6-2014

Development of Safety Performance Functions for Non-Motorized Traffic Safety

Hamidreza Ahady Dolatsara
Western Michigan University

DEVELOPMENT OF SAFETY PERFORMANCE FUNCTIONS FOR NON-MOTORIZED TRAFFIC SAFETY

by

Hamidreza Ahady Dolatsara

This thesis submitted to the Graduate College
in partial fulfillment of the requirements
for the degree of Master of Science in Engineering
Civil and Construction Engineering
Western Michigan University
June 2014

Thesis Committee:

Jun-Seok Oh, Ph.D., Chair
Valerian Kwigizile, Ph.D.
Joseph W. McKean, Ph.D.

DEVELOPMENT OF SAFETY PERFORMANCE FUNCTIONS FOR NON-MOTORIZED TRAFFIC SAFETY

Hamidreza Ahady Dolatsara M.S.E.

Western Michigan University, 2014

This study investigates the factors which affect the safety of non-motorized transportation within the influence area of intersections to enhance development of safety performance functions (SPFs). The scope of this study is limited to the four Michigan cities of Ann Arbor, East Lansing, Flint and Grand Rapids. Due to the current lack of research regarding the appropriate size of the influence area, this study investigates the distance of crashes relative to the center of 148 intersections to identify the most probable area of influence for different crash types. For motorized and non-motorized crashes, 240 ft. and 137 ft. are proposed, respectively. The proposed area of influence is adopted for developing the SPFs. Crash data (from 2008 to 2012) and geometric and exposure characteristics of the intersections are investigated to develop the SPFs. Results of the pedestrian SPF reveal that increased exposure, more left-turn lanes, presence of on-street parking, and bus stops at the intersections increase pedestrian crash frequency, while presence of speed signs decrease the number of pedestrian crashes. Results of the bike SPF demonstrate that exposure, presence of bicycle lanes, presence of bus stops and the number of left-turn lanes at intersections are positively associated with bicycle crashes. A structural equation model (SEM) is developed to decipher complex interrelationships among the variables affecting bike crash frequency. Results show that although the presence of bicycle lanes is significant in increasing bicycle crash frequency, bike lanes are correlated with bicycle volume, thus bicycle lanes do not endanger bicycle safety.

ACKNOWLEDGEMENTS

I would like to express my deepest appreciation to my academic advisor, Professor Jun Seok Oh for the academic journey over the past two years, concluding with this thesis. Without his tremendous encouragement, support, and comments, this dissertation would not have been completed.

I would like to thank Professor Valerian Kwigizile for his support and assistance in exploring my questions. He greatly developed my knowledge of statistics, information that I used widely in this study. I would also like to thank Professor Joseph McKean whose comments and guidance helped to work through the most challenging parts of this study.

I would like to appreciate my friends and research colleagues, Mr. Bryce T. Wegner and Mr. Matthew Levi Clark for their assistance in data collection and proof reading. I need to thank Ms. Jenica Moore who did a fantastic job in editing and proof reading of the material.

Finally, I also need to thank my dear friend Mr. Farhad Abasahl for his role in my academic journey for teaching me the tools and software packages that have been widely used in this study.

Hamidreza Ahady Dolatsara

TABLE OF CONTENTS

LIST OF TABLES

List of Figures

List of Figures-Continued

1. INTRODUCTION

In order to investigate the significant parameters affecting crash frequency, practitioners have conducted a range of safety studies to develop crash prediction models for estimating the number of crashes on roadways. The Federal Highway Administration (FHWA) introduced the safety performance functions (SPFs) in the Highway Safety Manual (HSM) as crash prediction models. SPFs are regression models for estimating the predicted average crash frequency of individual roadway segments or intersections (AASHTO, 2010).

In the HSM, a 250-foot radius around an intersection is considered the buffer size, or the influential area of an intersection. Accordingly, SPFs are developed based on this presumption. Although researchers have commonly adopted this size for safety studies, there is a lack of study in validating the 250-foot influence area size. Also, other studies adopted this buffer size for non-motorized safety studies without validating the appropriateness of a 250-foot buffer zone with regard to non-motorized safety. The first edition of the HSM lacked a bicycle SPF. Instead, the manual suggested estimation of bicycle crash frequency through application of a factor to the predicted vehicle crashes. Some researchers have developed bicycle SPFs, but they failed to consider crucial geometric characteristics of intersection. Thus, the mentioned SPFs do not help agencies and planners to find counter measures for improving the safety. Consideration of intersection geometric characteristics during SPF development will enhance their applicability for agencies in safety improvement efforts.

The vulnerability of pedestrians and cyclists during crashes causes a higher rate of severe injuries and fatalities. Cyclists account for about 2 percent of total crash fatalities (USDOT, 2012), while less than 1 percent of individuals use bicycles for commuting to work (US Census Bureau, 2013). About 6 percent of pedestrian injury crashes result in fatality, whereas just 1 percent of motorized vehicle injury crashes result in fatality (USDOT,2012). Meanwhile, just 2.8 percent of individuals walk to work and 91 percent of individuals drive vehicles or public transportation, however approximately 14 percent of total crash fatalities involve pedestrians (US Census Bureau, 2013). The previously mentioned facts demonstrate the urgency of providing a bikeable and walkable environment which promotes non-motorized safety (Winters et al., 2010). Table 1 illustrates traffic crash statistics in the state of Michigan. This table compares fatality rates between all kind of crashes (motorized crashes) and non-motorized crashes (pedestrian and bike) in the state of Michigan. The fatality rate of non-motorized transportation is higher than motorized transportation, since just 0.30 percent of motorized crashes result in a fatality, while 5.82 percent of pedestrian and 1.15 percent of bike crashes are fatal.

Table 1 : Michigan Traffic Crash Statistics (2008-2012)

Crash Type	Crash Number	Fatality Number	Fatality Percentage
All kind of crashes	1,771,224	5,280	0.30%
Pedestrian Crashes	13,274	772	5.82%
Bike Crashes	11,996	138	1.15%

Statistical analysis of crash frequency could lead to the identification of problematic locations and an understanding of the factors which endanger the safety of pedestrian and cyclists. 31 percent of fatal bicycle crashes and 25 percent of fatal pedestrian crashes happen at intersection areas, indicating intersections are important locations for safety study and improvement (USDOT, 2009). Some practitioners considered observed crash frequency to assess intersection safety; however a better measure of safety evaluation is expected crash frequency (Hauer et al., 1988) The number of crashes per exposure (ADT, pedestrian and bike volume) could also be a better measure for judging the safety performances of intersections (Kononov, 2004), but naively considering crash frequency per ADT as a safety performance function could lead to unproductive decisions regarding budgeting road projects.

This study proposes bicycle and pedestrian SPFs based on the determined influence area that could be used in the next edition of the HSM. Exposure, geometric and city characteristics of 164 intersections in the four Michigan cities of Ann Arbor, East Lansing, Flint, and Grand Rapids are investigated. Since these characteristics have complex interrelations among each other; a structural equation model (SEM) is employed to decipher the complexity of the relationship. Identifying the influential area of intersections for motorized and non-motorized transportation as well as the proposed SPFS are the products of this thesis which are expected to be beneficial for transportation researchers in further studies on intersection and midblock related crashes. Also, introducing some key intersection geometric characteristics as statistically significant factors which affect the safety of intersections could guide planners and designers in identifying targeted non-motorized safety countermeasures.

This study consists of 5 chapters. The first chapter introduces the subject and goals of the study. Chapter Two presents literature of non-motorized safety and the influential area. Chapter Three explains data collection and the spatial data base developed for the analytical part of the study. Chapter Four investigates the influential area of intersections for motorized and non-motorized transportation at different speed limits. Chapter Five develops the non-motorized (pedestrian and bike) SPFs based on the proposed area of influence. Chapter Six provides a summary of the results and contributions, and proposes future research developments.

2. LITERATURE REVIEW

Although identifying the influential area of intersection-related crashes is critical for intersection and midblock safety studies, there is a lack of published statistical approaches on identifying the influential area of an intersection. Practitioners usually use their own subjective definition for the influential area of an intersection (Wang et al., 2008). Generally practitioners define a virtual 250-foot buffer size around the intersections, and they associate crashes within the boundaries of the buffer to the intersection (National Research Council, 2008). Practitioners use different methods to perceive the influential area of intersections attributing to intersection-related crashes. Some of them consider crashes within the crosswalk of intersections, and many of them consider the buffer around intersections, which is typically a 250-foot radius from the center of the intersection (Box, 1970). In the first edition of the Highway Safety Manual (HSM), a 250-foot buffer size is adopted for collecting crash data and developing the safety performance functions (SPFs). The Department of Transportation (DOT) of each state assumes different buffer sizes, which differs from 132 feet in Missouri to 250 feet in California (Wang et al., 2008). Some of the crashes within the 250-foot radius of an intersection could happen regardless of the factors related to intersections, so a 250-foot buffer size applied to all crash types may result in some statistical errors in safety analysis (AASHTO, 2010). Regression models could be a potential tool for investigating the size of the influential area, but when using regression models, a causal statistical relationship among independent (X_i) and dependent (Y) variables should be assumed. In case of violation in the mentioned statistical assumption, the proposed models may incorrectly estimate the likelihood parameter (Chang, 2005).

The classification tree method is an old and subjective data mining tool which many practitioners widely utilized it for safety studies (Abdel-Aty et al., 2005; Kuhnert et al., 2000;

Karlaftis et al., 2002). Wang et al. (2008) applied the classification tree method to identify the influential area of an intersection for each approach. The study aided in understanding factors which might be considered for specifying the area of influence, but it relied on subjective police reports for identifying intersection related crashes and then employed a classification tree model to specify the influential area. This statistical model is an antiquated technique, and given that some branches of the tree that resulted in 2%, 3%, or 21% of the crashes, the tree may have not grown in a balanced way. Also, there is a lack of such studies on non-motorized crashes, whose nature is different than motorized crashes.

In short, practitioners measure the distance of crashes to intersections and discern those that are not correlated to the intersection, and then perform a safety study for the two separate sets of intersection related crashes and midblock crashes. Wang et al.(2008) depicted the distance of crashes to the intersections, and by observing the depicted figure, it could be inferred there are two sets of gamma distribution representing both intersection crashes and midblock crashes, separately. Many researchers have employed a gamma distribution for transportation studies. GÅRDER (1998) used the gamma distribution for safety studies on raised bike crossings. Polus (1979) employed the gamma distributions on a travel time study. Another method, the Expectation-Maximization algorithm (EM algorithm) deals with mixed distributions and simultaneously estimates the parameters (Moon, 1996). In one study, the EM algorithm was employed for fitting a mixture model (Bailey et al., 1994). Some researchers, such as Park et al. (2009) and Jin et al. (2011) applied this algorithm for transportation safety studies. "mixtools" is a useful package of the R software which is beneficial in distinguishing parameters of mixed distributions by employing EM algorithms (Benaglia et al., 2009). Discriminating two different and mixed distributions is a general classification problem. Welch (1939) developed a formula to minimize miss classification. The output of Welch's formula could be used as a border for specifying the boundary between two different distributions of intersection associated crashes and midblock associated crashes. Therefore, Welch's formula could be utilized to effectively pinpoint different buffer sizes for motorized and non-motorized crashes, enhancing the accuracy of crash prediction models and SPFs.

For safety studies of intersections, practitioners collect data of crashes inside the influence area for investigation. In order to develop crash prediction models, different sets of crash frequency or crash severity are commonly considered as dependent variables, while characteristics of the corridors are investigated as dependent variables. The Poisson regression model could be utilized to predict discrete dependent variables from significant parameters. However, researchers commonly observe data over dispersion (the variance is significantly different from the mean) e.g. Oh, et al. (2013). The heterogeneous nature of the data and possible errors in collecting exposures statistics (ADT and Bike Volume) may be responsible for the over dispersion (AASHTO, 2010). In the literature of safety studies, the over dispersion in crash data has been observed repeatedly e.g. Chang (2005). The over dispersion proves crash exposures are not Poisson-distributed, thus Poisson model is not recommended for regression analysis (Lord & Mannering, 2010). Many researchers have utilized the Negative Binomial (NB) regression model to deal with the over dispersion in the safety studies e.g. Hadi (1995). Additionally, some researchers have employed the NB model to develop SPFs, e.g. Oh et al. (2013).

Geometric and exposure (ADT, bike volume and pedestrian volume) characteristics of corridors are widely used to develop crash prediction models, such as SPFs. The American Association of State Highway Transportation Officials (AASHTO) recently used exposure measures to develop a pedestrian SPF in the first edition of the HSM, but geometric characteristics of the intersection were disregarded (AASHTO, 2010). Geometric characteristics could guide agencies to find safety countermeasures for improving non-motorized safety. Raford et al. (2005) considered the exposure term as a rate of contact with agents or events (ADT, pedestrian and bike volume) that are potentially harmful. In other words, pedestrian and bicyclist interactions with each other, vehicles and the intersection itself cause conflicts that result in bike and pedestrian crashes. Practitioners have investigated this effect on the crash frequency of intersections (Chin & Quddus, 2003; Turner et al., 2011), but the relationship between crash frequency and vehicle exposure (ADT) is non-linear (Tanner, 1953). Increasing volumes are associated with a decrease in crashes per vehicle exposure (Hadi et al., 1995; Mensah & Hauer, 1998). Nordback et al. (2014) adopted a 250-foot buffer size around intersections to collect bike crash data, and developed a bike SPF based on the bike volume and vehicle traffic (exposure), but geometric and city characteristics were not considered. In addition, no consideration was made for the complex interrelationship among the casual parameters. Higher level of traffic volume (ADT) is associated with a higher frequency of pedestrian crashes (Shankar et al., 2003). There is a significant relationship between the ratio of a minor road's ADT and a major road's ADT at intersection to pedestrian crash frequency at intersections (Harwood et al., 2008). This relationship was also investigated in a North Carolina study (Schneider et al., 2004).

Geometric configuration of facilities and design of corridors could affect crash frequency and crash severity. Some features, such as corridor width and presence of sidewalk and/or bike lane, may affect exposure (ADT, pedestrian volume and bike volume). Although it is expected that bike lanes improve safety, they attract increases in bike volume, thus cyclist crash frequency would be expected to increase (Oh, et al., 2013). Some studies revealed that the configuration of lanes is associated with crash frequency. Two-way left-turn lanes (TWLTL) have exhibited increases in the frequency of pedestrian crashes (Shankar et al., 2003). The configuration of the intersection (such as presence of a right turn lane) may increase crash frequency, while a raised median has been seen to reduce the number of crashes (Schneider et al., 2010). Since geometric specifications of the intersection may intensify or lessen conflicts and the resultant effect on safety. AASHTO (2010) used it to develop SPFs in different types of roadways. An increase in the number of lanes is associated with higher crash frequency (Noland & Oh, 2004). Turner et al. (2011) developed a SPF for bike crashes in New Zealand by including geometric characteristics, but there is a lack of such studies in the USA. Agencies might invest in geometric characteristics to increase the safety of non-motorized transportation.

The location of bike facilities, such as bike racks, is correlated with bike crashes, and on-street vehicle parking jeopardizes pedestrian safety (Moini & Liu, 2013). Location of bus stops is correlated with pedestrian crash frequency (Moini & Liu, 2013). In another safety study, a positive correlation between on-street parking and crash severity is investigated (Zahabi, et al., 2011). Although on-street parking provides more accessibility to businesses and retail, it is associated with higher crash frequency (Greibe, 2003). Since on-street parking narrows the width of the roadway, driver maneuvering space is decreased, which could reduce the driver's sight distance, required to distinguish and react to crossing pedestrians.

Some studies revealed that vehicle speed is associated with a higher risk of pedestrian crashes (Lee & Abdel, 2005). There are many studies concerning the effect of different traffic sign placement on crash frequency and drivers recognition (Monagle et al, 1955). The impact of an engineering program for improving traffic signs led to a 34 percent reduction in crash frequency (USDOT, 1989). This statistic demonstrates the role of traffic sign design and placement on crash frequency.

A higher density of intersections in an area is associated with increased frequency of non-motorized crashes in the area (Siddiqui et al., 2011). Presence of bike lanes in the corridors causes more bike crashes (Oh, et al., 2013), while a higher proportion of bike lanes on the corridor leads to a significant reduction in incapacitating-level injuries (Narayanamoorthy et al, 2013). Presence of a bike lane at the intersection is associated with higher bike volume, resulting in more conflicts between bikes and vehicles (Oh, et al., 2013). In another study, the relationship between lane configuration and crash frequency is investigated. A higher number of lanes on the corridor increase the pedestrian crash frequency, since pedestrians are required to cross more lanes (Harwood et al., 2008).

Movement of the non-motorized transportation could affect the motorized dynamics in the neighboring lane (Xie et al., 2009). Appropriate geometric design of a corridor’s facilities, such as on-street parking, could reduce conflicts between bikes and vehicles, thus improving cyclist safety (Barnes et al., 2013). The presence of a right turn lane at an intersection is associated with increased pedestrian crashes (Schneider et al., 2010). In one study, the impact of the number of lanes on non-motorized crash frequency is investigated (Oh, et al., 2013). Corridors with uncovered sidewalk are associated with higher pedestrian crash frequency (Schneider et al., 2004). Roadway lighting was found as a significant factor on non-motorized safety. Inappropriate lightening is associated with a higher frequency of pedestrian crashes, and may increase the severity of those crashes (Zahabi et al., 2011; Spainhour et al., 2006).

Although crash prediction models explain the relationship of significant independent variables to the dependent variable, these models can not reveal the complex interrelationship among the variables. The structural equation model (SEM) is a statistical tool which was developed to estimate and test casual inter-relationship between variables. This model categorizes variables into different groups, and demonstrates the level of covariance between two correlated variables.

Many practitioners employed the SEM to demonstrate complex interrelation among a group of variables in transportation studies. (Chung et al., 2002) investigated the interrelationships among socio-demographics, activity participation and driver behavior by developing a SEM. In another study, SEM was used to explain the interrelationship between travel demand, land use, socio-demographics, telecommunications and economic activities (Choo et al., 2007). Deutsch (2013) modeled travel behavior and the factors contributing to a sense of a place, utilizing a SEM. Hassan (2011) developed a SEM for analysis of driver behavior under reduced conditions. Hamdar (2013) developed a safety propensity index for both interrupted and uninterrupted flow scenarios using the structural equation method. In another application of SEM in safety studies, road factors, driver factors and environment factors are found as exogenous latent variables, and a factor representing the size of accidents is found as an endogenous latent variable (Lee et al., 2008). However, the literature review revealed an absence of SEM application in non-motorized safety studies.

In addition to the previously mentioned factors, some researchers revealed the effect of some miscellaneous factors such as neighborhood socioeconomics (income, poverty, etc.), weather conditions (seasons, snow, rain, etc.), education level (number of students, etc.), cultural issues (ethnics, etc.) and land use (business, campus, etc.) on crash frequency and crash severity e.g. Emaasit et al. (2013), Abasahl (2013), Narayanamoorthy et al. (2013), Kaplan et al. (2013), Wang et al. (2013), Wier et al. (2009), Wedagama et al. (2006), Harwood et al. (2008) and Spainhour et al. (2006). However, these parameters are not used typically used in the development of SPFs.

3. Data Collection

3.1 Crash Data

The scope of this study is limited to the four Michigan cities of Ann Arbor, East Lansing, Flint, and Grand Rapids. 164 signalized intersections are investigated for crash data collection. In order to analyze the influential area of an intersection and to develop Safety Performance Functions (SPFs), data of the crashes which happened inside of the buffer around the intersections is collected. FHWA attributed crashes within 250 feet around the center of the intersections to intersection-related crashes, and practitioners widely used this buffer size for safety analysis e.g. Vogt (1998).

In order to investigate the influential area of an intersection as the first goal of this study, a big buffer size around the intersection is considered for crash data collection. For this purpose, the longest stopping sight distance (SSD) from the intersection is considered as the buffer size. In this distance a driver could see an intersection and stop before it. Fambro et al., (1997) developed a formula for calculating the SSD. Equation 1 shows the SSD formula.

$$SSD = 1.47\, V * T + 1.075\, \frac{V^2}{a} \tag{1}$$

Where:

SSD: stopping sight distance (feet)

V: vehicle speed (mph)

T: reaction time (second)

a: deceleration rate (feet/second^2)

Fambro et al. (1997) considered 2.5 seconds as a perception and reaction time and 11.2 feet/second^2 as the deceleration rate. This assumption could be applied for 90 percent of drivers (Fambro et al., 1997).

Highest observed speed for the collected non-motorized crashes was 70 mph, so using Equation 1; SSD would come up as 497.2054 feet, which is rounded to 500 feet distance from intersection as the buffer around the intersections. This distance is adopted for crash data collection which is provided by Michigan State Police (MSP). Figure 1 depicts a 500-foot buffer around an intersection which is used for the crash data collection.

Since adjacent intersections may affect each other and the calculated buffer size is 500 feet, the 148 intersections which do not have any other intersections with in the 500-foot area from their center are selected for investigating the influential area of an intersection. Figure 2 to Figure 9 illustrate pedestrian and bike crashes in the scope of this study.

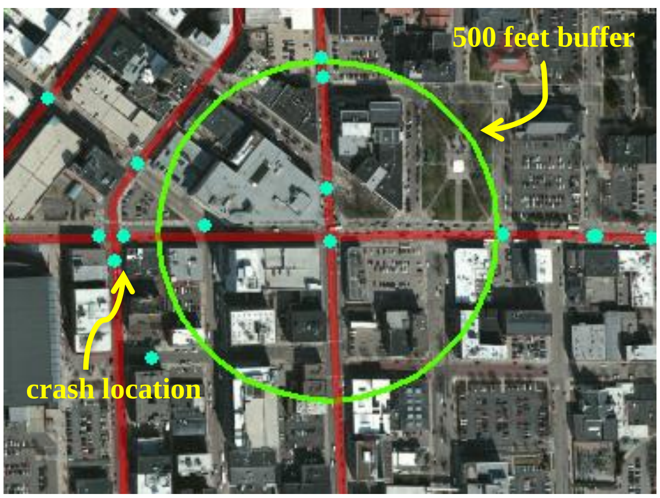

Figure 1 : 500-foot Buffer around Intersection and Crash Locations

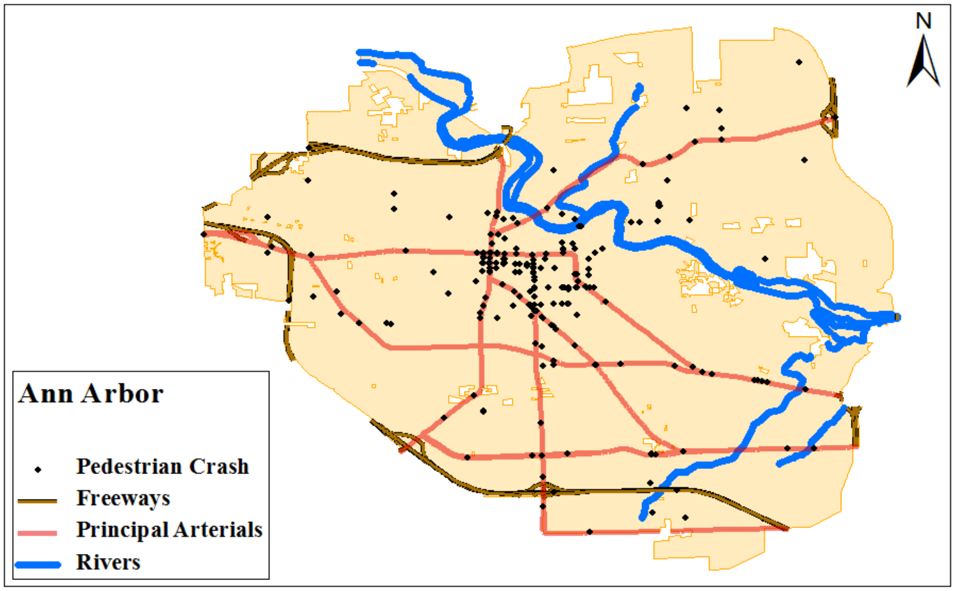

Figure 2 : Location of Pedestrian Crashes, Ann Arbor (2008-2012)

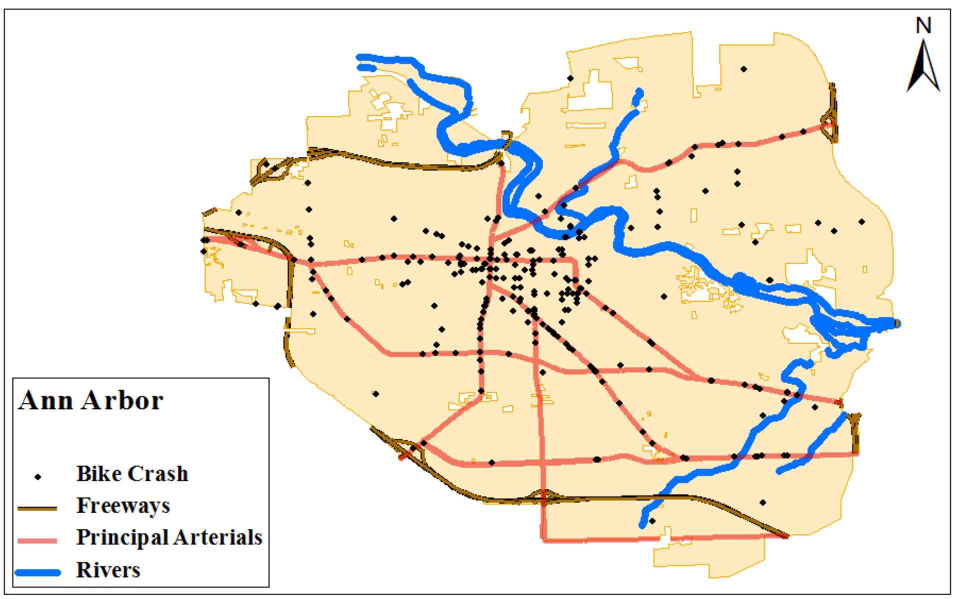

Figure 3 : Location of Bike Crashes, Ann Arbor (2008-2012)

Figure 2 and Figure 3 show that downtown Ann Arbor encompasses a higher number of pedestrian and bike crashes in comparison with other parts of the city. A big portion of Ann Arbor's downtown is dedicated to the University of Michigan.

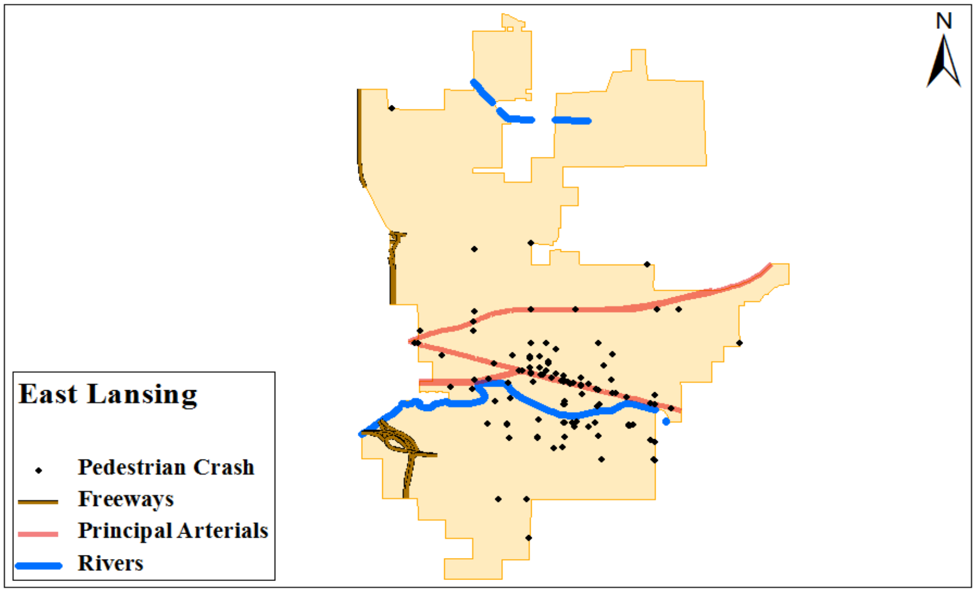

Figure 4 : Location of Pedestrian Crashes, East Lansing (2008-2012)

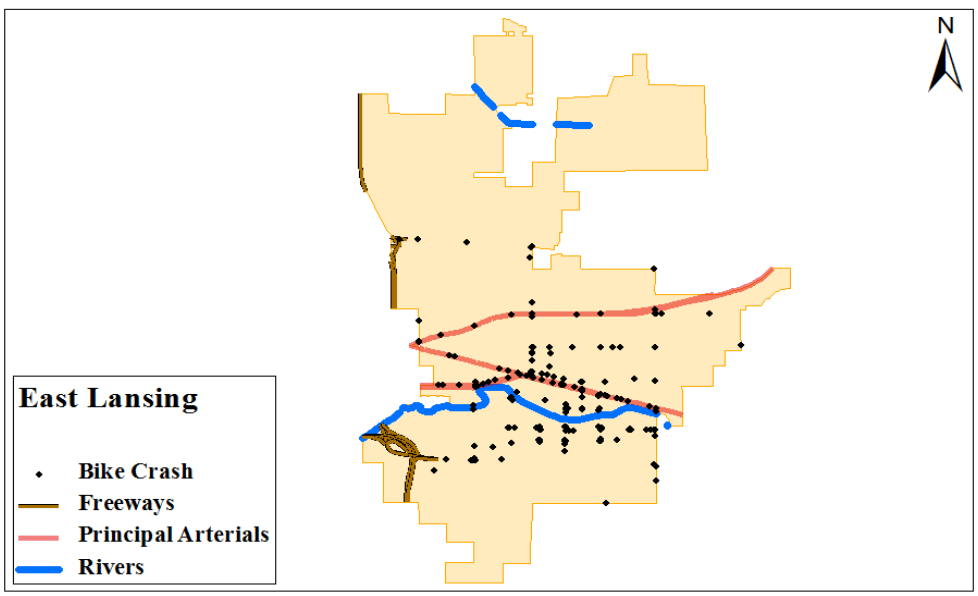

Figure 5 : Location of Bike Crashes, East Lansing (2008-2012)

The southern part of East Lansing is dedicated to Michigan State University, and as depicted above, most of the non-motorized crashes are located within the area.

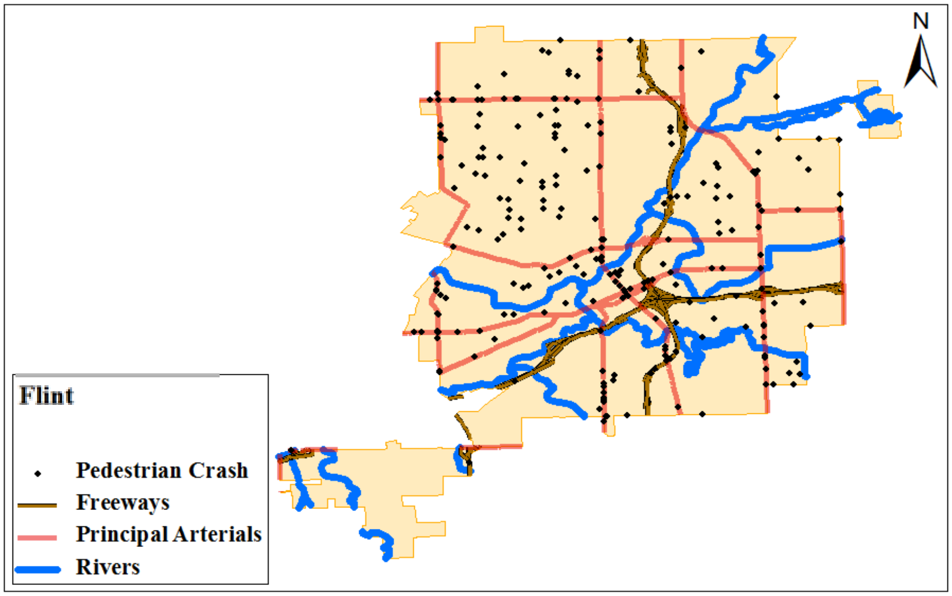

Figure 6 : Location of Pedestrian Crashes, Flint (2008-2012)

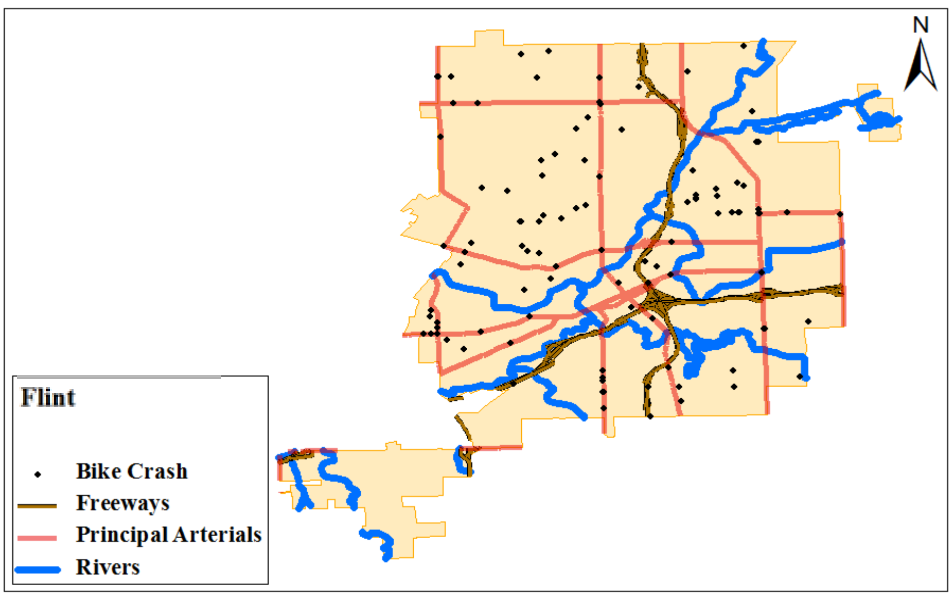

Figure 7 : Location of Bike Crashes, Flint (2008-2012)

As is illustrated in the above mentioned figures, locations of pedestrian and bike crashes are spread throughout Flint, and there are a greater number of pedestrian crashes than bike crashes.

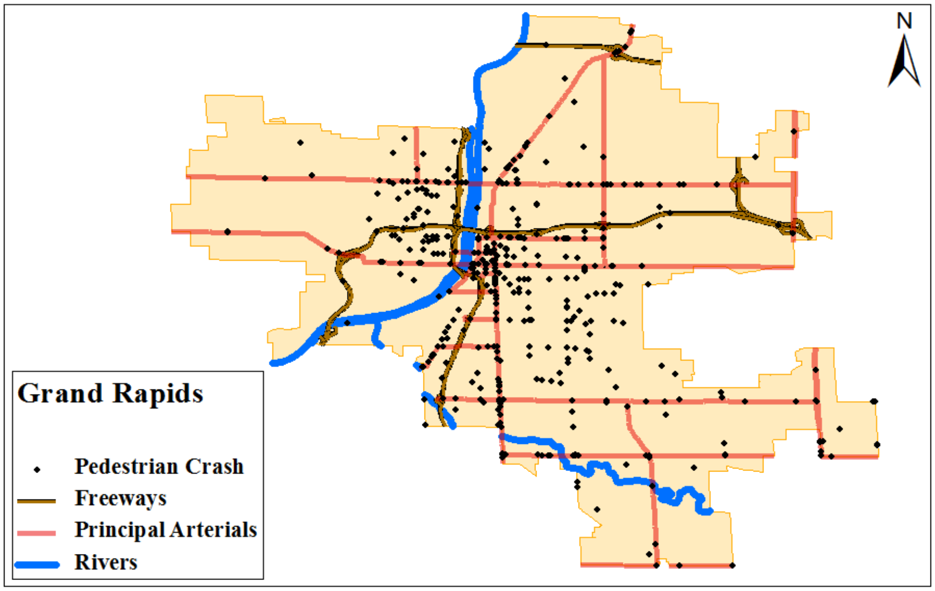

Figure 8 : Location of Pedestrian Crashes, Grand Rapids (2008-2012)

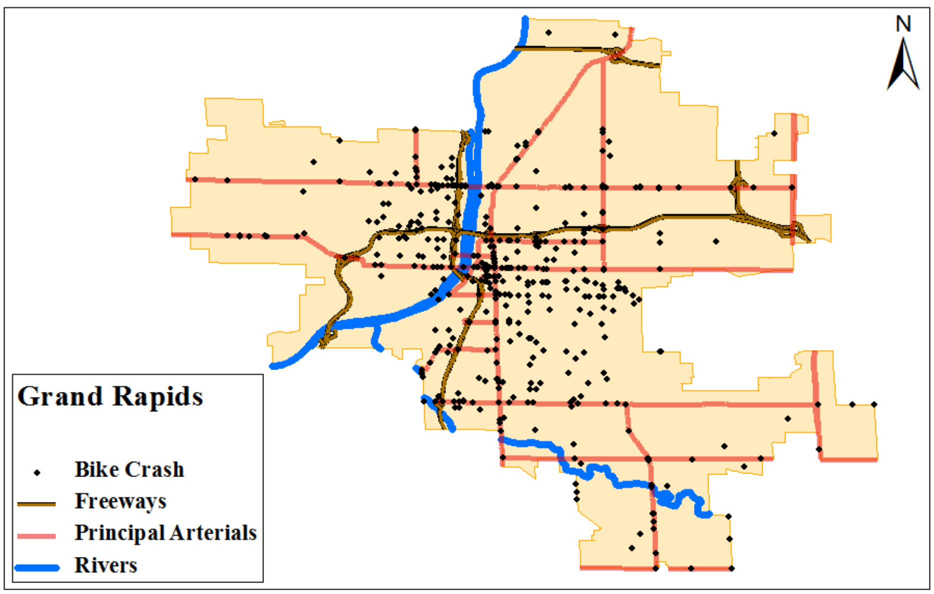

Figure 9 : Location of Bike Crashes, Grand Rapids (2008-2012)

According to the above mentioned figures, the density of non-motorized crashes in downtown Grand Rapids is higher than in other parts of the city.

ArcGIS 10.0 was employed to collect and manage spatial crash data within the buffer around the intersections.

Table 2 and Figure 10 illustrate crash pedestrian frequency in the four cities. Grand Rapids has the highest frequency of pedestrian crashes while East Lansing has the lowest frequency.

Table 2 : Pedestrian Crash Frequency

City	Pedestrian Crash Frequency					
	2008	2009	2010	2011	2012	Total
Ann Arbor	52	43	45	63	60	263
East Lansing	20	23	19	25	34	121
Flint	57	44	44	45	61	251
Grand Rapids	90	69	94	94	121	468

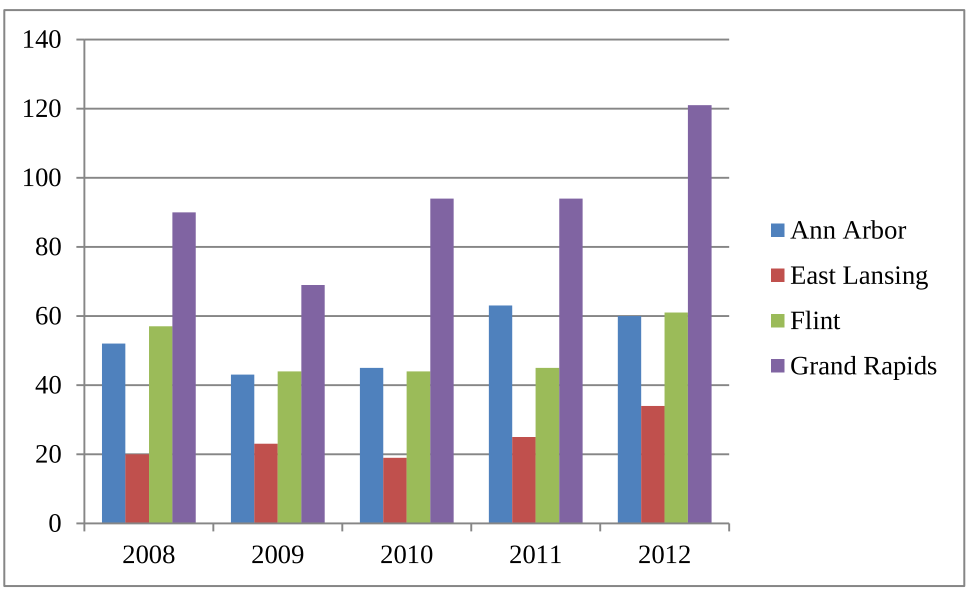

Figure 10 : Comparison of Crash Frequency in the Scope of the Study

Above mentioned figures provide brief statistics based on the annual crash frequency. Other characteristics of the cities are not provided. Further investigation for crash comparison is provided in Table 4.

Table 3 and Figure 11 illustrate crash pedestrian frequency in the four cities. Grand Rapids has the highest frequency of pedestrian crashes while East Lansing has the lowest frequency.

Table 3 : Bike Crash Frequency

City	Bike Crash Frequency					
	2008	2009	2010	2011	2012	Total
Ann Arbor	59	63	59	59	64	304
East Lansing	42	51	30	43	59	225
Flint	20	26	30	14	15	105
Grand Rapids	100	113	89	95	94	491

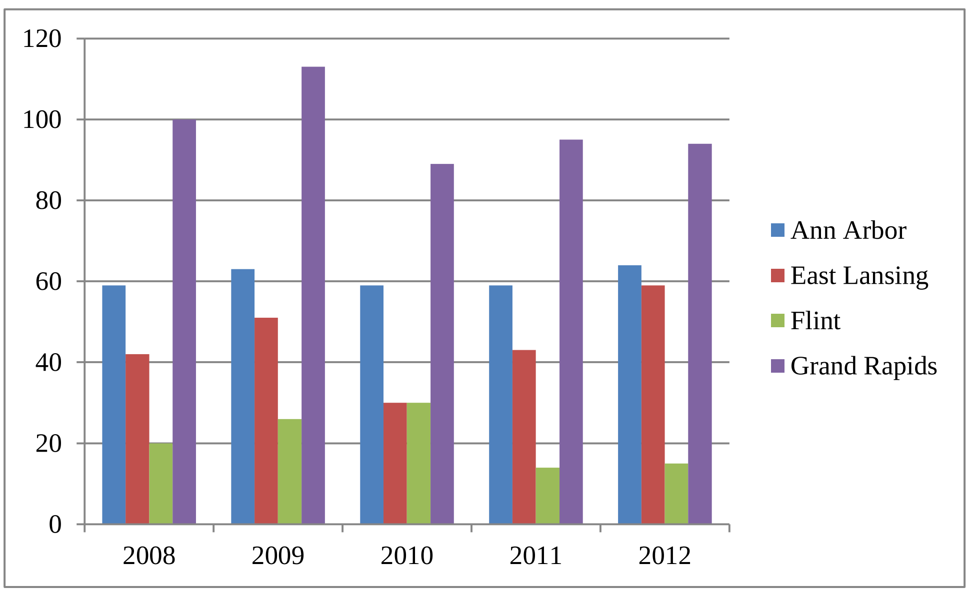

Figure 11 : Comparison of Crash Frequency in the Scope of the Study

For a better justification on crash statistics among the four cities, in Table 4 this study compares the average crash per population and the average crash per commuters among the four cities.

Table 4 : Comparison of Crash Frequency Rate

		City			
		Ann Arbor	East Lansing	Flint	Grand Rapids
Population (2010)		113,939	48,557	102,434	188,040
Means of Transportation to Work	Bike	1,728	1,468	20	764
	Walk	8,378	5,360	813	2,823
	Bus	5,292	1,227	1,362	2,967
1000*(Avg. Bike Crash)/ (Bike Commuters)		35.2	30.7	1050.0	128.5
1000*(Avg. Ped. Crash)/ (Walk Commuters)		6.3	4.5	61.7	33.2
1000*(Avg. Ped. Crash)/ (Bus Commuters)		9.9	19.7	36.9	31.5
1000*(Avg. Ped. Crash)/ (Walk & Bus Commuters)		3.8	3.7	23.1	16.2
10000*(Avg. Bike Crash)/ (Population)		5.3	9.3	2.1	5.2
10000*(Avg. Ped. Crash)/ (Population)		4.6	5.0	4.9	5.0

Comparison statistics on pedestrian crashes reveal that for the number of pedestrian crashes for commuters walking to work or to the bus stop, Flint is more dangerous while East Lansing is the safest place. But average pedestrian crashes per population in all of the cities are almost same. It shows that the number of bike, walk, and bus commuters in Flint is fewer than other cities.

Considering average bike crashes per bike commuters reveal that Flint is the most dangerous place for bike commuters and after that Grand Rapids. East Lansing seems to be the safest place for bike commuters. The least number of bike crashes occur in Flint. This fact is illustrated in Figure 12.

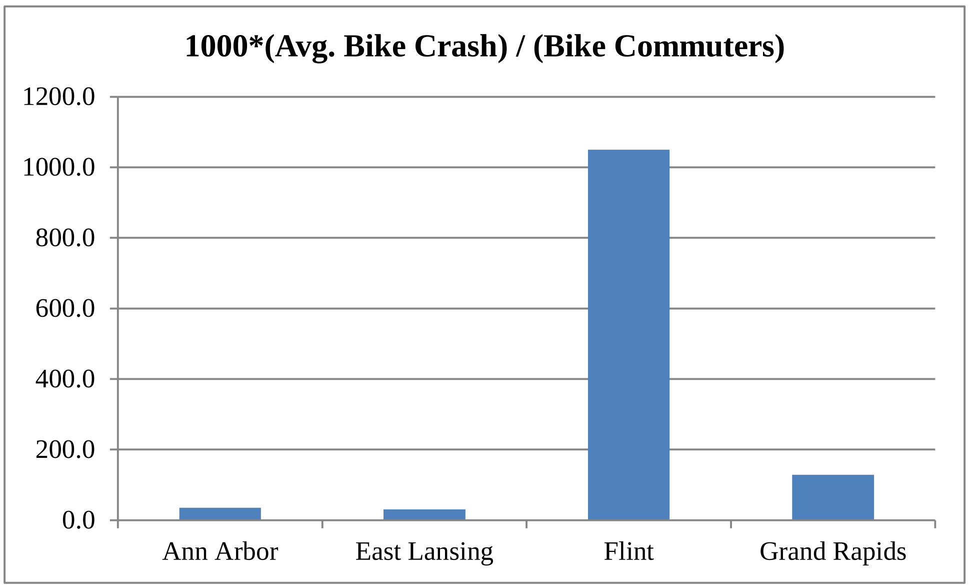

Figure 12 : Comparison of Bike Crash Rate per Commuters

In terms of the average bike crash per population, East Lansing has the highest rate, and Flint has the lowest. By considering the previous measure, it inferred that in Flint the average number of bike commuters per population has the lowest rate, and in East Lansing the average number of bike commuters per population has the highest rate. This fact is illustrated in Figure 13. A big portion of East Lansing is the Michigan State University campus, consequently a big portion of the population is students, and it is expected that there is a higher number of bike commuters.

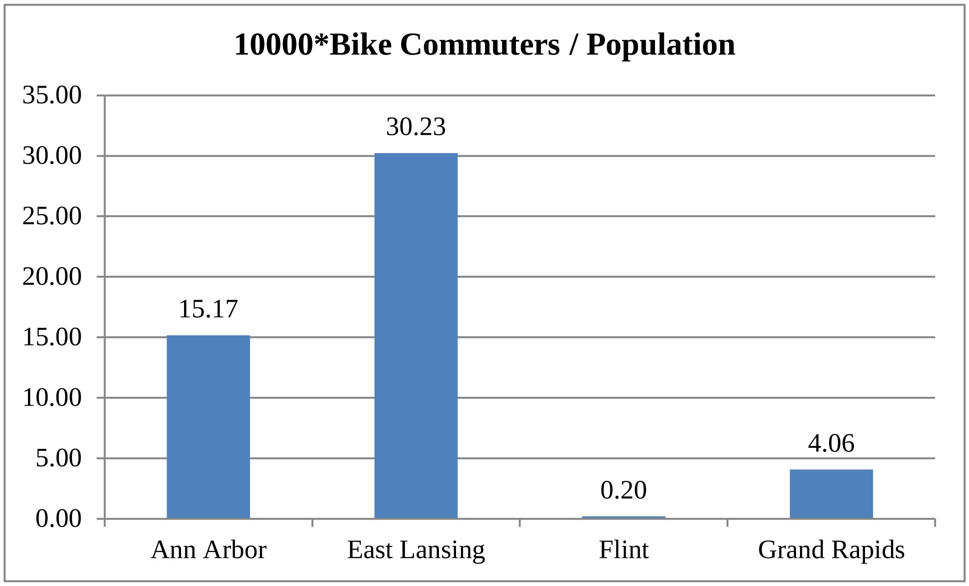

Figure 13 : Average Bike Commuters per Population

In terms of considering average pedestrian crashes per walk commuters, it could be inferred that Flint is a dangerous place for walk commuters, and East Lansing is the safest. Considering the average number of pedestrian crashes per total bus and walk commuters, Flint again is the most dangerous. Average pedestrian crashes per population and average bike crashes per population reveals that non-motorized transportation is not a popular way of transportation in Flint. This fact is depicted in Figure 14.

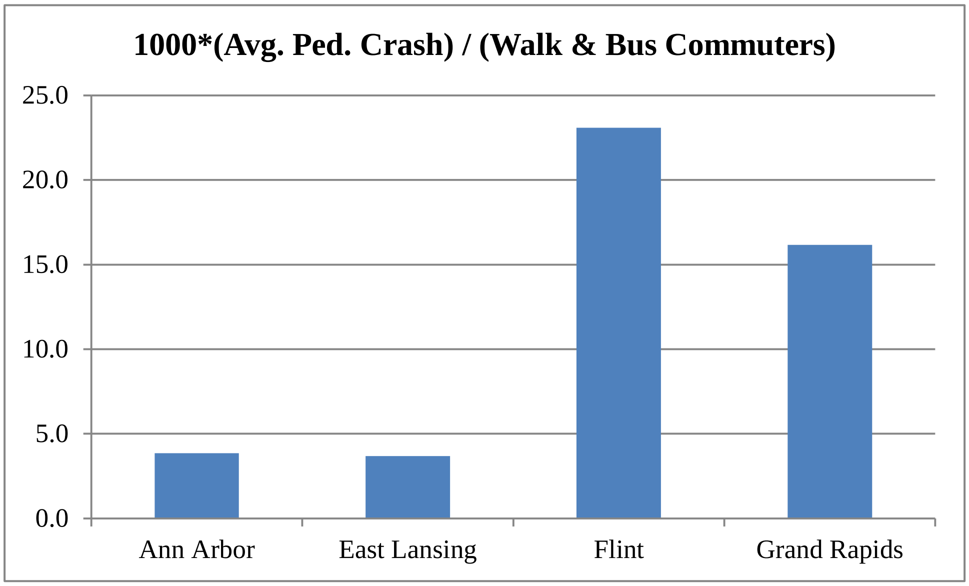

Figure 14 : Comparison of Pedestrian Crash Rate per Commuters

Since descriptive statistics show that the number of non-motorized commuters in Flint is fewer than the other four cities, it seems that Flint is not an attractive place for non-motorized transportation. Also, since the statistics show that the average number of crashes per commuters in Flint is higher than in other cities, then Flint is not an attractive or safe place for non-motorized transportation.

Oh et al. (2013) investigated the socio-economical characteristics of the cities. The results show that a higher crime rate is associated with less bike volume, and Flint has the most crime rate among the cities.

ArcGIS 10.0 is employed to calculate the distance of crashes to the center of intersections. For this purpose, UTM-NAD 83 as the geographic coordinate system is adopted for calculating the distance.

The histogram of distances is plotted in Figure 15. From this figure it could be inferred that this figure could consist of two mixed gamma distributions. By investigating this figure it seems that the first distribution starts from 0 feet and finishes around 250 feet, and then another gamma distribution starts. These adjacent distributions represent intersection-related crashes and midblock crashes. The R Software version no. 3.0.1 is employed to build the histogram of crashes.

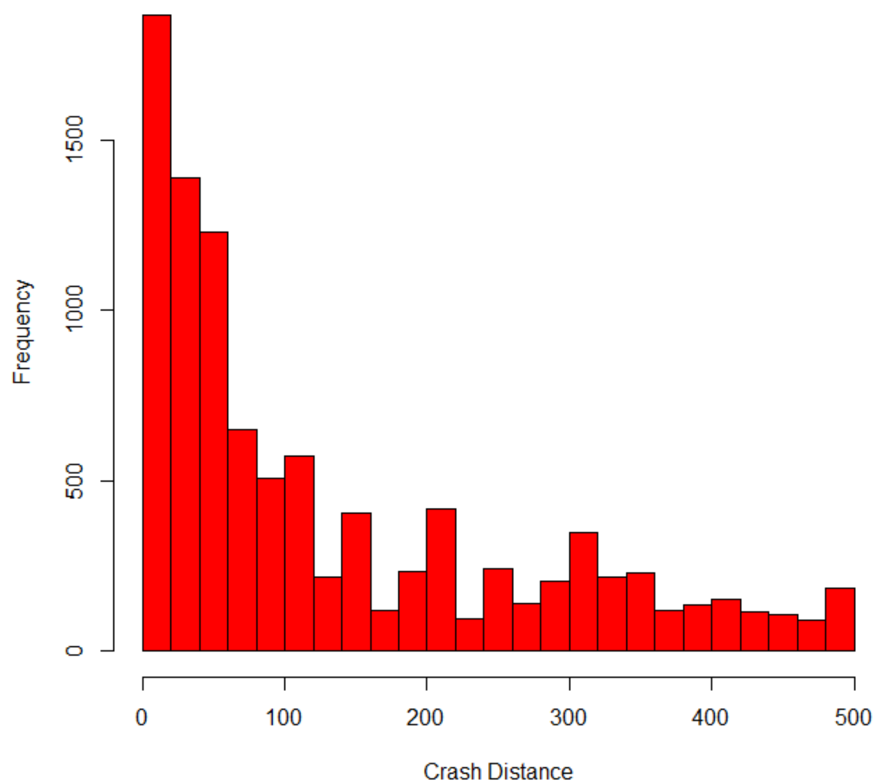

Figure 15 : All of Crashes' Histogram

The above mentioned histogram consists of all kinds of crashes. The next histograms show different distributions for the non-motorized crashes.

As illustrated in Figure 16, it seems that for non-motorized crashes (pedestrian and bike crashes), the distinguishing border between distributions is located approximately at 150 feet.

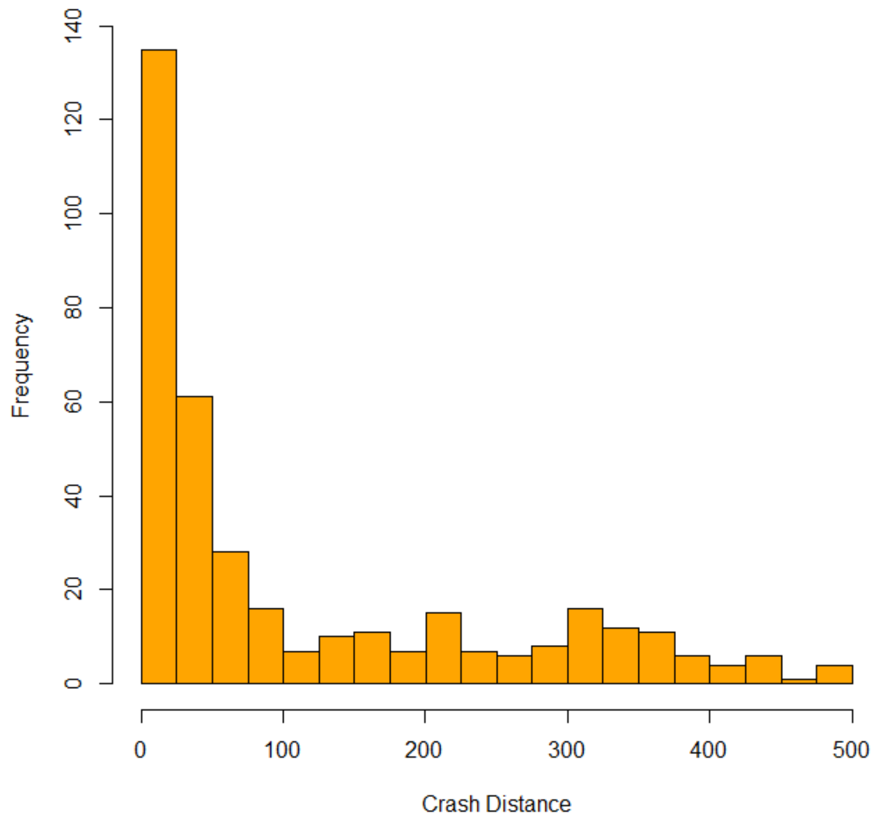

Figure 16 : Non-motorized Crashes' Histogram

Non-motorized crashes consist of pedestrian and bike crashes, and intersections might have a different area of influence for them. So, this study performed further investigations on each group of non-motorized crashes. In the next two figures, the histogram of pedestrian and bike crashes to the center of intersections is illustrated.

As in the histogram of the non-motorized crashes, it seems that for pedestrian crashes around 150 feet, a transition between distributions could be observed.

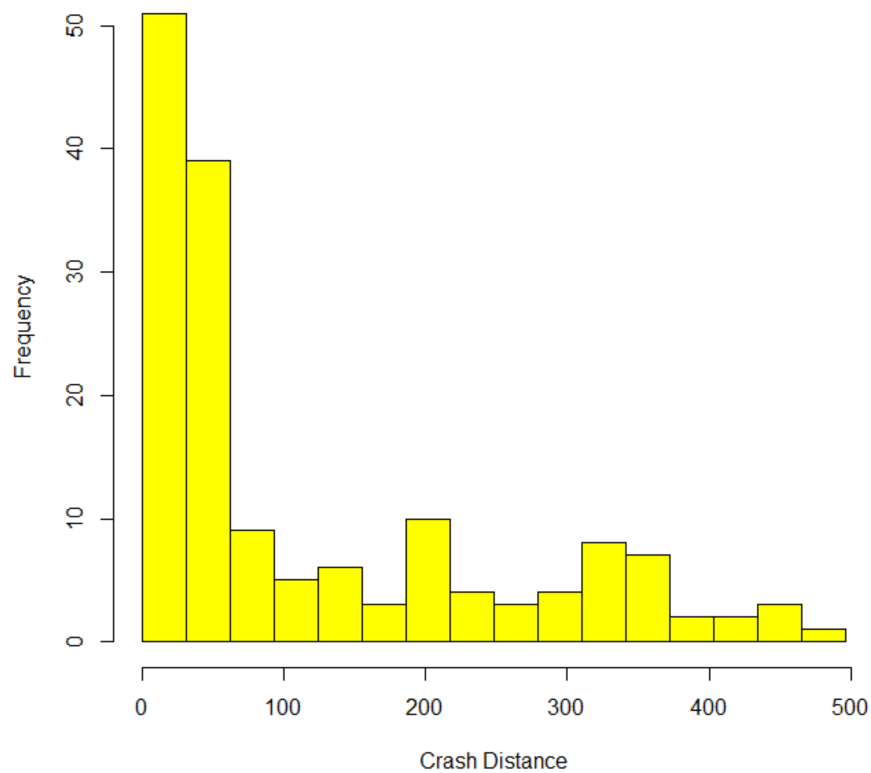

Figure 17 : Pedestrian Crashes' Histogram

The above mentioned figure shows that the pattern of non-motorized crashes and pedestrian crashes are similar, and the distinguishing border which represents the area of influence might be located in the same location. In order to compare the area of influence for bike crashes with the pedestrian and non-motorized crashes, the histogram of bike crashes is illustrated in the next figure.

As depicted in Figure 18, similar to the non-motorized and the pedestrian crashes, mixed distributions of bike crashes may have the same border near the 150-foot distance from center of the intersections.

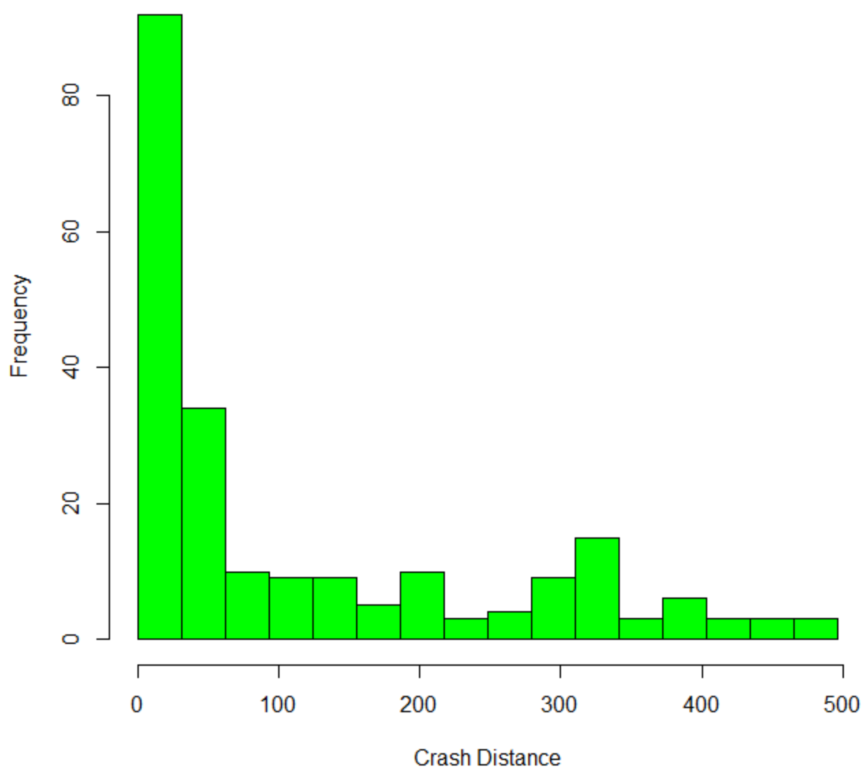

Figure 18 : Pedestrian Crashes' Histogram

The histogram of motorized, non-motorized, pedestrian, and bike crashes shows a difference in the intersection-related crashes for each type of transportation. Further investigation on finding the best border for specifying the area of influence is investigated in chapter 4 by employing the Welch’s formula.

3.2 Exposure Data

Exposures consist of the Average Daily Traffic (ADT) as the vehicle exposure and non-motorized exposure (pedestrian & bike volume). These exposures gathered from the 164 signalized intersections in the four Michigan cities of Ann Arbor (hosting the University of Michigan), East Lansing (hosting Michigan State University), Flint, and Grand Rapids. The Non-motorized exposures obtained from a recent non-motorized safety study of the Michigan Department of Transportation (MDOT) by Transportation Research Center for Livable Communities at Western Michigan University (Oh et al., 2013). The ADT of corridors are achieved from Google Earth Pro and updated to the most recent data considering the annual VMT changes in Michigan. Vehicle volume (ADT), bike volume, and pedestrian volume of intersections are considered exposure measures from each bound of the intersections.

Distribution of pedestrian volume and bike volume in the scope of this study is similar to the non-motorized crashes. Intersections with higher pedestrian and bike volume are associated with the more pedestrian and bike crash frequency. Oh et al. (2013) explained that pedestrian volume and bike volume are associated with the higher pedestrian and bike crash frequency.

The distribution of pedestrian volume on the scope of this study is demonstrated in Figure 19 to Figure 26.

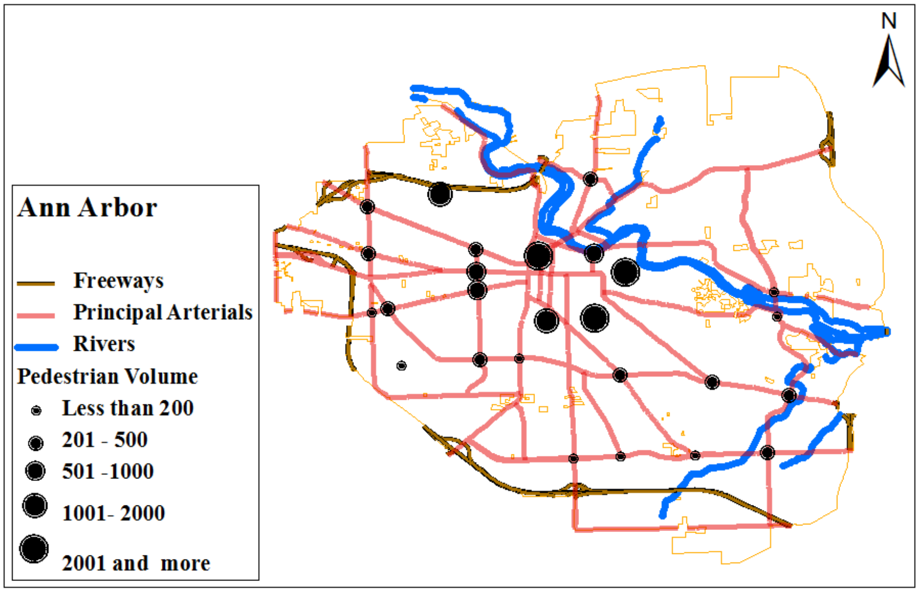

Figure 19 : Pedestrian Volume of Intersections at the Scope of the Study, Ann Arbor

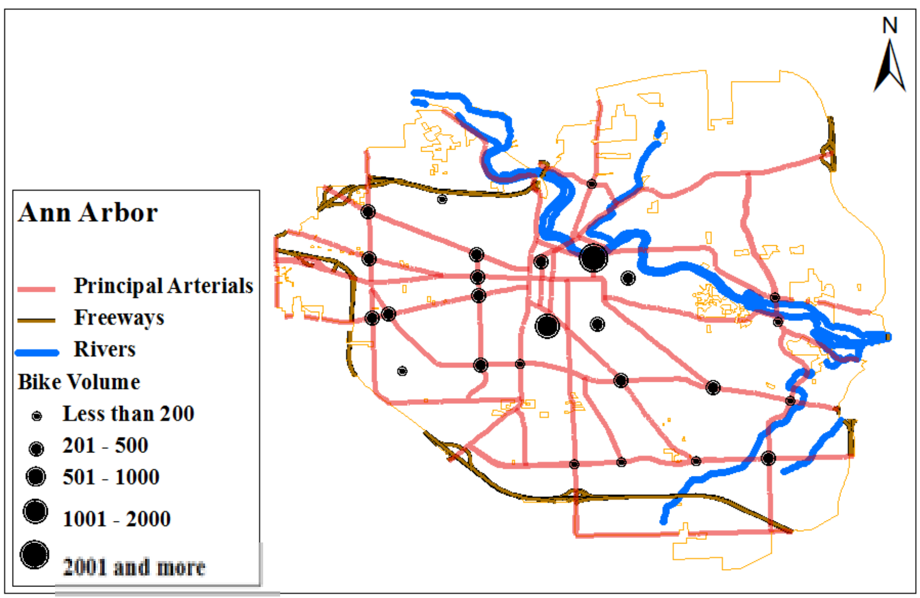

Figure 20 : Bike Volume of Intersections at the Scope of the Study, Ann Arbor

In comparison with other parts of Ann Arbor, higher bike and pedestrian is reported in the downtown.

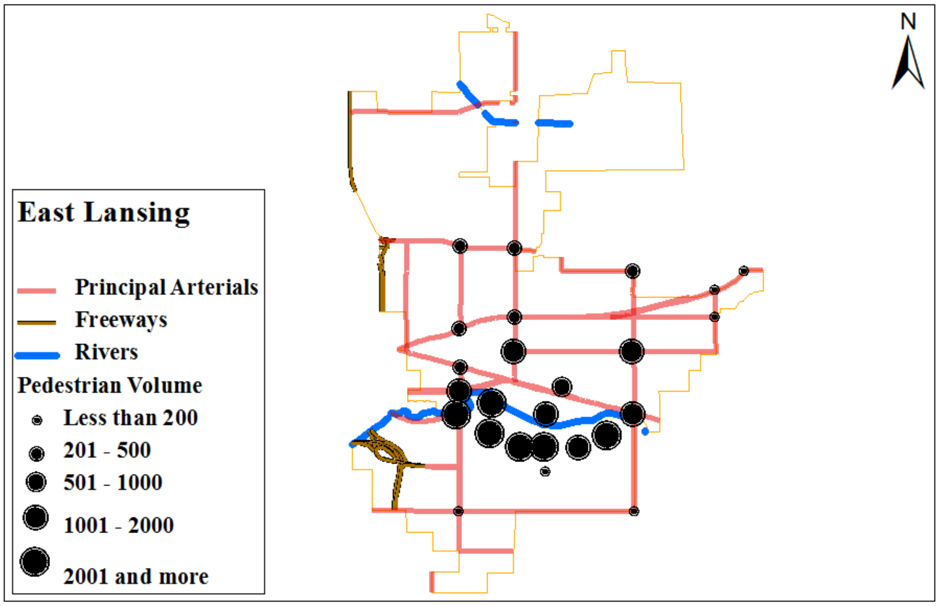

Figure 21 : Pedestrian Volume of Intersections at the Scope of the Study, East Lansing

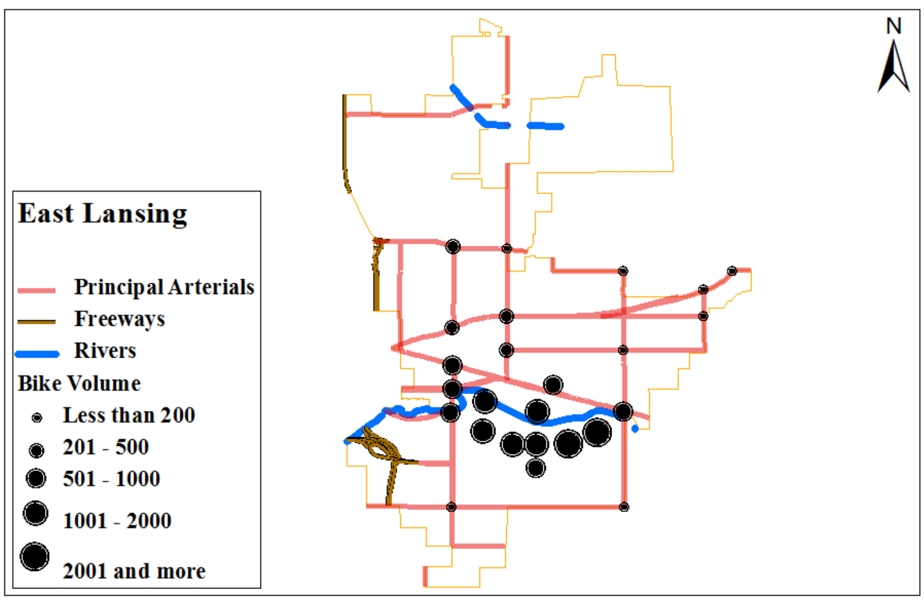

Figure 22 : Bike Volume of Intersections at the Scope of the Study, East Lansing

Campus of Michigan State University which is located in the southern part of East Lansing has higher non-motorized exposures in comparison with other parts of the city.

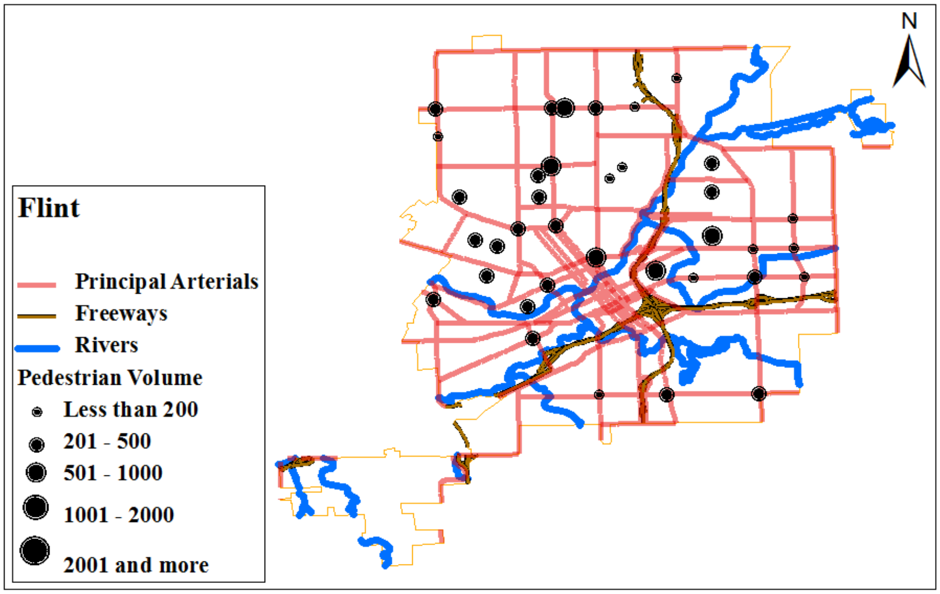

Figure 23 : Pedestrian Volume of Intersections at the Scope of the Study, Flint

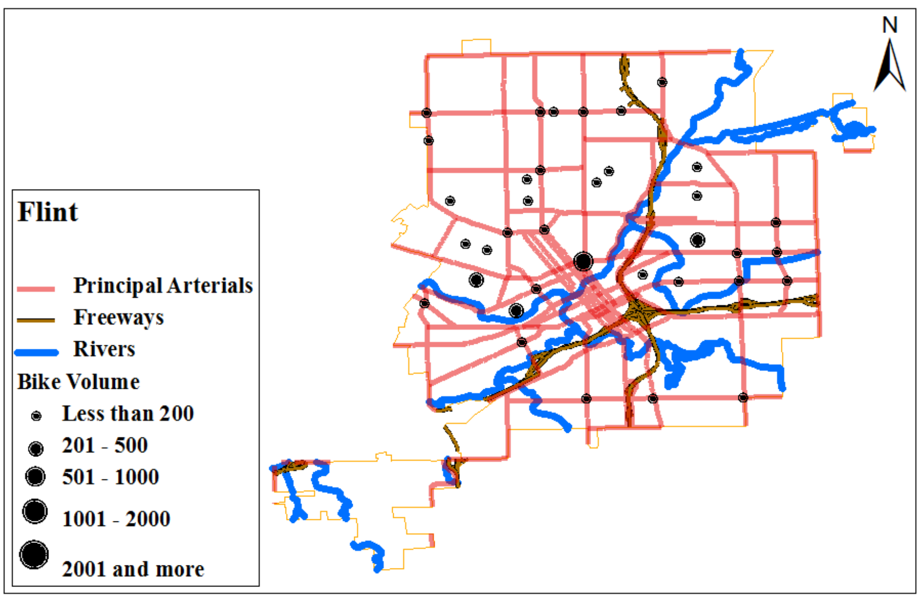

Figure 24 : Bike Volume of Intersections at the Scope of the Study, Flint

Figure 23 and Figure 24 demonstrate that pedestrian and bike volumes are spread throughout Flint and it is less than Ann Arbor and East Lansing.

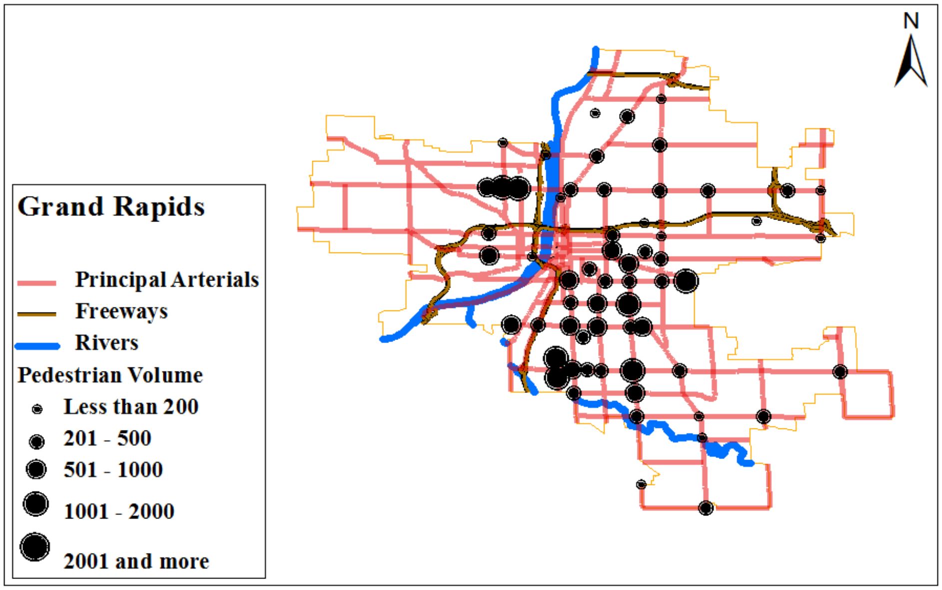

Figure 25 : Pedestrian Volume of Intersections at the Scope of the Study, Grand Rapids

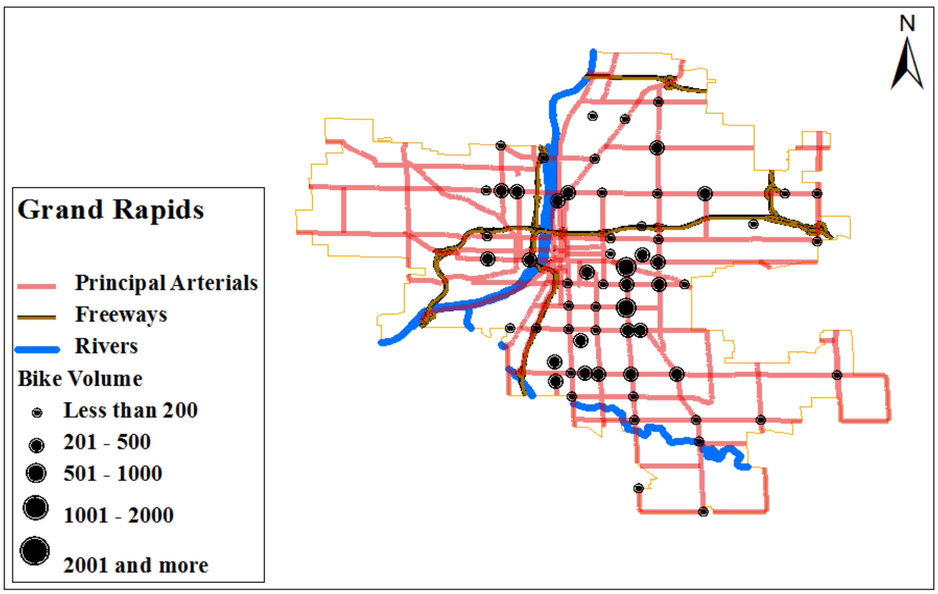

Figure 26 : Bike Volume of Intersections at the Scope of the Study, Grand Rapids

Figure 25 and Figure 26 illustrate that intersections of the downtown area have more non-motorized volume in comparison with the other parts of Grand Rapids.

ADTs of study corridors are collected by using Google Earth Pro, and the aggregated ADT of all approaches is regarded as a vehicular exposure of the intersection.

$$\mathrm{ADT}_{\mathrm{Intersection}} = \mathrm{ADT}_{\mathrm{North-South\ Corridor}} + \mathrm{ADT}_{\mathrm{East-West\ \ Corridor}} \tag{2}$$

In case of having no current ADT, an adjusting coefficient is applied to the collected ADT for updating them.

$$\mathrm{ADT}_{\mathrm{Current}} = \mathrm{C}_{\mathrm{i}} * \mathrm{ADT}_{\mathrm{year\ i}} \tag{3}$$

Where

$\mathrm{ADT}_{\mathrm{Current}}$: Current ADT

C_{i} : Adjusting Coefficient

$\mathrm{ADT}_{\mathrm{year\ i}}$: Obtained ADT from Google Earth Pro in year i.

C_{i} is calculated based on traffic changes during the years in the roadways of Michigan and updates ADT of the corridor from year i to the current date. In Figure 27 to Figure 30, the distribution of ADT in the intersections of Ann Arbor, East Lansing, Flint, and Grand Rapids is illustrated.

City characteristic of the intersections are investigated for significance on the crash frequency. The four study cities have different characteristics. East Lansing is a college city and encompasses Michigan State University just as Ann Arbor hosts the University of Michigan, meaning that a main part of these cities is dedicated to the campus. Grand Rapids is a major metropolitan region in west side of Michigan. Flint is a mid-range city which is recovering from a previous economic emergency situation and has a high crime rate (Harris, 2014).

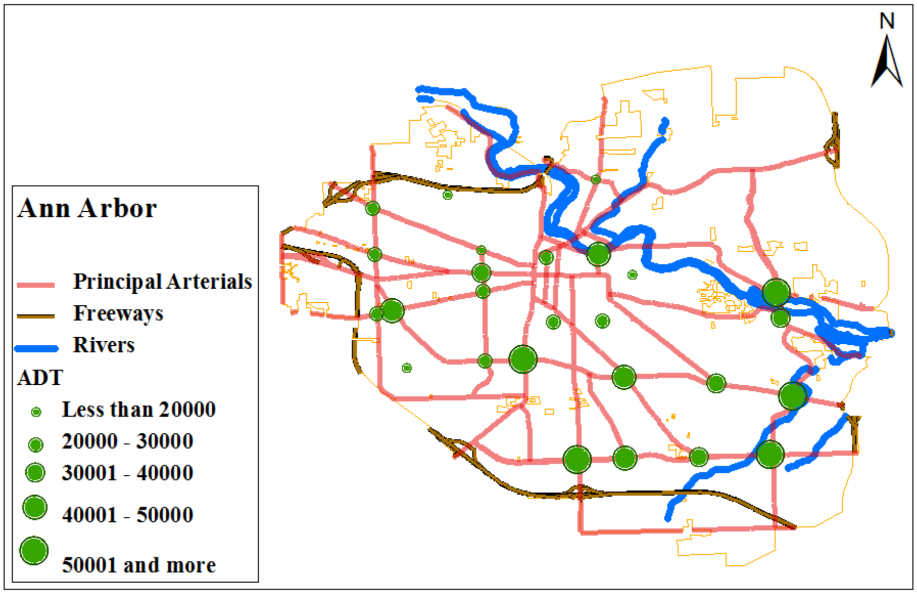

Figure 27 : ADT of Intersections at the Scope of the Study, Ann Arbor

The distribution of ADT in Ann Arbor reveals that the downtown area has less ADT although there is more pedestrian and bike volume. Figure 2 and Figure 3 already demonstrated that frequency of non-motorized crashes in the downtown is more than other parts of Ann Arbor. Oh et al. (2013) investigated the effects of land-use characteristics on the non-motorized volume and safety.

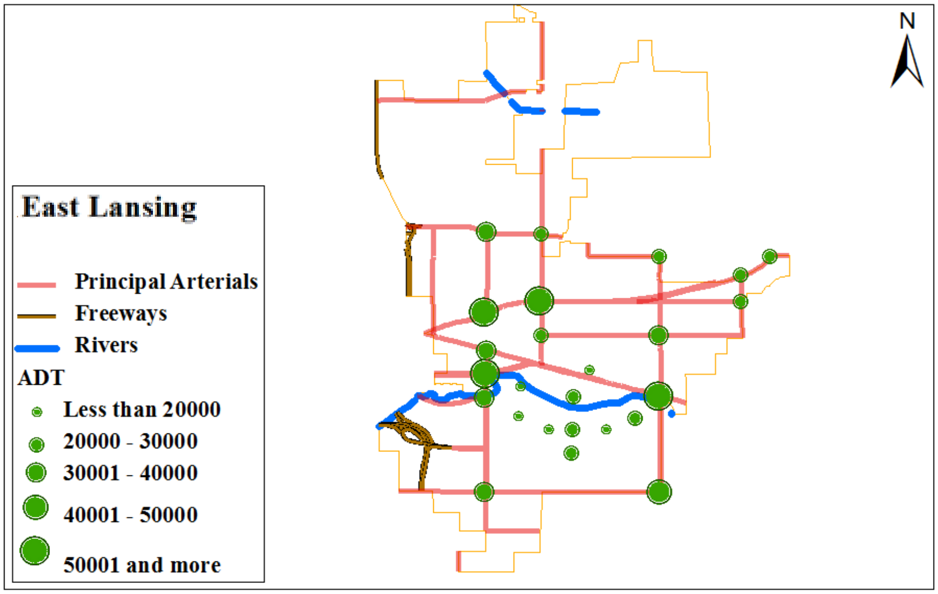

Figure 28 : ADT of Intersections at the Scope of the Study, East Lansing

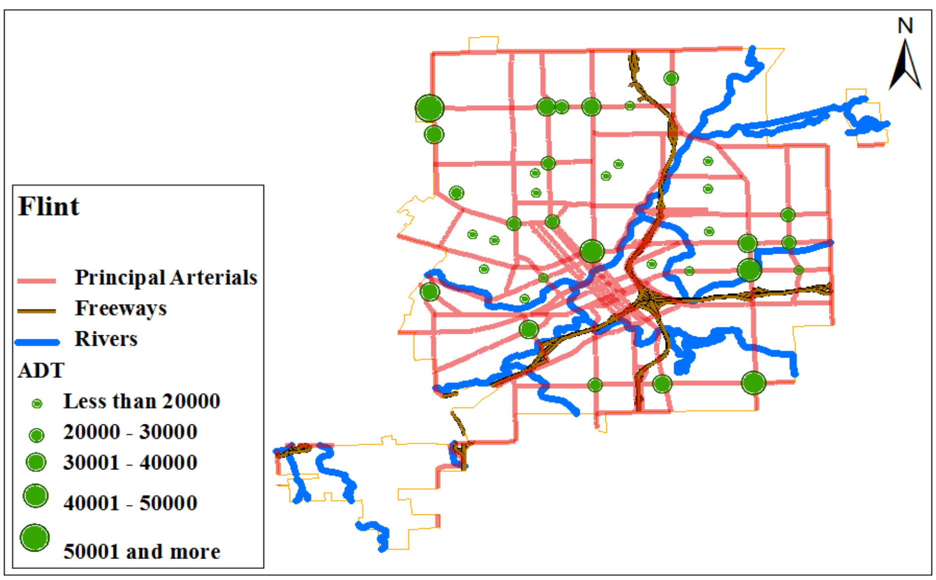

Figure 29 : ADT of Intersections at the Scope of the Study, Flint

Figure 28 and Figure 29 demonstrate downtown area has less ADT, although characteristics of the Flint's downtown are different from the downtown of East Lansing.

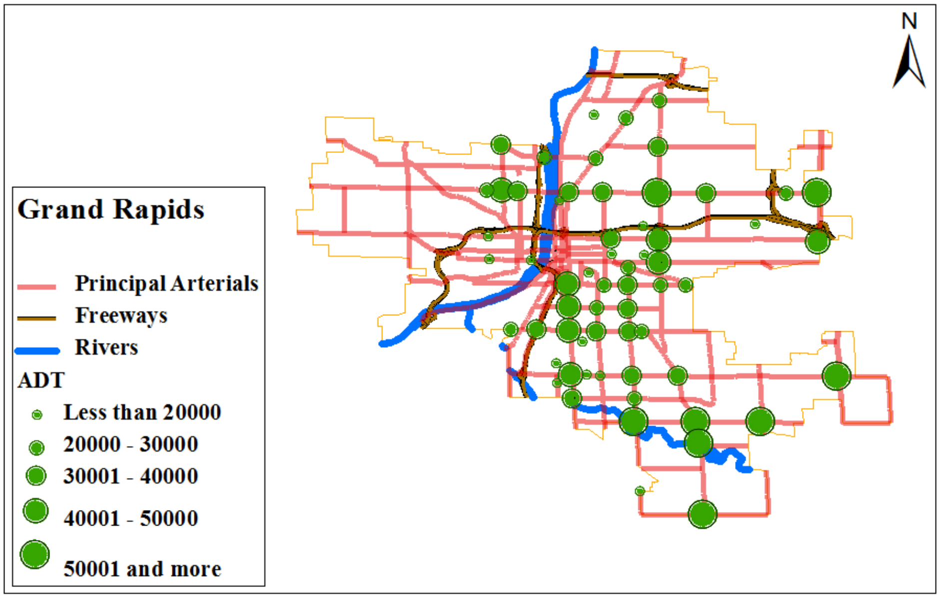

Figure 30 : ADT of Intersections at the Scope of the Study, Grand Rapids

Figure 30 reveals that similar to Ann Arbor, East Lansing and Flint ADT of downtown is less than the other parts of the cities. Figure 31 represents the average ADT of intersections in the scope of the study. Flint has the least amount of ADT on the intersections while East Lansing has the most.

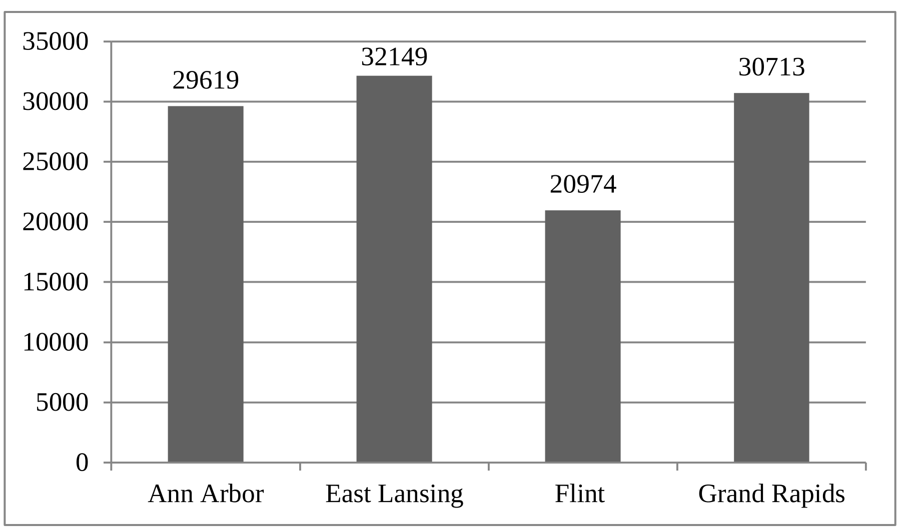

Figure 31 : Average ADT of Intersection in the Scope of the Study

3.3 Geometric Data

The design and physical characteristics of the intersections and corridors are considered as the geometric data. These data are investigated by visual observations, and Google Earth Pro is widely used for the data collection.

Table 5 shows the geometric data of intersections which is collected by observation. Then this data was applied in ARCGIS software as attributes of the intersections.

Table 5 : Geometric Data of Intersections

Geometric Characteristics
Number of Left Lanes
Number of Through Lanes
Number of Right Lanes
Total Number of Lanes
Presence of Bike Lane
Presence of Median
Width of Corridors
Length of Unpainted Crossing
Number of Access
Presence of On-Street Parking
Presence of Speed Sign
Posted Speed
Presence of Bus Stop

Figure 32 and Figure 33 present an example of on-street parking and speed sign in the scope of this study.

Figure 32 : An Example of On-Street Parking

Figure 33: An Example of Speed Sign

Collecting the geometric characteristics of the intersections was the most time consuming part of the study. It was required to observe each intersection and collect the data manually.

4. INFLUENTIAL AREA OF AN INTERSECTION

In order to define the influential area of an intersection, the distribution of the crash distance to the center of the intersection is investigated in this chapter. So far, most practitioners and agencies used their subjective idea based on the pattern of crash distribution, and none of them used a statistical approach to define the influential area. A 250-foot radius is the most common influential area, widely used for safety studies on intersections e.g. AASHTO (2010) and Nordback et al. (2014).

There is a rich literature of safety studies on intersection-related crashes. Since the nature of non-motorized transportation is different from motorized transportation, researchers studied them separately (Lee & Abdel, 2005; Harwood et al., 2008).

As illustrated in the previous chapter, the distribution of the crash distance may consist of two gamma distributions. In this chapter, parameters of the mixed distributions are investigated by employing the Expectation Maximization algorithm (EM algorithm). EM Algorithm is an iterative statistical method which searches for the maximum likelihood of distributions' estimates (Zhang et al., 2001). "mixtools" is a R software's package which employs EM algorithm to define parameters of the mixed distributions (Benaglia, 2009).

Gamma distribution is a continuous distribution which could be utilized to fit over Poisson distributed variables. This distribution is widely used in the Econometrics and the Bayesian statistics. Many researchers employed this distribution for traffic safety studies e.g. Bonneson & McCoy (1993) and Miaou (2003).

Probability density function and cumulative density function of the gamma distributions are presented below:

Probability density function:

$$g(x;\alpha,\beta) = \frac{x^{\alpha-1}\, e^{-x\beta}}{\beta^{-\alpha}\,\Gamma(\alpha)} \quad for\ (x \geq 0)\ \&\ (\alpha,\beta > 0) \tag{4}$$

Cumulative density function:

$$f(x;\alpha,\beta) = \int_0^x f(u;\alpha,\beta)du = \frac{\gamma(\alpha,\beta x)}{\Gamma(\alpha)} \tag{5}$$

Where:

α : Shape parameter

β : Scale parameter

$\Gamma(\alpha)$: Gamma function

$\gamma(\alpha,\beta x)$: Incomplete gamma function

In this study, the parameters of the mixed gamma distributions are investigated by utilizing the Expectation Maximization algorithm. This algorithm is a subset of the Majorize-Minimization algorithm. The EM algorithm consists of two iterative steps of Expectation and Maximization. The Expectation step evaluates the log-likelihood based on the current estimates of the parameters. The Maximization step maximizes the estimated log-likelihood of the Expectation step. The following equations explain the Expectation and the Maximization steps:

Expectation step:

$$Q\left(\theta|\theta^{(t)}\right) = E_{Z|X,\theta^{(t)}}[\log L(\theta;X,Z)] \tag{6}$$

Maximization step:

$$\theta^{(t+1)} = \arg max_{\theta}^{Q(\theta|\theta^{(t)})} \quad (7)$$

Where:

X : Observed data

Z : Latent data

θ : Vector of unkonwn variables

$L(\theta; X, Z)$: Likelihood function

For minimizing misclassification, Welch's formula is employed to define a border for specifying two different distributions of an intersection. Equation 7 represents Welch's formula for assigning a border.

$$\frac{f_1(x)}{f_2(x)} > \frac{\pi_2}{\pi_1} \quad (8)$$

Where

$f_i(x)$: Probability Density Function

π_i : Lambda Parameter of the mixed distribution

Newton's procedure is an iterative method which is employed to find an X satisfying Welch's formula. This procedure is illustrated in the following equations:

$$x_1 = x_0 - \frac{f(x_0)}{f'(x_0)} \quad (9)$$

$$x_{n+1} = x_n - \frac{f(x_n)}{f'(x_n)} \quad (10)$$

$$\Delta = (x_{n+1} - x_n) \tag{11}$$

Equation 9 iterates until $\lim_{n\to\infty} \Delta \sim 0$

In this chapter the influential area of an intersection in 4 levels is investigated:

1- For all kinds of Crashes together

2- Non-motorized Crashes

3- Pedestrian Crashes

4- Bike Crashes

Then for each kind of crash, this study performs an in-depth investigation to observe the effects of the corridor’s speed limit on the area of influence.

4.1 Influential area of an intersection for all kinds of crashes

By employing the Expectation Maximization algorithm, two different gamma distributions are specified. Parameters of the mixed distributions are presented in Table 6. In this table Lambda represents the portion of the data which belongs to each curve. Alpha and Beta represent the parameters of the gamma distribution for each curve.

Table 6 : Parameters of Mixed Gamma Distributions for all Kinds of Crashes

	Curve1	Curve2
Lambda	0.7947	0.2053
Alpha	0.886876	16.39488
Beta	89.5416	20.4736

Probability density curves of the mixed gamma distribution are illustrated in Figure 34. Near intersections, probability density is high. It decreases up to 258 feet, and then again it goes up, revealing a separate probability density representing midblock related crashes. It could be inferred that 250 feet is approximately a border between intersection-related crashes and midblock crashes.

Applying specifications of the gamma distribution to Welch’s formula and solving it by employing Newton’s procedure resulted in “X = 239.5513.” This result approves the previously assumed 250-foot buffer distance around the intersection as the intersection’s influential area, showing it to be a fairly good assumption.

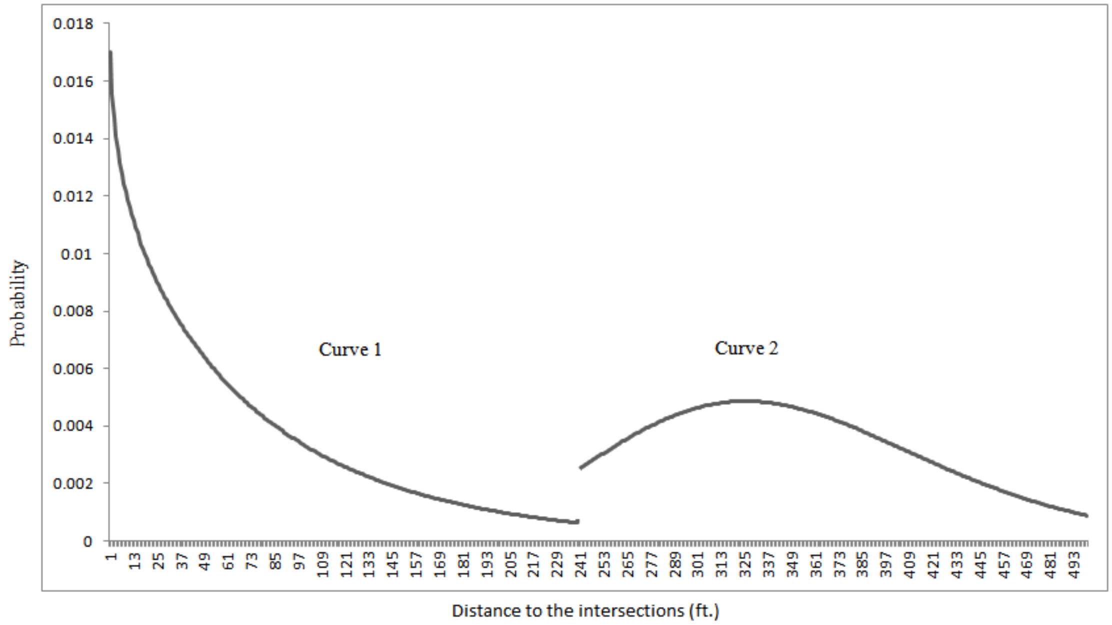

Figure 34 : Probability Density Curve of Distributions, All Kinds of Crashes

In further investigation, this study analyzes the effects of the corridors' posted speed on the influential area of an intersection for motorized crashes. This study categorized the posted speed in four groups by employing the Expectation Maximization algorithm, estimating the parameters of the mixed gamma distribution. Welch's formula is employed to find the distinguishing border for minimizing the misclassification which represents the influential area of an intersection. Table 7 demonstrates the different areas of influence for each category of the posted speed; it seems that the larger area of influence is associated with the higher posted speed.

Table 7 : Effect of Posted Speed on the Influential Area, Motorized Crashes

	Posted Speed (PS)			
	PS≤25	25<PS≤30	30<PS≤35	PS>35
Influential Area	233.0552	226.4812	217.4622	275.638

For investigating the effects of the corridor’s posted speed on the influential area of an intersection, the influential area on four levels of speed limits (25 mph, 30 mph, 35 mph, and more than 35 mph) is analyzed.

As demonstrated in the above mentioned table, it could be inferred that in the higher speed level (Speed Limit>35), the size of the influential area is more than in the lower speed levels. This means that if corridors have higher speed limits, vehicles are influenced from a longer distance.

4.2 Influential area of an intersection for non-motorized crashes

In this part, the influential area of an intersection for non-motorized transportation is investigated. The Expectation Maximization algorithm is applied for identifying the parameters of the mixed distributions, and the obtained results are presented in Table 8.

Table 8 : Parameters of Mixed Gamma Distributions for Non-motorized Crashes

	Curve1	Curve2
Lambda	0.683661	0.316339
Alpha	1.072581	8.721352
Beta	33.543145	33.218314

Probability density curves of the mixed gamma distributions for the non-motorized crashes are illustrated in Figure 35. Near intersections, probability density is high. It decreases up to 150 feet, and then again it goes up, showing a separate probability density that represents midblock-related crashes. It could be inferred that 150 feet is approximately a border between the intersection-related crashes and the midblock crashes.

Applying specifications of the gamma distributions to Welch's formula, and solving it by employing Newton's procedure resulted in "X = 137.2411."

This result shows that the influential area of an intersection for the non-motorized crashes is significantly different from the motorized crashes. In Chapter 5 the proposed influential area of intersections is considered for developing the SPFs for non-motorized crashes.

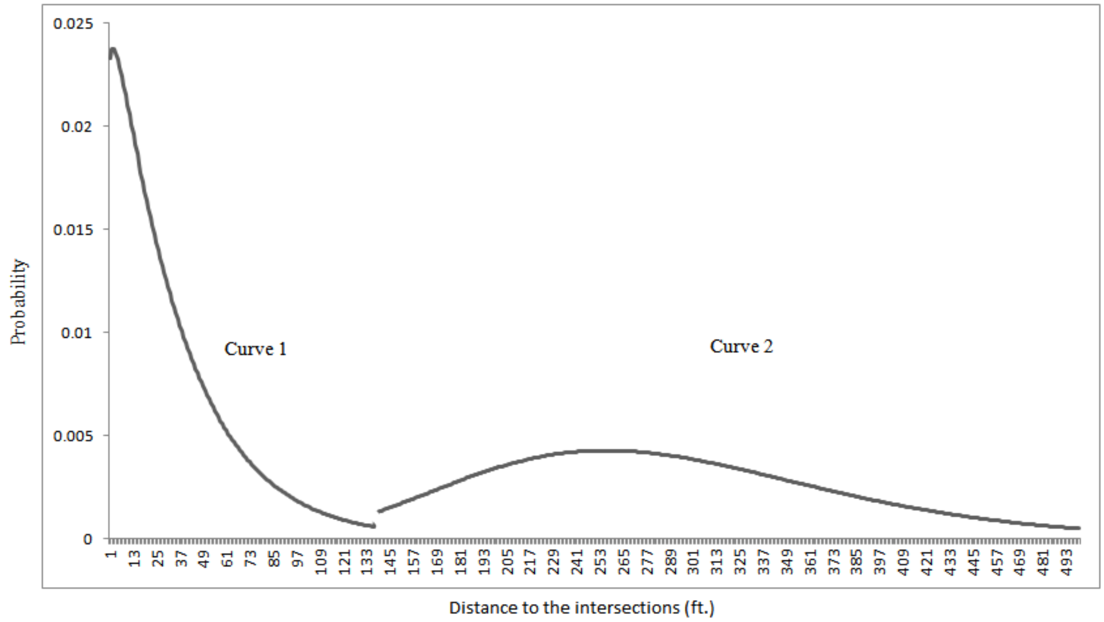

Figure 35 : Probability Density Curve of Distributions, Non-motorized Crashes

Since nature of the non-motorized crashes is different from the motorized crashes in this part of the study, the effects of corridors' posted speed on the influential area of an intersection and the influential area on four levels of speed limits (25 mph, 30 mph, 35 mph, and more than 35 mph) is analyzed. The next table demonstrates the parameters of the mixed gamma distribution in different speed levels for the non-motorized crashes.

Table 9 : Effect of Posted Speed on the Influential Area, Non-motorized Crashes

	Posted Speed (PS)			
	PS≤25	25<PS≤30	30<PS≤35	PS>35
Influential Area	148.5217	142.3603	75.78586	83.09108

The observed results reveal that corridors with higher speed limits are associated with a smaller influential area for non-motorized crashes. A non-linear but negative relation between the corridors' speed limit and the influential area of the intersections is observed. On the other hand, non-motorized transportation may avoid some dangerous movements, e.g. jaywalking, and they may just choose the intersection's crossing to cross a corridor. Since higher speed levels are associated with more sever crashes (Abdel-Aty, 2003), high speed vehicles may bring a feeling danger for the non-motorized transportation, causing non-motorized transportation to avoid conflict by using the intersection's crossings (Raford, 2005).

4.3 Influential area of an intersection for pedestrian crashes

In this section the influential area of an intersection for the pedestrian crashes is investigated. In the previous section, the influential area of an intersection was investigated, and the results demonstrated that the influential area for the non-motorized crashes is different from the motorized crashes. Cyclists and pedestrians are two groups of the non-motorized transportation. In the next two sections of this study, the influential area of an intersection for both types of the non-motorized transportation is investigated, and then it is compared with the motorized and the non-motorized transportation.

The Expectation Maximization algorithm is applied for identifying the parameters of the mixed distributions, and the results are presented in Table 10.

Table 10 : Parameters of Mixed Gamma Distributions for Pedestrian Crashes

	Curve1	Curve2
Lambda	0.6782865	0.3217135
Alpha	1.181826	9.025873
Beta	31.885086	31.562301

Applying parameters of the gamma distributions to Welch’s formula and solving it by employing Newton’s procedure resulted in “X = 137.0461.”

This result reveals that influential area of an intersection for pedestrian and non-motorized crashes is almost same.

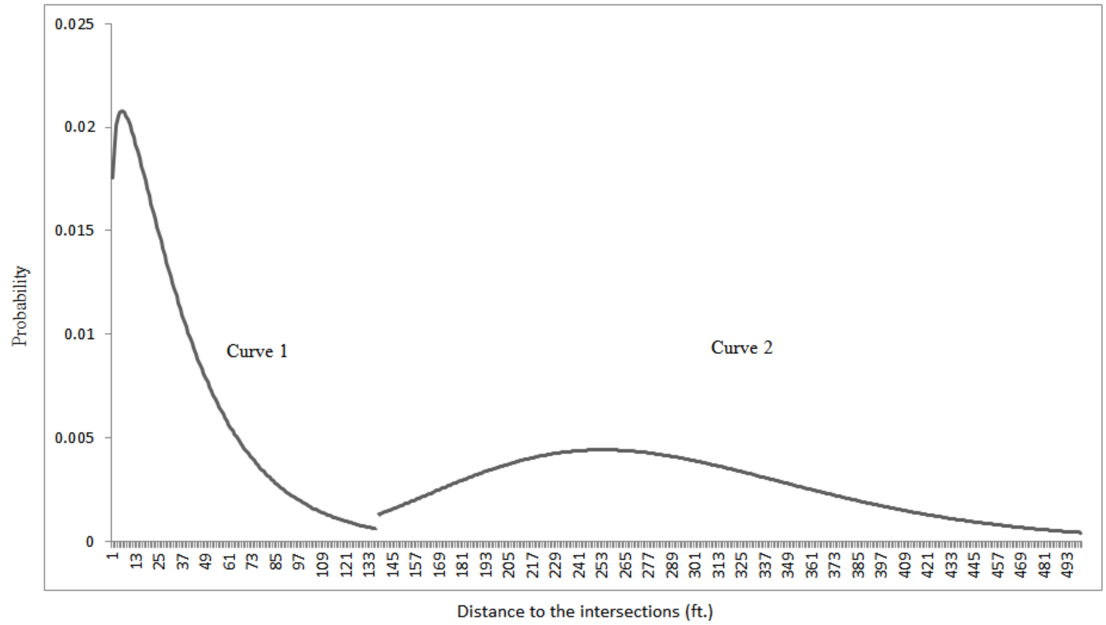

Figure 36 : Probability Density Curve of Distributions, Pedestrian Crashes

For investigating the effects of the corridor's posted speed on the influential area of an intersection, the area of influence on the four levels of speed limits (25 mph, 30 mph, 35 mph, and more than 35 mph) is analyzed. The next table demonstrates parameters of the mixed gamma distribution in different speed levels for the pedestrian crashes.

Table 11: Effect of Posted Speed on the Influential Area, Pedestrian Crashes

	Posted Speed (PS)			
	PS≤25	25<PS≤30	30<PS≤35	PS>35
Influential Area	159.6794	125.7069	99.20893	67.3896

Similar to the non-motorized case, the results of the Welch's formula reveal that corridors with higher speed limits are associated with a smaller influential area for pedestrian crashes. A negative relation between the corridors' speed limits and the influential area of intersections is observed. On the other hand, pedestrians may avoid some dangerous movements, e.g. jaywalking, and they may just choose the intersection's crossing to cross a corridor. Since higher speed levels are associated with more sever crashes (Abdel-Aty, 2003), high speed vehicles may bring a feeling danger for the non-motorized transportation, causing non-motorized transportation to avoid conflict by using the intersection's crossings (Raford, 2005).

4.4 Influential area of an intersection for bike crashes

This part of the study is dedicated to investigate influential area of an intersection for bike crashes and to compare it with the motorized, the non-motorized, and the pedestrian cases.

The Expectation Maximization Algorithm is applied for identifying the parameters of the mixed distributions, and results are presented in Table 12.

Table 12 : Parameters of Mixed Gamma Distributions for Bike Crashes

	Curve1	Curve2
Lambda	0.685203	0.314797
Alpha	1.016787	8.372507
Beta	33.805433	34.90417

Applying the parameters of the gamma distributions to Welch's formula and solving it by employing the Newton's procedure resulted in "X = 137.0461," which is exactly equal to the influential area of an intersection for the pedestrian.

This result reveals that the influential area of an intersection for pedestrians, bikes, and non-motorized crashes are almost same, and they are different from the motorized crashes. In order to develop the Safety Performance Functions, the proposed buffer size is utilized.

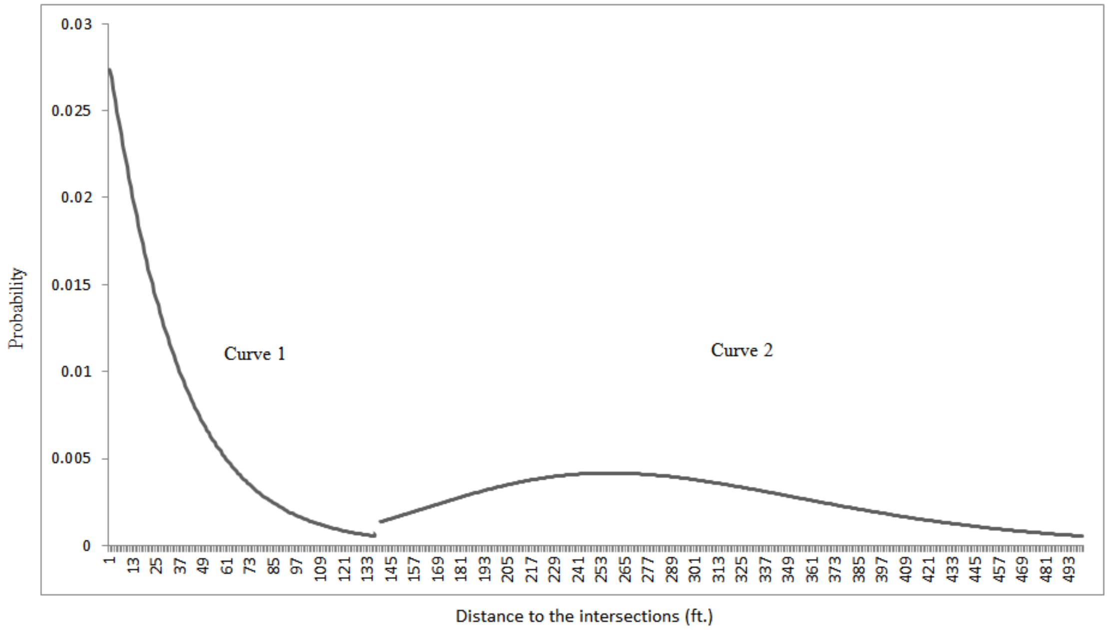

Figure 37: Probability Density Curve of Distributions, Bike Crashes

For investigating the effects of the corridor's posted speed on the influential area of an intersection, the influential area on two levels of speed limits (30 mph and more than 35 mph) is analyzed. The following table demonstrates parameters of the mixed gamma distribution in different speed levels for bike crashes

Table 13: Effect of Posted Speed on the Influential Area, Bike Crashes

	Posted Speed (PS)	
	PS≤30	PS>30
Influential Area	136.1297 (PS≤30)	72.6936 (PS>30)

Same as the non-motorized and the pedestrian cases, results of Welch's formula reveal that corridors with higher speed limits are associated with a smaller influential area for the bike crashes. A negative relation between corridors' speed limits and influential area of the intersections is observed. On the other hand, bikes may avoid some dangerous movements, choosing to use the intersection's crossing to cross a corridor. Since higher speed levels are associated with more sever crashes (Abdel-Aty, 2003), high speed vehicles may bring a feeling danger for the non-motorized transportation, causing non-motorized transportation to avoid conflict by using the intersection's crossings (Raford, 2005).

5. SAFETY PERFORMANCE FUNCTION

The American Association of State Highway and Transportation Officials (AASHTO) introduced Safety Performance Functions (SPFs) in the Highway Safety Manual (HSM) as crash prediction models for estimating crash frequency. Although, in recent years, many agencies have been promoting non-motorized transportation (cycling and walking) to promote health, decrease air pollution concerns, and stimulate a more sustainable transportation environment, there is unfortunately no SPF for bike crashes in the first edition of HSM. Also, the introduced pedestrian SPF in the HSM has ADT of the major and minor roads, pedestrian volume, and the maximum number of lanes as the significant factors (AASHTO, 2010). Other factors are not considered, and there is a lack of understanding about other factors for which agencies could create countermeasures for improving the safety of pedestrians.

Poisson Regression Model could be utilized to predict discrete dependent variables from significant parameters. However, due to the heterogeneous nature of data and possible errors in providing exposures data (ADT, pedestrian volume, and bike volume), their variance is very different from the mean, causing an over-dispersion to be inferred (AASHTO, 2010). In the literature of the safety studies, over-dispersion (i.e. variance is significantly beyond the mean) in crash data has been observed, repeatedly e.g. Chang (2005). The over-dispersion proves that exposures are not Poisson-distributed, and then the Poisson model is not recommended for regression analysis (Lord et al., 2010).

Many researchers utilized Negative Binomial Model to deal with over-dispersion in safety studies e.g. Hadi (1995).

Negative Binomial model is derived from the Poisson model. In Poisson model, probability of i crashes at a specific period of time is:

$$P(y_i) = \frac{{\lambda_i}^{y_i} \exp(\lambda_i)}{y_i!} \tag{12}$$

Where:

$P(y_i)$: probability of y crash at the intersection *i*

λ_i: Expected crashes frequency at a period

B: A vector for estimable parameters (coefficients)

It is presumed that Poisson regression models are fitted to the data by λ_i as a function of explanatory variables that

$$\lambda_i = \exp(\beta X_i) \tag{13}$$

Where:

X_i: A vector of explanatory variables

β: A vector of estimable coefficients

By employing standard maximum methods, β could be estimated with the likelihood function.

$$L(\beta) = \prod_i \frac{\exp[-exp(\beta X_i)]\,[\exp(\beta X_i)]^{y_i}}{y_i\,!} \qquad (14)$$

To deal with the over dispersion, an independent error parameter could be added to avoid the model's error. By applying the error term, the Negative Binomial regression model is derived as

$$\lambda_i = \exp(\beta X_i + \varepsilon_i) \qquad (15)$$

Where:

$\exp(\varepsilon_i)$: A gamma distributed error term (mean=1, variance=α)

By applying the error term, probability of y crash at the intersection *i* in the period would be changed as

$$P(y_i/\varepsilon) = \frac{\exp[-\lambda_i\, exp(\varepsilon_i)]\,[\lambda_i\, exp(\varepsilon_i)]^{y_i}}{y_i!} \qquad (16)$$

Integration of the error term in equation 6 lead to the unconditional distribution of crash frequency as

$$P(y_i) = \frac{\Gamma(\theta + \lambda_i)}{[\Gamma(\theta)\, \lambda_{i!}]}\, u_i^{\theta}(1 - u_i)^{y_i} \qquad (17)$$

Where:

u_i: $\theta/(\theta+y_i)$ and $\theta = 1/\alpha$

By maximizing the below mentioned likelihood function, coefficient estimates (α,β) would be achieved

$$L(\lambda_i) = \prod_{i=1}^{N} \frac{\Gamma(\theta + y_i)}{\Gamma(\theta) y_i!} \; [\frac{\theta}{\theta + \lambda_i}]^{\theta} \; [\frac{\lambda_i}{\theta + \lambda_i}]^{y_i} \quad (18)$$

Where:

N: Total number of samples

The below mentioned formulation let variance to differ from mean as

$$Var[y_i] \; = E([\;\; _i])[1 + \alpha E(y_i)] = E[y_i] + \; \alpha E(y_i)^2 \quad (19)$$

Where:

α: Variance of error and (measure of dispersion)

In this study for handling the over-dispersion of exposures' data a Negative Binomial Regression Model is used for pedestrian and bike crash frequency to develop the pedestrian and the bike SPF from the significant variables.

In the previous section, it is observed that a 137-foot radius from the center of an intersection is the influential area of an intersection for non-motorized (pedestrian and bike) crashes. This distance is rounded up to 150, and it is adopted for investigating significant factors that affect crash frequency to develop safety performance functions (SPFs).

5.1. Pedestrian safety performance function

Since in this study the pedestrian crash is a discrete variable, a Poisson Model could be employed to make a statistical fit for displaying a relationship between significant variables as the independent variables and crash frequency of the intersection as the dependent variable. However, due to heterogeneous data and some possible errors from extrapolation of ADT and estimation of pedestrian volume, the variance of some significant variables is different from their means, causing an over dispersion in the dataset (AASHTO, 2010). Table 14, shows the statistical parameters of the significant variables which were employed for developing the SPF.

Table 14 : Significant Variables for Pedestrian SPF

Definition	Min	Max	Median	Mean	Variance
ADT	3,438	73,374	27,509	28,284	192,629,905
Pedestrian Volume	29	29,365	308	1,083	8,253,070
Number of Left Lanes	0	6	2	2.34	2.72
Presence of Street Parking	0	1	0	0.45	0.25
Presence of Speed Sign	0	1	1	0.54	0.25
Presence of Bus Stop in 0.1 Mile Buffer around Intersections	0	1	1	0.79	0.17

On-street parking and the presence of speed signs are binary (dummy) variables in this study and change from 0 to 1. If there is an on-street parking or a speed sign these variables adopt 1. Segments of road from the signalized intersection to nearest intersection were investigated for the presence of speed signs. Presence of bus stops in the intersection is a dummy variable for existence of bus stops within a 0.1 mile buffer around the intersections.

Significant variables and their parameters for finding the SPF based on the proposed Negative Binomial Model are estimated by Stata 12. A t-test was used at a 95% confidence level of significance for the independent variables. As it discussed in the previous section, the total ADT of the intersections, pedestrian exposure (volume), total number of left turn lanes, the presence of on-street parking, the presence of speed signs, and the presence of the bus stops within 0.1 mile buffer around an intersection were found to be significant.

By using the mentioned variables for developing the Safety Performance Function, a SPF is developed which is demonstrated in the Table 15.

Table 15 : Proposed Safety Performance Function

Significant Variables.	Coefficient	Std. Err.	Z	P>|z|
ADT	0. 0000251	8. 22e-06	3. 05	0. 002
Bike Volume (Exposure)	0. 000091	0.0000344	2. 65	0. 008
Number of Left Turn Lanes	0. 2296894	0. 0821085	2. 80	0. 005
Presence of On-Street Parking	0. 5712769	0. 2418822	2. 36	0. 018
Presence of Speed Signs	-0. 4470537	0. 2291633	-1. 95	0. 051
Presence of Bus Stop	0. 9400843	0. 4297579	2. 19	0. 029
Constant	-2. 660768	0. 5042895	-5. 28	0. 000
Number of Observations = 164 LR Chi^2 (6) = 57.63 Prob >Chi^2 = 0.0000 Log likelihood = -178.8994 Pseudo R^2 = 0.1387				

The mathematical form of the proposed SPF is:

$$\text{Pedestrian Crash No @ Intersection} = \qquad (20)$$

$$exp\,(0.\ 0000251 * \text{ADT} + 0.\ 000091 * \text{Ped. Volume} + 0.\ 2296894 * \text{LL}_i + 0.\ 5712769 * SP_i - 0.\ 4470537 * \text{SS}_i + 0.9400843 * BS_i - 2.660768)$$

Where

$\text{Pedestrian Crash No @ Intersection}$: Number of pedestrian crashes at the intersection

ADT: Average daily traffic approaching the intersection

Ped. Volume: Number of pedestrians crossing the intersection

$\mathrm{LL_i}$: Total number of left turn lanes at the intersection

SP_{i}: Presence of the on-street parking on each corridor of an intersection (0: no On-street Parking, 1: if just 1 corridor of an intersection has on-street parking)

SS_{i} : Presence of speed signs on each corridor of an intersection (0: no Speed Sign, 1: if just 1 corridor of an intersection has Speed Sign)

BS_i : Presence of bus stops on each corridor of an intersection (0: no Bus Stop, 1: if just 1 corridor of an intersection has Bus Stop)

As it was stipulated in the literature review ADT, pedestrian volume, number of left turn lanes, presence of on-street parking, and bus stop were associated with a higher pedestrian crash frequency. Also, this study reveals that the presence of speed signs has a significant effect on pedestrian crash frequency although some previous literatures documented the role of a good traffic sign design in decreasing crash frequency.

The higher ADT, pedestrian volume, number of left turn lanes, and presence of on-street parking and bus stop lead to a greater crash frequency at intersections, while the presence of speed sign on the corridors of an intersection decreases the crash frequency.

Another advantage of this new SPF could be found in finding safety counter measures. Since this model demonstrates a positive correlation between crash frequency and the on-street parking, agencies could improve safety by removing on-street parking adjacent of the hazardous intersections. Additionally, by providing more speed signs a decrease in pedestrian crashes could be expected.

It is found that the presence of bus stops within a 0.1 mile buffer around intersections lead to more crashes. Therefore, by removing bus stops in that vicinity, a lesser number of pedestrian crashes would be expected.

5.2. Bike safety performance function

As in the pedestrian case, the Poisson Regression Model could be utilized to predict discrete dependent variables from significant parameters. However, due to the heterogeneous nature of data and possible errors in providing exposures data (ADT and bike volume), their variance is very different from the mean, causing an over-dispersion to be inferred (AASHTO, 2010). Therefore, the Negative Binomial Model could be employed to develop a bike SPF. Table 16 illustrates general statistics of the significant parameters which are used for developing the Bike SPF. As mentioned in the literature, an over-dispersion is observed in the ADT and bike volume. So, in order to develop Safety Performance Function for bikes, the Negative Binomial Regression model is utilized.

Table 16 : Significant Variables for Bike SPF

Definition	Min	Max	Median	Mean	Variance
ADT	3,438	73,374	27,509	28,284	192,629,905
Bike Volume	5	3,282	164	331	292,321
Number of Left Turn Lanes	0	6	2	2.34	2.72
Presence of Bike Lanes	0	1	0	0.37	0.24
Presence of Bus Stop in 0.1 Mile Buffer around Intersections	0	1	1	0.79	0.17
Presence of Intersection in Ann Arbor	0	1	0	0.20	0.16
Presence of Intersection in East Lansing	0	1	0	0.15	0.13
Presence of Intersection in Grand Rapids	0	1	0	0.40	0.24

Presence of bike lane is a binary (dummy variable) in this study and changes from 0 to 1. If any bike lane is observed in the intersection this variable would adopt 1. Presence of bus stop in the intersection is a dummy variable for the existence of the bus stops within a 0.1 mile buffer around the intersections.

The significant variables and their parameters for finding the SPF based on the proposed Negative Binomial model are estimated by Stata 12. A t-test was used at 90% confidence level of significance for the independent variables. The total ADT of the intersections, the city characteristics, the bike exposure (volume), total number of left turn lanes, and presence of bus stops within a 0.1 mile buffer around the intersection were found to as the significant variables.

By using the mentioned variables for developing the Safety Performance Function, a SPF is developed which is demonstrated in Table 17.

Table 17 : Significant Variables for Bike SPF

Significant Variables.	Coefficient	Std. Err.	Z	P>|z|
ADT	0. 0000317	7. 34e-06	4. 32	0. 000
Pedestrian Volume (Exposure)	0. 000531	0. 0001705	3. 12	0. 002
Number of Left Turn Lanes	0. 13213	0. 0693677	1. 90	0. 057
Presence of Bike Lanes	0. 4130955	0. 2083965	1. 98	0. 047
Presence of Bus Stop	1. 004044	0. 4155165	2. 42	0. 016
Ann Arbor	1. 409839	0. 5573415	2. 53	0. 011
East Lansing	1. 888949	0. 572683	3. 30	0. 001
Grand Rapids	1. 558988	0. 5308453	2. 94	0. 003
Constant	-4. 336118	0. 6451296	-6. 72	0. 000
Number of Observations = 164 LR Chi^2 (6) = 97.30 Prob >Chi^2 = 0.0000 Log likelihood = -162.91865 Pseudo R^2 = 0.2299				

The mathematical form of the proposed SPF is:

$$N_{Bike\ Crash\ No\ @\ Intersection} = \tag{21}$$

$$exp\ (0.0000317\ * ADT_i +\ 0.000531\ * Bicycle\ Volume_i +\ 0.13213\ * TL_i + 0.4130955\ *\ BL_i -\ 1.004044 *\ BS_i +\ 1.409839\ *\ AA_i\ +\ 1.888949\ *\ EL_i\ + 1.558988 *\ GR_i - 4.336118)$$

Where

ADT: Average daily traffic approaching the intersection

Bike Volume: Number of cyclists crossing the intersection

TL_i: Total number of left lanes at the intersection

BL_i: Total presence of bike lanes on each corridor of an intersection (0: no bike lane, 1: if just 1 corridor of an intersection has a bike lane)

BS_i: Presence of a bus stop on each corridor of an intersection (0: no Bus Stop, 1: if just 1 corridor of an intersection has a bus stop)

AA_i : A dummy variable for presence of an intersection in Ann Arbor (0: the intersection is located in Ann Arbor, 1: the intersection is not located in Ann Arbor)

EL_i : A dummy variable for presence of an intersection in East Lansing (0: the intersection is located in East Lansing, 1: the intersection is not located in East Lansing)

GR_i : A dummy variable for presence of an intersection in Grand Rapids (0: the intersection is located in East Lansing, 1: the intersection is not located in Grand Rapids)

Table 18: Bike Safety Performance Function (Without City Variable)

Significant Variables.	Coefficient	Std. Err.	Z	P>|z|
ADT	0. 0000419	7. 87e-06	5. 32	0. 000
Pedestrian Volume (Exposure)	0. 0008022	0. 0001504	5. 33	0. 000
Number of Left Turn Lanes	0. 1566364	0. 0730476	2. 14	0. 032
Presence of Bike Lanes	0. 5408297	0. 2150191	2. 52	0. 012
Presence of Bus Stop	0. 9032806	0. 4139433	2. 18	0. 029
Constant	-3. 377522	0. 4935776	-6. 84	0. 000
Number of Observations = 164 LR Chi^2 (6) = 81.72 Prob >Chi^2 = 0.0000 Log likelihood = -170.70794 Pseudo R^2 = 0.1931				

Higher ADT and bike volume cause more conflicts between cyclists and vehicle drivers which lead to more crashes. Total number of left turn lanes and the presence of bus stops within a 0.1 mile buffer around the intersections is associated with the higher frequency of bike crashes. Also, results reveal that the presence of bike lanes at the intersection is significantly related to more crashes, but it does not mean that bike lanes are problematic because they may bring more cyclists to the intersection, and as a result, the number of bike crashes would increase.

Findings reveal the effect of a city’s characteristics on the bike crash frequency, meaning that for each city a coefficient should be applied, and then the proposed SPF would be as below:

Then the mathematical form of the proposed SPF is:

$$N_{Bike\ Crash\ No\ @\ Intersection} = \tag{22}$$

$$[exp\ (0.0000419\ * ADT_i + 0.0008022 * Bicycle\ Volume_i + 0.1566364\ * TL_i + 0.5408297\ * BL_i - 0.9032806 * BS_i - 3.377522)] * C_i$$

Where

C_i : A coefficient for the city characteristics of the intersection.

Although bike lanes are considered one type of safety countermeasures, the results of the bike SPF demonstrated a positive relation among bike lanes and a higher bike crash frequency. So, for a deeper investigation on the bike lanes' role and other complex interrelations between the factors that cause bike crashes, a Structural Equation Model is employed for further investigation. Although, the previously approved 137-foot buffer around intersections is the influential area for non-motorized crashes, it does not mean that the previously considered 250-foot buffer may cause wrong results. It is expected that 137 feet lead to less misclassifications, decreasing error. A more complex model (SEM) is developed to handle the interrelationships among exposure, geometry, city and crash variables simultaneously by employing Stata 12.

A SEM usually consists of 3 major components: 1) exogenous (observed) variables 2) endogenous (latent) variables, and 3) sub-models which are connected together by virtual links.

This model is build used to estimates the regression based on the proposed relations between observed and latent variables. Although subsets of a latent variable (or an observed variable) may encompass an error but the latent variable (or the observed variable) doesn't. An example of a SEM is presented below.

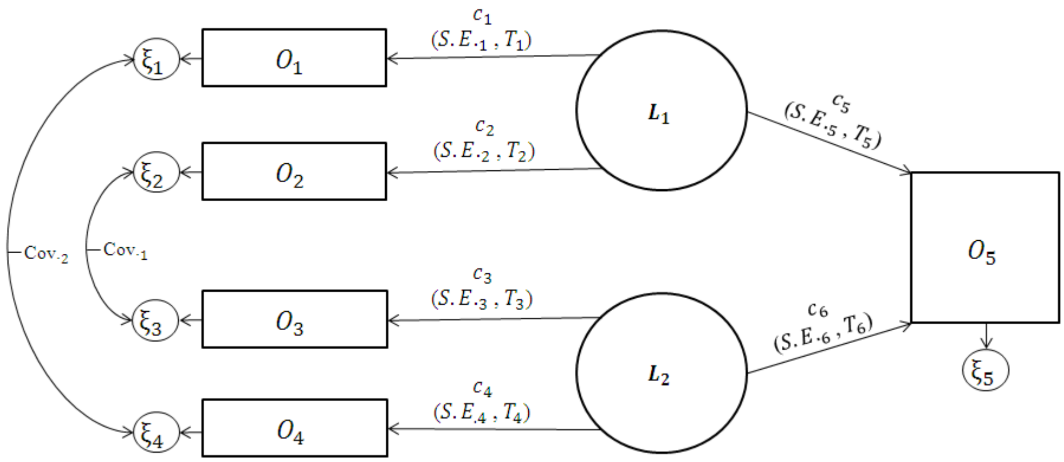

Figure 38 : A Sample Diagram of SEM

Where:

O_i : Observed Variable

L_i : Latent Variable

C_i : Coefficient

$S.E._i$: Standard Error

T_i : T-Test Result

ξ_i : Error

Cov.$_i$: Covariance

Comparative fit index (CFI) and Tucker-Lewis index (TLI) are other two measures for evaluating goodness of fit which are considered in this study. Equations (23) and (24) depict the general forms of CFI and TLI, respectively (Widaman & Thompson, 2003). CFI and TLI are between 0 and 1. A cutoff value close to 0.9 demonstrates an acceptable fit for CFI (Hu & Bentler, 1999). Some practitioners considered 0.9 for TLI as a cutoff value for the goodness of fit e.g. Mennaï (2011).

$$\mathrm{CFI} = 1 - \frac{\max[(\chi_k^2 - df_k),0]}{\max[(\chi_k^2 - df_k),(\chi_0^2 - df_0),0]} = 1 - \frac{\max\left[\left\{F_k - \left(\frac{df_k}{N-1}\right)\right\},0\right]}{\max\left[\left\{F_k - \left(\frac{df_k}{N-1}\right),\left(F_0 - \left(\frac{df_0}{N-1}\right)\right),0\right\}\right]} \tag{23}$$

$$\mathrm{TLI} = \frac{\left(\frac{\chi_0^2}{df_0}\right) - \left(\frac{\chi_k^2}{df_k}\right)}{\left(\frac{\chi_0^2}{df_0}\right) - 1} = \frac{\left(\frac{F_0}{df_0}\right) - \left(\frac{F_k}{df_k}\right)}{\left(\frac{F_0}{df_0}\right) - \left(\frac{1}{N-1}\right)} \tag{24}$$

Where:

χ_0^2: Chi-Square (null model)

df_0: Degrees of freedom (null model)

F_0: Minimum fit function value (null model)

χ_k^2: Chi-Square (substantive model of interest)

df_k: Degrees of freedom (substantive model of interest)

F_k: Minimum fit function value (substantive model of interest)

Several SEM were tested, but SEM at the 250-foot buffer size came up with accepted goodness of fit's measures. These results are illustrated in Figure 39, Table 19, and Table 20.

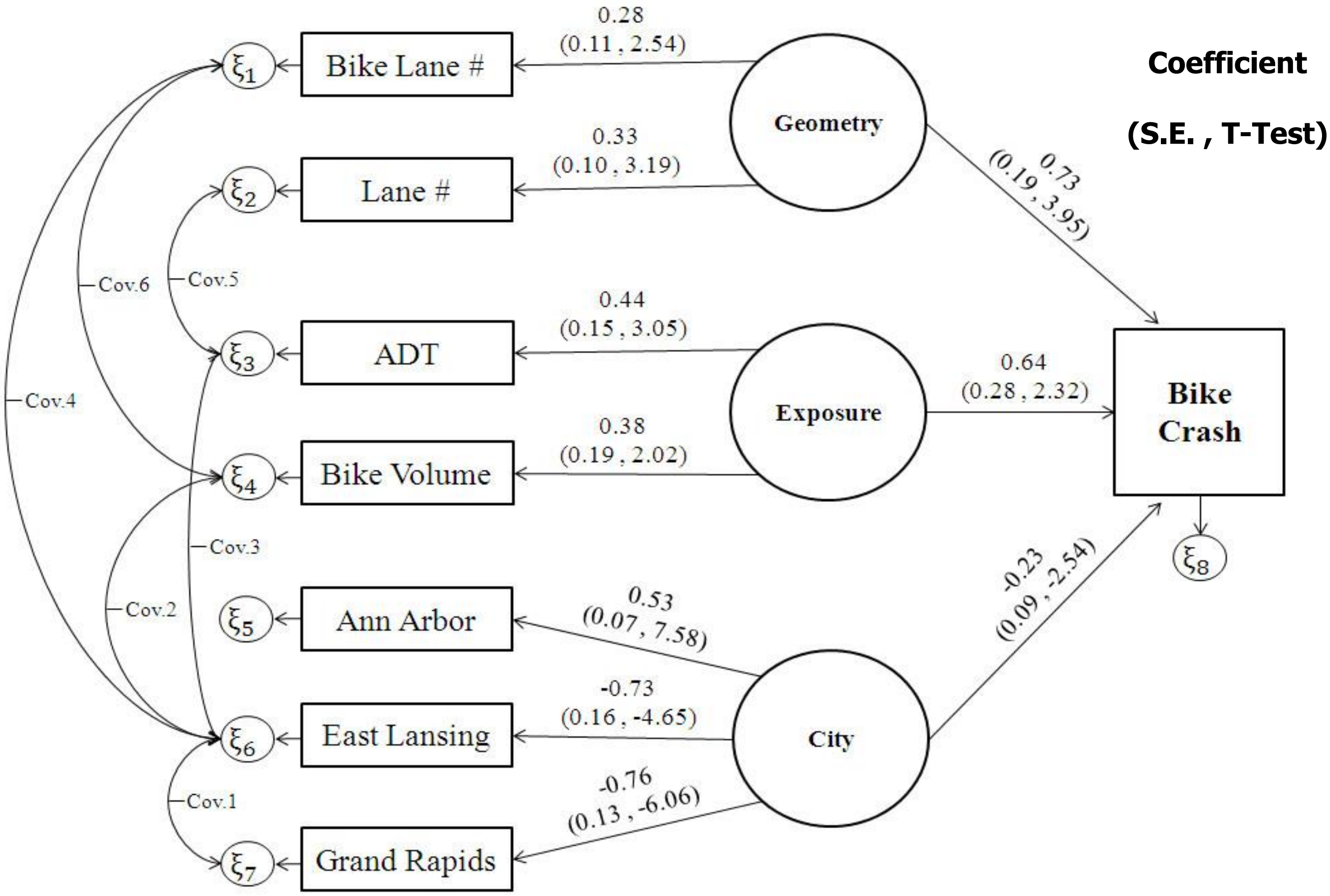

Figure 39 : Structural Equation Model, Bike Crash Frequency

ξ_i: Error term

Cov.i: Covariance Test Results

Results of the goodness of fit are illustrated in the Table 19, indicating an accepted goodness of fit for the SEM.

Table 19 : Goodness of Fit

Measure	Test Result
CFI	0.9
TLI	0.9

Since the standardized method was employed for the covariance test between variables, the result is used for conducting a correlation test among variables.

H0: There is no Correlation between variables

H1; There is Correlation

Table 20 shows the results of covariance test.

Table 20 : Covariance Test Results

	Covariance	Standard Error	Z – Test	P - value
Cov.1	-1.85	0.99	-1.87	0.062
Cov.2	0.57	0.16	3.52	0.00
Cov.3	0.34	0.10	3.50	0.00
Cov.4	0.38	0.11	3.35	0.00
Cov.5	0.69	0.10	6.83	8.42e-12
Cov.6	0.37	0.07	5.43	5.66e-8

Results of the covariance test reveal a significant correlation between the bike lane and the bike volume (Cov.6). Although bike lanes were found to be significant in the bike SPF, there is a significant correlation between the presence of bike lanes and bike exposure. Since the exposure causes the crashes, it could not be inferred that bike lanes endanger cyclists' safety.

Outcomes of the covariance test reveal a positive relation between higher levels of ADT and the number of lanes. Also, it could be inferred that bike lanes have more presence on roadways when there is a greater number of traffic lanes.

CONCLUSIONS

In this study, non-motorized safety performance functions for non-motorized transportation were developed. In order to understand the influential area of an intersection for non-motorized crashes, 144 intersections in four Michigan cities of Ann Arbor, East Lansing, Flint, and Grand Rapids were investigated. By achieving the influential area of intersections for motorized and non-motorized crashes, crash data on the influential area of 164 intersections were collected. ArcGIS 10.0 was employed to build a database from the collected data. This database contains crash data, exposure (ADT, pedestrian volume, and bike volume), and the geometric characteristics of intersections (such as lane configuration, road facilities, etc.). Using the database, safety performance functions for pedestrians and bikes were developed.

Distributions of crash distances to the center of intersections (all crashes, non-motorized crashes, pedestrian crashes, and bike crashes) consisted of two mixed gamma distributions representing intersection-related crashes and midblock crashes. A package of the R software ("mixtools") was employed to find parameters of the mixed distributions using the Expectation Maximization Algorithms. By utilizing Welch's formula (Welch, 1939) to minimize misclassification, the influential area of intersections was calculated. The result revealed that the influential areas of intersections for all crashes, non-motorized crashes, pedestrian crashes, and bicycle crashes were 239.55 ft, 137.24 ft, 137.04 ft, and 137.04 ft, respectively. In order to observe the effects of the posting speed on the influential area, corridors were classified in different groups based on the speed. Outcomes revealed that for all kinds of crashes, the influential area of intersections in higher speed corridors was larger than that in lower speed corridors. However, for non-motorized crashes, the influential area of intersections in lower levels of posted speed is larger. This result implies that pedestrians and cyclists in corridors with higher posted speeds tend to avoid dangerous movements (such as jaywalking, etc.), and they try to cross corridors from the intersections rather than the midblock part of a corridor. There may need further studies on other factors that affect the influential area.

Pedestrian and bike SPFs were developed based on crash data collected within the proposed influential area. The developed pedestrian SPF shows: 1) higher exposures (ADT and pedestrian volume) are associated with higher frequency of pedestrian crashes at intersections; 2) More total left turn lanes, presence of on-street parking, and presence of bus stops within a 0.1 mile distance to the center of the intersection as geometric characteristics of intersections are associated with higher frequency of pedestrian crashes; 3) Presence of speed signs as another geometric characteristic of intersections is significantly associated with lower pedestrian crash frequency at the intersection.

For the bike case, results of the safety study show: 1) higher exposure (ADT and bike volume) leads to more expectation of bike crashes; 2) presence of bike lanes, presence of bus stops within a 0.1 mile distance to the center of intersections, and an increase in the number of left lanes at the intersection (geometric characteristics) are associated with more bike crashes; 3) cities' specification significantly affects bike crash frequency. Therefore, the proposed bike SPF includes exposure, geometric specifications, and a city coefficient as its explanatory variables.

A Structural Equation Model (SEM) was developed for further investigation on the complex interrelationship among significant variables affecting bike crashes. The result reveals that bicycle lanes tend to increase bicycle-related crashes because of increased bicycle volume although bicycle lanes enhance bicycle safety.

REFERENCES

AASHTO. (2010). Highway Safety Manual. Washington, DC, 529.

Abasahl, F. (2013). Spatial Factors Impacting Non-Motorized Exposures and Crash Risks.

Abdel-Aty, M. (2003). Analysis of driver injury severity levels at multiple locations using ordered probit models. Journal of safety research, 34(5), 597-603.

Abdel-Aty, M., Keller, J., & Brady, P. A. (2005). Analysis of types of crashes at signalized intersections by using complete crash data and tree-based regression. Transportation Research Record: Journal of the Transportation Research Board, 1908(1), 37-45.

Bailey, T. L., & Elkan, C. (1994). Fitting a mixture model by expectation maximization to discover motifs in bipolymers.

Barnes, Emma et al. (2013). Improving a cylcist and pedestrian environment while maintaining vehicle throughput : A pre- and post-construction street analysis. Transportation Research Board 92nd Annual Meeting. Washington DC: Transportation Research Board.

Benaglia, T., Chauveau, D., Hunter, D. R., & Young, D. S. (2009). mixtools: An R package for analyzing finite mixture models. Journal of Statistical Software, 32(6), 1-29.

Bonneson, J. A., & McCoy, P. T. (1993). Estimation of safety at two-way stop-controlled intersections on rural highways. Transportation Research Record (1401).

Box, P. C. (1970). ntersections. Chapter 14, Traffic Control and Roadway Elements Their Relationship to Highway Safety, Revised. Washington, D.C.: Highway Users Federation for Safety and Mobility.

Chang, L. Y. (2005). Analysis of freeway accident frequencies: negative binomial regression versus artificial neural network. Safety science, 43(8), 541-557.

Chang, L. Y., & Wang, H. W. (2006). Analysis of traffic injury severity: An application of non-parametric classification tree techniques. Accident Analysis & Prevention, 38(5), 1019-1027.

Chin, H. C., Quddus, M. A. (2003). Applying the random effect negative binomial model to examine traffic accident occurrence at signalized intersections. Accident Analysis & Prevention, 35(2), 253-259.

Choo, S., & Mokhtarian, P. L. (2007). Telecommunications and travel demand and supply:Aggregate structural equation models for the US. Transportation Research Part A: Policy and Practice, 41(1), 4-18.

Chung, J. H., & Ahn, Y. (2002). Structural equation models of day-to-day activity participation and travel behavior in a developing country. Transportation Research Record: Journal of the Transportation Research Board, 1807(1), 109-118.

Deutsch, K., Yoon, S. Y., & Goulias, K. (2013). Modeling travel behavior and sense of place using a structural equation model. Journal of Transport Geography, 28, 155-163.

Emaasit, D., Chimba, D., Cherry, C. R., Kutela, B., & Wilson, J. (2013). A Methodology to Identify Factors associated with Pedestrian High Crash Clusters Using GIS Based Local Spatial Autocorrelation. Transportation Research Board 92nd Annual Meeting. Washington DC: Transportation Research Board.

Fambro, D. B., Fitzpatrick, K., & Koppa, R. J. (1997). Determination of stopping sight distances. Transportation Research Board, NCHRP Report 400.

GÅRDER, PER, LEDEN, LARS, PULKKINEN, URHO. (1998). Measuring the Safety Effect of Raised Bicycle Crossings Using a New Research Methodology. TRANSPORTATION RESEARCH RECORD 1636(1), 64-70.

Greibe, P. (2003). Accident prediction models for urban roads. Accident Analysis & Prevention,35, 273–285.

Hadi, M. A., Aruldhas, J., Chow, L. F., Wattleworth, J. A. (1995). Estimating safety effects of cross-section design for various highway types using negative binomial regression. Transportation Research Record, 1500, 169.

Hamdar, S. H., & Schorr, J. (2013). Interrupted versus uninterrupted flow: A safety propensity index for driver behavior. Accident Analysis & Prevention, 55, 22-33.

Harwood, D. W., Bauer, K. M., Richard, K. R., Gilmore, D. K., Graham, J. L., Potts, I. B., ... & Hauer, E. (2008, March). Pedestrian Safety Prediction Methodology - NCHRP (National Cooperative Highway Research Program). Retrieved from Transportation Research Board -: http://onlinepubs.trb.org/onlinepubs/nchrp/nchrp_w129p3.pdf

Hassan, H. M., & Abdel-Aty, M. A. (2011). Analysis of drivers' behavior under reduced visibility conditions using a Structural Equation Modeling approach. Transportation research part F: traffic psychology and behaviour, 14(6), 614-625.

Hauer, E., Ng, J. C., & Lovell, J. (1988). Estimation of safety at signalized intersections (with discussion and closure).

Hu, L. T., & Bentler, P. M. . (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. Structural Equation Modeling: A Multidisciplinary Journal, 6(1), 1-55.

Jin, S.; Qu, X.; Wang, D. (2011). Assessment of expressway traffic safety using Gaussian mixture model based on time to collision. International Journal of Computational Intelligence Systems, 4(6), 1122-1130.

Kaplan, S., Vavatsoulas, K., & Prato, C. G. (2013). Cyclist Injury Severity in a Cycling Nation: Evidence from Denmark. Transportation Research Board 92nd Annual Meeting. Wahington DC: Transportation Research Board.

Karlaftis, M. G., & Golias, I. (2002). Effects of road geometry and traffic volumes on rural roadway accident rates. Accident Analysis & Prevention, 34(3), 357-365.

Kononov, J., Allery, B.K. (2004). Level of service of safety: Conceptual blueprint and analytical framework. Transportation Research Record: Journal of the Transportation Research Board, 57-66.

Kuhnert, P. M., Do, K. A., & McClure, R. (2000). Combining non-parametric models with logistic regression: an application to motor vehicle injury data. Computational Statistics & Data Analysis, 34(3), 371-386.

Lee, C.; Abdel, A. M. (2005). Comprehensive Analysis of Vehicle-pedestrian Crashes at Intersections in Florida. (pp. 775-786). Accident Analysis and Prevention.

Lee, J. Y., Chung, J. H., & Son, B. (2008). Analysis of traffic accident size for Korean highway using structural equation models. Accident Analysis & Prevention, 40(6), 1955-1963.

Lord, D.; Mannering, F. (2010). The statistical analysis of crash-frequency data: a review and assessment of methodological alternatives. Transportation Research Part A: Policy and Practice, 44(5), 291-305.

Mennaï, M., & Rached, K. S. B. . (2011). Consumer-Based Brand Equity Scale: Improving the Measurement in Tunisian Context. China-USA Business Review,10(11), 1126-1138.

Mensah, A., Hauer, E. (1998). Two problems of averaging arising in the estimation of the relationship between accidents and traffic flow. Transportation Research Record: Journal of the Transportation Research Board, 1635(1), 37-43.

Miaou, S. P., Song, J. J., & Mallick, B. K. (2003). Roadway traffic crash mapping: a space-time modeling approach. Journal of Transportation and Statistics, 6, 33-58.

Moini, Nadereh; Liu, Rongfong Rachel. (2013). Geospatial analysis of pedestrian and cyclist crashes in an urbanenvironment: A case study. Transportation Research Board 92nd Annual Meeting. Washington DC: Transportation Research Board.

Monagle, Lauer and Mc. (1955). Do road signs affect accidents? Traffic Quarterly, 3, 322–329.

Moon, T. K. (1996). The expectation-maximization algorithm. Signal processing magazine, IEEE, 13(6), 47-60.

Narayanamoorthy, S., Paleti, R., & Bhat, C. R. (2013). On Accommodating Spatial Dependence in Bicycle and Pedestrian Injury Counts by severity level. Transportation Research Board 92nd Annual Meeting. Washington DC: Transportation Research Board.

National Research Council, C. o. (2008). Coordination in Highway Infrastructure, & Operations Safety. Safety Research on Highway Infrastructure and Operations: Improving Priorities, Coordination, and Quality (Vol. 292).

Noland, R. B., Oh, L. (2004). The effect of infrastructure and demographic change on traffic related fatalities and crashes: a case study of Illinois county-level data. Accident Analysis & Prevention, 36(4), 525-532.

Nordback, K., Marshall, W. E., Janson, B. N. (2014). Bicyclist safety performance functions for a US city. Accident Analysis & Prevention.

Oh, Jun-Seok; Kwigizile, Valerian; Van Houten, Ron; McKean, Joseph; Abasahl, Farhad ; Dolatsara, Hamidreza; Wegner, Bryce; Clark, Matthew. (2013). evelopment of Performance Measures for Non-Motorized Dynamics (No. RC-1603).

Park, B. J., & Lord, D. (2009). Application of finite mixture models for vehicle crash data analysis. Accident Analysis & Prevention, 41(4), 683-691.

Polus, A. (1979). A study of travel time and reliability on arterial routes. Transportation, 8(2), 141-151.

Raford, Noah, Ragland, David R. (2005). Pedestrian Volume Modeling for Traffic Safety and Exposure Analysis. Safe Transportation Research & Education Center. Institute of Transportation Studies (UCB), UC Berkeley.

Ragland, David R. et al. (2013). Roadway and infrastructure design and its relaation to pdestrian and bicyclist safety: Basic principles, applications, and benefis. Transportation Research Board 92nd Annual Meeting. Washington DC: Transportation Research Board.

Schneider, R. J. et al. (2010). Association Between Roadway Intersection Characteristics and Pedestrian Crash Risk in Alameda County, California. Transportation Research Board of National Academics, Journal of Transportation Research Board, No. 2198 , 41-51.

Schneider, R. J., Ryznar, R. M., & Khattak, A. J. (2004). An Accident Waiting to Happen: a Spatial Approach to Proactive Pedestrian Planning. Accident Analysis and Prevention, Vol. 36, No. 2 (pp. 193-211). Accident Analysis and Prevention.

Shankar, V. N., Ulfarsson, G. F., Pendyala, R. M., & Nebergall, M. B. (2003). Modeling crashes involving pedestrians and motorized traffic. Safety Science 41.7 (pp. 627-640). Safety Science.

Siddiqui, Chowdhury et al. (2011). Macroscopic spatial analysis of pedestrian and bicycle crashes. Accident Analysis and Prevention 45 (pp. 382-391). Accident Analysis and Prevention.

Spainhour, L. K. et al. (2006). Causative Factors and Trends in Florida Pedestrian Crashes. Transportation Research Record: Journal of the Transportation Research Board, 1982, 90-98.

Tanner, J. (1953). Accidents at rural three-way junctions. Inst. Highway Eng., 2 (11), 56–67.

Turner, S., Wood, G., Hughes, T., & Singh, R. (2011). Safety performance functions for bicycle. Transportation Research Record: Journal of the Transportation Research Board, 2236(1), 66-73.

US Census Bureau. (2013). SELECTED ECONOMIC CHARACTERISTICS 2008-2012. Retrieved from http://factfinder2.census.gov/faces/nav/jsf/pages/index.xhtml.

USDOT. (1989). Annual Report on Highway Safety Improvement Programs. United States Department of Transportation.

USDOT. (2009). The National Intersection Safety Problem. Washington, D.C.: Federal Highway Administration.

USDOT. (2012). Traffic Safety Facts. Washington, D.C.: National Highway Traffic Safety Administration.

Vogt, A., Bared, J. (1998). Accident models for two-lane rural segments and intersections. Transportation Research Record: Journal of the Transportation Research.

Wang Yiyi et al. (2013). A Conditional autoagressive model for spatial analysis of pedestrian crash counts across neighborhoods. Transportation Research Board 92nd Annual Meeting. Washington DC: Transportation Research Board.

Wang, X., Abdel-Aty, M., Nevarez, A., & Santos, J. B. (2008). Investigation of safety influence area for four-legged signalized intersections: nationwide survey and empirical inquiry. Transportation Research Record: Journal of the Transportation Research Board, 2083(1), 86-95.

Wedagama, D. P. et al. (2006). The Influence of Urban Land-Use on Non-Motorized Transport Casualties. Accident Analysis and Prevention, Vol. 38, No. 6 (pp. 1049-1057). Accident Analysis and Prevention.

Welch, B. L. (1939). Note on Discriminant Functions. Biometrika, 31(1-2), 218-218.

Widaman, K. F., & Thompson, J. S. (203). On specifying the null model for incremental fit indices in structural equation modeling. Psychological methods, 8(1), 16. Board, 1635(1), 18-29.

Wier, M. et al. (2009). An area-level model of vehicle-pedestrian injury collisions with implications for land use and transportation planning. Accident Analysis and Prevention.

Winters M., Brauer M., Setton E. M., Teschke, K. (2010). Built environment influences on healthy transportation choices: bicycling versus driving. Journal of urban health, 87(6), 969-993.

Xie, D. F., Gao, Z. Y., Zhao, X. M., & Li, K. P. (2009). Characteristics of mixed traffic flow with non-motorized vehicles and motorized vehicles at an unsignalized intersection. Physica A: Statistical Mechanics and its Applications 388.10 (pp. 2041-2050). Physica A.

Zahabi, S.A. et al. (2011). Estimating the Potential Effect of Speed Limits, Built Environment and Other Factors on the Pedestrian and Cyclist Injury Severity Levels in Traffic Crashes. Transportation Research Board 90th Annual Meeting. Washington DC: Transportation Research Board.

Zhang, Y.; Brady, M.; Smith, S. (2001). Segmentation of brain MR images through a hidden Markov random field model and the expectation-maximization algorithm. Medical Imaging, IEEE Transactions on, 20(1), 45-57.