. ir.M.F.A.M. van Maarseveen (Chair)
Dr. Tom Thomas (External Examiner, Centre for Transport Studies, ET-UT)
Dr. A.B. Grigolon
Ir. M.J.G. Brussel

ABSTRACT

This thesis seeks to study the spatial influence of environment on traffic accidents in Enschede, the Netherlands: its road segments/junctions with co-occurrence of traffic accident(accident hot zones), and the pivot toward the relationship between the urban environment and traffic accident distribution. Some suggestions for the improvement of road safety are given at the end.

The identifying of traffic accident hot zones relies on spatial analysis tools in ArcGIS. Fieldwork laid the foundation of the selection of environmental influence factors and allows for a qualitative assessment of the problem: why do specific accident locations have a high frequency of accidents but have not been identified as accident hot zones? Different model settings(buffer radius) were attempted, to explore the influence of different scale of environments on accident spatial concentration. Logistic regression modelling was utilised to evaluate the impact of environmental characteristics on the spatial concentration of traffic accidents.

Accident statistics in Enschede show that 98% of accidents involved the mode of auto. And the BPM (bicycle, pedestrian or moped) proves to be the most vulnerable road users. Besides, accident data show that more males were involved in accidents and younger people(0-24) are more likely to be involved in a crash. Hot zone analysis results in this research illustrate that hot zone analysis supplements hotspots analysis by identifying some neighbourhood with a moderately high number of accidents as hazardous locations. Most of the hot zones identified in Enschede were concentrated in certain districts near the city centre. And most of them were distributed along a 50km/h region of the road network. Buffer exploratory results show that the environmental characteristics in close proximity to a junction(100m buffer) have an influence on accident spatial concentration, whereas the road safety in road segments are mainly influenced by the surroundings within 150m of the road segment. Logistic regression results show that the road density within close proximity (100m buffer) around junctions is the first cause of hot zones. Residential areas have a slightly positive correlation to accident hot zones in the junction. Whereas the road segment with higher land diversity(150 m buffer) tends to have a higher risk of accident concentration. Some recommendations of urban safety development were given in the last part of the thesis.

Based on the results, the author suggests using the hot zone analysis to supplement the hotspot analysis to determine the dangerous accident locations. Besides, the acquired results support the planning of urban development to improve urban traffic safety.

Keywords:
Traffic Accident; Hot Zones; Road Environment; Binary Logistic Regression; Spatial Concentration

ACKNOWLEDGEMENTS

Working on this thesis after my study in ITC has been a crystallisation of what I learnt during the past year in ITC. I was able to complete my MSc thesis on time with the help of many during the school year at ITC.

I would like to sincerely thank my thesis supervisors Dr. A.B. Grigolon and Ir.M.J.G.Brussel for their patience, advice, and guidance. Even though after I come back to China, and with their tight calendars, they always made time for discussions and made sure I was on the right track. I am also deeply grateful for the help of Mr. Kees Lems, as a traffic expert in the Municipality of Enschede, who helped me a lot in understanding the background of my MSc topic. And I appreciate his providing all crucial data.

I am pleased to be a student at both University of Twente and Chang'an university at the same time through the cooperation program. Though this program, I had a chance to study at ITC, with advanced facilities and abundant resources and a chance to broaden my minds. A year study experience in ITC allows me to meet new friends at the same time of acquiring new knowledge. This valuable experience will no doubt benefit me for the rest of my life.

Thank you for my best friend LinFan, who always encouraged me during my depressed time even though we can only have video chat conversations.

Thanks to everyone who helped me during the last half year.

Finally, I am more than grateful for my family's enormous love and support.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

1. INTRODUCTION

1.1. Background and Justification

According to the (World Health Organization, 2015), more than 1.2 million fatalities and approximately 50 million victims at different degrees of severities are caused by road traffic accidents every year. Road safety is a serious social problem all over the world. Road accidents cause not only victims and deaths, but also considerable pecuniary damage (World Health Organization, 2015). In addition, the increasing number of victims lead to social problems, like the pressure on the social healthcare system, and the damage to public transport property. According to (Twisk, De Hair, & Otte, 2017; World Health Organization, 2015), road safety is a system term which involves investigating the interactions between the man, the vehicle and the environment (i.e. road infrastructure, traffic conditions and weather). To improve road safety, all these factors should be considered.

The Netherlands, having an injury rate of three victims per 100,000 residents (with a population of over 17 million) is a quite safe country regarding road safety (CBS, 2017). Road safety has been an important development priority during the past several decades in the Netherlands. As urban cycling paths began to increase in the 1970s, a great number of measures to ensure road safety have been taken and implemented by the municipality. For example, the legislation of traffic standards, protection of road users, and the improvement of transport infrastructure. Followed by the appearance of sustainable traffic safety at the beginning of the 1990s, this first example of a systematic approach organised by the national program for improving road safety works quite well (Nguyen, 2016). These improvement measures are implemented as a guide to sustainable safety. As described in (Wegman, Dijkstra, Schermers, & van Vliet, 2005), sustainable safety means preventing the probability of accidents to advance, thus, a pro-active approach to reduce traffic casualties. To reach sustainable safety, the understanding of traffic characteristics and causes is essential to ensure the targeted and effective improvement of road safety.

However, some indicators now show that the Dutch are losing ground. Since 2015, the number of traffic casualties has increased as opposed to previous years, and this trend continued in 2016 with more than six hundred people dead and countless victims caused by traffic accidents (ANWB, SWOV, Safe Traffic Netherlands, & Dutch Association of Insurer, 2017; SWOV, 2017). Thus, the Dutch continually want to improve the road safety because they do not want to live with preventable accidents. Besides, many traffic accidents happened on municipal roads in recent years especially on the municipal roads with speed limits of 50km/h (ANWB et al., 2017; SWOV, 2017; Wegman, Aarts, & Bax, 2006). With increasing urbanisation, a targeted approach in cities becomes even more relevant and urgent. Furthermore, about a third of the total road deaths and half of all serious accidents victims are cyclists, and cyclist deaths following motor vehicle crashes have decreased while deaths without motor vehicles have increased in the Netherlands (Schepers, Stipdonk, Methorst, & Olivier, 2017). So, it worthwhile to investigate which urban/road environments are much safer.

With more than 21,000 seriously injured in traffic, the Dutch are insufficiently aware of the cause and extent of the accidents they were involved in (ANWB et al., 2017). For example, they do not know exactly how many cyclists are seriously injured, where this happens (which type of road, at junctions or segments) and whether there is another person involved. Even in very serious accidents, it is not always guaranteed that there will be a systematic analysis of what has gone wrong, which factors in the environment, infrastructure, vehicle and behaviour has played a role and which safety lessons are to be drawn. In summary, it is therefore important to better understand both causes of accidents (accident prevention and monitoring) and ways to prevent accidents (via proactive and risk-driven traffic safety policy with a limited budget).

On the other hand, some traffic accidents are clustered in certain streets and some of the accident locations are at occasional random points due to people's behavior (Mr. Kees Lems, Traffic expert in the Municipality of Enschede, interview 5-Sept-2017), which illustrates that the spatial distribution of traffic accidents also needs to be investigated before exploring the causes of traffic accidents in certain sites. Spatial analysis of traffic accidents is the foundation of improving road safety in certain places, which helps to make targeted and effective investments. In terms of spatial analysis of traffic accident in the Netherlands, several pieces of research for identifying traffic accident hotspots were conducted in the early 20s in many big cities in the Netherlands (see. E.g. (Jovanis & Chang, 1986)) and subsequently, enormous targeted improvement measures were implemented and great effectiveness was achieved (Nguyen, 2016; Schagen et al., 2000). However, the results of the past are no guarantee for the future. What's more, there are few apparent serious traffic accident locations nowadays in the Netherlands, so using hotspots can no longer be efficient in Enschede. Instead, it is important to identify which zones have the most traffic accidents.

1.2. Research Problem

According to the SWOV (the national scientific institute for road safety research in the Netherlands), abundant statistical description of traffic accidents has been conducted to contribute to the vision of sustainable safety in the Netherlands. Generally, the causes of traffic accidents have been revealed in numerous research studies, and most of these studies were conducted based on the method of statistical analysis or qualitative analysis. However, the identification of traffic accidents from the perspective of urban design and urban environment is rarely investigated. Hence, this study attempts to explore the influence of the urban environment on the spatial distribution of traffic accidents, which contributes to a course of actions for both urban planning experts and transportation safety departments.

1.3. Research Objectives

The general objective of this research is to explore the spatial influence of the environment on traffic accidents in Enschede, the Netherlands.

To achieve the general objectives, two sub-objectives with corresponding research questions are developed, as follows:

1. Identifying the spatial distribution of traffic accidents in Enschede.
    a. What kind of method is suitable for identifying the spatial distribution of traffic accidents?
    b. What is the spatial distribution of traffic accidents in Enschede?

2. Exploring the influence of the urban environment on traffic accident distribution.
    a. What factors should be considered to evaluate the influence of road infrastructure/environment on traffic accidents?
    b. What factors should be considered to evaluate the influence of neighbourhood design/developments on traffic accidents?
    c. What kind of method, analysis unit and model are suitable to evaluate the influence of the environment on traffic accidents?
    d. Is there a (causal) relationship between the urban environment and traffic accident distribution? Can the urban environment account for the spatial distribution(concentration) of traffic

accidents?

e. What kind of suggestions can be given to help improve road safety in Enschede?

1.4. Research design

Both qualitative and quantitative analysis will be conducted to answer the formulated research questions proposed above. Spatial methods are the key quantitative methods throughout the whole research process, and spatial statistical methods consistent with qualitative analysis will be used to explore the influence of environment on traffic accident spatial distribution. The main work steps of the research methods illustrated above are shown in Figure 1, which will be used to achieve the objectives and answer each corresponding research question.

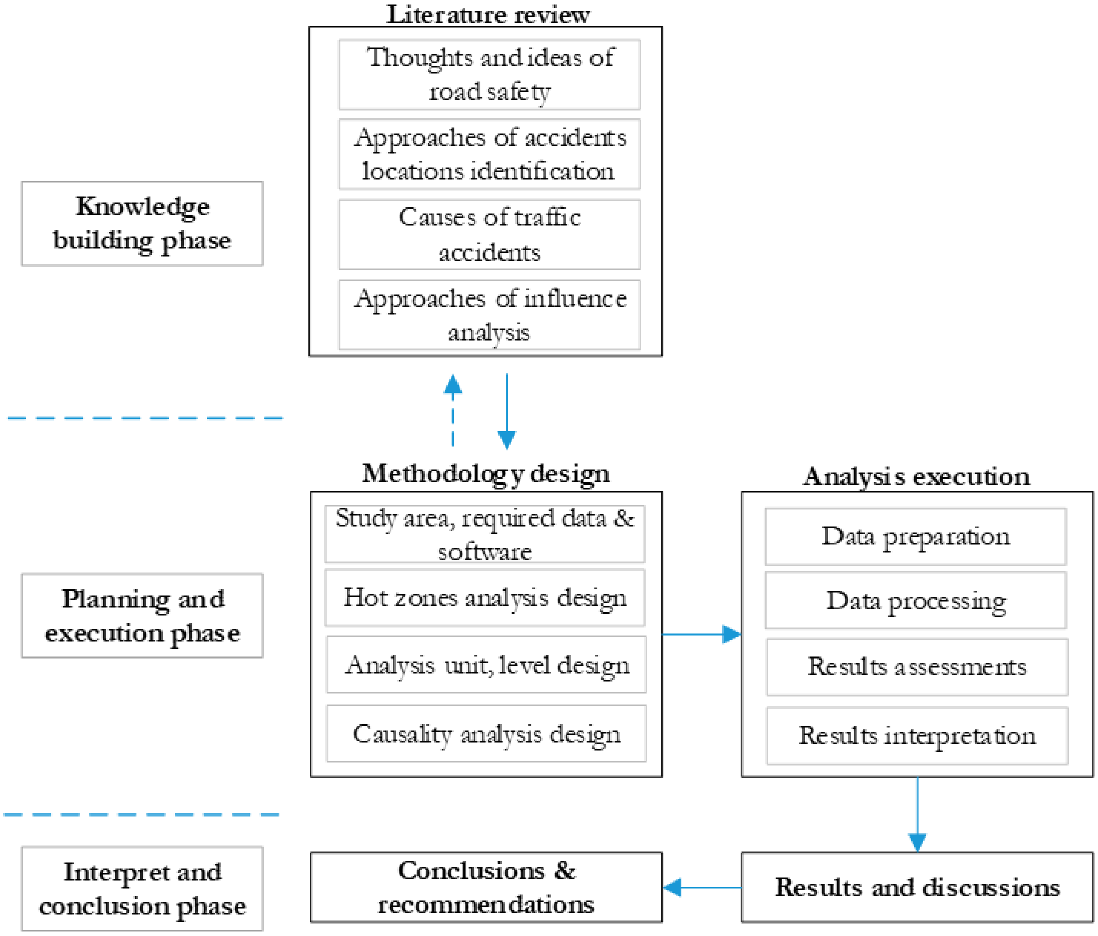


Figure 1: Methodological Framework

2. LITERATURE REVIEW

This literature review will critically discuss the concepts and thoughts of identifying traffic accident locations first, followed by a brief discussion about the causes of traffic accidents.

2.1. Identification of Traffic Accident Locations

"Which locations/area have a high-risk of traffic accidents?". This problem is known as geographical identification of traffic accident locations, which is the foundation for further investigation of high-risk traffic accident areas (Loidl, Traun, & Wallentin, 2016). The spatial analysis of traffic accidents allows for the explanation of its distribution characteristics or the prediction of its distribution trends.

2.1.1. The identification of traffic accident black sites

There are three typical concepts related to the problem of identifying high-risk traffic accident locations. The first is the black site methodology proposed in the 1980s (see, e.g.(Ogden, 1996; Silcock & Smyth, 1984)). The core idea of this concept is that the road junctions and connected surrounding roads within a certain distance are analysed first (Figure 2). Huang, Zhou, Wang, Chang, & Ma (2017) for instance, consider black sites at road junctions or segments, but this kind of focus will ignore the underlying influence of other elements of the road network. If criteria, such as the number of traffic accidents of a certain location is high, then this location is thought to be a black site that has a high-risk of traffic accidents. All the black sites are shown as "point" locations, which are the prototypes of hotspots methodology. Nevertheless, this black site methodology can only show the accident risk of a road within a certain distance around the junction, which is a rather limited analytical scale.

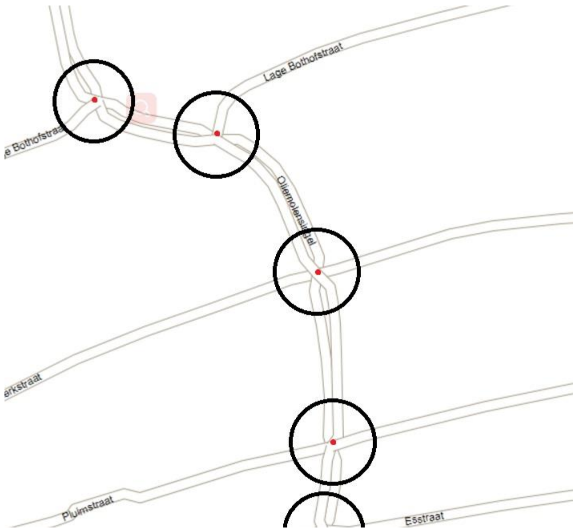


Figure 2: Virtual Road Network

2.1.2. The identification of traffic accident hotspots

Since more reliable and comprehensive spatial data are available, an advanced methodology called hotspot analysis was proposed later, in which the whole road network is considered (Loo, 2009). The identifying process of the hotspots is similar to the black site: counting the number of accidents within a fixed size circle and then testing its significance by some statistical analysis method (e.g. Monte Carlo method, see, e.g. Eckley et al., 2013). The basic idea of this methodology is dividing the whole network into basic linear units (see Figure 3, usually known as basic spatial units: BSUs, see, e.g. (Flahaut, Mouchart, Martin, & Thomas, 2003)). The length of the basic linear units should be suitable, which can be decided in different ways. For instance, Flahaut (2004) and Flahaut et al. (2003) divided it into segments with equal length. Others (e.g. Loidl et al., 2016) used the special algorithms to define the length of BSUs. Nevertheless, the resulting segment length should be long enough to distinguish the traffic accidents cluster and short enough to illustrate the distribution of accidents.

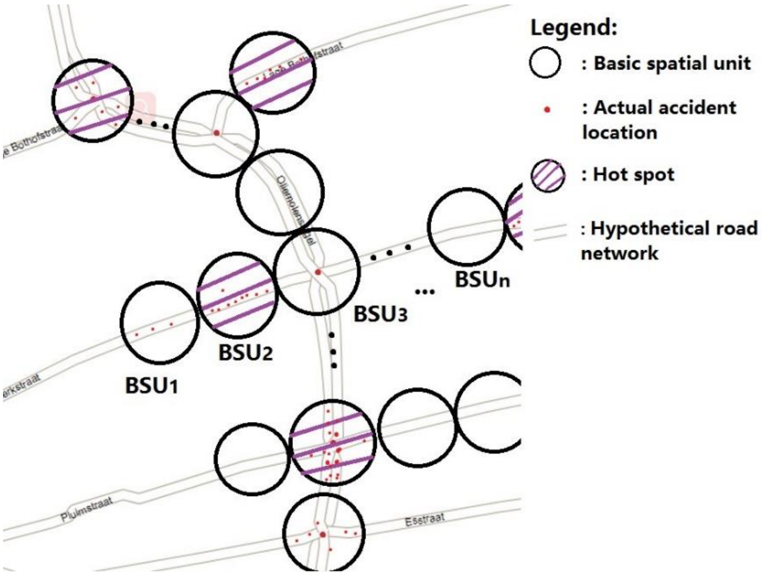


Figure 3: Basic Analysis Unit of Hotspots

The dangerous level of each road segments can be analysed and compared by the method of hotspots (Loo, 2009). In Figure 4, the hypothetical accident modes of one road are explained. Assume that a BSU will be identified as a hotspot if the number of accidents is larger than five, mode 1 has one hotspot at BSU 9. Mode 2 has two hotspots at BSU 1 and 10. However, is it conceivable that Mode 1 is safer than mode 2? On mode 1, the continuous BSUs from 8 to 12 with a high number of accidents obviously have a security risk which needs to be further inspected. It is maybe more dangerous for the road users when travelling in this continuous road segments than an individual hotspot. In other words, not only the hotspot location tends to have a risk of accidents, the whole neighbourhood can be (moderately) dangerous to the road users (Moons, Brijs, & Wets, 2009). Besides, there are possibilities that a high-risk hotspot can be a random result of traffic accidents caused by human behaviour like using a phone when cycling/driving, rather than caused by the deficiency of road/neighbourhood environment. This kind of hotspot has a high possibility of disappearing in the future even without intervention, which is called a phenomenon of randomness-to-regression (Elvik, 2006).

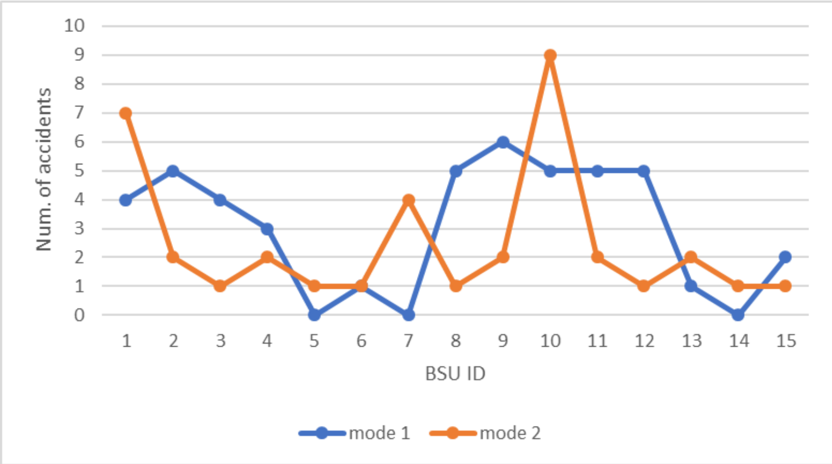


Figure 4: Two Different Types of Accident Modes in a Hypothetical Road

2.1.3. The identification of traffic accident hot zones

To decrease the rate of misjudgements of random traffic accident locations and increase the accuracy of the analysis results, a new concept called hot zones was initially proposed by (Black & Thomas, 1998). According to (Moons et al., 2009), high-risk traffic accident locations identified by the hot zone methodology consists of several continuous road units (defined by BSU). A positive spatial autocorrelation/co-occurrence of traffic accidents can result in the spatial clustering of more than one accident hotspot to form hot zones (Black et al., 1998). In this case, the high-risk traffic accident road streets are identified by a concentration of traffic accidents. In summary, when identifying high-risk traffic accident locations, the joint probability distribution (spatial autocorrelation) of BSU is considered in the hot zone methodology, which proves to be more superior and flexible to the hotspot methodology (Flahaut et al., 2003; Loo, 2009; Moons et al., 2009). Besides, the length of a hot zone is not fixed (shown in Figure 5), which is determined by the number of continuous hotspots.

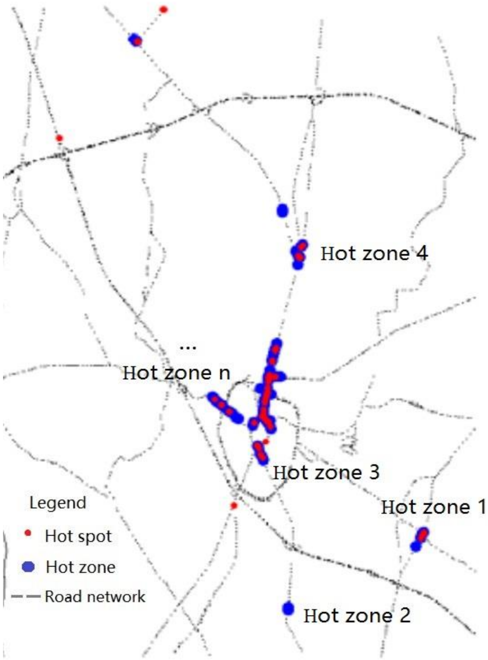


Figure 5: Hotspot vs. Hot Zone
Source: (Moons et al., 2009)

2.2. Causes of Traffic Accidents

The identification of traffic hot zones is a first step in improving road safety. Why do some locations have a higher risk of traffic accidents while others do not? To answer this question, it is essential to investigate the causes of traffic accidents.

As mentioned above, road safety is a system term which involves the interactions between the man, the vehicle and the environment. To improve road safety, all these factors should be taken into account since all these factors influence the frequency and severity of accidents simultaneously (Li, Zhu, & Sui, 2007). From the elements point of view, most of the research can be divided into these three categories to find the causes behind the frequency, collision or the severity of the traffic accidents. Other studies are tackling aspects of policy-making and evaluation (see, e.g. Dijkstra, Drolenga, & Maarseveen, 2007; Krishnakumar, Pulugurtha, & Nambisan, 2005; Vandenbulcke, Thomas, & Int Panis, 2014; Weijermars & Wesemann, 2013).

2.2.1. The role of human behaviour and vehicle on traffic accidents

Some of the interest is concentrated on the performance of road users including the psychological, behavioural and demographic aspects. For example, the severity of traffic accidents is influenced by demographic variables (such as age, gender, alcohol and drug use, etc.) due to their vulnerable characteristics or improper operation of vehicles (Li et al., 2007). The younger and old road user tend to be more vulnerable in an accident (see, e.g. CBS, 2017; Lourens, Vissers, & Jessurun, 1999; Wegman et al., 2008; World Health Organization, 2015). Besides, since more than half of the fatalities involve vulnerable road users like pedestrians and cyclists, much of the attention to improving road safety should move from car users towards these vulnerable groups when making transport design and urban planning measures. At the end of the nineteenth-century, measures to improve road safety concentrated on the aspect of car drivers but did not get expected results, which is not surprising because the measures should focus on the vulnerable road users rather than the car drivers (Wegman et al., 2008). Mandatory measures like the severe punishment of drivers with alcohol or drug use have received obvious benefits in many countries (World Health Organization, 2015).

In the Netherlands, most of the studies concerned with road safety are focused on this aspect according to the SWOV. For instance, the greatest number of deaths is found to occur among the younger and the older, especially cyclists aged 50 and over (ANWB et al., 2017; Wegman et al., 2006). This is mainly caused by distraction, unsafe behaviours like utilising drugs, alcohol or mobile phone. Besides, Netherlands, as a cycling friendly country, even having many policies and measures to ensure the relative safety of cycling, still has a third of accidental deaths and half of all serious traffic accident injuries to the cyclist. Since the behaviour of road users is hard to be controlled, the easiest way for the municipality to improve road safety is to increase the cyclist-friendly level of infrastructure and road environment.

On the other hand, the performance of vehicles also influences the possibility and severity of road accidents, especially motorised vehicles like cars and trucks. This aspect can be extended to the technical attributions of the vehicles. Sustainable safety version recommends that built-in attributions of the vehicles should be well designed to reduce traffic costs (Wegman et al., 2005). For example, seat belts must be used when driving a car; the design of the vehicles should be user-friendly with the majority of users to improve their sight observation or simplify the operation. For countries whose transport is mainly constituted of private transportation(car) have received significant improvements in road safety by harsh regulation and publishing notifications of improperly use of cars (Drummer et al., 2004; World Health Organization, 2015).Towards the objective of sustainable safety, the compulsory check of car conditions (especially those that have been used for more than two years) is one of the practical measures to improve road safety in the Netherlands.

However, the analysis unit of this research is the hot zone which contains many accidents within it. Thus indicators like human behaviour, the time of occurrence of a certain accident (hour, the day of week or season) cannot be considered in this research. This research focuses on a spatially oriented causality analysis: only accident location-related variables are considered here. Even though this is a limitation in this research, application of traffic hot zones has an advantage: more than one accident statistic in one black zone avoids the high possibilities of randomness when recorded as one event. Thus, this increases the level of reliability of the causality analysis results about the relationship between the urban environment and traffic accident distribution.

2.2.2. The influence of environment on traffic accidents

Environment, a broad term which includes road infrastructure (road design, quality, capacity, etc.), road environment (land use types, surrounding buildings and socio-economic environments) and weather (rainy or sunny etc.), influences road safety in different ways (Chen, 2015; Tasic & Porter, 2016). Abundant studies that investigated the causes of traffic accidents were concentrated on these aspects, which will be discussed in detail below.

i. Indicators related to road infrastructure

Many researchers explored the influence of road infrastructure on traffic accidents. This road infrastructure including the road types or road functions (e.g. road highway, arterial roads, collector roads), traffic flow, speed limitation, road geometry elements (e.g. slope and width), infrastructure quality (e.g. the road surface condition), traffic control (e.g. the separation of different modes of transport) etc. (see e.g.(Martin, 2002; Van Petegem & Wegman, 2014; Wang, Quddus, & Ison, 2013)). (Moeinaddini, Asadi-Shekari, & Zaly Shah, 2014) for example, concentrated on road network related variables like the number of blocks per area, road junctions per area and total road length per area. Traffic accident fatality rates have a positive relationship with these factors based on city level regression analysis. For the different length of the road network, a positive correlation has been found in many cases (Chen, 2015; Flahaut, 2004; L. Li et al., 2007; Rifaat, Tay, & De Barros, 2011; Y. Zhang, Bigham, Ragland, & Chen, 2015). Some studies results show that the number and pattern of road junctions are increasing the risk of traffic accidents(Chen, 2015; Flahaut, 2004; Moeinaddini et al., 2014; Shankar, Mannering, & Barfield, 1995; Y. Zhang et al., 2015). In the Netherlands, most of the accidents tend to happen at road network junctions according to (Nguyen, 2016), which demonstrated a need to be considered in this research.

As for the travel demand factors, traffic volume is the most common variable included (Flahaut, 2004; Li, Wang, Liu, Bigham, & Ragland, 2013), and almost all these researchers agree that higher travel volume results in more accidents. (Martin, 2002) for example discovered that high traffic flow areas tend to have a higher risk of traffic accidents, and the number of lanes(road width) can reflect the traffic volume and interaction degree among vehicles and can also influence the road safety, to some extent (Flahaut, 2004; Mannering, Shankar, & Bhat, 2016; Martin, 2002; Wang et al., 2013). For instance, (Martin, 2002) found that three-lane motorways have higher rates of accidents than two-lane motorways. Since a high percentage of road victims are vulnerable road users like pedestrians and cyclists, whether there is a separate non-motor lane (bike lanes and sidewalks) or not is a factor of great importance among researchers (Nguyen, 2016; P Schepers et al., 2017; Shankar et al., 1995). Separated non-motor lanes can significantly decrease the likelihood of accidents occurring in some cases(Nguyen, 2016) because it can lessen the interaction between different transport modes. This factor is very important in the context of the Dutch, where have a high percentage of modal share(P. Schepers et al., 2017).

Regarding speed limitation, traffic accident rates do not positively associate with the allowed maximum speed, especially in the Dutch context (Nguyen, 2016). In the Netherlands, streets with the speed of 50 km/h, which usually have more interaction between vehicles and non-motor road users, tend to have a high-risk of accidents (Wegman et al., 2005). Seldom has research taken speed intervention into account. Speed intervention, as a speed limitation measure, is usually conducted on a particular road segment, especially highway and road junctions(Mannering et al., 2016). The road design speed has been widely used to evaluate the influence of speed limitation(see, e.g. Jones et al., 2008; Van Petegem et al., 2014; Wang et al., 2013). However, the results achieved by using the actual operating speed may differ from the results achieved by using the road design speed (Mannering et al., 2016). Speed limited signs related to weather are examined by (Shankar et al., 1995) even though no significant relationship was found.

For road curvature which reflects the design of the road network was also examined by some researchers. For instance, (Rifaat et al., 2011) explained the impact of road patterns on the frequency of accident fatalities. In their research, the city can be classified into four different patterns corresponding to four street patterns: grid-iron, loops, lollipops and warped parallel. And loops and lollipops blocks tend to have a higher rate of accidents compared to the grid and straight patterns. Contrary, (Kaygisiz, Senbil, & Yildiz, 2017) claim that the occurrence of pedestrian-vehicle accidents has no apparent relationship to the road type(straight, curved or warped road segments) when there are plenty of pedestrians. So, it would be interesting to examine the impacts of these indicators in the Dutch context.

In many cases, the road quality was also considered (Flahaut, 2004; Nguyen, 2016; P. Schepers et al., 2017; Wang et al., 2013). The road quality usually refers to the road surface quality that depends on the road pavement materials, which has also been studied in some research (Flahaut, 2004). Other infrastructure indicators usually have a strong correlation relationship with some others. For example, the number of road lanes may have a strong correlation with the road width in the exploratory analysis.

Other factors like weather were added by some researchers when determining the influence indicators (see, e.g. (Navin, Zein, & Felipe, 2000)). The statistical characteristics of these factors are described, and the influence of these factors on traffic accidents are explored by different aggressive models (see, e.g. (Mannering et al., 2016; Nguyen, 2016; Van Petegem et al., 2014)). To improve road safety, the design of road infrastructure needs to be modified to harmonise with other elements to decrease interaction influence of man, vehicle and environment (Wang et al., 2013).

ii. Indicators related to neighbourhood environment

Another type of environmental influence factor is usually considered from the perspective of the adjoined environment of the road. In relation to neighbourhood environments, it refers to the land use type and socio-economic background in this research. Land use, as a motivation for commute and daily activities, has been demonstrated to have a known relationship to traffic accidents. Some studies found that traffic analysis zones with higher land use diversity tend to have higher bicycle-involved accidents (see, e.g. Chen, 2015) due to the higher traffic volume attracted by mixed land use. Whereas in (Jones et al., 2008)'s research, variations in rural land use have a very limited impact on road death numbers. In addition, (Kaygisiz et al., 2017) claim that road safety improvement measures related to land use should concentrate on taking precautions in areas with accident-increasing factors instead of simply decreasing the land use diversity. (Fan,

2009) claims that mixed land use which can contribute to economic vitality is an urban design principle (e.g. transit-oriented development, TOD) in many countries. (Singh, Fard, Zuidgeest, Brussel, & Maarseveen, 2014) have a similar point of view; higher land diversity decreases vehicular trips, which can give rise to road safety to some extent. Since inconsistencies exist in the influence of land use diversity; it would be meaningful to explore the relationship between land use diversity and traffic safety.

As main land use categories, the building density/non-built up(green area, water or farmland) must be considered according to (Herold, Couclelis, & Clarke, 2005). They claim that building density can reflect the building patterns of the urban environment, thus can have different transport system and socio-economic characteristics related to traffic safety. Land use in an urban environment can be classified into three categories including buildings, vegetation and other (transport, water or soil). High building density urban regions tend to have a higher risk of accidents due to the higher number of human activities (Amorim, Ferreira, & Couto, 2017).To illustrate the influence of different land use categories on road safety in an urban environment, building density was chosen here.

Except above generalized land use indicators, the types of land use have been classified and explored in more detail in various studies. First, leisure places which are highly correlated with daily activities, contribute greatly to travel volume and have demonstrated to have a positive correlation with accidents in many previous studies(Herold et al., 2005; Miranda-Moreno, Morency, & El-Geneidy, 2011). Leisure places like bars/pubs were included due to their connection with traffic accidents, by increasing the opportunity for alcohol consumption (Dai et al., 2016; Mathijssen, 2005). For example, (LaScala, Gerber, & Gruenewald, 2000) found that accidents involving pedestrian victims have a high percentage of drunks leaving or spending time around these entertainment venues. Second, schools, another establishment that can attract lots of trips are also considered by many researchers(Dai et al., 2016; Hadayeghi, 2009; Hadayeghi, Shalaby, & Persaud, 2010), even though no significant influence was found. Third, many previous studies have found a positive correlation between the number of transport transit stops and the risk of accidents, especially those involving pedestrian(Clifton, Burnier, & Akar, 2009; Dai et al., 2016; Fan, 2010). For example, according to (Kaygisiz et al., 2017), the risk of accidents increases when drivers try to find a place to park. Other researchers like (Amorim et al., 2017; Flahaut, 2004; Nguyen, 2016) also took this into account. These factors influence traffic accidents through indirect human activities, in other words, traffic volumes.

Population density can be considered as a substituted variable of travel volume in the urban road network, which has been illustrated in many previous studies. For instance, (Amorim et al., 2017) found that high population density has a negative correlation with accidents. This was because high population density tends to occur in residential areas, which tend to have fewer human activities during most of the day compared to commercial or leisure areas. (Siddiqui, Abdel-Aty, & Choi, 2012) found that high population density has a higher positive correlation with pedestrian and bicycle accidents in the urban environment. So, it would be interesting to explore the influence of population density in traffic accidents in the Dutch context.

Income and education, which can reflect the types of travel mode to some extent, was found to be negatively correlated with traffic accidents in many previous studies (Nguyen, 2016; Rifaat et al., 2011; Tasic at al., 2016; Yao, 2013; Y. Zhang et al., 2015). Lower income regions are more likely to have accidents involving non-motor users, as found by (Siddiqui et al., 2012; Rifaat et al., 2011). According to (Wegman et al., 2006), the lifestyle of double-income households gives rise to the higher use of motor vehicles due to the combination of commuting and picking-up of school children. (LaScala et al., 2000) found that higher education neighbourhoods tend to have lower accidents because lower education residents usually work outside the community, where there are higher roadway hazards.

Other aspects like urban development have also been explored (see, e.g. (Amorim et al., 2017; Henning-Hager, 1986; Mohan, Bangdiwala, & Villaveces, 2017)). Specifically, (Mohan et al., 2017) concluded that a city with a higher number of wider roads tends to have higher traffic accident death rates. Examples of such research include the urban vs. rural areas (Ossenbruggen, Pendharkar, & Ivan, 2001), city comparison(Mohan et al., 2017), national comparisons(Yang & Otte, 2007). (Ossenbruggen et al. 2001), for example, found that rural areas tend to have a lower risk of traffic accidents compared to shopping and commercial areas, which implies that land use types can influence the rate of traffic accidents.

This short review of literature illustrates that road infrastructure and urban environmental factors have great influence on traffic accidents. In addition to other influence factors like the population density, the socio-economic characteristics of a neighbourhood are also considered in some studies, which are mainly focused on motor-pedestrian accidents (see, e.g. (Amorim et al., 2017; Cottrill & Thakuriah, 2010; LaScala et al., 2000)). Generally, neighbourhoods with lower income, old buildings, and less educated residents, tend to have a higher risk of pedestrian-involved accidents.

The common characteristic of these studies is the attribution of the dependent variable concentrated on the traffic accident itself, like the frequency or the severity of a BSU. It may be efficient to use the frequency of traffic accidents as the dependent variable for the causes of accident hotspots because the basic analytical unit of these studies is BSU. However, as discussed in section 2.1, traffic accidents occurring in certain hotspots may have the possibility of randomness. Thus the number of real traffic accident dangerous locations were overestimated by using the methodology of the hotspot, which will decrease the accuracy of causality analysis results. What's more, the spatial autocorrelation between the hotspots was ignored in the consulted studies. Thus, the present research will explore the influence of the environment on traffic accident spatial distribution (spatial concentration/co-occurrence) at the level of the road network. Compared to the traditional dependent variable of accident-related characteristic, a binary variable (the basic analysis unit will be given the value of one if an analysis unit belongs to the road hot zones, zero otherwise) will be used as a dependent variable in this research, which can help illustrate the influence of environment on the spatial concentration of traffic accidents instead of the risk of accidents on the roads.

3. METHODOLOGY

To improve road safety in Enschede, this thesis will investigate the spatial patterns of traffic accidents, followed by exploring the influence of the urban environment on traffic accident distribution; and some corresponding suggestions for improving road safety will be given at the end. This section describes how this research will achieve the general objectives. Details include the description of the study area, the explanation of data, and the types of methods corresponding to each research question.

3.1. Study Area

The study area is the city of Enschede, an active university city located at the eastern part of the Netherlands in the province of Overijssel, with about 160,000 inhabitants (Province of Overijssel, 2017), which is illustrated in Figure 6. Regarding urban development, the patterns of urban growth were rather irregular until an official expansion plan was formulated in 1907 (Wikipedia, 2017). The city centre zone constructed at an earlier stage (approximately before 1930) compares to the other regions. The southwest part of the city was built during the period 1950 to 1960 and the northeast of the city developed in the year around 1960 (personal communication, Mr Kees Lems, Traffic expert in the Municipality of Enschede, interview 5-Sept-2017). Thus, the characteristics of the road environment vary a lot in the different regions of Enschede.

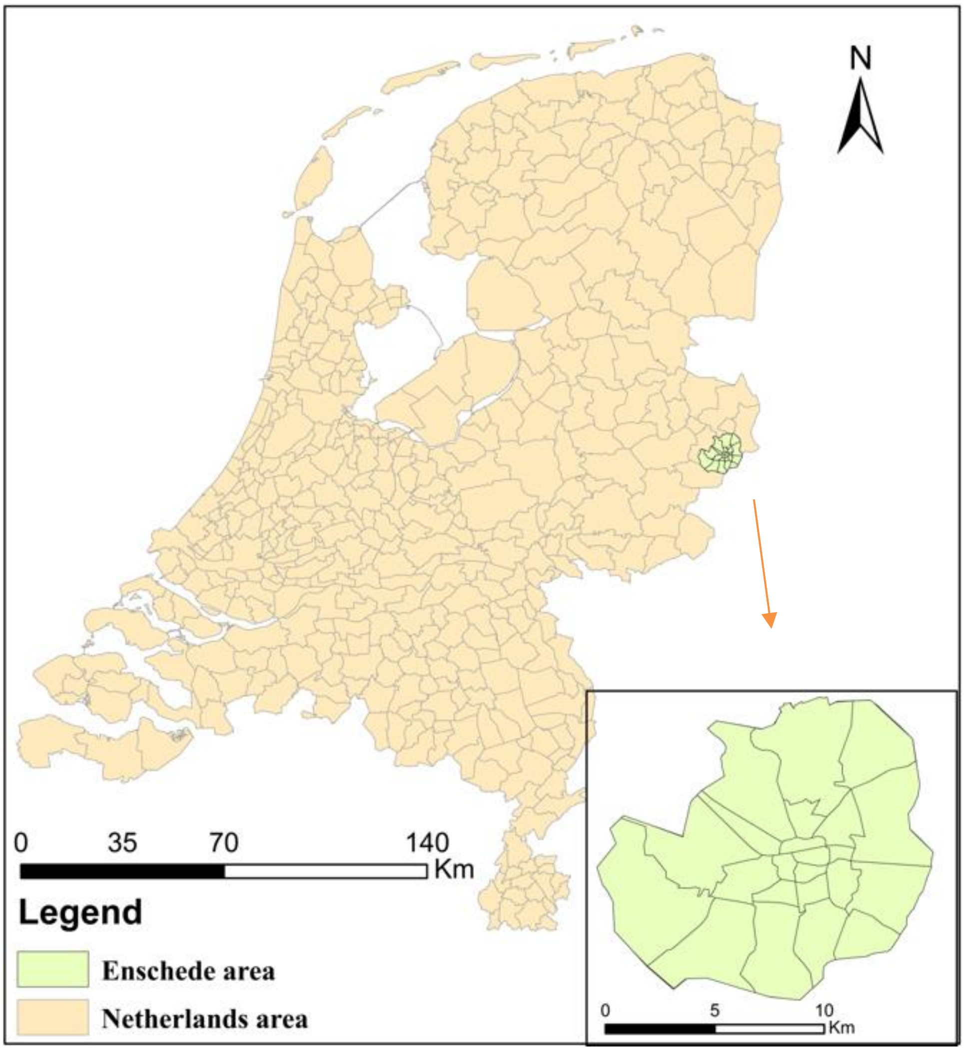


Figure 6: Study Area

The transport system in Enschede includes four train stations, a reliable bus system and good cycling infrastructure. To explore the influence of the urban environment on traffic accidents, only the accidents happening on municipal roads are included in this research. The good cycling infrastructure in Enschede makes it a livable and active city which was awarded as one of the best cycling cities in the Netherlands in 2014 with more than 65% of short distance travellers cycling (Bicycle Dutch, 2014). However, there are still many traffic accidents concentrated in several streets of Enschede as shown in Figure 7. Additionally, the city council of Enschede intends to conduct "Enschede Fietsstad 2020" to make it a bike city (Koolhoven, 2017). For this reason, it is meaningful to investigate the influence of the environment (infrastructure & road neighbourhood) on the spatial distribution of traffic accidents in Enschede to create a sustainable safe urban environment for all road users, especially the cyclist and pedestrian who are the most vulnerable road users.

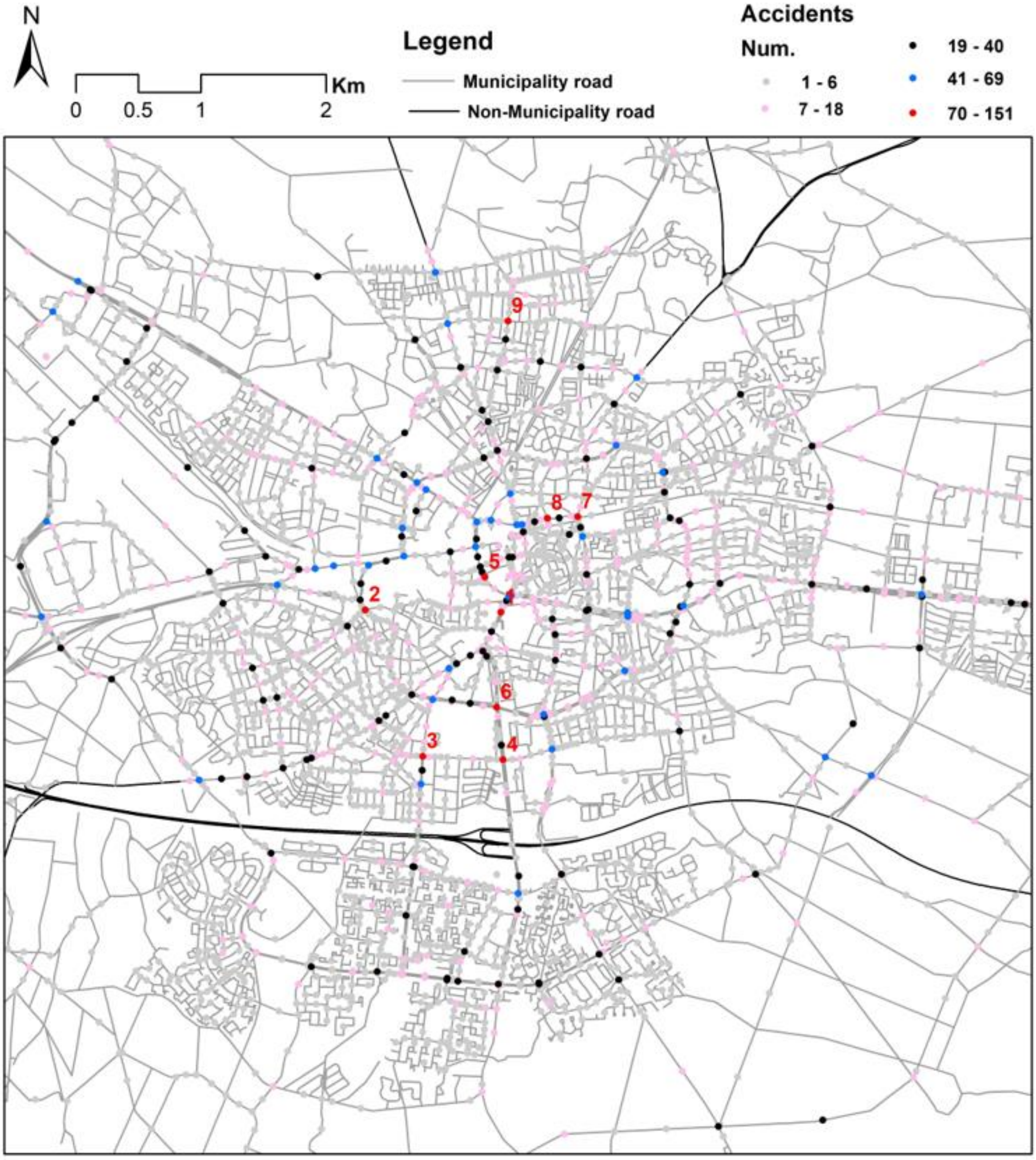


Figure 7: Traffic Accident Locations in Enschede from 2001 to 2015(2009 is missing)

3.2. Data Description and Statistical Analysis

3.2.1. Sample characteristics of the accident data

In this research, secondary data provided by Enschede Municipality will be used. On the one hand, accident data exported from VIA software[1] was offered by the municipality, including the georeferenced information of recorded accident locations and the number of accidents in each recorded point (3987 points including

[1] A road safety platform which incorporating various data sources including accidents, speed, audit results or reports. This web-based software is also allowing for analysis data online or exporting data after authorized.

18,074 accidents in total recorded in the municipality road network). However, there is a repeat count problem of accident location when exporting the data from the VIA software. To be more specific, some accident locations have different objective ID/identification; they even have the same coordinate, which may lead to wrong analysis results. Thus, the number of accidents in these duplicate locations were summarized by the following step: the coordinate for each location were added first in the *Auto.shp*, which was then exported to Excel. A column named *Coordinate1* was added, and the value of *Coordinate1* is equal to the value of *Coordinate*. Then deleted the duplicate value by column *Coordinate1*.
Added a new column which was named ONG1, calculated by the function *SUMIF (Coordinate, Coordinate1, ONG)*. Finally, this *Auto.xls* was joined to the *Auto.shp* in ArcGIS by *Coordinate1* and only match records were kept in this process.

The recorded accident location was not the actual location of accidents because the police often only determine the nearest house number to record the accident location for accidents happening on the community street. For each recorded location, the number of accidents and the number of victims was counted according to the severity of accidents from 2001 to 2015 (the year 2009 is missing), which includes ONG (Total number of accidents recorded in each point), UMS (The number of accidents that have no victims), ONGSLA (The number of accidents that have victims), ONGDOD (The number of accidents that involve dead people), ONGZKHGEW (The number of accidents that involve people sent to the hospital), ONGOVGGEW (The number of accidents that involve slightly injured people), ONGERNST (The number of serious accidents).

After validating the traffic accident on the road network dataset, these recorded points can be classified into two categories. One category is located at the junction, which aggregated all accidents happening at the road intersection. Another type is the point that is located on the medium of the road segment, which aggregated all accidents happening on that road segments. That means that even though the actual accident locations may not be the same, they will be recorded at the same point when they are geographically close to each other. The file format of this traffic accident data is Shapefile, which can be input to ArcGIS directly, as shown in Figure 7. There are 18,074 traffic accidents recorded with approximately 61% happening at the junction, and 39% happened on the road segment from 2001 to 2015 (the year 2009 is missing) in the municipality road network, as shown in Table 1. The sample characteristics of accidents that happened on the road junction and road segment were statistically described in Table 1. Accidents that did not happen on the municipality roads will be excluded from this research.

Table 1: Sample Characteristics in Road Junction and Road Segments

Variables	**Classes**	**Explanation**	**Junction** N=11,104(%)	**Segments** N=6,970(%)
Transport mode	Auto	statistics of accidents as long as involved the mode of auto (The passenger car, van, lorry, motor or other vehicles)	98%	96%
	BPM	statistics of accidents as long as involved the mode of BPM (bicycle, pedestrian or Moped)	26%	19%
	Bicycle	statistics of accidents involving the mode of bicycle	14%	6%
	Auto+BPM	statistics of accidents happening between Auto and BPM	24%	15%
	Auto+bicycle	statistics of accidents happening between Auto and bicycle	0%	2%
	Auto+pedestrian	statistics of accidents happening between auto and pedestrian	10%	8%
	Auto+Moped	statistics of accidents happening between Auto and Moped	98%	96%
Severity	Non-victims	statistics of accidents that have no victims	85%	86%
	Slight injured	statistics of accidents in which the road users were slightly injured	10%	9%
	Hospital	statistics of accidents that involve road users who went to the hospital	4%	5%
	Fatal	statistics of fatal accidents	0%	1%
Variables	**Classes**	**Explanation**	**Junction** N=2,240(%)	**Segments** N=1,630(%)
Gender	Male	statistics of victims whose gender is male	55%	59%
	Female	statistics of victims whose gender is female	44%	40%
	Not filled in	statistics of victims whose gender is not known	1%	1%
Age	0-24	statistics of victims whose age is 0-24	36%	35.5%
	25-39	statistics of victims whose age is 25-39	23%	27.7%
	40-49	statistics of victims whose age is 40-49	13%	12.1%
	50-59	statistics of victims whose age is 50-59	11%	11.3%
	60-69	statistics of victims whose age is 60-69	6%	6.2%
	70 and older	statistics of victims whose age is 70 or older	8%	6.4%

3.2.2. Attribution information of the accident data

To analyse the transport mode involved in each accident, the mode of transport was classified into six categories including Auto, BPM, Auto+Bicycle, Auto+Pedestrian, Auto+Moped, Auto+BPM. The category of auto including the passenger car, van, lorry, motor and other vehicles, but not including rail vehicles (only two accidents involved rail vehicles during the last 14 years). The category of BPM including the bicycle, pedestrian and moped. The number of accidents involving the fixed/loose objective or animal is zero, which will not be shown here. Since most of the accidents involved at least one mode, one accident has the high possibility of being repeat counted when counting the number of accidents according to the categories of mode. Thus, the total number of accidents (22,307) is not equal to the simple sum of accidents counted according to the transport mode. After subtracting the repeated number of accidents in each mode, the statistical results are shown in Table 1, which illustrates the percentage of accidents in the different category of transport mode.

As shown in Table 1, the percentage of accidents in different mode group at the junction is slightly higher than that of on road segments, in general. Enschede, as a cycling friendly city, most accidents happened there involved cars (auto), with about 98% at the junction and 96% on road segment. Surprisingly, the same percentage of accidents exists in the mode group of Auto+Moped. The high-risk of moped riders are mainly caused by the high speed combines with little physical protection. In this research, the BPM (bicycle, pedestrian or Moped), as one of the major activities in the urban environment, must be considered. The percentage of accidents involving the mode BPM is about one fourth, which is much better compares to that of Auto.

Looking at the demographic characteristics (gender and age) of victims, accident data show that more males were involved in accidents, no matter the severity level (indicated in Table 1). For the age which can reflect experience sensitivity of reflexes, accident data illustrates that younger people(0-24) are more likely to be involved in an accident. This is mainly caused by lack of experience and skills when participating in the traffic as a cyclist, moped rider, motorcyclist, and these transport modes tend to have a higher risk of accidents especially the Moped(Wegman et al., 2005).

When taking the severity of the accident into consideration, the percentage of accidents of different severity calculated under the different mode of transport is illustrated in Figure 8 (junction) and Figure 9 (road segments). Whether road junctions or road segments, for accidents which involved auto, approximately 10% of victims were slightly injured and 4.00% of them went to the hospital. Whereas, for accidents which involved BPMs (bicycle, pedestrian or moped), the situation is rather worse (11% and 26% respectively at road junctions and even higher on road segments). For accidents involving the bicycle in the cycling city Enschede, which continues to be worse, more than 12% of victims went to the hospital, and 1% perished. The low percentage of total accidents but the high percentage of victims illustrated that the BPM (bicycle, pedestrian or moped) are the most vulnerable road users compared to those travelling by auto.

Since almost all accidents involving BPM happened on the city roads which also allow car users, any accidents involving the auto will be included in this research. As can be seen from Table 1, more than 96% of the accidents involve the mode of the auto, and accidents do not involve auto will be excluded from the accident data set. For the statistical number of accidents in each point, the total number of accidents is adopted, instead of the number of fatal accidents or serious accidents, because the number of fatal accidents or serious accidents in most recorded points has a value of 0. Besides, the low number of fatal accidents or serious accidents in the past does not mean that the recorded location is not dangerous; after all, there were accidents happening in that place before. Last but not least, there is no temporal information in the dataset. Thus, the temporal scale was not considered in this research. After excluding the accidents that did not happen in the municipality road, and accidents not involving the auto, there are about 17563 accidents recorded in 3903 points in the municipality road, with a mean value of five accidents per point. For the location of the accidents, 10860 accidents recorded at 1704 points happened in a road junction and 6703 accidents recorded at 2199 points happened in the road segment.

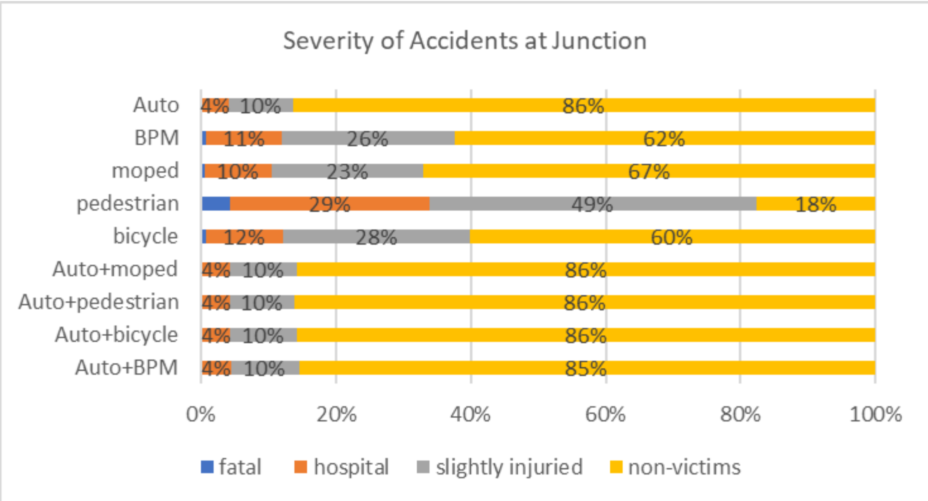


Figure 8: The Severity of Accidents at Junction Over the Transport Mode Group

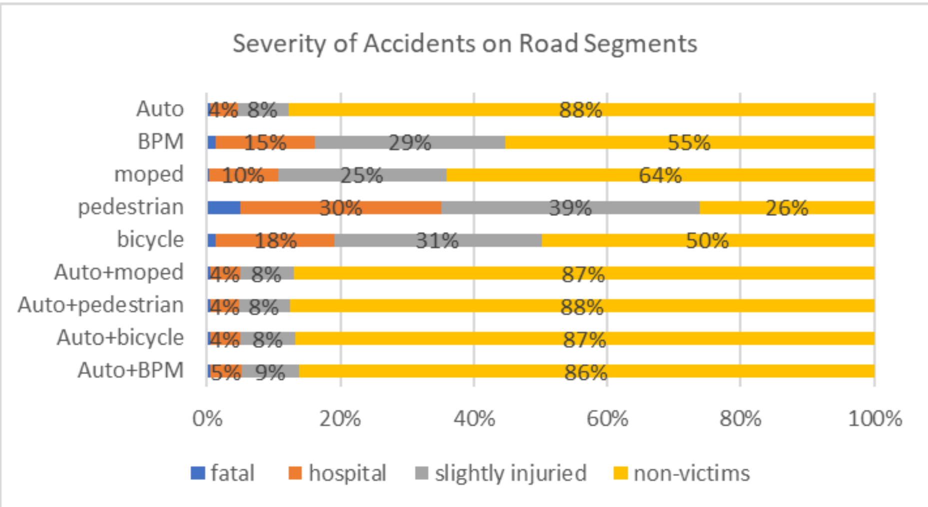


Figure 9: The Severity of Accidents on Road Segment Over the Transport Mode Group

3.2.3. Land use and socio-economic data

On the other hand, road network dataset, land use data set and socio-economic dataset were also provided by the Enschede municipality. The total length of the road network is about 1051 kilometres. The road network dataset only has the ID and length of each street. Whereas other road-related characteristics (such as road widths and the shape of junctions) are not included in the dataset. Thus, many factors (e.g. traffic volume, the number of lanes, separate bicycle lane) related to road infrastructure were not included in this research.

The land use data is derived from BAG (*Basisregistraties Adressen en Gebouwen*), which contains details of each building including address, year of construction, surface area and purpose of use. The geometry type of land use data is the point, and there is a separate build footprint for each building. Since it is possible for a building to have several accommodation objectives, several points in one location may have different attributes and thus indicate a different address (house number). Many points have the same address but in different records due to the non-continuous *Document Date* (Registration Date) in the municipality. These repeated records should be deleted before further analysis.

The feature class *pand.shp*(polygons) contains the footprints of the buildings. The data contains history, so buildings that no longer exist are still there (recognisable by the attributes in the columns *EndDate, Status* and *IndicationRecordInactive*). Properties can also overlap (history and current). For accident causes analysis only buildings in use between 2001 and 2015 are relevant. Thus, records with buildings abandoned for use before 2001 or with data beginning after 2015 should be excluded from the land use dataset. It is also possible for a location to have several overlay footprints due to the re-expression of the expand/multi-floor buildings at the same locations. These duplicated polygons should also be excluded before further analysis.

The dataset from the CBS (Dutch Statistics Bureau) include considerable attribution information, including the number of people in different age group, gender group and marital status, but only the information on total population is relevant for the causality analysis.

3.3. Methods of Identifying Traffic Accident Hot Zones

Spatial identification of traffic accident locations allows accounting for the spatial characteristics of the accident locations in comparison to traditional statistical (Poisson) regression models to determine high-risk traffic accident sites (Moons et al., 2009). Besides, the map generated in the spatial analysis is more intuitive and direct to non-professionals in explaining high-risk traffic locations. Thus, the spatial analysis will be conducted to investigate the spatial distribution of traffic accidents. As discussed in section 2.1, traditional traffic accident spatial analysis methods (hotspot analysis) may be impacted by random causes incorporating human behaviour like using a phone when cycling/driving, rather than being caused by the road environment. To explore the influence of the road environment on traffic accident spatial distribution, the spatial analysis of hot zones is more suitable for this research.

Although continuous high-risk traffic accident locations (hot zones) are dangerous to road users, compared to accident hotspot identification, there is not too much systematic methodology to identify the traffic accident hot zones on the road network (Loo & Yao, 2013). Usually, there are three main stages to identify the traffic accident hot zones based on geographic information system (GIS). The first stage is the validation of a traffic accident on the road network dataset. In this research, the accident dataset from the VIA software was exported to ArcGIS. Following this, usually, the basic spatial unit (BSUs) is defined and the number of accidents in each BSU will be calculated (see, e.g. (Flahaut et al., 2003; Loo, 2009; Loo et al., 2013; Moons et al., 2009)). However, almost all this research used the road hectometers for the definition of basic analysis unit because the study object of their study is the highway or the urban arterials that have hectometres. That method is not suitable for my research because this research is focused on accidents happening on urban municipality roads with no hectometers. Instead, recorded accident points (3903 in total, where accident data are spatially available) are used as the basic analysis unit. Most of these recorded accident points located in the centre of the road segment or the junction of the road, which records the total number of accidents, happened close, in geographical proximity. For quantitative analysis of the impact of road environment on traffic accidents, the transport mode should be considered in the process of determining the influence factors. As discussed in section 3.2.2, accidents that do not involve the auto will not be considered in this research. And the total number of accidents is calculated and used for the statistic of accidents in this research.

The final step is the identification of traffic accident hot zones. Hot zone identified in this research refers to the road segment/junction which has spatially aggregated recorded accident points. These recorded accident points are close by in terms of geographical proximity, and the number of accidents recorded should be high (calculated using ArcGIS according to the studied point and neighbouring accident points) as discussed in section 2.1.3. To assess the level of spatial concentration of traffic accidents, the local indicators of spatial autocorrelation (LISA)/ local Moran's I is used in this paper, which can evaluate the degree to which the accident number of the studied recorded accident point is similar to or different from its neighbouring observations (Flahaut et al., 2003). According to (Anselin, 1995), the LISA for the studied location i is

$$I_i = (x_i - \bar{x}) \sum_j W_{ij} \left(x_j - \bar{x}\right)$$

Where x_i is the total number of accidents at location i, $\bar{x}$ is the mean value of studied location i and its neighbours. W_{ij} is the weight value reflecting the spatial network proximity relationship between location i and neighbouring location j. The value of I_i can be positive, negative or zero. The positive value indicates that there are similar values (high-high value or low-low value) at the studied location i and its neighbours.

It is negative in the case of opposite values (high-low value or low-high value), and 0 when the value is randomness.

The main steps of hot zone identification are shown in Figure 10. To be more detailed, the concept of the contiguous neighbourhood should be defined between each recorded point. The network spatial matrix tool in ArcGIS is used to define the proximity relationship between the point (proximity weight matrix w_{ij}). According to (Flahaut et al., 2003; Loo, 2009), the easiest way to define w_{ij} is the 0-1 matrix (a value 1 when the point i and j are continuous, and 0 otherwise). But this method is not appropriate for this research due to the variable characteristics of distance between continuous accident points. Instead, INVERSE function with a parameter of 1 is used, expressed by d_{ij}^{-1} (that is $w_{ij} = d_{ij}^{-1}$). During the generation of the w_{ij}, both d_{ij}^{-2} and d_{ij}^{-1} were used to reflect the weight which drops off quicker with distance as this exponent value increases. Set the parameter value as d_{ij}^{-2} and d_{ij}^{-1} separately, the hot zones identified were almost the same in terms of the number of hot zones and the location of hot zones. The tiny difference is that the number of hot zones in parameter d_{ij}^{-1} is slightly higher than that of d_{ij}^{-2}. To identify as many hot zones as possible, d_{ij}^{-1} was chosen in this research.

To define the most suitable threshold value for the autocorrelation analysis, 20, 15, 11, 10, 9, 8, 5 neighbourhoods (i.e. number of neighbour accident locations to a studied accident point) were considered and set separately as the threshold value in the process of *Generating Network Spatial Weight Matrix*. The maximum number of streets intersecting at the same junction is five, and considering that at least two neighbouring points should be considered as continuous in each street, 10 was chosen as the reference value. And the hot zones result with threshold values of 20, 15, 11, 10, 9, 8, 5 were conducted separately to allow comparison. In this way, (*X*= 20, 15, 11, 10, 9, 8, or 5) closely continuous neighbourhoods around the studied point i.

Based on the number of accidents per point and the proximity weight matrix (w_{ij}), the LISA can be computed for the 3903 points and utilized for evaluating the hazardousness of a road. The concentration level of the hazard is determined by the value of LISA. The higher the value of LISA, the higher the concentration of traffic accident locations with a higher number of accidents. Besides, for each studied location i, the parameter Z value and P value were all calculated correspondingly. The positive Z value represents the spatial association of similar high values of i and its neighbourhoods (high-high concentration), and negative Z value represents the spatial concentration of low values (low-low concentration). For the identification of traffic hot zones, positive Z values with a positive LISA value are the only ones considered in this research, that is the high-high concentration. In this research, the significance level of 5% is used as the cut-off value to determine the traffic accident hot zones ($P<0.05$). After this, high-high LISA points were analysed separately for road junctions and road segments, because the way of accidents recorded is different at junction and road segments, as described in section 3.2.1. Finally, the *accident hot zone map* will be created by spatially joining them with the *municipality road.*

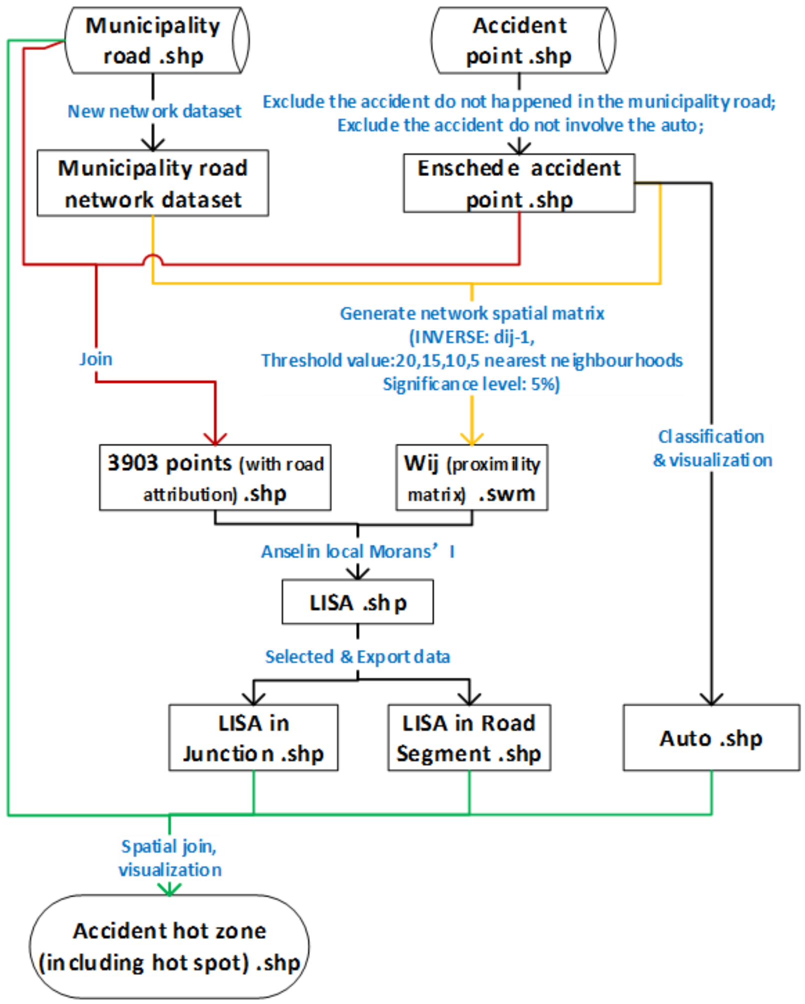


Figure 10: Flowchart of Hot Zone Analysis

3.4. Analysis Unit, Level of Causal Relationship Analysis

The urban environment which offers an opportunity for daily activities, at the same time also offers the risk of traffic accidents. The influence of the urban environment on road safety is a fact. The evaluation of the urban environment is complex due to the special characteristics of network-constrained accidents. Many previous researchers explored the influence of road safety in the urban environment, and nowadays the interest has shifted to spatial analysis due to the spatial component inherent to accident locations, compared to statistical analysis. The spatial analysis unit of the influence of environment on traffic accidents can be divided into two categories: area-based unit and network-based unit.

Most of the research concerned with the influence of environment on traffic safety focus on the area-based unit. The analysis level of area-based unit variables can be traffic analysis zones (Flahaut, 2004; Jaarsma, Louwerse, Dijkstra, De Vries, & Spaas, 2011), neighbourhoods (Martin, 2002; Nguyen, 2016;

Ossenbruggen et al., 2001; Tasic et al., 2016; Zhang, Chen, & Tu, 2013), census-based units (Cottrill et al., 2010; Siddiqui et al., 2012; Tasic et al., 2016), cities (Jones et al., 2008; Moeinaddini et al., 2014; Mohan et al., 2017; Van Petegem et al., 2014) and countries (Z. Li et al., 2013; Mohan et al., 2017; Yannis, Papadimitriou, & Antoniou, 2008).

The analysis level of the area-based analysis unit depends on the data availability and research objectives, with indicators chosen correspondingly. There is a trade-off when an aggregation level is selected. Larger analysis unit has more stable analytical results but lower accuracy, and some localised factors can only be inspected on a small level (Chen, 2015). For example, the influence of urban structure can only be detected at the city level, while land use types are usually conducted at the neighbourhood level. For the city level, the characteristics of different urban development/design are illustrated by some "macro-level indicators" like the road network density, road network diversity, population density, land use diversity (see e.g. (Jones et al., 2008; Mohan et al., 2017; Van Petegem et al., 2014)).

Several studies aggregate accident data in smaller areas(traffic analysis zone level or neighbourhood level). (Chen, 2015) for example found that traffic analysis zones with higher land use diversity, more road signals and road parking signs tend to have higher bicycle-related accidents by using the method of hierarchal Bayesian estimation. Census tracts with higher road network density and junctions were likely to cause more non-motorist traffic accidents as (Zhang et al., 2015) suggested when using the method of geographically weighted regression modelling. For the area-based neighbourhood analysis, independent variables like road network total length, road network density/diversity, number of junctions, population density, traffic volume, travel demand/generation, vehicle density, land use density, land use diversity, employment and income are usually selected (Martin, 2002; Ossenbruggen et al., 2001; Tasic et al., 2016; J. Zhang et al., 2013).

Fewer studies have been conducted on a more macro-level (city level or country level). The accident-related model at the city-based macrolevel informs urban planners in the process of road safety planning or management. The general objective of this research is to identify the influence of macro road network characteristics (like road network density, street connectivity or patterns) on traffic accident outcomes. For example, (Mohan et al., 2017) found that a city with a higher percentage of wider roads or bigger city block size is likely to have a higher risk of fatal accidents, after statistical analysis of the influence of road types and road density in 16 U.S. cities. However, even though these analyses considered urban design factors, these influence factors are not strong enough to reflect the influence of microscopic urban road features on road safety (Moeinaddini et al., 2014). Most importantly, this processing approach has the drawback of ignoring the heterogeneity of the road environment under the same analysis unit which covers broad areas (Li et al., 2007).

This problem can be solved by using smaller analysis units and different levels of spatial aggregation (Páez & M.Scott, 2004). For area-based units, accidents are aggregated per spatial polygon, so it is essential to take the influence of spatial correlation into consideration, which has been argued and added in some previous studies. (Tasic et al., 2016) found strong spatial relationships between traffic accidents and multimodal transportation availability in census tract units in the City of Chicago, United States. A multimodal that combines Local Moran's I and negative binomial models was utilised in their research to explore the relationship between urban environment and traffic safety. (Siddiqui et al., 2012) attempted Bayesian spatial aggregation method to model the influence of roadway characteristics, demographic and socio-economic factors on nonmotor accidents at the census tract level. In this case of considering the spatial correlation, the "unobserved heterogeneity" can be explained by utilising influence factors in each analysis unit (Tasic et al., 2016). However, there is still a technical problem. The spatial correlation among accidents that happened on the road network cannot be simply represented by the method of planar aggregation (Euclidean distance), which is not consistent with the reality of the unique nature of network-constrained accidents compare to the road network distance.

In conclusion, the area-based method makes it easier to quantify the road environment (for example, the number of public transit stops in a neighbourhood). However, there are two drawbacks to this area-based analysis unit. Firstly, accidents which happened at cross-boundary among two neighbourhoods/census-based units would have to be assigned to one of them, which may not be a reasonable assumption. For example, neighbourhoods are usually distinguished by arterial roads, where a higher rate of accidents may occur. This may lead to unreliable conclusions due to the error of counting the number of accidents for each area and the quantitative analysis result being unstable under different assisted method (Chen, 2015). Second, the way of quantitative indicators in a Euclidean distance mismatch with the reality of the unique nature of network-constrained accidents (Dai & Jaworski, 2016).

To lessen these challenges, many pieces of research aggregated the accidents into a network-based unit, which is demonstrated to be more realistic and superior (Dai et al., 2016; Martin, 2002; Shankar, et al., 1995; Zhang et al., 2013). Hot zone analysis, for example, is a network-based quantitative analysis that tries to identify the road segments with highly concentrated traffic accidents (Flahaut, 2004; Loo et al., 2013; Yamada & Thill, 2010). Many network-based methods have been proposed to deal with accident-related spatial analysis; for instance, the network-based kernel density estimation (Anderson, 2006; Hashimoto et al., 2016; Sabel, Kingham, Nicholson, & Bartie, 2005), G* statistics (Loo & Anderson, 2015; Yamada et al., 2010) and Local Moran's I (Flahaut, 2004; Loo, 2009; Moons et al., 2009; Yamada et al., 2010).

Following the objective of this research, to find high-risk accidents hot zones and explore the influence of road environment on the spatial distribution of road accidents, a buffer was therefore defined as the analysis unit. As the buffer represents the surrounding environment, the radius of the buffer is quite important in the whole analysis. To the best of our knowledge, there is no literature talking about this problem under the environment of an urban road network, which is much denser than urban arteries or highways. Similar to the hot zone analysis, an exploratory analysis will be conducted first under different buffer radius to decide the most suitable. Considering that there are two types of accident locations (road junction and road segment), the influence of the environment/buffer will be explored separately for junctions and segments. Thus, these two types of units(buffer around the junction and road segment) should be examined separately. The buffer of accident location was selected here for three reasons. First, accidents taking place on the road network have the unique network-constrained characteristics in network-constrained spatial correlation. Second, road segment unit decreases the influence of the mismatch between the accident original impact site and the location where the vehicle finally stopped or where the injuries were identified (Dai et al., 2016). Third, the neighbourhood level matches with the statistical information of land use, socio-economic information and the travel demand profiles.

3.5. Operationalization of Influence Factors

After traffic accident spatial analysis, the influence factors of traffic accident distribution should be identified and analysed. The selection of influential factors in this research is based on literature review, fieldwork observation and data availability. Final deterministic indicators for both junctions and road segments were presented in the former part of this section; the corresponding quantitative method for each indicator was explained in the latter part of this section.

Some variables like road quality (road surface material) were excluded after fieldwork and data consultation, because the majority of roads in Enschede are in good quality and composed of asphalt. Other indicators (e.g. shape of the junction, traffic volume, income) were not considered due to lack of data. The selected influence factors are organized into three categories, for both junction and road segment: (1) road infrastructure, (2) land use and (3) socio-economic environment, which are summarized in Table 2. Explanatory variables regarding road infrastructure, land use and socio-economic factors were selected, with corresponding indicators.

Table 2: Influence Factors

Category	Variable	Quantitative method	Units	Sources
Road infrastructure	Road network density	calculate *Road network density* for each buffer: *Road network density* =total length of road/buffer area;	m/m^2	https://www.rijkswaterstaat.nl/apps/geoservices/geodata/dmc/bron/
Land use	Building density	*Building density*, calculated for each buffer: *building density*=building area/buffer area;	1	BAG
	Land use diversity	*Land diversity*, calculated for each buffer by counting the number of land use types within a certain buffer	counts	
	Residential density	*Residential density*, calculated for each buffer: Residential area/buffer area	1	
	Leisure place	Calculated for each buffer by counting the *number of leisure place, schools, transport transit stops* etc. within a certain buffer	Counts	
	School			
	Healthcare			
	Industry			
	Office			
	Transport transit stops			
Socio-economic	Population density	When a buffer overlaps two or more neighbourhoods, a weighted average will be calculated based on the area of the buffer in each neighbourhood (neighbourhood population density = population/neighbourhood area)	counts/ m^2	CBS

The quantitative process of road environment indicators adopted in this research is different from the previous research due to the spatial analysis unit (road segment and junction buffer separately). Geospatial statistical methods in Geo-Information systems (GIS) were utilised to quantify the road environment characteristics. The value of continuous variable road network density was calculated for each buffer, which is equal to total road length within the buffer divided by the buffer area.

The building density, residential density and land use diversity were also quantified relative to the buffer area. The building density of each neighbourhood is the value of the building area divided by buffer area. The quantified process of residential density is similar to that of building density, which was quantified by the ratio value of the residential area to buffer area. Regarding land use diversity, most of the consulted studies were calculated based on the area of building (e.g. Singh et al., 2014). Since the geometry type of land use data set is the point, this method is not suitable in this research. Instead, the total number of land use types in the studied buffer was counted and expressed per buffer area.

Since land use types are one of the main factors reflecting the urban environment, it is quite important to quantify them appropriately. Many studies quantified land use types by the patch density in the area-based analysis (Henning-Hager, 1986; Kaygisiz et al., 2017; Yang et al., 2007; Yao, 2013). However, sometimes it is difficult to distinguish residential buildings from commercial buildings due to the multifunction of an individual building. Thus, considering that the analysis unit of this research is a road segment and junction, the number of special buildings (schools, bars/pubs, leisure place, parking lots) within a certain distance was also used. The transport transit stops were also calculated by this method. Similarly to many previous studies (Dai et al., 2016; Miranda-Moreno et al., 2011), the number of leisure places, bars/pubs, schools, transport transit stops and parking lots can be counted within a certain buffer of a road segment. A 1km buffer was adopted in many previous studies. However, this radius is too large to be applied to the background of this research due to the tight spaces between each street in the urban municipality road compare to the highway/urban arterial. Since the area around the intersection/road segment (especially close to the city center) is high density, with a high mixed types of land use, high building density and high road density, this research utilized a smaller buffer radius than those applied by (Clifton et al., 2009; Dai et al., 2016). The 50m buffer radius was applied to study the influence of the immediate surroundings around the intersection/road segment on accident spatial distribution (whether belonging to a hot zone or not). Buffer radius of 50, 100, 150 and 200m was adopted to examine how close surroundings affect the occurrence of hot zones. The 300m buffer radius was also examined to explore how the environment at a walking/cycling distance influence the accidents spatial concentration.

Finally, population density was expressed by joining the value of a neighbourhood where the buffer was located, similarly to the method described in (Dai et al., 2016). When a buffer overlaps two or more neighbourhoods, a weighted average was calculated based on the area of the buffer in each neighbourhood.

Each indicator was computed using ArcGIS, and the process is summarized in Figure 11. The operational process of residential density is similar to that of road density. Note that since the accidents dataset is from 2001 to 2015, records with building use ending before 2001 or for beginning in use after 2015 were excluded from the land use dataset. Finally, the attribution table/dBASE table which contains the value of each indicator were joined into one dBASE table. Then, this table was exported to SPSS for further analysis.

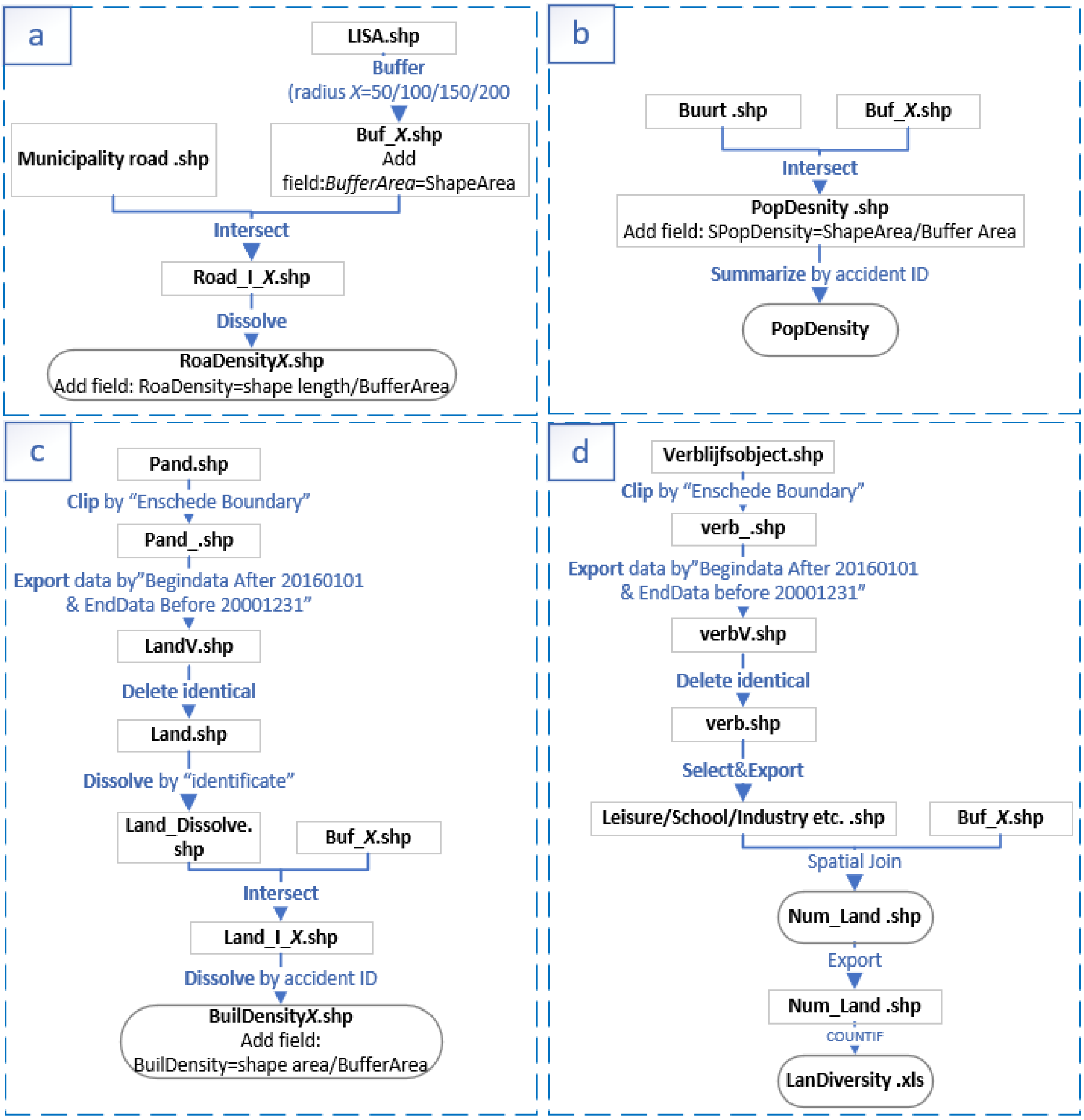


Figure 11: Flowchart of Indicators Calculation

3.6. Modeling The Influence Of Urban Environment On Traffic Accidents

Regression analysis is the most commonly used method to identify the causality between different factors. (Mohan et al., 2017) for example investigated the impact of road type and road junction density on road safety by means of descriptive analysis and regression models. Their study concluded that the increasing length of the road network and a higher percentage of highways and main roads were associated with higher accident death rates. Demographic variables are also taken into account by the method of simple linear regression. (Benedek, Ciobanu, & Man, 2016) found that alcohol use is a main cause of accidents, whereas (Moeinaddini et al., 2014) found that accident numbers increase with the number of blocks, junctions and road total length per area. Another common method successfully used in causality analysis of traffic accidents is the Poisson and negative binomial (NB) regression, which can split the accident statistic (frequency/severity) match with certain classifications. For example, traffic accidents can be classified by the mode of transport, severity of injury or the time period (see, e.g. Chen, 2015; Tasic et al., 2016).

(Hadayeghi, 2009) found that the Poisson model is suitable for the modelling of accident data for an individual location, but may lead to overdispersion under different locations, and this can be improved by the method of NB regression. However, these regression models fail to adopt spatial variables like land use types into consideration, which may amplify the influence of non-spatial influence factors. Even though the spatial characteristics were considered in some of these studies, the spatial autocorrelation of the accident location was not considered.

Realizing the importance of spatial correlation among accidents, few studies have integrated the regression model with spatial statistics instead of statistical analysis or spatial analysis alone. These regression models integrated with spatial autocorrelation explore the traffic accident data at different spatial aggregation level. (Tasic et al., 2016) for example, integrate the spatial correlation of accidents into NB models in the urban multimodal environment. (Li et al., 2013) utilized geographically weighted Poisson regression to illustrate the influence of countrywide variables on road safety, while Flahaut (2004) explored the impact of road environment on traffic safety by using logistic modelling with spatial autocorrelation. These spatial statistical methods capture the spatial heterogeneity that exists among fatal traffic accidents. These methods have been demonstrated to be more realistic.

In this research, the spatial unit of a hot zone, instead of an accident point is used as a dependent variable based on the hot zone analysis results. Thus, the NB regression-based model is not suitable here because the conduct of a regression model is based on "sufficient quantity" of accident hot zones; otherwise, the causality analysis results will not be significant. A logistic model, which proved to be efficient in causality relationship modelling in the case of limited data (Flahaut, 2004), is used here due to the limited number of hot zones identified. Besides, the dependent variable in this research is a binary variable as mentioned above. As the road segment/road junction is being used as a dependent variable in this research, whether a road subjective is safe or not can be decided by whether a road segment/road junction belongs to a hot zone or not. If a road segment/road junction belongs to a hot zone, then it will be considered unsafe and be given a value of one. Otherwise, it will be considered safe and be given a value of zero. Thus, linear regression is not suitable for this research because the value of the dependent variables varies from minus infinity to positive infinity. Due to the discrete and categorical characteristics of the dependent variable, the logistic regressions can be used to identify the spatial concentration of traffic accidents under the influence of the urban environment. The form of logistic modelling follows the general principle of a linear regression model. The basic idea of a linear regression model is the mean value of the dependent variable to the independent variable:

$$\mathrm{E}(Y/x) = a_0 + a_1 x + a_2 x + \cdots + a_n x$$

Where Y is the dependent variable, x is a value of independent variable, a_0,a_1, a_2…are the parameters of the model.

Similar to (Al-Ghamdi, 2002; Flahaut, 2004), the form of logistic regression model used in this research is:

$$P(\text{Belong to hot zone}) = E(Y/x) = \frac{e^{g(x)}}{1 + e^{g(x)}}$$

And thus

$$P(\text{Not belong to hot zone}) = 1 - P(\text{Belong to hot zone}) = 1 - E(Y/x) = \frac{1}{1 + e^{g(x)}}$$

And

$$g(x) = ln\left(\frac{E(Y/x)}{1 - E(Y/x)}\right) = a_0 + a_1 x + a_2 x + \cdots + a_n x$$

Where $g(x)$ known as *logit transformation*, which makes it possible to go back to the desirable properties of the linear regression: $g(x)$ is linear with its parameters may be continuous or category varies from minus infinity to plus infinity. And $g(x)$ also stands for the equation between the dependent variable and independent variable. More detailed theoretical explanation is found in (Al-Ghamdi, 2002; Hosmer, Stanley, & Lemeshow, 2000).

Different from other regression models, the goodness fit information of the logistic model is given by the *Hosmer and Lemeshow test* instead of *Adjusted R-square*. Hosmer and Lemeshow test results are obtained by applying a Chi-square test on a contingency table (Hosmer et al., 2000). Similar to the statistic measure R-square, the higher the value of the Hosmer and Lemeshow, the better the regression line approximates the real data points. And Hosmer and Lemeshow significance test of one also indicates that the regression line perfectly fits the data. Besides, the -2 loglikelihood is used to evaluate the overall significance of the model (Flahaut, 2004). For each developed model, the significance level of each covariate was evaluated. The Wald significance test, together with significance(sig.) was used as the evaluation criteria to involve or remove the independent variables in the equation. The higher the value of Wald tests, and the lower the value of sig., the greater significance of the independent variable contributions to the dependent variable (Hosmer et al., 1989). Besides, Z-value with a significance value of 5% was adopted.

Interpretation of any regression model needs to compare the predicted value of the model under the situation of with/without a certain variable in the equation. Exp(B), also known as Odds Ratio(OR), provides the information about how much the dependent variable is expected to change (increase/decrease) with a unit increase in the independent variable. The odds ratio was defined as the ratio of the probability that a certain independent variable will be present to that of not being present (Hosmer et al., 1989). In this research, OR greater than one increases the probability of an area belonging to the accident hot zones, for independent variables having a positive relationship with the hot zone. Vice versa, OR below one decreases the probability of an area belonging to an accident hot zone, and the independent variables have a negative relationship with the hot zone.

In this research, SPSS software which has built-in routines to calculate various parameters of the logistic model was used for the logistical regression modelling process. To avoid an overestimation of the significance of some influence actors, two separate logistic models were constructed to explore the degree of traffic accident distribution related to every indicator. Model 1 examined the influence of the urban environment including land use type-related factors and socio-economic-related factors. Model 2 was constructed including all factors (land-use, socio-economic and road infrastructure), as shown in Figure 12. These two models will be constructed for both *Junction* and *Road segment* separately under different buffer radius. And this research follows the widely used forward stepwise(conditional) method as the parameter in SPSS(see, e.g.(Flahaut, 2004; Karacasu et al., 2014)), which proves to be more priority in the field of road safety. In that way, within each model, an individual significance of each indicator is calculated based on the score test (detailed theory please see (Hosmer et al., 2000)). And all indicators are introduced to the model in the order of significance and the highest one is the first to be introduced. If the whole model becomes not significant after adding a new indicator, it will be removed from the model. Each time that a new indicator is introduced, the significance of the remaining indicators is re-calculated and re-ordered. Finally, all modelling results will be interpreted and compared on the next chapter.

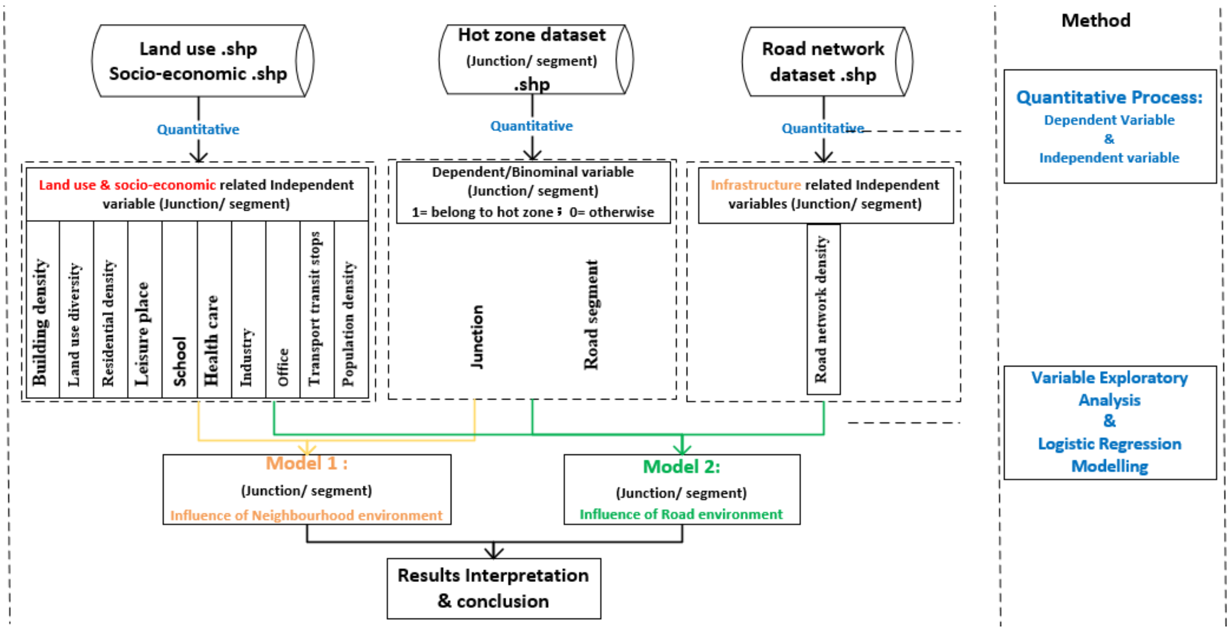


Figure 12: Main Steps of Causality Analysis

4. RESULTS AND DISCUSSION

4.1. Traffic Accident Hot Zones

From 2001 to 2015 (2009 is missing), 17,563 traffic accidents, aggregated to 3903 junctions/segment points happened in Enschede. As discussed in section 3.3, hot zones results under different threshold values (number of neighbouring points) were analysed. Table 3 shows that different threshold values have almost no influence on the uncovered number of hot zones, as values range from 217 and 191. As a result, 10 neighbouring points was the threshold value adopted in this research. For the final hot zone results, the number of accidents in individual points varies between 5 and 151, with a mean of 18.1 accidents per point, which is obviously higher than that of the entire road network (4.55 per point). 202 hot zones have been identified with 37.6% of them happening on the road junctions and 62.4% of them happening on the road segment, which is a reverse trend when compared to the analysis of traffic accident points. This is mainly because some locations with the high number of accidents (blue and black dots in Figure 13) are not identified as accident hot zones, instead the whole neighbourhood with a moderately high number of accidents (pink dots in Figure 13) are identified as hot zones. And most of the blue and black dots located at the junction and pink dots which represent the moderately high number of accidents are located on road segments.

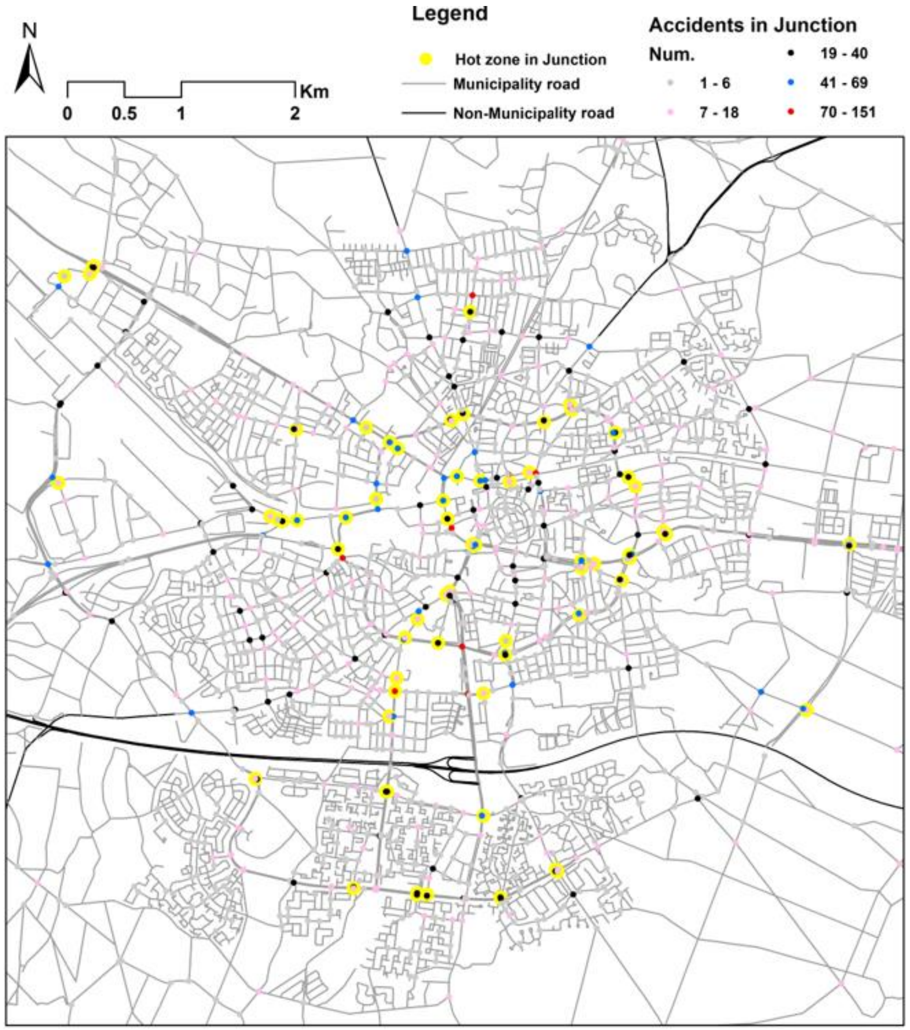
Table 3: Statistics of Accidents in Hot Zones

	Entire network	Hot zones with different threshold value						
		20	15	11	10	9	8	5
Num. of points	3903	217	209	210	202	208	199	191
Num. of Accidents	17563	3825	3796	3794	3653	3854	3595	3267
Mean num. of accidents per point	4.55	17.6	18.2	18.1	18.1	18.5	18.1	17.1
Minimum num. of accidents	1	5	5	5	5	5	5	5
Maximum num. of accidents	151	151	151	151	151	151	151	151

Compared to the number of accidents before clustering to hot zones, almost all points with a high number of accidents (blue and red dots shown in Figure 13[2]) are identified as hot zones. However, there are minority hotspots (several red dots and blue dots) not embedded in the hot zones, which implies that even though there is a high number of accidents in these points, the number of accidents in the surrounding neighbourhoods is low. Although the hotspot methodology superior in identifying the most dangerous

[2] The total number of accidents in each recorded point was classified in five categories and shown in Figure 13. The number of classes should be more to illustrate the details and less to be legible (too much classification is hard to see and print reliably) according to (Axis Maps, 2010), thus five classes were chosen here. And the natural break was used because it is optimal to minimize within-class variance and maximize between-class differences (Axis Maps, 2010).

locations, hot zone methodology supplements the hotspot methodology by identifying some neighbourhood with a moderate-high number of accidents as hazardous locations. Many points were identified as hot zones but do not have a high number of accidents (black and red dots). This is because the joint probability distribution (spatial autocorrelation) between different points is considered in the process of hot zone analysis. Before quantitatively exploring the influence of environment on hot zone distribution and making road safety improvement suggestions, the entire hot zone will be investigated first through fieldwork, which will be explained later in section 4.2.



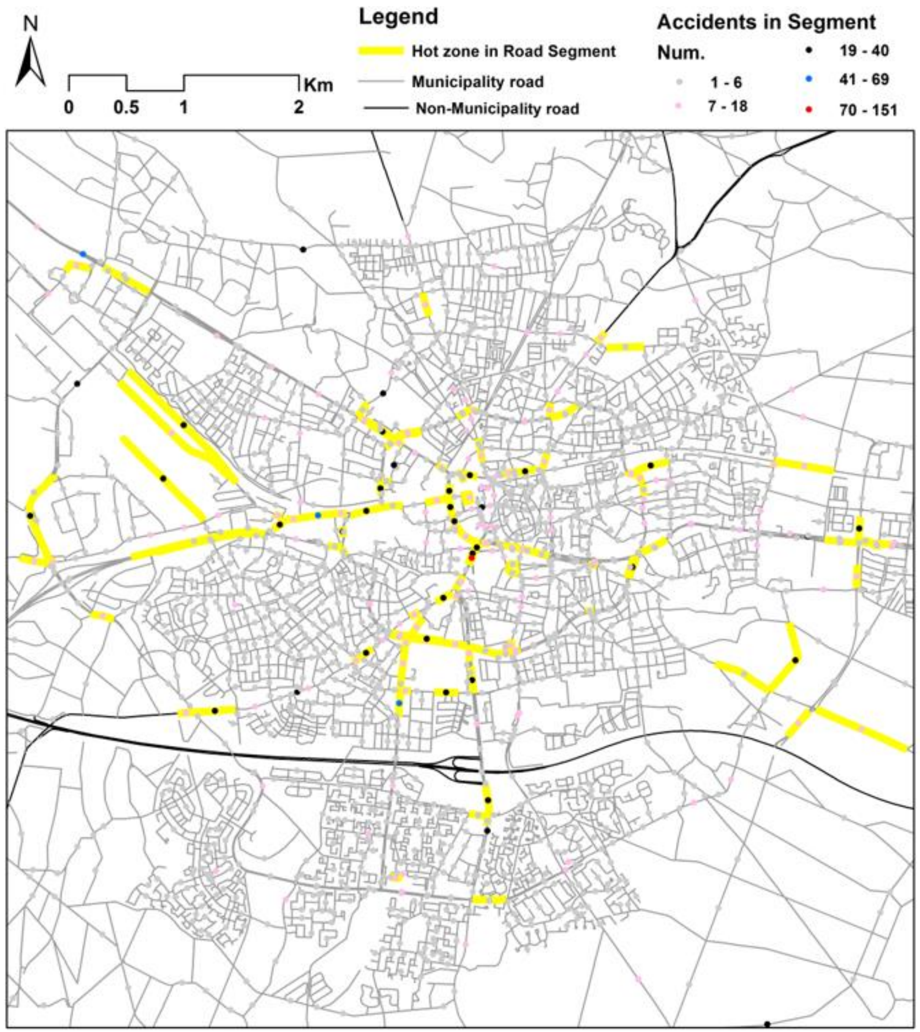


Figure 13: Accident Hot Zones in Junction (left) and Road Segment (right)

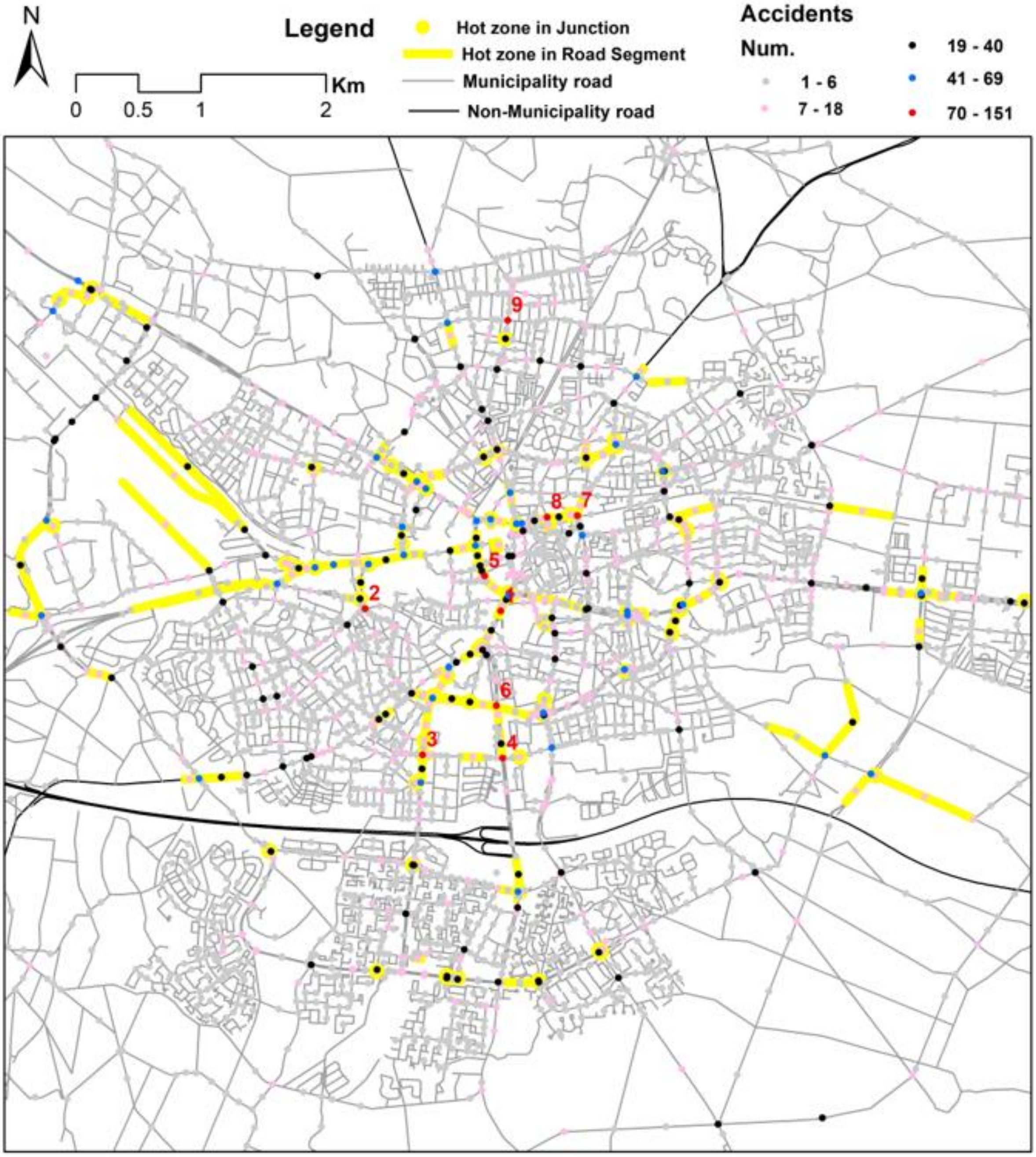


Figure 14: Accidents Hot zones of Enschede (junctions and segments combined)

Figure 14 shows the distribution of accident hot zones, and most of them are concentrated in certain districts near the city centre, with high population density and building density. After checking the road attribution information from the VIA software, most of the hot zones especially those near the city centre have a speed limit of 50km/h. This conclusion is consistent with previous research work in the Netherlands (ANWB et al., 2017; SWOV, 2017; Wegman et al., 2006). This urban arterial network of 50 km/h usually refers to the roads between dense areas and the highway in the Netherlands (Wegman et al., 2006), and usually there are separate cycle paths along the 50 km/h. Mixed use of different transport modes in the same physical space may lead to high rates of accidents and a speed limit of 50km/h may be still too high for the traffic combination of auto and venerable road users. A small number of hot zones are located in the eastern part of Enschede, where the speed limit is 30km/h. The high concentration of traffic accidents on these streets is not a coincidence. As can be seen from the map, the road network and neighbourhood density of the hot zone area near the city centre are obviously higher compared to the non-hot zone areas, which implies that the characteristics of the road environment are related to the spatial distribution of traffic accidents. Nevertheless, fieldwork was undertaken to observe which other characteristics of the road environment would explain why minority hotspots (several red dots and blue dots) not embedded in the hot zones.

4.2. Hot Zones Audit and Qualitative Analysis

The observation of road environments through fieldwork is necessary to complement the findings from the hot zone analysis and the literature review. Nine red dots were chosen for the road environment audit based on the hot zone analysis results. One of them is not embedded in the hot zone and eight of them are located at the road junction. Fieldwork-based observation allows for a broader understanding of the local environment, which was unavailable in the secondary dataset. It reveals significant characteristics of certain variables in the hot zones. For instance, fieldwork analysis of red dot number 9 reveals the lack of a traffic sign or junction light (Figure 15), while the other 8 red dots have a median strip near the road junction, as illustrated in Figure 16 for red dot number 3.

Figure 15: Red Dot 9
Source: Author

Figure 16: Red Dot 3
Source: Author

Each red dot was assessed in relation to the indicators selected based on literature review, and are summarised in Table 4. Results show that only one street near a red dot has a speed limit of 30km/h, and others are located at 50km/h zones. High spatial concentration of accidents near road junctions is common since all red dots are at junction. For the road segment intersected in the red spot, seven of them have separate bicycle lanes. The road material observed for all fieldwork locations is asphalt, which implies that the road material does not have a clear relationship with traffic accidents in the context of Enschede. As for traffic calming, no related engineering measures (e.g. speed humps, chicanes, curb extensions, living street) were found in all red dots. About half of the junction does not have a junction light (such as Figure 16) even though the perceived traffic volume is high for some of them.

Table 4: Road Environment Audit of all 9 Red Dots

Road environment/sub-environment		Num. of hotspot
The forms of junction	T	4
	+	5
Speed limitation	30km/h	1
	50km/h	8
Non-motor lane	1=yes	7
	0 =no	2
Road surface quality (pavement material)	Stone	0
	Concrete	0
	Asphalt	9
Road curvature	straight	6
	warped	0
	curved	3
Traffic calm	1=yes	0
	0 =no	9
Median strip	1=yes	8
	0 =no	1
Junction light	1=yes	4
	0 =no	5
Land use type	Residential	2
	Commercial	6
	Green area	1

4.3. Modeling Results of Causality Analysis

Two models were analysed to explore the degree of traffic accident distribution related to the different category of indicators. Model 1 seeks to understand the relationship between road environment (land-use and socio-economic) and traffic accidents spatial concentration, while Model 2 examined the influence of all factors (land-use, socio-economic and infrastructure) on accidents spatial concentration. Different buffer radius (50m, 100m, 150m, 200m and 300m) were analysed with the purpose of defining the most suitable analysis unit to represents the surrounding environment.

Table 5 discusses the goodness of fit and overall significance statistics of two models for both junction and road segment, while Table 6 and Table 7 presents the estimated coefficients of the most significant models on junction and road segment.

Table 5: Model Fit Statistics under Different Buffer Radius

Criteria		Hosmer-Lemeshow Test Sig. (Goodness of fit)		-2Log likelihood	
Location	Buffer radius	Model 1	Model 2	Model 1	Model 2
Junction	50m	0.807	0.000	583.674[a]	553.074[a]
	100m	**0.658**	**0.560**	**569.175[a]**	**543.028[a]**
	150m	0.650	0.550	573.932[a]	554.607[a]
	200m	0.126	0.177	571.640[a]	567.713[a]
	300m	0.064	0.064	570.660[a]	570.660[a]
Segment	50m	0.663	0.055	926.911[a]	922.511[a]
	100m	0.122	0.122	914.538[a]	914.538[a]
	150m	**0.814**	**0.814**	**916.811[a]**	**916.811[a]**
	200m	0.851	0.041	892.590[a]	890.648[a]
	300m	0.366	0.243	860.298[a]	855.497[a]

As can be seen from Table 5, all model was evaluated to be overall significant because all the -2 loglikelihood value were found to be larger than 3.841, which represents an χ^2 distribution with one degree of freedom at the significance level of 0.05. As for the goodness fit of the model, when the buffer radius increased from 50m to 300m, the goodness fit of the model 1 for junction decreased gradually from 0.807 to 0.064, suggesting that at junctions, the spatial concentration of accidents is mainly influenced by the immediate surroundings. As for the road segments, no obvious pattern (increase or decrease in model fit) was found for increasing buffer sizes, which is logical as segments have varying extensions.

Model 2 adds a road infrastructure-related indicator into consideration. Results of the goodness of fit statistics show overall smaller numbers than for Model 1, which illustrates that adding road infrastructure-related indicator does not contribute to overall model fit. Notice that the goodness fit value of model 2 is the same as model 1, this is because that the variables are entered into the model one by one according to the individual significance value and the road infrastructure indicator not been included in the final logistic regression equation. The dependent variable has no obvious relationship with the explanatory variables under the buffer of 50m, which was the buffer with the best result for Model 1 (junctions). Combined with the results of model 1, this illustrates that the road infrastructure related indicator(road network density） within the immediate proximity of junction/road segment do not influence the occurrence of accident hot zones. A goodness fit value of 0.064 for both model 1 and model 2 under the buffer 300m on junction implies that 300m distance from the junction has limited influence on the accidents spatial concentration. In summary, the model fit is not improved after introduction of the infrastructure related indicator (road network density), which illustrates that the concentration of accidents is mainly caused by the environment related indicators.

The model with the highest goodness fit was selected for the final discussion. The step-wise logistic regression model with a buffer of 100m for junction and 150m for road segment are presented in Table 6 and Table 7. Moreover, note that the model fit is only slightly improved when incorporating road infrastructure related indicators, but several land use-related indicators become no longer significant.

Table 6 shows the coefficients estimates and significance associated with each variable for the model of the junction. Model 1, which do not include infrastructure related indicators, shows that there are six variables (see Table 6a) with statistically significant effects on the location of accident hot zones. Results suggest that junctions in areas with higher population density, in fact, experienced more car accidents. This is mainly because high population density areas tend to have higher traffic volumes. The second highest coefficient is for land use diversity, clearly illustrates that encouraging land use diversity, which is one of the strategies of promoting transit-oriented environments (e.g. Singh et al., 2014) is expected to significantly increase the possibility of car accidents at junctions. This study utilised the total number of land use types as a substitute for land diversity utilising the quantitative method according to the area of build footprint introduced in other literature (e.g. (Chen, 2015; Singh et al., 2014)). Even though the variables are a little different, the conclusion is somewhat consistent with (Chen, 2015). The influence of office and residential was found to have a slightly positive correlation with traffic accidents; this may be caused by heavier work-travel demands. Whereas the building density and school have a negative influence on the concentration of traffic accidents. This is mainly because road users tend to be more careful (e.g. lower driving speeds) in high building density area and school, or there are clearly traffic control measures like traffic signs near the school. However, the remaining land uses(leisure, healthcare, industry and transport transit stops) have no direct impact on the occurrence of hot zones of traffic accidents at road junctions.

Moreover, note that the model fit is only slightly improved (from 0.658 to 0.560) after incorporating an infrastructure related indicator (road density) in model 2. This implies that the influence factors of accident hot zones can be mainly explained by the environment related indicators. However, road density has the greatest influence on the occurrence of an accident hot zone. Thus, road density has the greatest impact on the hot zones compared to the environment related variables.

Table 6: Logistic Regression Results for Junction (100m)

(a). Model 1 ***Statistically significant at 5%

Variables	B	S.E.	Wald	Sig.	Exp(B)/OR	Rank
Population Density	127.570	64.902	3.863	0.049	2.52E+55	1
Land Diversity	0.518	0.098	28.274	0.000	1.679	2
Office	0.053	0.021	6.438	0.011	1.055	3
Residential Density	0.040	0.021	3.695	0.055	1.041	4
School	-1.031	0.434	5.644	0.018	0.357	5
Building Density	-8.646	1.715	25.415	0.000	0.000	6
Constant	-3.702	0.323	131.145	0.000	0.025	

Goodness-of-fit: Hosmer-Lemeshow sig.=0.658

(b). Model 2 ***Statistically significant at 5%

Variables	B	S.E.	Wald	Sig.	Exp(B)/OR	Rank
Road Density	105.866	17.490	36.637	0.000	9.48E+45	1
Land Diversity	0.478	0.095	25.060	0.000	1.612	2
School	-0.872	0.408	4.559	0.033	0.418	3
Building Density	-6.535	1.457	20.108	0.000	0.001	4
Constant	-5.532	0.510	117.455	0.000		

Goodness-of-fit: Hosmer-Lemeshow sig.=0.560

The results of the regression analysis for the accidents located around 150m from the centroid of road segments shows that land diversity, school and population density are significant at the 95% confidence level (see Table 7a). However, land diversity has a positive effect on the occurrence of an accident hot zone, whereas school and population density have a negative effect. This suggests that the road segment within mix land use areas tends to be at higher risk of becoming an accident hot zone, whereas the risk of a road segment being a hot zone decreases when near schools (due to decreased travel speeds). As for population density results are contrary to those obtained for the analysis of road junctions. The negative coefficient suggests that for lower population densities, there is an increasing potential of an area becoming an accident hot zone. This may be due to high speed or careless driving on these roads.

When analysing model 2 (adding an infrastructure-related indicator) for the segments, model coefficients do not change, which implies that road density has no impact on the occurrence of a hot zone in the case of road segments.

Table 7: Logistic Regression Results for Road Segment (150m)

(a). Model 1 ***Statistically significant at 5%

Variables	B	S.E.	Wald	Sig.	Exp(B)/OR	Rank
Land Diversity	0.398	0.063	40.618	0.000	1.489	1
School	-0.596	0.193	9.531	0.002	0.551	2
Population Density	-162.114	45.135	12.901	0.000	0.000	3
Constant	-3.824	0.281	185.264	0.000		

Goodness-of-fit: Hosmer-Lemeshow sig.=0.814

(b). Model 2 ***Statistically significant at 5%

Variables	B	S.E.	Wald	Sig.	Exp(B)/OR	Rank
Land Diversity	0.398	0.063	40.618	0.000	1.489	1
School	-0.596	0.193	9.531	0.002	0.551	2
Population Density	-162.114	45.135	12.901	0.000	0.000	3
Constant	-3.824	0.281	185.264	0.000		

Goodness-of-fit: Hosmer-Lemeshow sig.=0.814

The common point about the cause of hot zones on junction and road segment is that the location with school tends to have a lower risk of being accident hot zones, this may because sufficient road safety strategies have been conducted at school neighbourhoods. Besides, the positive correlation between land diversity and hot zones in both junction and road segment suggests that road safety should also be considered in pursuing the idea of mixed land use in the urban planning process. This is because land use mixture may result in conflicts among concentrated human activities in destinations with various land use purposes. The road density is the first reason that causes hot zones in junctions, whereas it has no direct influence on the hot zones on road segment. This implies that road safety attention for areas with high road density should be directed towards road junctions. Besides, the opposite effect of population density on the spatial co-occurrence of accidents on junction and road segments suggests that road safety improve attention on high population areas should also be directed towards road junctions.

5. CONCLUSIONS AND RECOMMENDATIONS

5.1. Conclusions

This study analysed the influence of urban environment on traffic accidents concentration using an accident dataset from 2001 to 2015(2009 is missing). Results show that almost all recorded accidents in Enschede involve car (approximately 98%), with 0.35% of them involved fatalities, 4.41% required hospitalisation and 85.78% with no victims. Approximately 96% of total accidents happen between auto and moped in Enschede, which implies that extra attention needs to be given to the high-risk road user categories especially the mopedists. The BPM modes (bicycle, pedestrian and moped), being the most vulnerable road users, accounts the lowest percentage of total accidents (26%) but the highest percentage of victims. Looking at the demographic characteristics (gender and age) of victims, accident data show that more males were involved in accidents and younger people (0-24) are more likely to be involved in an accident.

A spatial analysis with the purpose of identifying the area with high concentration of accidents was conducted. An initiative exploratory analysis concerning the sensitive of hot zone results under different threshold value was conducted in this research. Results show that different threshold values have almost no influence on the results of hot zones. Although the hotspot methodology superior in identifying the most dangerous accident locations, hot zone analysis results in this research illustrate that it supplements hotspots analysis by identifying some neighbourhood with moderately high number of accidents as hazardous locations. Besides, the spatial autocorrelation between adjacent accident locations is considered in the hot zone methodology , which is more consistent with the reality. Therefore, the author suggests using the hot zone analysis to supplement the hotspot analysis to determine the dangerous accident locations.

The hot zone results illustrate that most of the hot zones were concentrated in certain districts near the city centre. Most of them were distributed among the 50km/h region of the road network, which is consistent with previous research in the Netherlands (ANWB et al., 2017; SWOV, 2017; Wegman et al., 2006). These roads are mainly the urban arterials connecting the city centre (which was constructed in an earlier stage, approximately before 1930) and the southwest part of the city (which was constructed from 1950 to 1960). Thus, road safety improvements should be strengthened on these roads, especially the 50km/h roads. Only a limited number of hot zones were uncovered in the eastern part of Enschede, where the speed limit is 30km/h.

Observational analysis of nine hot dots (point locations with the highest number of accidents) implies that almost all hotspots in Enschede are located at a road junction. During fieldwork, it was noticed that 5 out of 9 of these junctions do not have clearly marked traffic signs, even though the perceived traffic flow is high. Thus, new traffic control measures like traffic signs need to be applied in these places. Besides, the priority on these intersections should also be assigned(for example by traffic signs). Several of them not been identified as hot zones likely because of the special surroundings like a tall building that makes the sight worse. On the other hand, even though they are traffic accident hot dots, their neighbouring accident point locations were not high enough to be classified as a hot zone. This illustrates that the hot zone methodology results complement the hotspot results again, which is consistent with past research (Flahaut et al., 2003; Loo, 2009; Moons et al., 2009; Yao, 2013).

Influential factors related to infrastructure, land use and population were selected in this research based on literature review, fieldwork observation and data availability. Different buffer radius from the centroid of accident locations (both at junctions and segments) were explored. Results show that the environmental characteristics in close proximity to the junction (100m buffer) have an influence on the accident spatial concentration, whereas the road safety in road segments are mainly influenced by a buffer area within 150m.

Logistic regression results show that the high road density within close proximity(100m) around a junction is the first reason that causes hot zones whereas has no direct influence on the hot zones on road segments. Similarity, accidents in high population areas tend to concentrate on the junction. These imply that in high-density areas(population density or road density), attention should be focused more on the road junction than the road segment. Compared to the environment related variables, the infrastructure related indicator(road network density) tend to have a greater impact on the concentration of accidents on junctions, while has no obvious effect on the concentration of accidents in road segments. This may be due to a higher possibility of vehicle encounters on the junction in higher road network density area. Residential areas have a slightly positive correlation to accident hot zones in the junction, which may be caused careless attention at residential junctions. These locations tend to have high road density. Thus, we should try to reduce the number of intersections within the residential area, which can be achieved by planning large residential areas without distributors, because the residential street is safer than a distributor according to (Hummel, 2001)). However, the upper limit of the size of residential areas should also be controlled to meet some guides/strategies of the urban planning. For example, the accessibility to urban facilities within certain distance travelled by walking or cycling should be planned.

Additionally, results suggest that the main cause of concentration of accidents in the road segment is when the location has high land use diversity. A combination of different land use types in close proximity promotes the use of multi-modes especially cycling, due to the short distance. However, the positive correlation between the land diversity and auto accidents in Enschede reminds us of the need to pay attention to road safety when increasing land diversity. The spatial allocation of office locations also needs to be deliberate in the early planning stage of urban planning. One thing that needs to be kept in mind is that the application of certain land use planning ideas does not, on its own, influence road safety. For example, the allocation of offices should also consider the location of residential areas, because work-travel demands caused by the office location may also increase the possibility of accident concentration. However, no significant correlation between the other types of land uses (e.g. leisure, transport transit stops and healthcare) and the accident spatial concentration was found. Similarly, building density, which reflects the urban design (e.g. compact or disperse urban structure) also does not correlate with the accident hot zone distribution in Enschede.

5.2. Limitations and recommendations

This study concentrates on the identification of traffic accident hot zones and treats everyone equally without priority, which is not practical to give treatments to every hot zone due to the limited resources(e.g. budget or labour). In most cases, policy-makers maybe more interesting in the most dangerous hot zone instead of all hot zones. In this research, 202 hot zones have been identified as accident hot zones, together with hotspots, all these accident locations are dangerous to road users. Thus, the question follows is: which one needs to be treated with top priority? Hence, in order to answer a question like this, it is important to rank the hot zones based on some criterion (e.g. the length, the severity) in the future.

Introducing more indicators related to road infrastructure (e.g. traffic volume, the number of lanes, separate bicycle lane) would allow a more thorough investigation on whether these factors have an impact on accident hot zones. These were analysed during fieldwork, but only for nine locations with the highest number of accidents (9 hot dots). Although this evaluation may be subjective, having these variables quantified for the whole city would ensure more objective and reliable conclusions.

Since this research is mainly concentrated on the spatial co-occurrence of traffic accidents instead of those accidents, this leads to a spatially oriented modelling. In that case, indicators selected in this research are mainly focused on the spatial variables, which reflects the characteristics of basic analysis units rather than that of accidents. Hence many non-spatial factors were not examined in this research. Other environment-related indicators (e.g. traffic volume, street type, separate bicycle lane) were not included due to the substantial unavailability of data. Future research could involve non-spatial factors (e.g. road users). Such research direction is applicable, and maybe by the method of multi-level modelling. For example, model 1 may be used to examine the influence of spatial variables; model 2 was used to evaluate the influence of road uses, model 3 concerns with vehicles and model 4 including all indicators with a certain combine method.

Determining whether the studied influence factors in this research hold in other geographic backgrounds in the Netherlands is important; this is a challenging task for future research. The methodology developed in the present research could be applied to other cities and then compare the different results.

LIST OF REFERENCES

Al-Ghamdi, A. S. (2002). Using logistic regression to estimate the influence of accident factors on accident severity. *Accident Analysis and Prevention*, *34*(6), 729–741. https://doi.org/10.1016/S0001-4575(01)00073-2

Amorim, M., Ferreira, S., & Couto, A. (2017). Road safety and the urban emergency medical service (uEMS): Strategy station location. *Journal of Transport* & *Health*, *6*, 60–72. https://doi.org/10.1016/j.jth.2017.04.005

Anderson, T. K. (2006). Comparison of spatial methods for measuring road accident "hotspots": a case study of London. *Journal of Maps*, *3*(1), 55–63. https://doi.org/10.1080/jom.2007.9710827

Anselin, L. (1995). Local indicators of spatial association—LISA. *Geographical Analysis*, *27*(2), 93–115. https://doi.org/10.1111/j.1538-4632.1995.tb00338.x

ANWB, SWOV, Safe Traffic Netherlands, & Dutch Association of Insurer. (2017). *Verkeersveiligheid: een nationale prioriteit* [Road safety: a national priority]. *ANWB*. The Hague. Retrieved from https://www.swov.nl/publicatie/verkeersveiligheid-een-nationale-prioriteit

Benedek, J., Ciobanu, S. M., & Man, T. C. (2016). Hotspots and social background of urban traffic crashes: A case study in Cluj-Napoca (Romania). *Accident Analysis and Prevention*, *87*, 117–126. https://doi.org/10.1016/j.aap.2015.11.026

Bicycle Dutch. (2014). Enschede, nominee for best cycling city. Retrieved July 29, 2017, from https://bicycledutch.wordpress.com/2014/02/20/enschede-nominee-for-best-cycling-city/

Black, W. R., & Thomas, I. (1998). Accidents on Belgium's motorways: A network autocorrelation analysis. *Journal of Transport Geography*, *6*(1), 23–31. https://doi.org/10.1016/S0966-6923(97)00037-9

CBS. (2017). Traffic deaths include more young and old drivers. Retrieved July 27, 2017, from https://www.cbs.nl/en-gb/news/2017/18/traffic-deaths-include-more-young-and-old-drivers

Chen, P. (2015). Built environment factors in explaining the automobile-involved bicycle crash frequencies: A spatial statistic approach. *Safety Science*, *79*, 336–343. https://doi.org/10.1016/j.ssci.2015.06.016

Clifton, K. J., Burnier, C. V., & Akar, G. (2009). Severity of injury resulting from pedestrian-vehicle crashes: What can we learn from examining the built environment? *Transportation Research Part D: Transport and Environment*, *14*(6), 425–436. https://doi.org/10.1016/j.trd.2009.01.001

Cottrill, C. D., & Thakuriah, P. (2010). Evaluating pedestrian crashes in areas with high low-income or minority populations. *Accident Analysis and Prevention*, *42*(6), 1718–1728. https://doi.org/10.1016/j.aap.2010.04.012

Dai, D., & Jaworski, D. (2016). Influence of built environment on pedestrian crashes: A network-based GIS analysis. *Applied Geography*, *73*, 53–61. https://doi.org/10.1016/j.apgeog.2016.06.005

Dijkstra, A., Drolenga, H., & Maarseveen, M. Van. (2007). Method for assessing safety of routes in a road network. *Journal of the Transportation Research Board*, (2019), 82–90.

Drummer, O. H., Gerostamoulos, J., Batziris, H., Chu, M., Caplehorn, J., Robertson, M. D., & Swann, P. (2004). The involvement of drugs in drivers of motor vehicles killed in Australian road traffic crashes. *Accident Analysis and Prevention*, *36*(2), 239–248. https://doi.org/10.1016/S0001-4575(02)00153-7

Eckley, D. C., & Curtin, K. M. (2013). Evaluating the spatiotemporal clustering of traffic incidents. *Computers, Environment and Urban Systems*, *37*(1), 70–81. https://doi.org/10.1016/j.compenvurbsys.2012.06.004

Elvik, R. (2006). New approach to accident analysis for hazardous road locations. *Transportation Research Record: Journal of the Transportation Research Board*, *1953*, 50–55. https://doi.org/10.3141/1953-06

Fan, J. X. (2009). Mixed land use and walkability: Variations in land use measures and relationships with BMI, overweight, and obesity. *Health Place*, *15*(4), 1130–1141. https://doi.org/10.1016/j.healthplace.2009.06.008.Mixed

Flahaut, B. (2004). Impact of infrastructure and local environment on road unsafety: Logistic modeling with spatial autocorrelation. *Accident Analysis and Prevention*, *36*(6), 1055–1066. https://doi.org/10.1016/j.aap.2003.12.003

Flahaut, B., Mouchart, M., Martin, E. S., & Thomas, I. (2003). The local spatial autocorrelation and the kernel method for identifying black zones:A comparative approach. *Accident Analysis* & *Prevention*,

35(6), 991–1004. https://doi.org/10.1016/S0001-4575(02)00107-0
Hadayeghi, A. (2009). *Use of advanced techniques to estimate zonal level safety planning models and examine their temporal transferabilty*. Doctoral dissertation. University of Toronto, Canada.
Hadayeghi, A., Shalaby, A. S., & Persaud, B. N. (2010). Development of planning level transportation safety tools using Geographically Weighted Poisson Regression. *Accident Analysis and Prevention*, *42*(2), 676–688. https://doi.org/10.1016/j.aap.2009.10.016
Hashimoto, S., Yoshiki, S., Saeki, R., Mimura, Y., Ando, R., & Nanba, S. (2016). Development and application of traffic accident density estimation models using kernel density estimation. *Journal of Traffic and Transportation Engineering (English Edition)*, *3*(3), 262–270. https://doi.org/10.1016/j.jtte.2016.01.005
Henning-Hager, U. (1986). Urban development and road safty. *Accid. Anal.* & *Prev*, *18*(2), 135–145. Retrieved from http://ac.els-cdn.com/0001457586900588/1-s2.0-0001457586900588-main.pdf?_tid=6b08ae2c-9565-11e7-b10b-00000aab0f6c&acdnat=1504965028_e903ca33d45ff61d59c177cdd1f5d695
Herold, M., Couclelis, H., & Clarke, K. C. (2005). The role of spatial metrics in the analysis and modeling of urban land use change. *Computers, Environment and Urban Systems*, *29*(4), 369–399. https://doi.org/10.1016/j.compenvurbsys.2003.12.001
Hosmer, D. W., Stanley, L., & Lemeshow, S. (2000). *Applied logistic regression* (2nd ed.). New York: Wiley-interscience. Retrieved from https://books.google.nl/books/about/Applied_Logistic_Regression.html?id=Po0RLQ7USIMC&redir_esc=y
Huang, H., Zhou, H., Wang, J., Chang, F., & Ma, M. (2017). A multivariate spatial model of crash frequency by transportation modes for urban intersections. *Analytic Methods in Accident Research*, *14*, 10–21. https://doi.org/10.1016/j.amar.2017.01.001
Hummel, T. (2001). *Land use planning in safer transportation network planning*. The Netherlands. Retrieved from https://www.swov.nl/sites/default/files/publicaties/rapport/d-2001-12.pdf
Jaarsma, R., Louwerse, R., Dijkstra, A., De Vries, J., & Spaas, J. P. (2011). Making minor rural road networks safer: The effects of 60 km/h-zones. *Accident Analysis and Prevention*, *43*(4), 1508–1515. https://doi.org/10.1016/j.aap.2011.03.001
Jones, A. P., Haynes, R., Kennedy, V., Harvey, I. M., Jewell, T., & Lea, D. (2008). Geographical variations in mortality and morbidity from road traffic accidents in England and Wales. *Health and Place*, *14*(3), 519–535. https://doi.org/10.1016/j.healthplace.2007.10.001
Jovanis, P. P., & Chang, H.-L. (1986). Modeling the relationship of accidents to miles traveled. *Transportation Research Record*, *1068*, 42–51. Retrieved from http://onlinepubs.trb.org/Onlinepubs/nchrp/cd-22/references/regression6.pdf
Karacasu, M., Ergül, B., & Yavuz, A. A. (2014). Estimating the causes of traffic accidents using logistic regression and discriminant analysis. *International Journal of Injury Control and Safety Promotion*, *21*(4), 305–312. https://doi.org/10.1080/17457300.2013.815632
Kaygisiz, Ö., Senbil, M., & Yildiz, A. (2017). Influence of urban built environment on traffic accidents: The case of Eskisehir (Turkey). *Case Studies on Transport Policy*, *5*(2), 306–313. https://doi.org/10.1016/j.cstp.2017.02.002
Koolhoven, W. (2017). Bike association meeting - Enschede bicycle city 2020. Retrieved July 27, 2017, from http://enschedefietsstad.nl/bijeenkomst-fietsersbond/
Krishnakumar, V. K., Pulugurtha, S. S., & Nambisan, S. S. (2005). Identification and ranking of high Pedestrian crash zones using GIS. In *Annual ESRI International User Conference* (pp. 1–10). https://doi.org/doi:10.1061/40794(179)70
LaScala, E. A., Gerber, D., & Gruenewald, P. J. (2000). Demographic and environmental correlates of pedestrian injury collisions: A spatial analysis. *Accident Analysis and Prevention*, *32*(5), 651–658. https://doi.org/10.1016/S0001-4575(99)00100-1
Li, L., Zhu, L., & Sui, D. Z. (2007). A GIS-based bayesian approach for analyzing spatial-temporal patterns of intra-city motor vehicle crashes. *Journal of Transport Geography*, *15*(4), 274–285. https://doi.org/10.1016/j.jtrangeo.2006.08.005
Li, Z., Wang, W., Liu, P., Bigham, J. M., & Ragland, D. R. (2013). Using geographically weighted poisson regression for county-level crash modeling in California. *Safety Science*, *58*, 89–97. https://doi.org/10.1016/j.ssci.2013.04.005

Loidl, M., Traun, C., & Wallentin, G. (2016). Spatial patterns and temporal dynamics of urban bicycle crashes-A case study from Salzburg (Austria). *Journal of Transport Geography*, *52*, 38–50. https://doi.org/10.1016/j.jtrangeo.2016.02.008
Loo, B. P. Y. (2009). The identification of hazardous road locations: a comparison of the blacksite and hot Zone methodologies in Hong Kong. *International Journal of Sustainable Transportation*, *3*(3), 187–202. https://doi.org/10.1080/15568310801915583
Loo, B. P. Y., & Anderson, T. K. (2015). *Spatial analysis methods of road traffic collisions*. (Taylor & Francis Group, Ed.). New York: CRC Press. Retrieved from http://ezproxy.utwente.nl:3782/lib/itc/reader.action?docID=4003043
Loo, B. P. Y., & Yao, S. (2013). The identification of traffic crash hot zones under the link-attribute and event-based approaches in a network-constrained environment. *Computers, Environment and Urban Systems*, *41*, 249–261. https://doi.org/10.1016/j.compenvurbsys.2013.07.001
Lourens, P. F., Vissers, J. A. M. M., & Jessurun, M. (1999). Annual mileage, driving violations, and accident involvement in relation to drivers' sex, age, and level of education. *Accident Analysis and Prevention*, *31*(5), 593–597. https://doi.org/10.1016/S0001-4575(99)00015-9
Mannering, F. L., Shankar, V., & Bhat, C. R. (2016). Unobserved heterogeneity and the statistical analysis of highway accident data. *Analytic Methods in Accident Research*, *11*, 1–16. https://doi.org/10.1016/j.amar.2016.04.001
Martin, J.-L. (2002). Relationship between crash rate and hourly traffic flow on interurban motorways. *Accident Analysis and Prevention*, *34*, 619–629. Retrieved from www.elsevier.com/locate/aap
Mathijssen, M. P. M. (2005). Drink driving policy and road safety in the Netherlands: A retrospective analysis. *Transportation Research Part E: Logistics and Transportation Review*, *41*, 395–408. https://doi.org/10.1016/j.tre.2005.03.001
Miranda-Moreno, L. F., Morency, P., & El-Geneidy, A. M. (2011). The link between built environment, pedestrian activity and pedestrian-vehicle collision occurrence at signalized intersections. *Accident Analysis and Prevention*, *43*(5), 1624–1634. https://doi.org/10.1016/j.aap.2011.02.005
Moeinaddini, M., Asadi-Shekari, Z., & Zaly Shah, M. (2014). The relationship between urban street networks and the number of transport fatalities at the city level. *Safety Science*, *62*, 114–120. https://doi.org/10.1016/j.ssci.2013.08.015
Mohan, D., Bangdiwala, S. I., & Villaveces, A. (2017). Urban street structure and traffic safety. *Journal of Safety Research*, *62*, 63–71. https://doi.org/10.1016/j.jsr.2017.06.003
Moons, E., Brijs, T., & Wets, G. (2009). Identifying hazardous road locations: Hot spots versus hot zones. *Trans. Computational Science*, *6*, 288–300. https://doi.org/10.1007/978-3-642-10649-1_17
Navin, F., Zein, S., & Felipe, E. (2000). Road safety engineering: an effective tool in the fight against whiplash injuries. *Accident Analysis & Prevention*, *32*(2), 271–275. https://doi.org/10.1016/S0001-4575(99)00077-9
Nguyen, H. H. (2016). *Multidimensional analysis of road safety case study of Amsterdam*. Master thesis. Université Paris Diderot. Retrieved from https://mastergeoprisme.files.wordpress.com/2017/01/rapportsta_nguyensept.pdf
Ogden, K. W. (1996). *Safer roads : a guide to road safety engineering*. Avebury Technical. Retrieved from https://trid.trb.org/view.aspx?id=459597
Ossenbruggen, P. J., Pendharkar, J., & Ivan, J. (2001). Roadway safety in rural and small urbanized areas. *Accident Analysis* & *Prevention*, *33*(4), 485–498. https://doi.org/10.1016/S0001-4575(00)00062-2
Páez, A., & M.Scott, D. (2004). Spatial statistics for urban analysis: A review of techniques with examples. *GeoJournal*, *61*(1), 53–67. https://doi.org/10.1007/sGEJO-004-0877-x
Rifaat, S. M., Tay, R., & De Barros, A. (2011). Effect of street pattern on the severity of crashes involving vulnerable road users. *Accident Analysis and Prevention*, *43*(1), 276–283. https://doi.org/10.1016/j.aap.2010.08.024
Sabel, C. E., Kingham, S., Nicholson, P. A., & Bartie, P. (2005). Road Traffic Accident Simulation Modelling - A Kernel Estimation Approach. In *The 17th Annual Colloquium of the Spatial Information Research Centre* (pp. 1–5). Dunedin, New Zealand: University of Otago.
Schepers, P., Stipdonk, H., Methorst, R., & Olivier, J. (2017). Bicycle fatalities: Trends in crashes with and without motor vehicles in the Netherlands. *Transportation Research Part F: Traffic Psychology and Behaviour*, *46*, 491–499. https://doi.org/10.1016/j.trf.2016.05.007
Schepers, P., Twisk, D., Fishman, E., Fyhri, A., & Jensen, A. (2017). The Dutch road to a high level of

cycling safety. *Safety Science*, *92*, 264–273. https://doi.org/10.1016/j.ssci.2015.06.005

Shankar, V., Mannering, F., & Barfield, W. (1995). Effect of roadway geometrics and environmental factors on rural freeway accident frequencies. *Accident Analysis* & *Prevention*, *27*(3), 371–389. https://doi.org/10.1016/0001-4575(94)00078-Z

Siddiqui, C., Abdel-Aty, M., & Choi, K. (2012). Macroscopic spatial analysis of pedestrian and bicycle crashes. *Accident Analysis and Prevention*, *45*, 382–391. https://doi.org/10.1016/j.aap.2011.08.003

Silcock, D., & Smyth, A. W. (1984). Methods used by British highway authorities to identify accident blackspots. *Traffic Engineering and Control*, *25*(11), 542–545. Retrieved from https://www.safetylit.org/citations/index.php?fuseaction=citations.viewdetails&citationIds[]=citjournalarticle_63468_12

Singh, Y. J., Fard, P., Zuidgeest, M., Brussel, M., & Maarseveen, M. van. (2014). Measuring transit oriented development: A spatial multi criteria assessment approach for the city region Arnhem and Nijmegen. *Journal of Transport Geography*, *35*, 130–143. https://doi.org/10.1016/j.jtrangeo.2014.01.014

SWOV. (2017). Traffic deaths in the Netherlands. Retrieved September 4, 2017, from https://www.swov.nl/feiten-cijfers/factsheet/verkeersdoden-nederland

Tasic, I., & Porter, R. J. (2016). Modeling spatial relationships between multimodal transportation infrastructure and traffic safety outcomes in urban environments. *Safety Science*, *82*, 325–337. https://doi.org/10.1016/j.ssci.2015.09.021

Twisk, D. A., De Hair, S. H. H. M., & Otte, D. (2017). Editorial for special issue – “Improving cyclist safety through scientific research.” *Accident Analysis and Prevention*, *105*(2017), 1–3. https://doi.org/10.1016/j.aap.2017.05.021

Van Petegem, J. W. H., & Wegman, F. (2014). Analyzing road design risk factors for run-off-road crashes in the Netherlands with crash prediction models. *Journal of Safety Research*, *49*, 121–127. https://doi.org/10.1016/j.jsr.2014.03.003

Vandenbulcke, G., Thomas, I., & Int Panis, L. (2014). Predicting cycling accident risk in Brussels: A spatial case-control approach. *Accident Analysis and Prevention*, *62*, 341–357. https://doi.org/10.1016/j.aap.2013.07.001

Wang, C., Quddus, M. A., & Ison, S. G. (2013). The effect of traffic and road characteristics on road safety: A review and future research direction. *Safety Science*, *57*, 264–275. https://doi.org/10.1016/j.ssci.2013.02.012

Wegman, F., Aarts, L., & Bax, C. (2006). *Advancing sustainable safety*. (F. Wegman & L. Aarts, Eds.), *Safety Science* (SWOV Insti). Netherlands: SWOV, Leidschendam. https://doi.org/10.1016/j.ssci.2007.06.013

Wegman, F., Aarts, L., & Bax, C. (2008). Advancing sustainable safety. National road safety outlook for The Netherlands for 2005-2020. *Safety Science*, *46*(2), 323–343. https://doi.org/10.1016/j.ssci.2007.06.013

Wegman, F., Dijkstra, A., Schermers, G., & van Vliet, P. (2005). Sustainable Safety in the Netherlands : the vision , the implementation and the safety effects. In *3rd International Symposium on Highway Geometric* (pp. 1–35). Chicago, Illinois: SWOV, Leidschendam.

Weijermars, W., & Wesemann, P. (2013). Road safety forecasting and ex-ante evaluation of policy in the Netherlands. *Transportation Research Part A: Policy and Practice*, *52*, 64–72. https://doi.org/10.1016/j.tra.2013.06.001

Wikipedia. (2017). Enschede. Retrieved September 10, 2017, from http://www.enschede.nl/gemeente/BenW/00001/#.UzIY5VdN0tg

World Health Organization. (2015). *Global status report on road safety: time for action*. *WHO Library*. Switzerland: World Health Organization. https://doi.org/10.1136/ip.2009.023697

Yamada, I., & Thill, J. (2010). Local indicators of network-constrained clusters in spatial patterns represented by a link attribute. *Annals of the Association of American Geographers ISSN:*, *100*(2), 269–285. https://doi.org/10.1080/00045600903550337

Yang, J., & Otte, D. (2007). *A comparison study on vehicle traffic accident and injuries of vulnerable road users in China and Germany* (No. 07–0417). Retrieved from http://www-nrd.nhtsa.dot.gov/pdf/nrd-01/esv/esv20/07-0417-O.pdf

Yannis, G., Papadimitriou, E., & Antoniou, C. (2008). Impact of enforcement on traffic accidents and fatalities: A multivariate multilevel analysis. *Safety Science*, *46*(5), 738–750. https://doi.org/10.1016/j.ssci.2007.01.014

Yao, S. (2013). *Advances in spatial analysis of traffic crashes : the identification of hazardous road locations*. Doctoral dissertation, The University of HongKong, Pokfulam, Hong Kong SAR.
Zhang, J., Chen, X., & Tu, Y. (2013). Environmental and traffic effects on incident frequency occurred on urban expressways. *Procedia - Social and Behavioral Sciences*, *96*, 1366–1377. https://doi.org/10.1016/j.sbspro.2013.08.155
Zhang, Y., Bigham, J., Ragland, D., & Chen, X. (2015). Investigating the associations between road network structure and non-motorist accidents. *Journal of Transport Geography*, *42*, 34–47. https://doi.org/10.1016/j.jtrangeo.2014.10.010