EVALUATION OF PEDESTRIAN SAFETY AROUND BUS STOPS USING GEOGRAPHIC INFORMATION SYSTEMS

A THESIS SUBMITTED TO
THE GRADUATE SCHOOL OF NATURAL AND APPLIED SCIENCES
OF
MIDDLE EAST TECHNICAL UNIVERSITY

BY

İREM YÜKSEKOL

IN PARTIAL FULFILLMENT OF THE REQUIREMENTS
FOR
THE DEGREE OF THE MASTER OF SCIENCE
IN
CIVIL ENGINEERING

SEPTEMBER 2012

Approval of the thesis:

EVALUATION OF PEDESTRIAN SAFETY AROUND BUS STOPS USING GEOGRAPHIC INFORMATION SYSTEMS

submitted by **İREM YÜKSEKOL** in partial fulfillment of the requirements for the degree of **Master of Science in Civil Engineering Department, Middle East Technical University** by,

Prof. Dr. Canan Özgen _______________
Dean, Graduate School of **Natural and Applied Science**

Prof. Dr. İsmail Özgür Yaman _______________
Head of Department, **Civil Engineering**

Asst. Prof. Dr. Hediye Tüydeş Yaman _______________
Supervisor, **Civil Engineering Dept., METU**

Examining Committee Members:

Assoc. Prof. Dr. Murat Güler _______________
Civil Engineering Dept., METU

Asst. Prof. Dr. Hediye Tüydeş Yaman _______________
Civil Engineering Dept., METU

Assoc. Prof. Dr. Ela Babalık Sutcliff _______________
City and Regional Planning Dept., METU

Dr. Soner Osman Acar _______________
Civil Engineering Dept., METU

Zerrin Ardıç Eminağa, MSc. _______________
Instructor, Civil Engineering Dept., METU

Date: **12.09.2012**

I hereby declare that all information in this document has been obtained and presented in accordance with academic rules and ethical conduct. I also declare that, as required by these rules and conduct, I have fully cited and referenced all material and results that are not original to this work.

Name, Last name : İrem YÜKSEKOL

Signature :

ABSTRACT

EVALUATION OF PEDESTRIAN SAFETY AROUND BUS STOPS USING GEOGRAPHIC INFORMATION SYSTEMS

Yüksekol, İrem
M.Sc., Department of Civil Engineering
Supervisor: Asst. Prof. Dr. Hediye Tüydeş Yaman

September 2012, 175 pages

Pedestrians are the most vulnerable road users in terms of traffic safety. Public transit users mostly have a pedestrian trip before and/or after the transit one. Thus, pedestrian activity is produced at transit stops naturally. The main focus of this study is pedestrian safety problems around transit stops, more specifically bus stops. The proposed methodology first includes Geographic Information Systems (GIS) analyses of the pedestrian safety along the study corridors and around bus stops on them; this enables determination of accident-prone corridor segments and bus stops, respectively. Later, two analyses are studied to understand their correlation. Finally, linear regression analyses are performed to find the significant factors affecting pedestrian safety. These analyses use parameters created in the GIS analyses in the first part, as well as others (i.e. built environment, traffic network, etc.) that have potential impact on pedestrian movement or safety. In corridor safety models, the number of pedestrian accidents or accident density (or some transformation of them) is used as the dependent variable; while it is selected as the total number of accidents within a selected buffer zone in the bus stop safety models. The case study corridors are selected based on the high density of pedestrian accidents in Ankara, including the Central Business District (CBD) and four main arterials serve from CBD to different regions. The bus stops on corridors with high motorized and pedestrian flows are found to be more critical than others.

Keywords: Pedestrian safety, Geographic Information Systems, Regression analysis, Accident density, Bus stop safety

ÖZ

YAYA GÜVENLİĞİNİN OTOBÜS DURAKLARI ÇEVRESİNDE COĞRAFİ BİLGİ SİSTEMLERİ KULLANILARAK DEĞERLENDİRİLMESİ

Yüksekol, İrem
Yüksek Lisans, İnşaat Mühendisliği Bölümü
Tez Yöneticisi: Yrd. Doç. Dr. Hediye Tüydeş Yaman

Eylül 2012, 175 sayfa

Yayalar, trafik güvenliği açısından en hassas kara yolu kullanıcılarıdırlar. Toplu taşım kullanıcıları, toplu taşıma yolculuğundan önce ya da sonra genellikle yaya olarak yolculuk ederler. Bu nedenle, yaya aktivitesi toplu taşım duraklarında doğalından oluşur. Bu çalışmanın ana odağı toplu taşıma durakları, özellikle otobüs durakları çevresindeki yaya güvenliği problemleridir. Önerilen yöntemin birinci bölümü, seçilmiş koridorlar boyunca yapılan bir dizi Coğrafi Bilgi Sistemleri (CBS) analizi içermektedir. Bu method kazaya meyilli koridor kesimlerinin ve otobüs duraklarının belirlenmesine olanak sağlar. Sonrasında, aralarındaki bağıntının anlaşılması amacıyla bu iki analiz üzerinde çalışılmıştır. Doğrusal regresyon analizleri nihayetinde yayaların güvenliği açısından önemli olan faktörleri bulmak amacıyla gerçekleştirilmiştir. Bu analizlerde methodun ilk bölümünde oluşturulmuş, yaya hareketleri ve güvenliği üzerinde potansiyel etkisi olabilecek parametreler (başka bir deyişle yapılı çevre, trafik ağı vb.) kullanılmıştır. Koridor güvenliği modellerinde , yaya kazaları sayısı veya kaza yoğunluğu (veya bunların bazı dönüştürülmüş halleri) bağımlı değişken olarak kullanılırken, bağımlı değişken otobüs durağı güvenliği modellerinde etki alanına düşen toplam kaza sayıları kullanılmıştır. Durum çalışma koridorları, Ankara ilindeki yaya kazalarının yüksek yoğunlukta olduğu yerler temel alınarak seçilmiştir. Bu koridorlar Merkezi İş Alanı (MİA) ve MİA'dan değişik bölgeler hizmet veren dört ana koridoru içermektedir. Yüksek taşıt ve yaya akışının olduğu koridorlar üzerşnde bulunan otobüs durakları diğerlerine gore daha kritik bulunmuştur.

Anahtar Kelimeler: Yaya Güvenliği, Coğrafi Bilgi Sistemleri, Regresyon Analizi, Kaza Yoğunluğu, Otobüs Durağı Güvenliği

To My Family

ACKNOWLEDGMENTS

I would like to express my sincere appreciation to my supervisor Asst. Prof. Dr. Hediye Tüydeş Yaman for her support, guidance, encouragement and patience throughout this study.

I would like to express my gratitude, firstly to the Turkish Security General Directorate Traffic Education and Research Department, and specifically the Head of the Department, Mr. Yusuf Avan and related personnel for their contributions and the data provided. Secondly, I would like to thank to the Turkish Highway Traffic Safety Council Subcommittee members that were assembled to study pedestrian safety around bus stops in Ankara, including the representative of Ankara Metropolitan Municipality.

I would like to express my sincere thanks to Dündar Ünsal for his supports to this study. I would also like to express my special thanks to Demet Yücel, Orhan Mete Işıkoğlu, Bahar Şimşek, Açelya Yıldırım, Çağıl Çakman for both technical supports and motivations. I feel lucky to have them all near by me.

Also, I would like to express my special thanks to my family for their endless patience and supports, not only during this study but also throughout my life.

TABLE OF CONTENTS

LIST OF TABLES

TABLES

LIST OF FIGURES

FIGURES

LIST OF ABBREVIATIONS

GDH	General Directorate of Highway
GDS	General Directorate of Security
GIS	Geographic Information Systems
CBD	Central Business District
UTZ	Urban Transition Zone
N_A	Number of Accidents on a Link
N_{IA}	Number of Injury Accidents on a Link
N_{FA}	Number of Fatality Accidents on a Link
N_I	Number of Injuries on a Link
N_F	Number of Fatalities on a Link
D_A	Accident Density on a Link
D_I	Injury Density on a Link
D_F	Fatality Density on a Link
SN_A	Number of Accidents on a Segment
SN_{IA}	Number of Injury Accidents on a Segment
SN_{FA}	Number of Fatality Accidents on a Segment
SN_I	Number of Injuries on a Segment
SN_F	Number of Fatalities on a Segment
SD_A	Accident Density on a Segment

SD_I	Injury Density on a Segment
SD_F	Fatality Density on a Segment
$B_{50m}\_N_A$	Number of Accidents in 50m-radius Buffer Zone of a Bus Stop
$B_{100m}\_N_A$	Number of Accidents in 100m-radius Buffer Zone of a Bus Stop
$B_{400m}\_N_A$	Number of Accidents in 400m-radius Buffer Zone of a Bus Stop
$B_{50m}\_N_I$	Number of Injuries in 50m-radius Buffer Zone of a Bus Stop
$B_{100m}\_N_I$	Number of Injuries in 100m-radius Buffer Zone of a Bus Stop
$B_{400m}\_N_I$	Number of Injuries in 400m-radius Buffer Zone of a Bus Stop
$B_{50m}\_N_F$	Number of Fatalities in 50m-radius Buffer Zone of a Bus Stop
$B_{100m}\_N_F$	Number of Fatalities in 100m-radius Buffer Zone of a Bus Stop
$B_{400m}\_N_F$	Number of Fatalities in 400m-radius Buffer Zone of a Bus Stop
AA_	Annual Average

CHAPTER 1

INTRODUCTION

The number of motor vehicles increases in metropolitan cities and the infrastructure investments are generally made to provide more comfortable travels for motor vehicles. However, it is not a real solution, adding more lanes and road capacity is generally the first choice of the authorities to meet the motorized demand. However, this trend threatens pedestrian activities, as there is a strong relationship between the pedestrian mode and motorized flows in urban traffic. The real solution would be making investments to the public transit and creating pedestrian friendly environments to encourage people to walk, or to use public transit.

The public transit users are still pedestrians, while they access the transit services. For public transit, users mostly have a pedestrian trip before and/or after the transit one. Thus, transit stops are the natural pedestrian attraction locations and the effect of the transit stop in pedestrian safety should be evaluated to keep pedestrians from involving accidents. On the other hand, pedestrians are the most vulnerable road users in terms of traffic safety. The pedestrian safety is always a critical issue that has to be studied, as the conflict of pedestrians with motor vehicles always (bus or car) creates a risk for the former. Thus, if not designed adequately, transit stops may cause safety problems for the pedestrians at or around them.

In Turkey, according to the statistics from General Directorate of Security (GDS), Traffic Services Department, there is a big traffic safety problem with more than 4000 fatalities and 150,000 injuries that resulted in approximately 710,000 accidents annually (GDS, 2012). When the question of pedestrian safety around transit stops was raised in the Highway Traffic Safety Commission (HTSC), a subcommittee was formed to study this issue around bus stops in the City of Ankara, which was the main motivation behind this research. As the main public transit mode is bus, scope was defined as the safety around bus stops. In this study, a spatial analysis was performed to study the possible relation between the location of bus stops and pedestrian accidents. The study included bus stops along major corridors in the Urban Transition Zones (UTZs) and the Central Business District (CBD). The study included the pedestrian accidents for the years 2007-2009, as well as run-off road accidents, which were considered as “potential” pedestrian accidents, by the HTSC subcommittee.

The proposed methodology five major steps; the first two are Geographic Information Systems (GIS) analyses of the pedestrian safety along the study corridors and around bus stops on them; this enables determination of accident-prone corridor segments and bus stops, respectively. Later, two analyses were studied to understand their correlation. Finally, two linear regression analyses, are performed to find the significant factors affecting pedestrian safety along the corridors and around the bus stops. These regressions require a GIS database including information on traffic flow (speed and pedestrian volume), traffic network (number of lanes, segment length, etc.), built environment (land-use characteristics, existence of pedestrian barrier, existence of underpass or overpass, existence on metro underpass and number of bus stops), and pedestrian safety measures developed in the first GIS analyses. In corridor safety models, the number of pedestrian accidents or accident density (or some transformation of them) is used as the dependent variable; while it is selected as the total number of accidents within a selected buffer zone in the bus stop safety models.

The layout of this thesis is as follows: In Chapter 2, general information about traffic safety analysis methods are presented, with a special focus on a) GIS and regression analyses, b) pedestrian safety analysis, c) the relation between pedestrian safety and public transit especially bus public transit; furthermore, facilities and bus public transit system in Ankara is introduced briefly. The proposed methodology used in this study is presented in Chapter 3. Chapter 4 includes the brief introduction of study region and pedestrian accident statistics for the study area. Chapter 5 presents results of the case study. Lastly, conclusions and recommendations for future research are presented in Chapter 6.

CHAPTER 2

LITERATURE REVIEW

To provide a background for this study, some concepts and methods for pedestrian safety in urban regions have to be reviewed first. To give a more organized summary of the literature, the issues concerning the traffic safety analysis methods will be given first, which will be followed by a review of studies on pedestrian safety analysis and it is relationship with built environment. A general overview of GIS and its use in pedestrian safety evaluations will be presented later. The following section will summarize the literature on pedestrian safety and public bus transit. The last section in this chapter includes a brief introduction of public bus transit system in Ankara.

2.1 Traffic Safety Analysis Methods

Due to the differences in the infrastructure and traffic flow characteristics of highways and urban roads, traffic safety can be categorized into two main titles, as highway traffic safety and urban traffic safety. Review of the literature showed that highway (also urban freeway) safety generally included analysis of the vehicle-to-vehicle crashes (Golob et al., 2003; Golob and Recker, 2003). When it was compared with urban roads, highways have less complex networks and more continuous flow than urban roads. Still, there are multiple factors that affect the traffic safety in highways, like roadway geometry, driver’s behavior, traffic conditions and environmental factors . However, urban roads have more complex and interactive factors, higher rates of vulnerable road users. For these reasons, the traffic safety analyses used in highways and urban roads are different from each other.

The complexity of traffic safety issue in an urbanized environment comes mostly from the contribution of pedestrian activity that is higher than highways. According to the US National Highway Traffic Safety Administration (2001), although, the proportion of the vehicle to pedestrian crashes is small in the USA (2 to 3% of all crashes), the proportion of fatalities of these crashes is high (around 11-13% of all fatality crashes). A study by Hebert-Martinez and Porter (2004) showed that the most of pedestrian crashes (75-80% of all pedestrian crashes) that occurred between years of 1990-1999 in Virginia, US, happened in urbanized environment; thus, pedestrian safety is more vital, especially for the urbanized environment in the traffic safety analyses. To solve the pedestrian safety problems in urban environment with direct investments, Clifton et al. (2009) suggested that the relationship between pedestrian accidents and the attributes of the urbanized environment like, land-use, urban form and transit facility characteristics should be clarified. Thus, this chapter concentrated on pedestrian safety and effecting factors in urbanized environment in the next sections.

2.1.1 Highway versus Urban Traffic Safety Analysis

The safety analyses of the highways and urban roads show some differences, some of which can be seen in the datasets that were evaluated. Depending on the road type of the accident, the responsible authority and data collection methods may vary significantly between countries; so, are the studies performed with the available accident data. Whenever possible, researchers used Geographic Information Systems (GIS) to evaluate accident distributions, as well as regression analyses tools to search for affecting factors.

In the USA, Federal Highway Administration (FHWA) determined the data needed for highway traffic safety analyses within the scope of Highway Safety Improvement Program (HSIP) (see Table 2.1). This data is evaluated through some tools provided by the FHWA (2012). These analytical tools are used to quantify the safety performance of highway infrastructures (Highway Safety Manual), to evaluate the geometric design of the highway in terms of safety (Interactive Highway Safety Design Model), to detect the problematic locations and to take precautions (FHWA GIS Safety Analysis Tools).

Table 2.1 The Listing of Fundamental Data Elements for HSIP (FHWA, 2012)

Roadway Segment	**Intersection**
Segment Id*	Intersection ID
Route Name*	Location
Alternate Road Name*	Intersection Type
Route Type*	Date Opened to Traffic
Area Type	Traffic Control Type
Date Opened to Traffic	Major Road AADT
Start Location*	Major Road AADT Year
End Location*	Minor Road AADT
Segment Length*	Minor Road AADT Year
Segment Direction*	Intersection Leg ID
Roadway Class*	Leg Type
Median Type	Leg Segment ID
Access Control*	**Ramp/Interchange**
Two-Way vs One-Way Operation*	Ramp ID*
Number of Through Lanes*	Date Opened to Traffic
	Start Location
Interchange Influence Area on Mainline Freeway	Ramp Type
AADT*	Ramp/Interchange Configuration
	Ramp Length
AADT Year*	Ramp AADT*
	Ramp AADT Year

*Highway Performance Monitoring System full extent elements are required on all Federal-aid highways and ramps located within the grade-separated interchanges, i.e., National Highway System (NHS) and all functional systems excluding rural minor collectors and locals

In Turkey, the General Directorate of Highways (GDH) uses the rate-quality control method in the evaluation of highway traffic safety. This statistical method analyzes the highways by dividing it into 1 km of segments. For each segment, three parameters are calculated; (i) accident rate, (ii) accident frequency, (iii) severity index; and each of these parameters are compared with the different critical values. When all of these three values of a segment exceed the critical values, that segment is considered a “black spot” (Sjolinder et al., 2001).

One of the urban traffic safety methods is crash rate calculations that aim to improve the safety in unsafe locations. Iowa State University Center for Transportation Research and Education (CTRE) (2012) explained the parameters used in ranking the accident locations by using, number of crashes, the severity of crashes and the crash rate per traffic volume. To find the source of the problems in the accident locations, the field observation reports that have information about physical and operational characteristics of the locations have to be evaluated. The same method can be used for an intersection. The required data included the number of crashes at the intersection for the time period of the study, the number of years in the study and the annual average daily traffic (AADT) for each leg of the intersection. The crash rate can be calculated by using the following equation:

$$R_i = \frac{2 \times c \times 1000000}{\sum AADTs \times Y \times 365} \qquad \text{(Eqn. 1)}$$

where

R_i= crash rate per million entering vehicles,

c= number of crashes and

Y= number of years analyzed (Iowa DOT, 1989).

Another traditional urban traffic safety analysis method is creating push-pin maps for determining the locations where the crash density is high (FHWA, 2012). This method can be defined as the early version of the GIS-based maps that only indicates the location of the crash.

While urban traffic safety analysis is discussed, various actors contribute to the urban traffic like pedestrians, cyclist, public transit vehicles, etc. The analysis should cover the specific characteristics of these participants. Thus, the studies beyond the traditional urban traffic safety analyses focused on certain problems in urban traffic safety.

2.1.2 GIS Analysis for Traffic Safety

A common tool that was used in traffic safety analysis is GIS environment. GIS are the computer-based systems that capture, store, manipulate, display, and analyze the geographical information (Thill, 2000). Besides the geo-visualization capabilities of GIS software, it can store the attributes of entities that are geo-referenced in a database. Thus, the relationship between spatial and topological properties and other attributes of the entities can be stored. The storage of the data can be built only one layer as well as different layers. Thus, GIS based software can generate new database for further analysis by matching and combining data that are different layers. Due to these features, GIS use in various disciplines like city and regional planning, resource management, environmental engineering, archaeology, geography, geology, transportation etc.

In the traffic safety analysis, GIS can visualize the locations of accidents (coordinate/GPS location/ latitude longitude of accidents) and store the attributes of accidents like, time of the accident, number of injuries, number of fatalities, characteristics of the roads that accidents occur, land-use characteristics that the roads take part, etc. Furthermore, it is considerably explanatory in traffic safety analysis to relate the accident locations with the attributes of surroundings to find the reasons behind the occurrence of accidents.

There are many studies in the literature that GIS was used recently. One of these studies was conducted by Ardıç Eminağa (2008) to find the impact of the speed on the traffic safety of urban corridors. The developed methodology was used GIS based data like time-dependent average link speeds and location of accidents that occurred on the links. The aim of the study is to find the accident prone locations on urban arterials by generating GIS based maps like traffic accident maps, severity index maps, time dependent accident map and time dependent severity index maps.

Hijar et al. (2003) made a series of analysis that included also GIS analysis to investigate the reasons behind the pedestrian injuries in the city of Mexico. By using GIS tools, maps were prepared in three different levels of aggregation to designate the high concentration of the pedestrian accidents. Later, the high-risk environments were determined. Results showed that wide avenues with high traffic flow and high pedestrian mobility, and locations where inappropriate pedestrian bridges exist, were at high risk of pedestrian injuries.

Spatial Analysis

The difference between any database management system and GIS is that GIS can spatially manipulate the data. When GIS tools have the ability to do spatial analysis, they can also manipulate and manage the spatial database. Thus, the real power of GIS is coming from the ability of doing spatial analysis (O'Sullivan and Unwin, 2010). Spatial models in traffic safety analysis commonly designates the spatial effects that were unmeasured, like weather, population etc. in the occurrence of accidents (Valverde and Jovanis, 2005).

In the literature, the usage of spatial data shows differences. The spatial data can be defined such as point, line, area and field (O'Sullivan and Unwin, 2010). Thus, in the traffic safety analysis the accidents can be defined as point events or different area levels, such as along the road segments with a defined width, counties or cities. The methodology of spatial analysis can change according to the data type. However, the aim of the spatial analysis is to detect the hot spot locations, where the occurrence of traffic accidents was high, and that is not related with the data type.

Levine et al. (1995) developed a spatial model that assigned each of the motor vehicle accidents that happened between 1990 and 1991 to the nearest intersection and described the spatial variations in the accidents. The study focused on the determination of systematic variations between the accidents according to hour of the day, weekday or weekend, type of crash, number of vehicles etc. In addition, for each time period, the model was associated with the population, employment, road and land use characteristics. Results showed that the higher the employment density was, the higher the crash concentration. Also, in residential areas, there were more alcohol-related crashes.

Some traffic safety analyses were conducted in area levels. For example, Amoros et al. (2002) developed a model that compared the traffic safety between the counties of France and tried to explain the differences in terms of different road type, distribution and different socio-economic characteristics. The results suggested an interaction between road types and traffic safety of counties in terms of severity, but this was meaningful in just ranking the counties in terms of safety, and it did not explain the traffic safety differences between counties. This strong relationship made the researchers to investigate the road types in more detail to find the differences before analyzing the socio-economic factors.

Kernel Density Estimations

There is a pattern for every point located events and this pattern allows explaining the areal density that is preferable in spatial analysis. The idea behind the Kernel density Estimation (KDE) is to estimate the density of pattern at any location in the study region. The simple approach of KDE is to define a circle centered at the location of interest and count the events that are falling into the circle and divide this number by area of the circle. Intensity is estimated for a point where the circle is centered (O'Sullivan and Unwin, 2010). For more complex KDE, kernel functions are created which produces a smooth curved surface around the point according to the determined simple bandwidth.

Thus, for traffic safety analysis, KDE is used to determine the "hot spots" where density of traffic accidents is high. Erdogan et. al (2007) estimated risky locations on the highways within the Afyonkarahisar administrative borders, by using two different methods: KDE and repeatability analysis. For both analyses, almost the same locations were found as hot spots where most of them locate at cross roads to the villages and small cities.

Another study in Turkey was made by Keskin et. al (2011) to discover the problematic locations, where traffic incidents clustered in the Middle East Technical University (METU) campus by using both Nearest Neighbor Distance and KDE. Besides searching the locations, they tried to figure out the changes in traffic events for different seasons, days and time periods. For both of the methods, the results were compared and the hot spots zones where the traffic incidents clustered on the roads were determined. Also, the analysis is made for days and time periods. It is determined that the traffic incidents were clustered mostly in Monday. Most of the accidents occurred within the time period of 12:00 and 19:00. The study included a suggestion of police or security member control at high-risk locations during this time period.

2.1.3 Regression Analysis

Statistical methods can also be useful in associating spatial data with any descriptive statistics, which can be stored in dataset (O'Sullivan and Unwin, 2010). It is very common to use statistical models to explain the factors that are effective in the occurrence of traffic accidents. The simplest form of regression analysis is the linear regression, the formula of which is given for the i^{th} observation below;

$$y_i = \beta_1 x_{i1} + \beta_2 x_{i2} + \ldots + \beta_p x_{ip} + \varepsilon_i \qquad i = 1,\ldots..,n, \qquad \text{(Eqn. 2)}$$

where

y = dependent variable,

x_p= an independent variable

p= number of independent variables,

ε=error term

The regression model can be more complex, such as probit models, if the study focuses on determination of other aspects related to traffic safety, such as, accident risk at a location.

In the literature, there are various types of statistical models developed to identify the influencing factors of the traffic safety. Clifton et al. (2009) used first cross tabulations and then an ordered probit model to estimate effective factors in the severity of injuries in the pedestrian–vehicle crashes. Severity of injury was used as a dependent variable in three levels as no injury, non-fatal injury and fatality. They developed three models; in Model 1, personal characteristics (age-child, age-older, sex) were used as independent variables. In Model 2, environmental conditions (daylight, weather) and in Model 3, built environment (road condition, road facility type etc.) were added to Model 1, as independent variables. The results of cross tabulation related age and sex as follows; people who were age of between 16-64, were involved more in pedestrian vehicle crashes than children and people age of 65 and older. Males involving ratio in a crash was higher than females. Pedestrians who wore light coloured clothes were exposed to more crashes than wearing dark clothes; probably because 70% of crashes were in daylight hours. The results of ordered probit model for Model 1 show that child pedestrians (0-15 years of age) had a positive association with injury. Males had more risk than females to sustain injuries. Also, pedestrians over 65 years of age had positive association with fatal injuries. For Model 2, only variable that was statistically significant, was crashes that occurred after sunset. The results of Model 3 related with built environment will be discussed in Section 2.2.1

Miranda-Moreno et al. (2011) used log-linear and standard negative binomial regression models to investigate the relationship between parameters of built environment like,

- land use characteristics (employment density, commercial etc.),
- demographics (population density, children density etc.),

- transit supply (kilometers of bus lanes, number of transit stops), and
- road network characteristics (kilometers of streets and major roads, number of intersections, speed limits etc.)

with pedestrian activity and pedestrian accident occurrence at the intersections. The first developed model suggested that some built environment characteristics, like population density, commercial land use, number of jobs, number of schools, presence of metro station, number of bus stops, percentage of major arterials and average street length were significantly associated with pedestrian activity. A second model showed that reasons of the pedestrian crashes at the intersections were mainly pedestrian activity and traffic volume.

Ossenbruggen et al. (2001) used logistic regression analysis to find the statistically effective factors that predicted the probability of occurrence of the accidents and injury accidents in rural and small-urbanized areas. Land use activity, roadside design, traffic control, merging and crossing traffic, speed and crash type were used as independent variables in the logistic regression model for three different sites, village, shopping and residential. The results showed that the risk of pedestrian crashes was higher in residential and shopping zones than villages. The reason behind this situation was explained by the fact that the village sites had sidewalks more than shopping and residential sites, and it had a pedestrian friendly atmosphere. Also, due to the high pedestrian volume and low speed limits, drivers were more careful when compared with drivers of shopping and residential sites.

2.2 Pedestrian Safety Analysis

Pedestrians are the most vulnerable road users with cyclist in terms of safety. Thus, beside the vehicle-to-vehicle crashes in traffic safety analysis, some studies focus on only pedestrian-to-vehicle crashes. The same analysis methods that were explained in previous sections, GIS and regression analysis, can be used while studying vehicle-to-vehicle and pedestrian-to-vehicle accidents.

One of the earlier studies for pedestrian safety (Hijar et al., 2003) focused on the pedestrian accidents in Mexico City for the period of 1994-1997. The study tried to identify the determinants of deaths due to vehicle-to-pedestrian crashes by using quantitative and qualitative methods. In quantitative method, simple mortality rates were calculated in accordance with age and sex for 16 different districts. The results show that rate of death due to the pedestrian accident was highest in two semi-rural and one urbanized regions. The mortality rates of males were higher than females. Secondly, spatial mortality analysis was conducted to the geocoded deaths to figure out in which environment the concentration of deaths was high. Ten of neighborhoods and six intersections in these neighborhoods were determined as the highest density of deaths. With these two analyses, the critical locations in terms of pedestrian safety were determined. However, none of these two explained the traffic characteristics of the hazardous locations. Thus, third step of the method was the physical space analysis to observe the real condition of the traffic. The observations showed that all risky locations were on large avenues with high pedestrian mobility. There existed the pedestrian bridges. However, pedestrian did not use them, instead, tried to cross the street despite a refuge about 1 m high. The last analysis was in-depth interviews with the injured pedestrians and drivers involved in an accident, to identify the social determinants of pedestrian death. Most of the pedestrian interviewees reported that they did not know the traffic rules, traffic signals and use the pedestrian bridges ever.

Many other studies investigate the behaviors of pedestrians in traffic to get the safety precautions. A study by Bernhoft and Carstensen (2008) focused on especially between older (age of 70 and above) and younger (age of 40-49) female and male pedestrians and cyclists. A questionnaire about general behaviors in traffic in their daily lives was prepared and sent to the pedestrians and cyclists. The results of regression analysis showed statistically significant differences between traffic behaviors of older and younger people as follows; older people used the routes that had sidewalks, pedestrian crossings, signalized intersections, least traffic and good street lighting more than younger. Another study carried by Tiwari et al. (2007), used video cameras placed at seven intersections in Delhi, India to observe the pedestrians crossing behaviors. Using survival analysis statistical methods, it was found that mean waiting time of females were more than males. However, when the waiting time was getting longer, both females and males try to cross before the traffic lights change to green for pedestrians. This increased the risk for them to get into a crash by a motor vehicle. Thus, the waiting time in cross-sections for pedestrian should be minimized.

2.2.1 Pedestrian Safety and Built Environment

More than the studies focused on understanding pedestrian behaviors solely, for this study, it is more crucial to focus on pedestrian safety and its relation to public transit and built environment, etc. It is only natural to think that the density of the pedestrian accidents would be high where the pedestrian activity is high. Most of the studies start with a hypothesis that mobility of pedestrian was high due to the built environment characteristics and this was directly related with the occurrence of pedestrian accidents. However, it is important to quantify such expectations as much as possible. Although the studies in the literature use land use information in different analysis, the common aim is to associate the built environment and the frequency or sometimes severity of pedestrian crashes with the pedestrian accidents.

The study that focused on pedestrian-vehicle crashes in the City of Baltimore, USA (Clifton et al., 2009) also investigated the impact of personal characteristics of pedestrians, environmental characteristics and built environment on severity of injuries. The study included a rather thorough literature review of studies focusing on impact of built environment measures on pedestrian safety. The pedestrian-vehicle crash data included information of pedestrian characteristics (age, sex, clothing worn etc.), crash characteristics (road facility type, road condition, etc.), level of injury (fatality, injury, no injury). Each of crash record was geocoded to the nearest intersection. The built environment information of crash locations are gathered by drawing 0.25 mile buffer zone around the nearest intersection. For the ordered probit model the variables were determined for 0.25 mile buffer zones such as, transit access (number of bus stops), connectivity (number of intersections and cul-de-sacs), population density, median income, presence of schools, land use types (commercial, residential, mixed). Out of three models constructed, the model was related with built environment showed that the impacts of pedestrian connectivity and transit access have positive impacts on injury crashes, but not as much significant on fatality ones. Clifton et al. explained these relationships as follows: In these regions due to the grid-like street network and developed transit access created high pedestrian mobility. In addition to this, the speed of the motor vehicles is generally low due to the speed limits. Thus, the severity of crashes did not kill pedestrians.

The study by Retting et al. (2003) suggested that the risk and severity of pedestrian-vehicle crashes could reduce with the modifications in built environment that give the priority to pedestrians. Three engineering modifications were discussed in the light of the literature: reducing the speed limit, separating the pedestrians from vehicles by time or space and increasing the visibility of pedestrians. The results showed that these applications reduced the risk of occurrence and severity of injuries.

Clifton and Kreamer-Fults (2007) studied the relationship between the crash occurrence and severity of injuries and the built environment characteristics around schools. The built environment data included the volume and speed of traffic, the type and mix of land uses, and the socio-economic characteristics of neighborhood residents. A multivariate model was developed between built environment data and crash severity and crash risk exposure. The results show that there was a negative relationship between existing of driveway or turning bay on the school entrance and the crash occurrence and injury severity. On the other hand, the presence of recreational facilities on the school site increased the crash occurrence and injury severity.

Beside the built environment characteristics, Lee and Abdel-Aty (2005) added traffic characteristics to their study that analyzes vehicle-pedestrian crashes at intersections of Florida. The log-linear and ordered probit models were used to identify the association between these two characteristics and high pedestrian crashes and severity of injuries. The results showed that for non-passenger vehicles severity of injury of pedestrians was higher than passenger vehicles. On the other hand, passenger cars are more involved in the crashes than vans, trucks and buses. Also, the conditions of high frequency and injury severity of pedestrian crashes for drivers' and pedestrians' fault were the same. Thus, the demographic factors of pedestrian and driver, and road geometric, traffic and environment conditions had positive effects on frequency and injury severity of pedestrian crashes.

Loukaitou-Sideris et al. (2007) tried to find the impacts of socio-demographic, land use, density, urban form, and traffic characteristics on the involvement of pedestrians in vehicle crashes. Four different methods were used in the study. From the results of the regression analysis, density of pedestrian accidents was found higher in the commercial and residential areas, whereas it was quite low in the industrial areas.

2.2.2 GIS-based Pedestrian Safety Evaluations

Spatial analysis is very useful in pedestrian safety evaluations. To see the accidents on the maps give the chance to relate the accidents with any characteristics of entities that can be defined in GIS environment. In most of the studies aforementioned, a spatial analysis was conducted to create a database to use it in further analysis mostly statistical ones. Thus, spatial analysis gives a great help in clarifying the reasons behind pedestrian accidents. In this section, the spatial analysis process and results of these studies will be reviewed in detail.

Clifton et al. (2009) used GIS tools as a part of their analysis. Due to the missing information of the exact locations of pedestrian accidents, each of them was assigned to the nearest intersection. They produced the pedestrian accidents map that shows fatality, non-fatal injury and no injury accidents of Baltimore City with the accident data for years 2000-2004. From this map, the problematic locations of fatality accidents, non-fatal injury accidents and no injury accidents can be observed. A database was produced also from the spatial analysis to use it in the statistical analysis part. Around each of the accident to get the characteristic of the environment a 0.25-mile (approximately 400 m) buffer zone was drawn. The most of attributes were related with built environment. The results showed that most of the crashes occurred at locations without pedestrian signals. Almost 60% of the crashes are on the roadway whereas 22% are at the crosswalks.

Loukaitou-Sideris et al. (2007) made a spatial analysis in their study with not only the locations of the pedestrian accidents but also they aggregated the land use, socioeconomic characteristics and traffic variables to the maps also. Thus, a visual correlation could be done between pedestrian accidents and these variables. The results of spatial analysis showed that pedestrian accidents did not spread equally along the city of Los Angeles. However, the pattern of the accident sites was linear along the major arterials. From the visual correlations, they found that a positive relationship between pedestrian collisions and Latino population, the percentage of the population that fell below poverty, percentage of commercial land use and the level of average annual daily traffic.

2.3 Traffic Safety and Public Transit System

The general acceptance about the existence of public transit system is that it improve road safety due to the decrease in the number of motor vehicles (Brenac and Clabaux, 2005). However, the statistics and the studies that made before show that public transit systems can cause the safety problems. Any of the transit systems (or their vehicles) can be involved in accidents directly or indirectly. Direct involvement means that being involved in the collision itself, on the other hand indirect involvement means playing a role in the occurrence of the collision (e.g. as a sight obstruction) (Brenac and Clabaux, 2005).

Wahlberg (2004) analyzed the 2237 low speed accidents (in areas with 50 kph speed limit or lower) that buses involved during the years 1986-2000 in the city of Uppsala, Sweden. The aim of the study was to identify the relation between various characteristics of accident involvement and to investigate the ways to reduce the number of effective characteristics in the involvement. 17 variables were used in the analysis and the following variables were dependent ones; hit object, hit person, injury inside bus, injury outside bus, shunt by bus, shunt by other vehicle, side contact and single (not involvement of any other road users). The rest of the variables were independent ones. The remarkable results especially related with the bus stops showed that, from the 286 hit object accidents, 16.1%; from the 129 injury in bus accidents, 28.7 %; from 219 shunt by bus driver accidents, 24.2%; from 164 shunt by other driver accidents, 62.8% and from 845 side contact with other vehicle accidents, 34.7% of them occurred at the bus stops. Thus, the bus stops can be evaluated as accident-prone locations for various types of accidents according to the study. To reduce the accidents in the bus stops Wahlberg (2004) suggested re-building bus stops that are completely isolated from the rest of the traffic, providing longer safety distance to other buses and also keeping away parked cars from the bus stops.

In the study of Cameron et al. (2001), from 90 patients that had injuries related with tram accidents, 41 patients were cyclists, 23 pedestrians, 12 motorists or motorcyclists and 14 workers in the construction site of tram. As it can be seen from the results the most vulnerable users are the cyclists and pedestrians. Thus, the next subsection will focus on the relationship between pedestrian safety and public bus transit systems.

Pedestrian Safety and Public Bus Transit

The study of Brenac and Clabaux (2005) analyzed the indirect involvement of buses in the traffic injury accidents. The study focused on the accidents for Boulogne –Billancourt region in France, for which the available accident database included searchable text elements, such as transcriptions of the statements from the people involved and witness and a summary of the facts by the police officers. By simply searching for the word “bus” or its synonyms in these text parts, the researchers were able to compile 56 accidents involving bus directly (17 cases) or indirectly (26 cases). The indirect involvement was further classified into three categories as a) bus constituting a sight obstruction (13 cases), b) pedestrian hurriedly crossing the street to catch the bus (12 cases), and c) other cases (1 case). The more detailed and remarkable results about the characteristics of indirect involvement accidents as follows; in the category of bus constituting a sight obstruction, 3 of the 13 accidents occurred while bus boarding or alighting at a bus stop. In the category of pedestrian hurriedly crossing the street to catch the bus, all of the 12 accidents naturally happened at bus stops. Also, the study included some suggestions against indirect involvement of bus stops in traffic injury accidents such as, installment of pedestrian island to the pedestrian crossings located at the intersections or near the bus stops, so that the other vehicles overtaking the bus do not crash the pedestrians that pass in front of the bus. It was predicted that 5 of the 13 obstruction cases could be avoided. Other suggestions included installing a pedestrian guardrail on the sidewalk facing the bus stops and increasing the frequency of the bus services to prevent people from hurried-crossing behavior.

A more specific study conducted by Unger et al. (2002) focused on the most vulnerable group of pedestrians, children. A retrospective study that included 30 children pedestrian injuries at tram and bus stops in Austria was conducted to clarify the reasons behind the accidents. A questionnaire was prepared and implemented. The content of the questionnaire included physical conditions of the accident location (road characteristics, presence or absence of zebra crossings or traffic lights, sign posted speed limits), traffic flow conditions (speed and position of the vehicle, type of vehicle, traffic density), existence of obstructions for driver or pedestrians, the wearing of reflecting clothes, existing of accompanying persons like, other children, parents or other adults at the scene etc. The results of the survey showed that from the 30 cases, 26 cases were in high traffic density, 3 cases, in moderate traffic density, and 1 case in low traffic density. In 4 cases, traffic lights and zebra crossings were present and for 10 cases only zebra crossings were present, while for 16 cases, neither was present. In 2 cases, the children were death due to the severe head and neck injuries. The cause of the fatal injury was the crash speed that was 75 and 80 km/h, and accident location was a bus stop alongside a country road, where higher speed travel should be expected. As a conclusion, they determined that bus stops along country roads were hazardous locations without effective speed limits and safe crosswalks and continuous pavements, for children to get on and off buses. Recommendations included installment of sign-posted speed limits lower than 50 kph at bus stops along the country roads, cross walks in front of incoming bus, continuous pavement in the area of bus or tram stop, barriers at bus or tram stops to prevent children from crossing behind the buses or trams and applying parking restrictions combined with regular controls at bus or tram stops.

2.4 Public Bus Transit System in Ankara

As a case study, the pedestrian safety at bus stops in the City of Ankara, Turkey was analyzed in this study. Thus, to form a base information about the Ankara and bus public transit system in the city, a brief introduction is provided here. Ankara is the capital city of Turkey with 3 965 232 inhabitants (almost 4 000 000) according to the statistics of 2011 (TUİK, 2012a). This number corresponds to 5.3% of population of Turkey. Due to the characteristics of capital cities, the density of government offices is considerably high. Besides, there are commercial regions, shopping sites etc. There are 1 010 182 registered motor vehicles in the city according to May 2012 statistics (TUİK, 2012b).

The current modes used in Ankara can be categorized as follows: bus transit systems (municipality buses, private bus services), demand responsive services (minibuses), rail systems (Metro, Ankaray, Suburban Rail), shuttle/service systems, service vehicles (private and public institutions), private cars and taxis. The distribution of usage of these modes can be seen in Figure 2.1. According to the study that was conducted by Çubuk et al. (2003) approximately 75% of all trips is made by public transport vehicles. This means that the public transit systems are still the major mode in Ankara. Although the road design, traffic signals, and traffic management policies are not specifically designed for bus transit system, buses carry approximately 1.315.000 commuters every day (Çubuk et al., 2003).

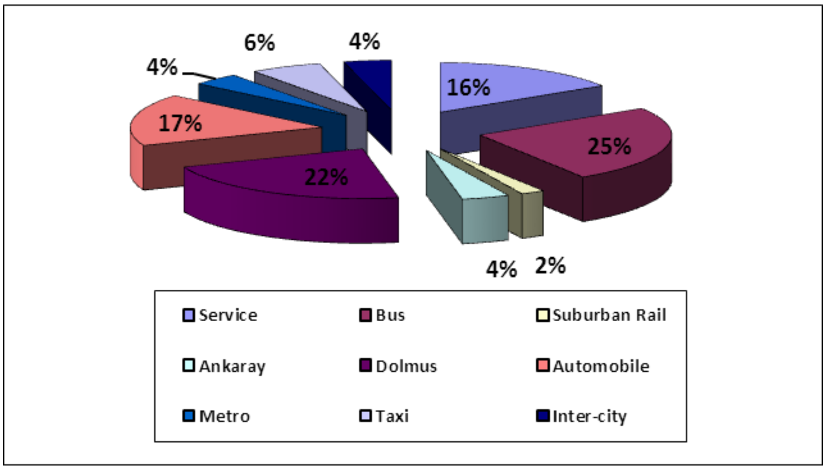


Figure 2.1 The Proportion of Trips Made by Motorized Modes in Ankara (Çubuk et al., 2003)

The daily production of travel per capita is 1.96 in Ankara. In a working day, there are made more than 4 million trips in average inside the city. Approximately 750 000 of these trips are made by the buses that belong to the EGO, 200 000 ones are with the private buses, 900 000 ones are with the minibuses and the rest is with services, taxis, private cars, suburban train, Metro and Ankaray (Ankara Büyükşehir Belediyesi, 2000). More than ¾ of all trips are made with public transit vehicles. Between these vehicles, buses come in the first line. Buses carried the highest number of passengers (see Table 2.2).

EGO Public Bus System

EGO that is an institution of Ankara Metropolitan Municipality provides the electricity, natural gas and bus services to the public in the city of Ankara. The bus service is currently given with 1789 buses (the distribution of number of buses according to their models can be seen in Table 2.3). EGO bus services are developed for 5 main regions (see Figure 2.2). The number of bus lines and bus stops according to their types are given in Table 2.4 (EGO, 2012b).

Table 2.2 Distribution of Number of Passengers for City Transit Modes in Ankara (Ankara Büyükşehir Belediyesi, 2000)

Modes	Number of passenger	Public Transit Share (%)	General Modal Split (%)
Public Transit	**3 340 000**	**100**	**76.78**
Bus	1 315 000	36.38	27.93
Minibus	990 000	29.64	22.76
Service	685 000	20.51	15.75
Suburban Rail	100 000	2.99	2.3
Metro	175 000	5.24	4.02
Ankaray	175 000	5.24	4.02
Private Transport	**1 010 000**		**23.22**
Private Car	750 000		5.98
Taxi	260 000		17.24
Total	**4 350 000**		**100**

Table 2.3 The Model and Number of the Buses that EGO has (EGO,2012a)

MODEL	**NUMBER**
BELDE BMC	2
MAN -NORMAL	1172
MAN -ARTICULATED	50
MERCEDES	222
MERCEDES - ARTICULATED	70
İKARUS-NORMAL	237
İKARUS - ARTICULATED	36
TOTAL	**1789**

Table 2.4 The Number of Lines and Bus Stops for each Region (EGO,2012b)

	Region 1	Region 2	Region3	Region 4	Region 5	Total
Public Bus Line Statistics						
No. of Lines	93	74	92	53	63	375
Public Bus Stop Statistics (by Stop Type)						
Open Stop	1020	776	938	679	316	3729
Closed Stop	326	298	289	314	387	1614
Total	1346	1074	1227	993	703	5343

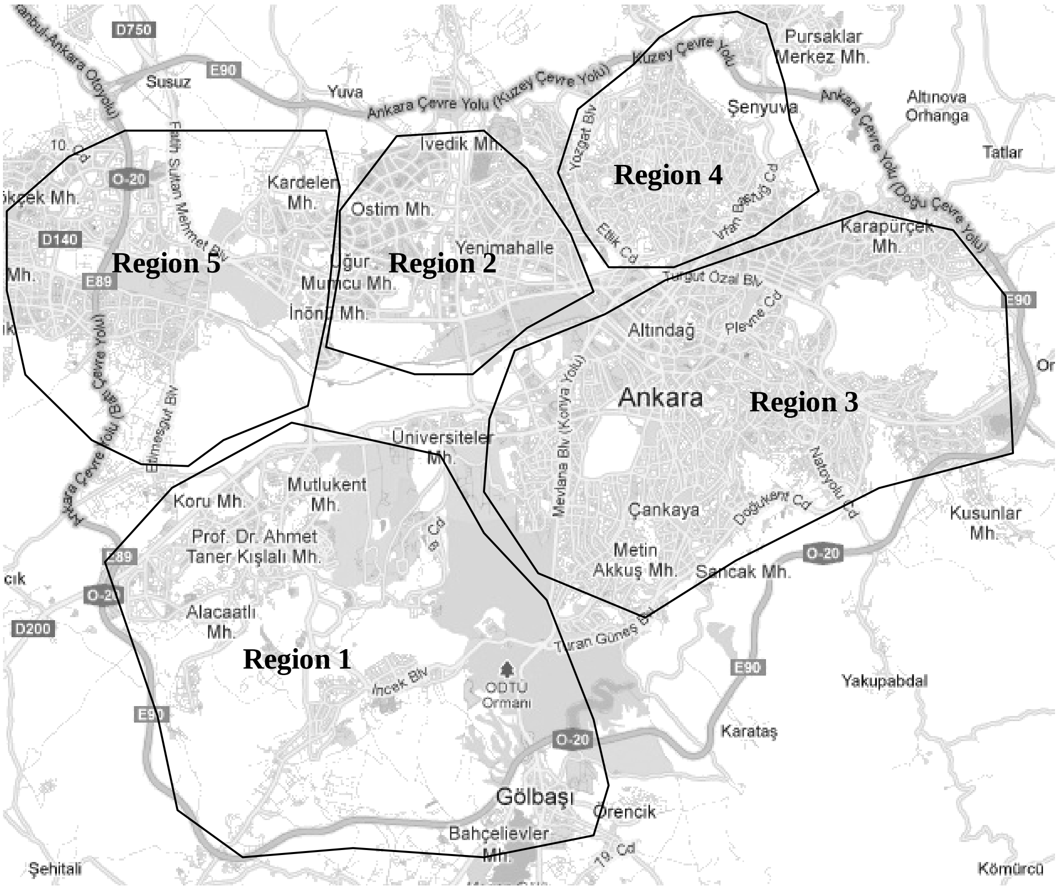


Figure 2.2 The Bus Service Regions (Map of Ankara City Center, 2012)

CHAPTER 3

METHODOLOGY

In case of pedestrian safety analysis in Turkey, it is not possible to retrieve original accident reports and search for “bus” or “bus stop” keywords directly. In addition, the current digitized accident data in Turkey does not include any information about bus services, unless the vehicle that gets involved in the accident is a bus. But, even in that case, the description of the accident that may include any further information about the location of the accident being near or at a bus stop is not digitized, thus, not searchable. Furthermore, a similar problem is valid for the bus stops locations of Ankara. Due to the lack of the geographic locations of bus stops, the GPS data were collected for the major urban corridors and CBD region manually within the limits of time and budget.

As the focus of this study is pedestrian traffic safety around bus stops, and instead of a regional perspective bus stops along urban corridors will be analyzed based on the location of pedestrian accidents and bus stops data, a specific analysis method has been developed to achieve the goal of this study. As the details of pedestrian accidents were not available and neither was their relation to the bus system, the proposed methodology will simple depend on the analysis of the “accident distribution around the bus stops” along the selected corridors , which can be shortly summarized as a “accident distribution analysis” method. There are five major steps of this proposed method shown in Figure 3.1, and explained briefly as follows:

Step 1: Data Collection

An accident distribution analysis method requires more detailed accident reports about accidents and bus stops. In fact, physical characteristics of road that pedestrian accidents occurred, time of accidents, weather conditions etc. could help identifying the reasons of pedestrian accidents. However, in the absence of all such details, it is imperative to have at

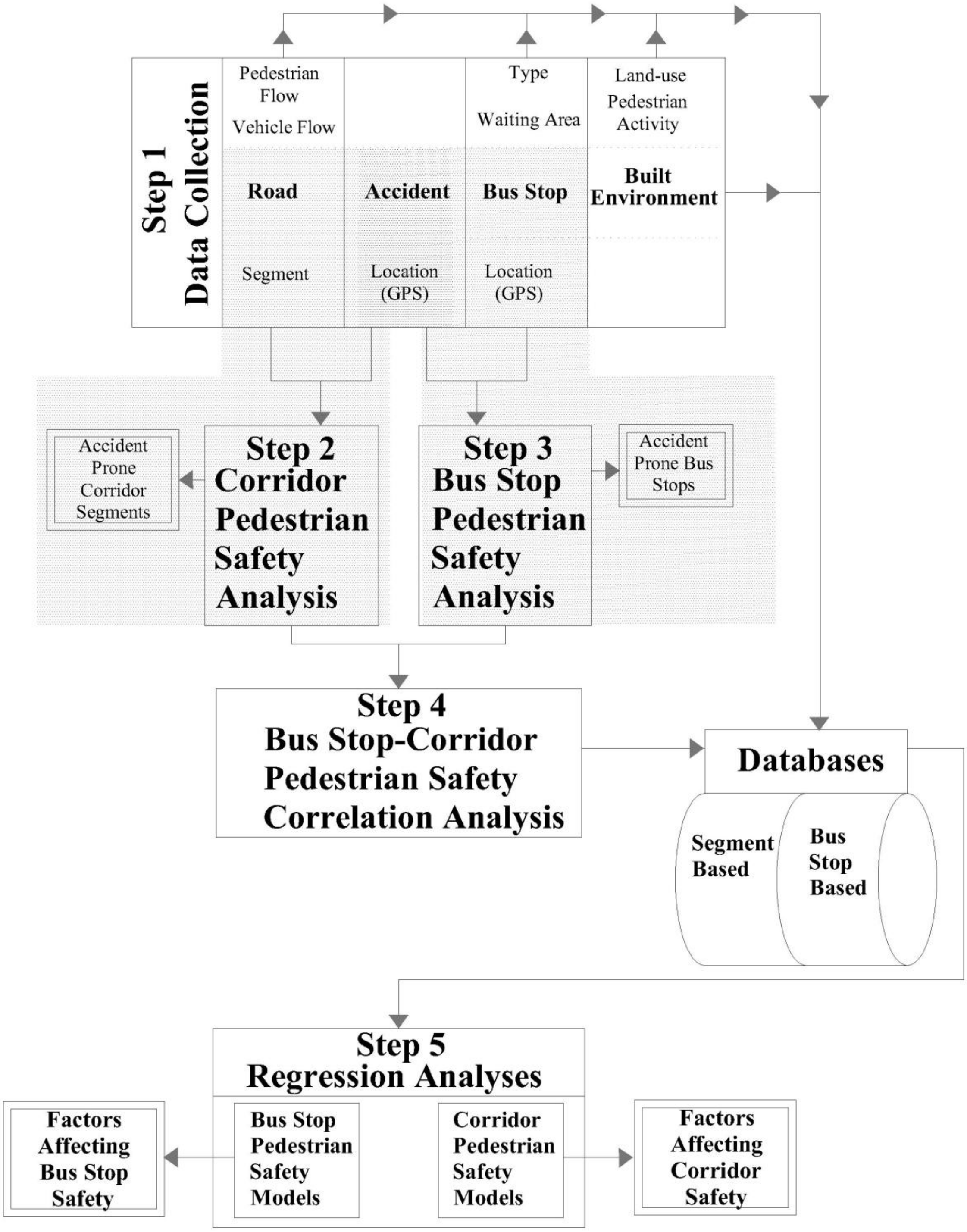


Figure 3.1 Framework of the Proposed Methodology for Pedestrian Safety around Bus Stops

least the location information for the accident (geocoded or as an address entry) for this method.

Similarly, the locations of the bus stops must be digitized and/or geocoded. To investigate the effect of bus stops on pedestrian accidents, it is better to know the physical conditions of bus stops such as, type and waiting area. All such details, if available, should be included in a GIS database and linked to the bus stop location.

Also, the characteristics of segments that the bus stops take place on are important. They may be traffic flow parameters (pedestrian flow level, vehicular flow speed limits, etc.), traffic network parameters (number of lanes, number of intersections, number of legs, etc.) and built environment ones (land use characteristics, number of bus stops, type of urban region, etc.). The more these parameters are defined using measurements, etc. whenever possible, the more their impact can be determined.

Step 2: GIS-based Corridor Pedestrian Safety Analysis

As bus stops would not be the only factor causing pedestrian accidents, it is important to look at the distribution of the pedestrian accidents along the corridor that the bus stops take place. For this purpose, road corridors must be digitized and matched with the accident data. In traffic analysis, generally, road segments (also called links) are defined as parts along which geometry or accessibility does not change, such as parts between intersections. If links are too long, such as in highways, accident analyses for black spots, group accident data for every kilometer to create a uniform road segment; this way total number of accidents also reflects the accident density on that road segment.

For an urban region, especially with Central Business Districts (CBDs), link lengths can be very uneven and affect the accident analysis based on number of accidents on a link. Creating a uniform distance, such as 1km portioning in highways, is not a straightforward process. As an improvement, it is possible to work with accident density (average accident per distance), as this measure also considers segment length inherently. However, this may still be not enough, as there may be localized “accident-prone locations” along long links, which could be lost during the average density calculation. It is important to create a uniform segmented structure for the urban road networks, which requires a pre-processing of the road network. The details of this preprocessing will be discussed later in Section 3.1.2.

It should be noted that the corridor-based approach discussed here can be easily expanded to a regional safety analysis, where all the pedestrian accidents are matched against all the bus stops; however, this requires data collection within the whole region.

Step 3: GIS-based Bus Stop Pedestrian Safety Analysis

Due to the direct and indirect involvement of buses in pedestrian accidents, and the data reliability problems, it is not meaningful to seek accidents exactly at the bus stops. Instead, number of pedestrian accidents within a zone around the bus stops (which can be called "bus stop buffer zone") should be calculated. The crucial issue is the decision of the size of this zone. There is no predefined limit on the "impact distance" of bus stops on pedestrian safety. As a result, it is possible to work with a range of impact distance to capture different interactions between the bus stops and pedestrian safety. These distances can be as small as 20m-50m to look at the bus stop safety in microscopic way, or as large as 400m, which is assumed as the walking distance in urban networks.

Step 4: Bus Stop-Corridor Pedestrian Safety Correlation Analysis

Once the pedestrian safety measure of the corridors and bus stops are determined, it should be checked if there is any correlation between the two. A simple correlation plot of number of accidents around a bus stop and the pedestrian accident density of the corridor segment on which the bus stop is located can be created. Plotting multiple graphs for the same location and bus stop based on different definitions of "bus stop buffer zone" will give an idea about the sensitivity of the correlation to the assumed parameter of impact distance. The form of this correlation can be further sought after by trying to fit some trend lines between the two factors.

Step 5: Regression Analysis

Finally, all the corridor and bus stop analyses' results should be added to the data collected at

Step 1 to create a pedestrian safety database. It would be possible to develop regression models as a follow-up study using this database. Safety measures (total number of accidents, accident density, etc.) are the dependent variables, while the characteristics of the segments of corridors are the independent ones in the regression analyses. Eventually, the focus of this step is to select parameters that can explain the pedestrian safety problem for a region in a statistically significant way. In the following sections, all of these steps will be discussed in detail.

3.1 Corridor Pedestrian Safety Analysis

In corridor pedestrian safety analysis, it is imperative to define a) the corridor itself and b) analysis units to study it, as well as c) the measures needed to assess the pedestrian safety along it. A road itself has a width, which definitely requires a definition of the "thickness" of the study corridor. Furthermore, pedestrian accidents on the sidewalks have been seen when an out-of-control vehicle had a "run-off road" type of an accident. In addition, problems with the accident location geocoding may also require this kind of corridor stripe definition as shown in Figure 3.2. In this study, accidents along a corridor are selected based on the street information in the data first; consulting with the GDS experts, a buffer zone with a thickness of 100 m on both sides of the centerline of the road is assumed to define the corridor itself to tolerate the error margin of the GPS device.

Even though the corridor and distribution of accidents along can be defined as continuous variables, it is not practical to develop any continuous pedestrian safety measure for a long corridor or a regional study. That is why it is necessary to digitize the corridor into smaller analysis units (which are called "segments") and develop safety measures for them. Later, the values of these segment safety measures can be thematically mapped to identify accident-prone segments along the corridor. In a sense, similar to "black spots" on highways, the focus of this analysis can be summarized as the determination of "black segments" along the corridor. For better explanation of this analysis, first, we have to review the potential measures that can be used in corridor pedestrian safety maps, and then discuss the issues regarding the segmentation of an urban corridor.

Figure 3.2 A Generic Urban Corridor Stripe and Accidents on its Links

3.1.1 Traditional Safety Measures

In the literature, there are traditional measures used in GIS analyses, such as, number of accidents, injury measures, severity indexes, etc. Measures used in traffic safety analyses by GDS and General Directorate of Highways (GDH) in Turkey are presented in Table 3.1. In addition, values of the traditional measures for generic corridor can be seen in Table 3.1. The aim of Step 2 of the proposed methodology is to capture the accident-prone corridor segments in terms of traditional safety measures. The thematic maps of the generic corridor that highlights the accident-prone segments can be seen in Figure 3.3. While the number of accidents (N_A) would capture locations with high number of accidents, thematic maps of the number of injury accidents (N_{IA}) and the number of fatality accidents (N_{FA}) would highlight locations, which are prone to respectively, injury and fatality accidents specifically.

The major differences between critical segments are in the number of injuries (N_I) and the number of fatalities (N_F) maps. Their difference can be clearly observed in thematic maps (see Figure 3.3 (d), (e) and (f)) prepared for the generic road corridor. Although the severity index (SI_A and SI_B) is defined as traditional measure, calculation of it is still a debated subject. GDH in Turkey accepted the calculation of SI_A. However, the Property Damaged Only (PDO) accidents are not reported since 2001 in Turkey. Thus, by modifying these two equations severity can be calculated leaving out the PDO and Property Damaged Vehicles (PDV) terms (Sjolinder et al., 2001).

Table 3.1 Traditional Measures Used in Traffic Safety Analyses in Turkey

Safety Measure	Traditional Definition	
Number of Accidents	$\mathbf{N_A}$	Total # of accidents on a road segment/link
Injury Measures		
Number of Injury Accidents	$\mathbf{N_{IA}}$	Total # of injury accidents on a road segment/link
Total Injuries	$\mathbf{N_I}$	Total # of injuries in all the accidents on a road segment/link
Fatality Measures		
Number of Fatality Accidents	$\mathbf{N_{FA}}$	Total # of fatality accidents on a road segment/link
Total Fatalities	$\mathbf{N_F}$	Total # of injuries in all the accidents on a road segment/link
Severity Indexes	SI_A*	$PDO+3*N_{IA}+9*N_{FA}$
	SI_B**	$PDV+3*N_I+9*N_F$

*PDO: number of property damage only accidents
**PDV: number of damaged vehicles

Table 3.2 Values of Traditional Measures for the Generic Corridor

Link		**Number of Accidents**			**Total Injuries (N_I)**	**Total Fatalities (N_F)**	**Severity (SI)**
ID	**Length (km)**	**Total (N_A)**	**Injury (N_{IA})**	**Fatality (N_{FA})**			
1	0.150	6	1	5	30	5	45
2	0.330	20	17	3	18	8	42
3	0.970	30	23	7	23	7	44
4	0.200	12	11	1	11	10	41

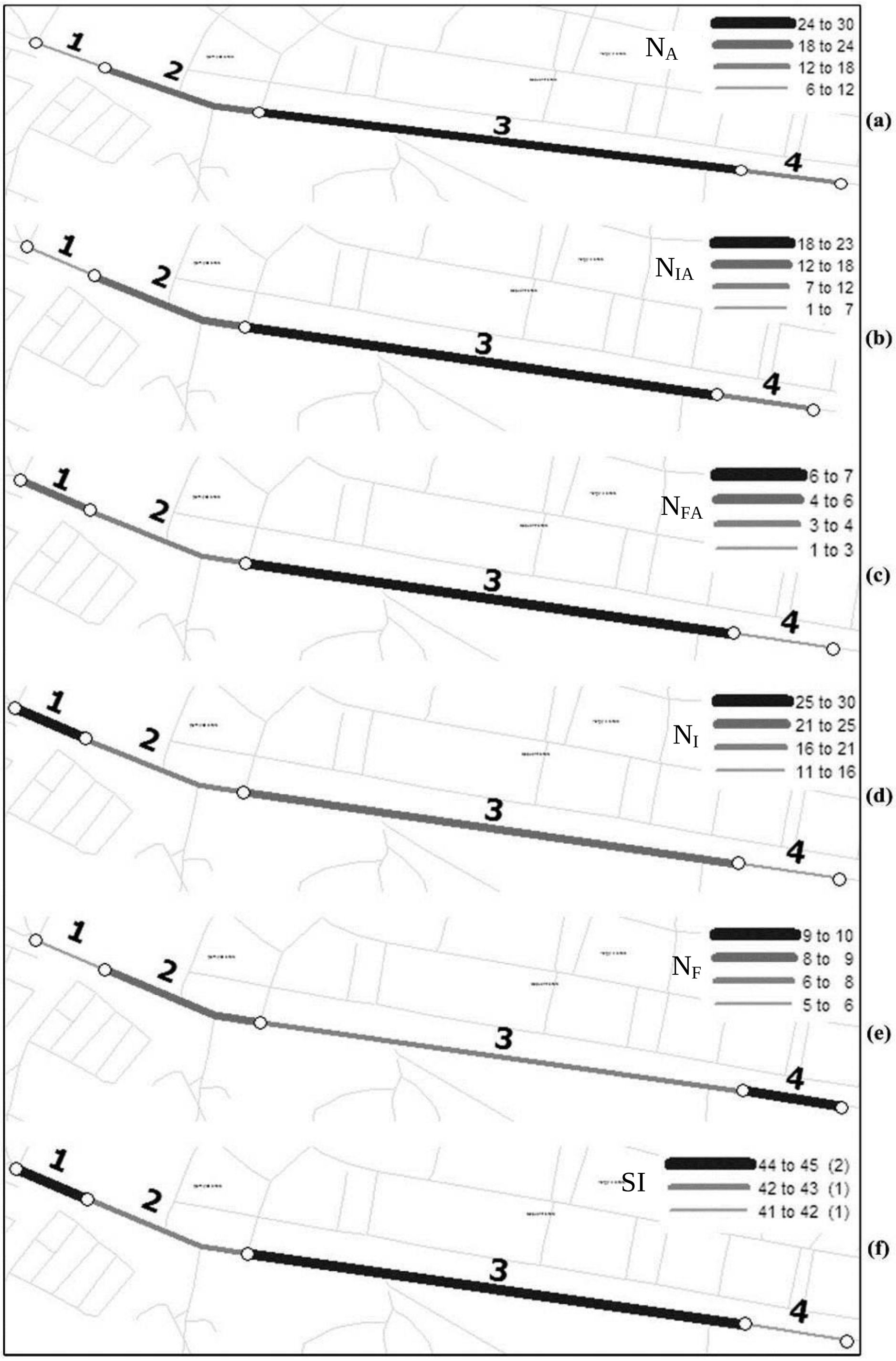


Figure 3.3 Thematic Maps of (a) the Number of Accidents, N_A, (b) the Injury Accident, N_{IA}, (c) the Fatality Accident, N_{FA}, (d) the Number of Injuries, N_I, (e) the Number of Fatalities, N_F, (f) the Severity Index, SI

3.1.2 A Simple Improvement for Safety Mapping: Density Mapping

In a road network, segments are defined as parts along which geometry or accessibility does not change, such as parts between intersections. Thus, the length of segments can be various. The same situation can be observed for the generic road corridor that is given in Figure 3.2. Lengths of some segments of the corridor are longer than the other ones. The thematic maps were created according to the number of accidents, injury accidents, etc. that took place along the segments. Thus, this resulted in the same critical levels for both of long and short segments that have the same number of accidents. To avoid this situation, the density measures are created by dividing the traditional measures by segment length.

The definitions and formulas of density measures can be seen in Table 3.3. Eventually, using density measures makes the shorter segment more critical than the longer one, which both of them has the same number of accidents. Thus, the density measures are calculated and thematic maps are created for generic road corridor again to visualize the situation clearly. Both of them can be observed in Figure 3.4 and Table 3.4. The change in the accident-prone segments can be identified by comparing the Figure 3.3 and Figure 3.4.

Table 3.3 The Definitions of Density Safety Measures

Traditional Measure	Density Measure		Definition
N_A	$D_A = N_A / l_i$	Accident Density	Total # of accidents per unit length
N_{IA}	$D_{IA} = N_{IA} / l_i$	Injury Accident Density	Total # of injury accidents per unit length
N_{FA}	$D_{FA} = N_{FA} / l_i$	Fatality Accident Density	Total # of fatality accidents per unit length
N_I	$D_I = N_I / l_i$	Injury Density	Total # of injuries per unit length
N_F	$D_F = N_F / l_i$	Fatality Density	Total # of fatalities per unit length

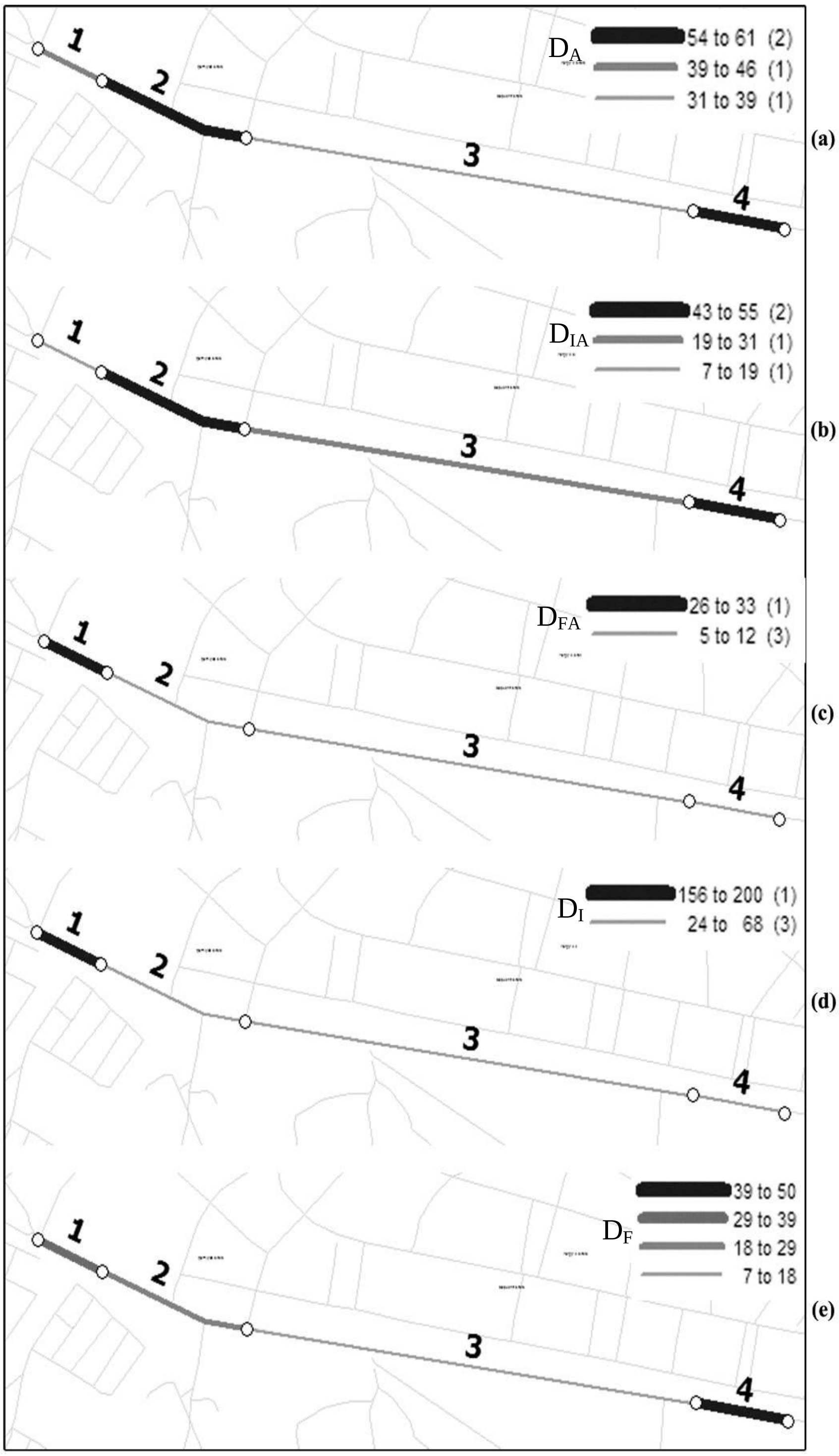


Figure 3.4 Thematic Maps of (a) the Accident Density, D_A, (b) the Injury Accident Density, D_{IA}, (c) the Fatality Accident Density, D_{FA}, (d) the Injury Density, D_I, (e) the Fatality Density, D_F

Table 3.4 The Values of Density Measures for the Generic Corridor

Link		Accident Density			Injury Density (D_I)	Fatality Density (D_F)
ID	Length	Total (D_A)	Injury (D_{IA})	Fatality (D_{FA})		
1	0.150	40	7	33	200	33
2	0.330	61	52	9	55	24
3	0.970	31	24	7	24	7
4	0.200	60	55	5	55	50

3.1.3 Safety Mapping on Urban Traffic Networks

Although, more realistic evaluation was created by using density measures in determining the accident prone segments, there is still a persisting problem in the analysis due to the fact that the geographical distribution of accidents are not regular along the corridor. This means that for long road links, although most of the accidents cluster in a region along the link, whole link would be highlighted in the thematic maps as critical one. This unequal distribution of accidents results underestimation for long segments in terms of density measures.

Thus, to overcome this problematic situation, the division of links into unit segments that have the same length as far as possible is needed. The length of unit segment is determined as 0.150-0.300 km. For a real network, the unit segmentation of links means that the long links are divided into the length of unit segment (disaggregation) and the short ones are compounded (aggregation). For generic road corridor segments, there is no link shorter than 0.150 km. Thus, only disaggregation process was carried out. The new segmentation of generic road corridor, and the values of traditional and the density measures of new segments can be observed in Figure 3.5 and Table 3.5, respectively. In the thematic maps of new segments, it is clearly seen that criticality of the new-segmented long links changes. By this method, the real accident-prone segments can be identified in terms of traditional and density measures (see Figure 3.6 and Figure 3.7).

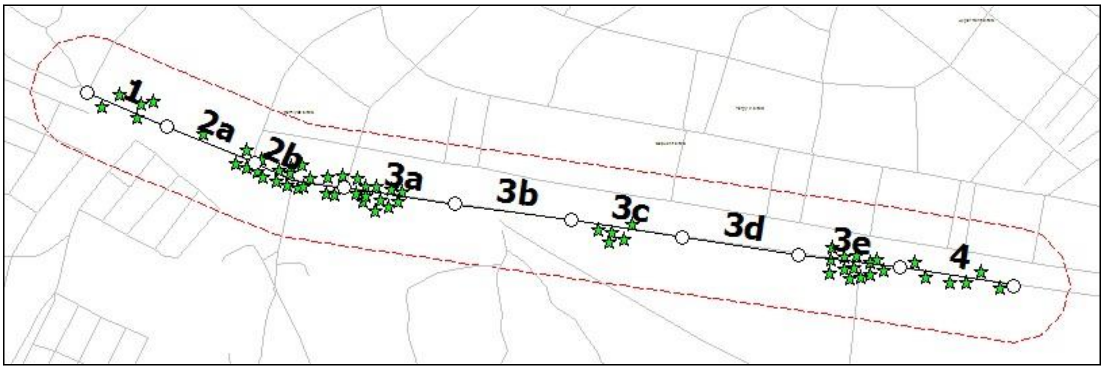


Figure 3.5 The Segmentation of the Generic Corridor

Table 3.5 The Values of Traditional and Density Measures for the Segmented Generic Corridor

Traditional Measures						
Segment		**Number of Accidents**			**Segment Total Injuries (SN_I)**	**Segment Total Fatalities (SN_F)**
ID	**Length (km)**	**Segment Total (SN_A)**	**Segment Injury (SN_{IA})**	**Segment Fatality (SN_{FA})**		
1	0.150	6	1	5	30	5
2a	0.165	4	2	0	2	0
2b	0.165	16	15	3	16	8
3a	0.200	12	6	1	6	1
3b	0.200	0	0	0	0	0
3c	0.200	5	4	0	5	0
3d	0.200	0	0	0	0	0
3e	0.180	13	13	6	12	6
4	0.200	12	11	1	11	10
Density Measures						
Segment		**Accident Density**			**Segment Injury Density (SD_I)**	**Segment Fatality Density (SD_F)**
ID	**Length (km)**	**Segment Total (SD_A)**	**Segment Injury (SD_{IA})**	**Segment Fatality (SD_{FA})**		
1	0.150	40	7	34	202	34
2a	0.165	24	12	0	12	0
2b	0.165	97	91	18	97	48
3a	0.200	60	30	5	30	5
3b	0.200	0	0	0	0	0
3c	0.200	25	20	0	25	0
3d	0.200	0	0	0	0	0
3e	0.180	72	72	33	67	33
4	0.200	60	55	5	55	50

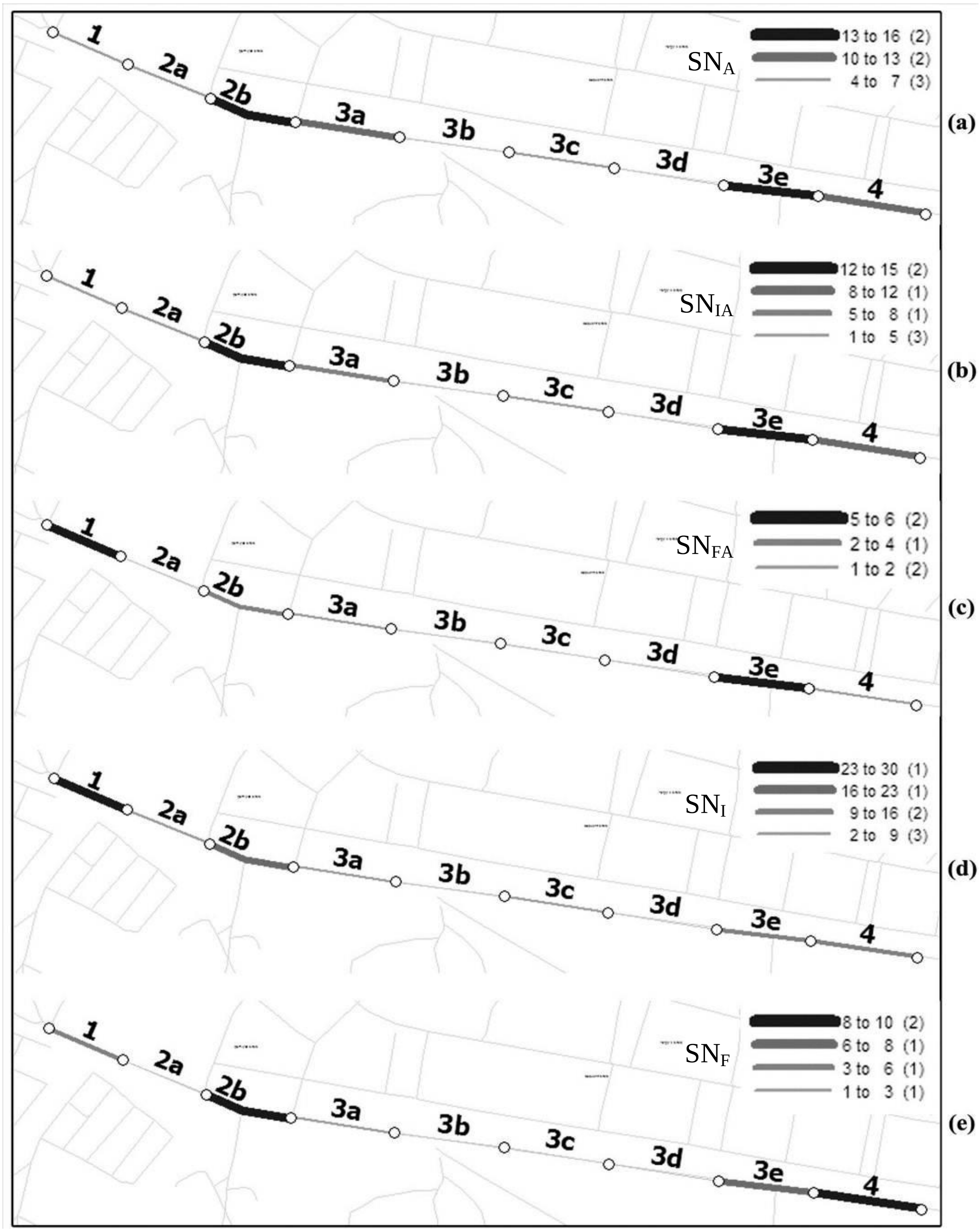


Figure 3.6 Thematic Maps of Segment (a) the Number of Accidents, SN_A, (b) the Injury Accidents, SN_{IA}, (c) the Fatality Accident SN_{FA}, (d) the Number of Injuries, SN_I, (e) the Number of Fatalities, SN_F

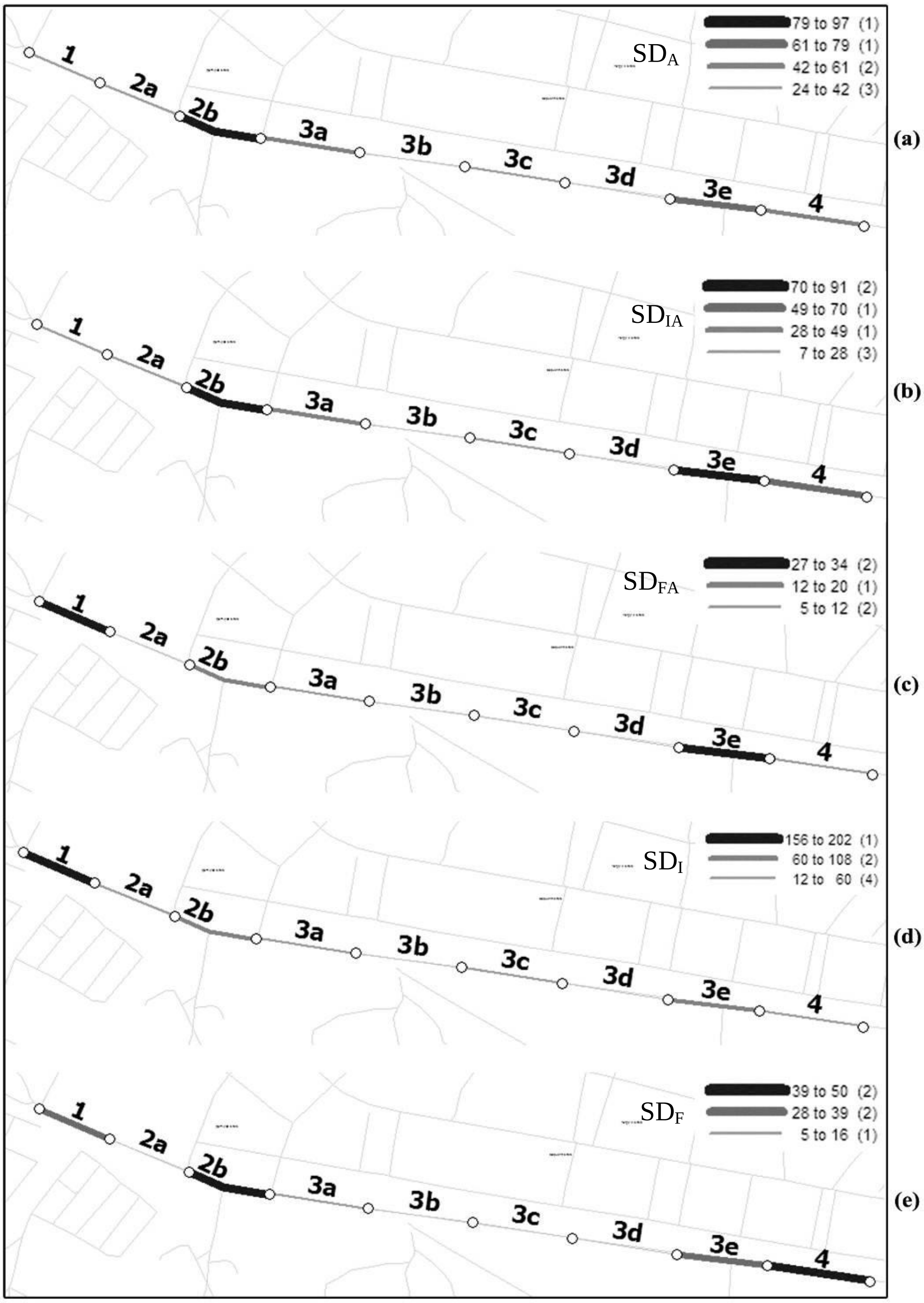


Figure 3.7 Thematic Maps of Segment (a) the Total Accident Density, SD_A, (b) the Injury Accident Density, SD_{IA}, (c) the Fatality Accident Density, SD_{FA}, (d) the Injury Density, SD_I, (e) Fatality Density, SD_F

When the unit segmentation process is carried on, the traditional and density values change according to segment lengths. How this change is also an important issue to support the idea of unit segmentation process. Thus, for the change parameters, various notations are created in Table 3.6. After the unit segmentation of generic road corridor, the change in the value of traditional and density measures are represented in Table 3.7. Minus sign describes the decrease in the value of traditional and density measures and zero value defines there is no change in the value after unit segmentation process. For example, Link 3 had 30 accidents, which was representing the total link length. However, after segmentation, Segment 3b had no accidents, that showed a “-30” accident change in Table 3.7.

Sometimes, the size of change is not important, and it is necessary to look at the percent change in the values before or after the segmentation, as shown in Table 3.8. The segments that have 100% changes do not have any accident anymore after unit segmentation; these are non-accident prone locations that were originally labeled due to the use of traditional link-based definitions.

Table 3.6 The Definition of Parameters of Change due to Segmentation

Change in Number	Definition	Percent Change	Definition
ΔN_A	$= SN_A - N_A$	δN_A	$= \Delta N_A / N_A * 100$
ΔN_{IA}	$= SN_{IA} - N_{IA}$	δN_{IA}	$= \Delta N_{IA} / N_{IA} * 100$
ΔN_{FA}	$= SN_{FA} - N_{FA}$	δN_{FA}	$= \Delta N_{FA} / N_{FA} * 100$
ΔN_I	$= SN_I - N_I$	δN_I	$= \Delta N_I / N_I * 100$
ΔN_F	$= SN_F - N_F$	δN_F	$= \Delta N_F / N_F * 100$
ΔD_A	$= SD_A - D_A$	δD_A	$= \Delta D_A / D_A * 100$
ΔD_{IA}	$= SD_{IA} - D_{IA}$	δD_{IA}	$= \Delta D_{IA} / D_{IA} * 100$
ΔD_{FA}	$= SD_{FA} - D_{FA}$	δD_{FA}	$= \Delta D_{FA} / D_{FA} * 100$
ΔD_I	$= SD_I - D_I$	δD_I	$= \Delta D_I / D_I * 100$
ΔD_F	$= SD_F - D_F$	δD_F	$= \Delta D_F / D_F * 100$

Table 3.7 Change in Pedestrian Safety Measures along the Generic Corridor due to Segmentation

ID	ΔN_A	ΔN_{IA}	ΔN_{FA}	ΔN_I	ΔN_F	ΔD_A	ΔD_{IA}	ΔD_{FA}	ΔD_I	ΔD_F
1	0*	0	0	0	0	0	0	0	0	0
2a	-16	-15	-3	-16	-8	-37	-40	-9	-43	-24
2b	-4	-2	0	-2	0	36	39	9	42	24
3a	-18	-17	-6	-17	-6	29	6	-2	6	-2
3b	-30	-23	-7	-23	-7	-31	-24	-7	-24	-7
3c	-25	-19	-7	-18	-7	-6	-4	-7	1	-7
3d	-30	-23	-7	-23	-7	-31	-24	-7	-24	-7
3e	-17	-10	-1	-11	-1	41	48	26	43	26
4	0	0	0	0	0	0	0	0	0	0

* "0" represents no change in value
** Minus sign represents the decrease in value

Table 3.8 Percent Change in Pedestrian Safety Measures along the Generic Corridor due to Segmentation

ID	δN_A	δN_{IA}	δN_{FA}	δN_I	δN_F	δD_A	δD_{IA}	δD_{FA}	δD_I	δD_F
1	0	0	0	0	0	0	0	0	0	0
2a	-80	-88	-100	-89	-100	-61	-77	-100	-78	-100
2b	-20	-12	0	-11	0	59	75	100	76	100
3a	-60	-74	-86	-74	-86	94	25	-29	25	-29
3b	-100	-100	-100	100	-100	-100	-100	-100	-100	-100
3c	-83	-83	-100	-78	-100	-19	-17	-100	4	-100
3d	-100	-100	-100	100	-100	-100	-100	-100	-100	-100
3e	-57	-43	-14	-48	-14	132	200	371	179	371
4	0	0	0	0	0	0	0	0	0	0

In Figure 3.8, the thematic maps of traditional and density safety measures are visualized before and after unit segmentation process. It can be clearly seen that the critic segments are changed after a new measure were introduced to the analysis, with a more precision.

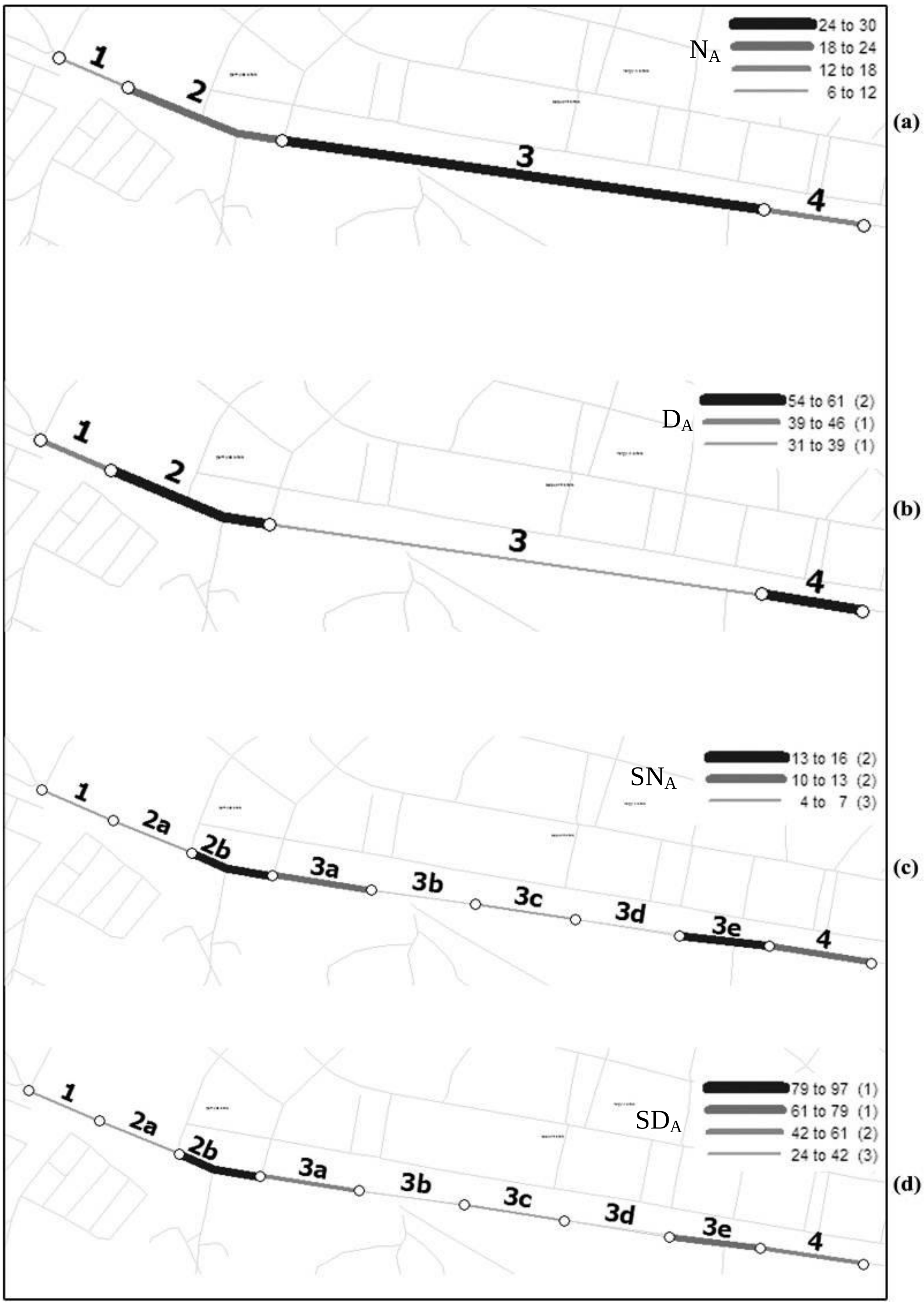


Figure 3.8 Comparison of the Thematic Maps of (a) Number of Accidents, (b) Accident Density, (c) Segment Number of Accident, and (d) Segment Accident Density for the Generic Corridor

Although the unit segmentation process was discussed throughout the disaggregation of links into unit segment due to the selection of the unit segment length, the original network required the disaggregation process, in some cases aggregation could be a need especially for links with small lengths. In this situation, original network links can be aggregated to obtain unit segment length.

3.2 Bus Stop Pedestrian Safety Analysis

The aim of bus stop pedestrian safety analysis is to investigate the accident-prone bus stops along the selected corridors by using GIS based thematic maps in terms of traditional safety measures. In a GIS based analysis proposed, to measure pedestrian safety around bus stops, it is important a) to define the bus stop itself geographically, first and b) to associate the stop with the pedestrian safety measures around it. From this point on, whenever the term "stop" is used, it should be perceived as "bus stop". It should be noted that, despite some differences in design and the right-of-way, the proposed methodology could be easily extended for the GIS-based analysis of pedestrian safety around tram stops, too.

To define the bus stops in a GIS environment, stop locations with GPS coordinates have to be determined. This can be imported from any other database, if available. If not, bus stop locations have to be collected manually. As discussed in the literature review, the type and design of the bus stop itself may be important in the safety of pedestrians at or around it. Thus, it is also important to collect data on the type of the bus stops (just a post, post with a shelter, stop with bus pockets, etc.), size of the bus stop (shelter area), and service rate of the bus stop (number and frequency of bus lines served by the stop). The details may be more important in defining the pedestrian activity at a stop in regression analysis.

3.2.1 Bus Stop Buffer Zones

Due to various problems in bus stop and accident data geocoding, and indirect involvement of bus and bus stops in pedestrian accidents, it is not realistic to define the bus stop as a single point on the map. Instead, a "bus stop buffer zone" has to be defined to capture the possible "impact" of a bus stop on pedestrian safety. In addition, it is not easy or clear to choose one “buffer zone” that relates bus stops to pedestrian accidents; instead, a series of buffer zones have to be defined to see what kind of safety relation in terms of different sizes of buffer zones. Shortly, the focus of the proposed GIS-based bus stop pedestrian safety analysis can be summarized as “the selection of the bus stops, which have high number of accident/injury/fatality” by determining the accident/injury/fatality numbers that fall into the selected buffer zones.

Selection of the size of the buffer zones is obviously a critical issue, which can directly influence the number of accidents associated with the stop. According to the research by Planning Commission Transit-Oriented Committee of Fairfax County of Virginia (2012), in urban areas, people are expected to walk 400-450 m to a transit stop. In addition, Clifton et al. (2009) determined 400 m as a buffer zone radius around each accident to capture the built environment characteristics of the study area. Miranda-Moreno et al. (2011) used different buffer zone sizes of 50m, 150m, 400m and 600 m, around the intersections to investigate the impact of buffer dimension.

The smaller the size of the buffer zone is the more immediate safety problems around the stop will be studied. The buffer zone definition and its impact area can be seen in Figure 3.9. The relationship between the accidents and the stops may not be directly investigated, when the buffer zones of adjacent or closer stops overlap (see Figure 3.9 (b), (c) and (d)). In such cases, an accident may be associated with more than one bus stop, and counted for multiple stops, which may be normal, as the injured or killed pedestrians may not necessarily be at their destination stop at the time of the accident. This may be a bigger issue for bus stops in CBD regions with overlapping bus routes, at even smaller buffer zones (50m or 100m). In this study, stops studied were located in two different regions: the CBD region that shows urban pattern with high pedestrian activity and arterials serving the urban transition zones (UTZ) that shows more suburban characteristics. The distances between stops on the CBD region are, thus, short; the 400m-buffer zones of CBD stops are carrying the accident information of many accidents around too many stops to a rather away stop, and are not meaningful. However, the buffer zones of bus stops for CBD and UTZ are chosen as 50 m, 100 m and 400 m.

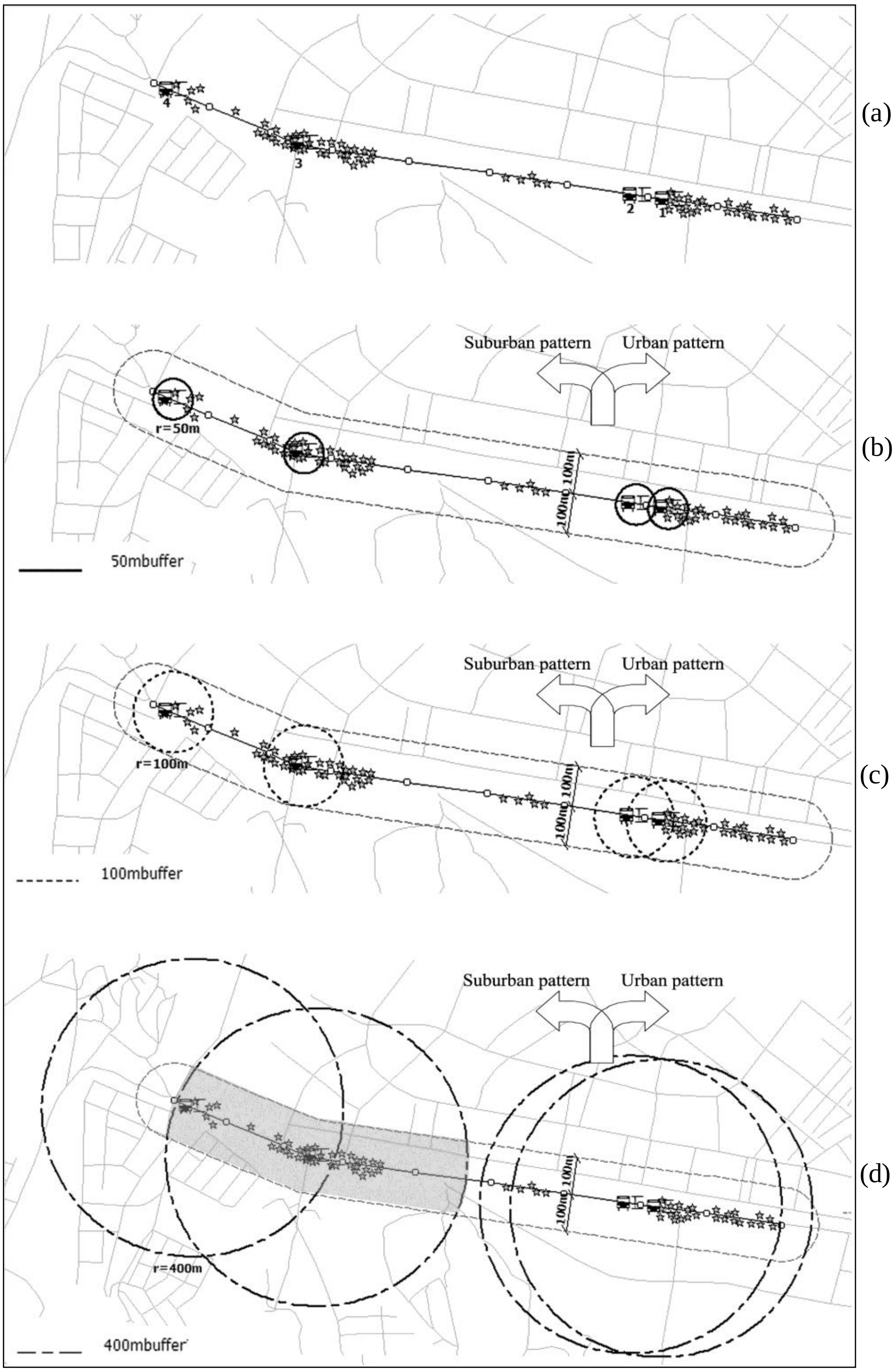


Figure 3.9 Bus Stop (a) Locations and Buffer Zones with impact distances of (b) 50m (B_{50m}), (c) 100m (B_{100m}), (d) 400m (B_{400m})

When the radiuses of the buffer zones are 50 and 100 m, there is no problem with the impact areas because they still in the impact area of corridor buffer zone with radius of 100 m. However when the radius of buffer zone becomes 400 m, the border of the buffer zone of corridor limits the impact area of the buffer zone with radius of 400 m (Figure 3.9 (d))

3.2.2 Bus Stop Pedestrian Safety Measures and Maps

For the generic road corridor, the distribution of bus stops is visualized in Figure 3.9 (a). While the bus stops 1 and 2 take place in urban region, the others, 3 and 4 are in suburban region. Most of the segments of UTZ corridors are in the suburban regions. Thus, generally the distance between the stops in suburban regions are longer than ones in the CBD. The numbers of accidents on the buffer zones of the stops are identified in Table 3.9. Accident-prone bus stops are identified by using thematic maps of traditional safety measures. Accident that place in the buffer zones are counted and assigned to each bus stop. In addition, the number of injury and fatality is determined and thematic maps are created according to these values too. In thematic maps, bus stops that have highest number of accidents to lowest are colored in scale to determine the accident-prone bus stops. In Figure 3.10, thematic maps of bus stops on the generic road corridor can be observed. As it can be seen, when the radius of buffer zone changes, the criticality of bus stops also changes. However, still the most accident-prone bus stop is the stop 3 in the suburban zone.

Table 3.9 Bus Stop Safety Statistics for the Generic Corridor

Bus Stop ID	No. of Accidents in Buffer Zone with Radius of 50m ($B_{50m}\_N_A$)	100m ($B_{100m}\_N_A$)	400m ($B_{400m}\_N_A$)
1	8	13	29
2	0	3	30
3	10	18	38
4	4	6	24

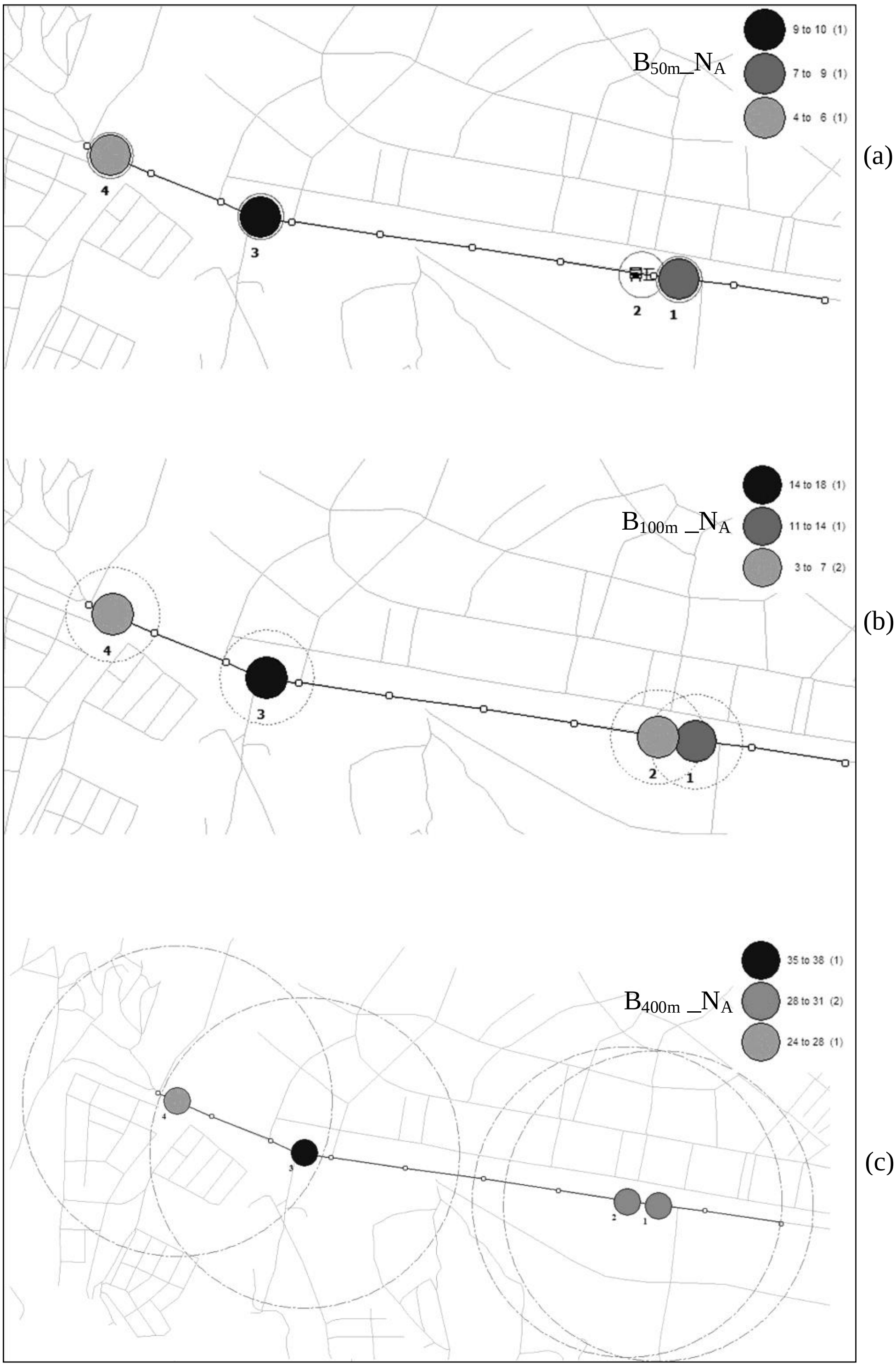
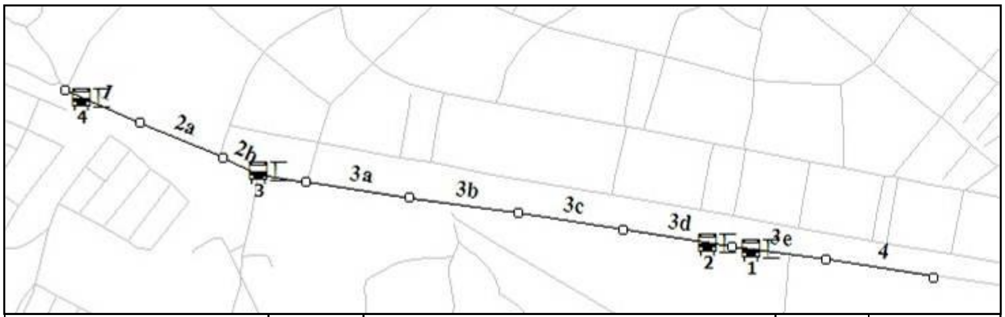


Figure 3.10 Thematic Maps of Bus Stops for the Buffer Zones of (a) 50m, $B_{50m}\_N_A$, (b)100m, $B_{100m}\_N_A$, (c) 400m, $B_{400m}\_N_A$

3.3 Bus Stop-Corridor Pedestrian Safety Correlation Analysis

To investigate the relationship between the bus stops and the segments of the corridors that the bus stops take places, a correlation analysis is conducted. In Figure 3.11, the distribution of bus stops and which bus stops belong to which segments and also safety measures of the segments and bus stops can be observed. The aim of the correlation analysis is to identify how the relationship between segments of corridor and bus stops in terms of the number of accidents. Road segments' accident density and the number of accidents in the buffer zones of corresponding bus stops are the parameters of the correlation. The results of the correlation are visualized in Figure 3.12 as a scatter graph. For the number of accidents in 50 m buffer zones of stops and corresponding corridor segments, there is almost 100% linear relationship with the value of R^2 0.9955. The same correlation is valid for 100 m buffer zones although the linear equation has smaller R^2 value, 0.9513. R^2 value of the correlation of 400 m buffer zones is rather small because 400 m buffer zones cover wider areas than the buffer zones of segments that they will be correlated with. Thus, number accidents that fall into 400 m buffer zones of bus stops can be higher even though, the segment that bus stops take place has no accidents on it. This is the reason of weakness of the relationship between number of accidents in 400 m buffer zones of stops and the accident density on the segment of stops take place. As a result, for such a small generic road corridor, there is a strong linear relationship between the corridor and bus stop pedestrian safety.



Segment			Bus Stop ID	Bus Stop Accidents			$\Delta_{100m-50m}$	Nearest Bus Stop Distance (km)
ID	SD_A	Length (km)		$B_{50m}\_N_A$	$B_{100m}\_N_A$	$B_{400m}\_N_A$		
1	40	0.150	4	4	6	24	2	0.086
2a	24	0.165		0	0		0	
2b	97	0.165	3	10	18	38	8	0.086
3a	60	0.200		0	0		0	
3b	0	0.200		0	0		0	
3c	25	0.200		0	0		0	
3d	0	0.200	2	0	3	30	3	0.358
3e	72	0.180	1	8	13	29	5	0.358
4	60	0.200		0	0	0	0	

Figure 3.11 The Location and Safety Statistics for the Bus Stops on the Generic Corridor

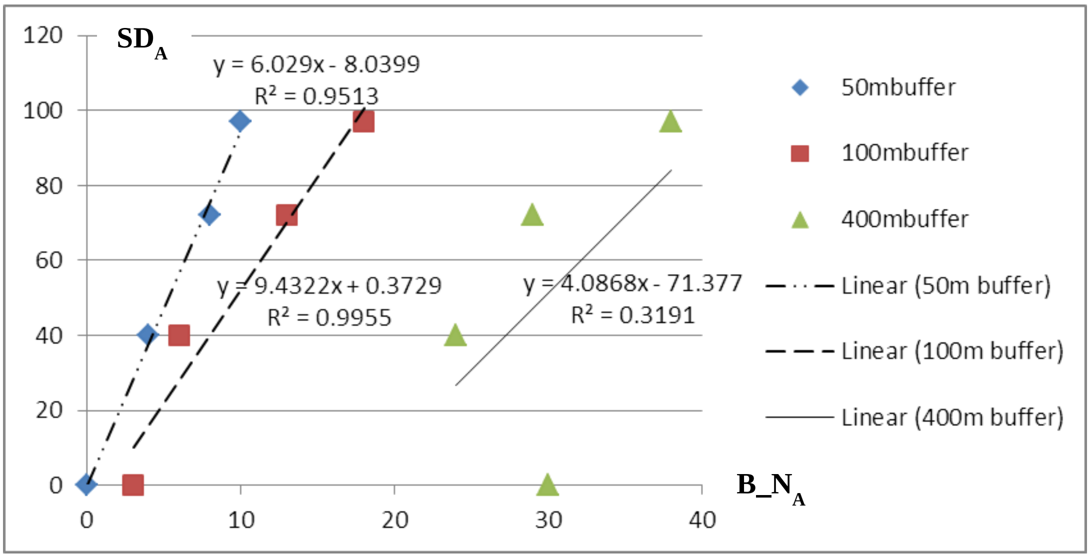


Figure 3.12 The Graph of Corridor and Bus Stop Pedestrian Safety Correlation Results

3.4 Pedestrian Safety Regression Analysis

The simple form of a regression analysis, a linear regression, is presented in Section 2.1.3. The application of the format for the pedestrian safety problem around the bus stops can be in different ways based on the selected dependent and independent variables. First, number of accidents of segments or accident densities can be modeled as a function of traffic flow, built-environment and network parameters as follows:

$$SN_A = \beta_o^{N_A} + \beta_1^{N_A} x_{i1} + \beta_2^{N_A} x_{i2} + ... + \beta_p^{N_A} x_{ip} + \varepsilon_i^{N_A} \quad i = 1,.....,n_{segment} \qquad \text{(Eqn. 3)}$$

$$SD_A = \beta_o^{D_A} + \beta_1^{D_A} x_{i1} + \beta_2^{D_A} x_{i2} + ... + \beta_p^{D_A} x_{ip} + \varepsilon_i^{D_A} \quad i = 1,.....,n_{segment} \qquad \text{(Eqn. 4)}$$

Secondly, the number of accidents associated with the bus stops (with *m* representing the selected buffer zone radius) can be modeled as a function of the selected parameters as follows:

$$B_m\_N_A = \beta_o^{B_m} + \beta_1^{B_m} x_{i1} + \beta_2^{B_m} x_{i2} + ... + \beta_p^{B_m} x_{ip} + \varepsilon_i^{B_m} \quad i = 1,.....,n_{segment} \qquad \text{(Eqn. 5)}$$

Ultimately, statistically significant variables in the developed models would yield information about their relation to the selected pedestrian safety on segments or at bus stops.

In the literature, the dependent and independent variables vary according to the study region and the data that is available. Clifton et al. (2009) used severity of injuries as dependent variables in their statistical analysis. In addition, personal characteristics of injured people, weather and road conditions, road facility type, existence of transit access, connectivity, population information, land use characteristics (commercial, residential and mixed) etc. were used as independent variables. Also, another study that was conducted by Miranda-Moreno et al. (2011) used independent variables of land use characteristics (commercial, residential, industrial etc.), demographic information, transit characteristics (presence of metro station, number of bus stops etc.), road network connectivity (road length, number of street segments, number of intersections etc.).

As it can be seen from the studies in the literature, if the data exists, several variables can be used in the regression models. The independent variables that will be used in regression analyses of this study were shown in Table 3.10. In addition to these variables shopping land use characteristic at the one side of the segments and existence underpass or overpass on the segments were excluded from the analyses that they were not statistically significant in any of the models. Actually, for analyses that are more reliable some other variables that could have been included in the regression models. Based upon the studies in the literature, these variables should be sidewalks, bus stop waiting area, existence of traffic control systems at the intersections (signalized or not), quantified pedestrian volume, quantified speed, personal characteristics, accident characteristics, weather and road conditions, bus stop service statistics etc.

Table 3.10 Independent Variables Used in Pedestrian Safety Regression Models

Variable Name	Type	Description
Traffic Flow		
D_HighSpeed	Dummy	High speed (If speed limit<50; 0; otherwise, 1)
D_ HighPedVol	Dummy	High pedestrian volume, decided by observation and land use.
Traffic Network		
No_Lanes	Integer	Total number of lanes in the road cross-section in the segment or two-way road
Seg_L_km	Real	Length of the segment in km.
No_Intersections	Integer	Number of intersections in the road segment
No_Legs	Integer	Number of legs (including the segment) that cut the segment
Built Environment		
D_Urban zone	Dummy	
D_Shop	Dummy	Existence of shopping zones in one of the sides of the segment ***eitherside***
D_Shop_2sides	Dummy	Existence of shopping zones on **both sides of road segment**
D_BusComInd	Dummy	Existence of **business, commercial or industrial zones** in either or both sides of the road segment
D_PedBarr	Dummy	Existence of a barrier (traffic island, pedestrian guard rails, obstacles) along the road segment
D_Overpass	Dummy	Existence of overpass(es) along the road segment.
D_Underpass	Dummy	Existence of underpass along the road segment
D_Under_Over_pass	Dummy	Existence of under- or over-pass(es) along the road segment.
No_Bus_Stops	Integer	Number of bus stops along the segments of the corridors

CHAPTER 4

CASE STUDY: PEDESTRIAN SAFETY IN MAJOR ARTERIALS IN ANKARA

As a case study, the main corridors and the CBD region of Ankara was selected due to the high number of pedestrian accidents that had been occurred between the years 2007-2009 (Figure 4.1). The distribution of total pedestrian and run-off road accidents throughout the city is shown in Figure 4.2. Total number of accidents in three years is 8749 out of which 7833 accidents are pedestrian accidents. The statistics about the number of injuries and fatalities in these accidents were summarized in Table 4.1. In 7833 pedestrian accidents, 8410 pedestrian were injured and 120 died (GDS, 2012). Actually, in safety analyses that were conducted in the literature, accident rate values that are calculated by using traffic counts were used. However, the traffic counts are not available for the city of Ankara. Thus, the number of accidents will be used for this study. In addition, all of the maps in this study are produced with the MapInfo Professional 10.0 software.

The selection of study corridors were made according to the density of the pedestrian accidents along the corridors. Due to the characteristics of high pedestrian and motor vehicle volume, the roads in the CBD region were the first choices for the case study. However, when the pedestrian accidents distribution was examined, in some corridors, the number of these accidents is considerably high. Thus, in addition to the roads of the CBD region, four main arterials were also investigated in the case study (Figure 4.3).

A second criterion in the selection of study corridors, especially main arterials, aimed to study corridors with different characteristics. Corridors that have different features were tried to be selected for the analysis. In this chapter, the analysis that were explained in the previous chapter will be applied and the results will be discussed in next chapter

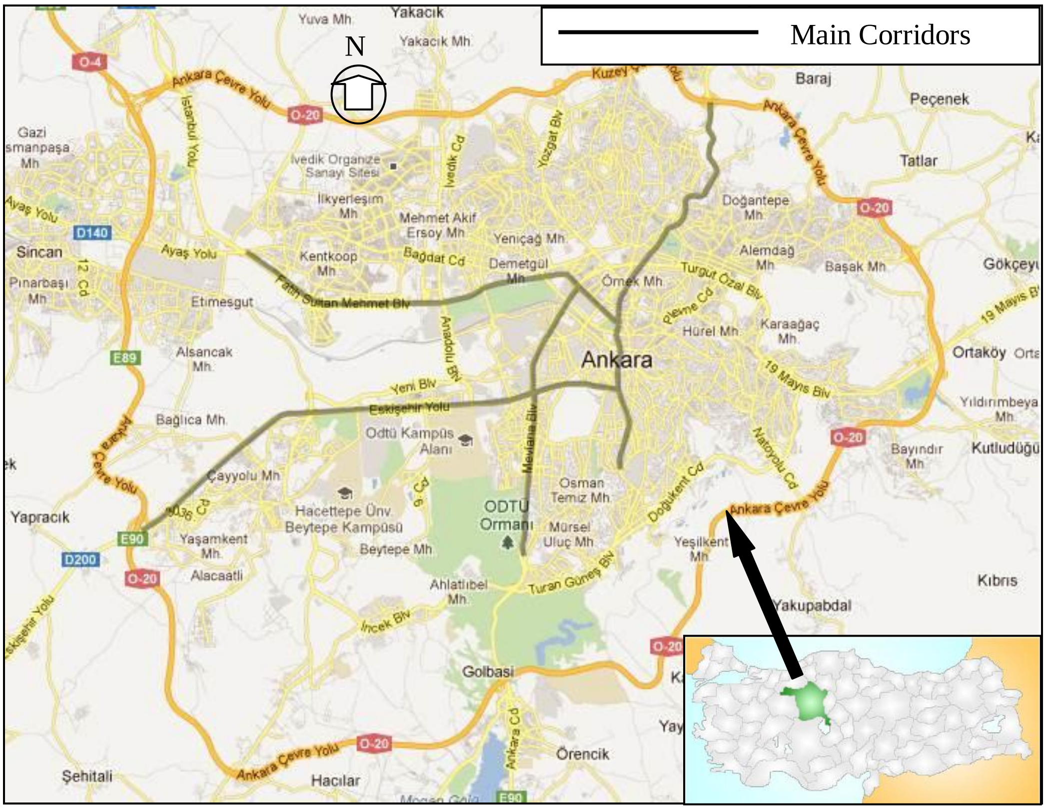


Figure 4.1 City Map of Ankara within the Beltway (Map of Ankara City Center, 2012)

Table 4.1 Number of Pedestrian Related Accidents in Ankara between 2007-2009

	Number of Accidents		**Number of Injuries in**		**Number of Fatalities in**	
Year	**Pedestrian**	**Run-off Road**	**Pedestrian Accidents**	**Run-off Road Accidents**	**Pedestrian Accidents**	**Run-off Road Accidents**
2007	2763	336	2943	626	53	11
2008	2488	265	2656	469	38	8
2009	2582	315	2811	546	29	10
Total	**7833**	**916**	**8410**	**1641**	**120**	**29**

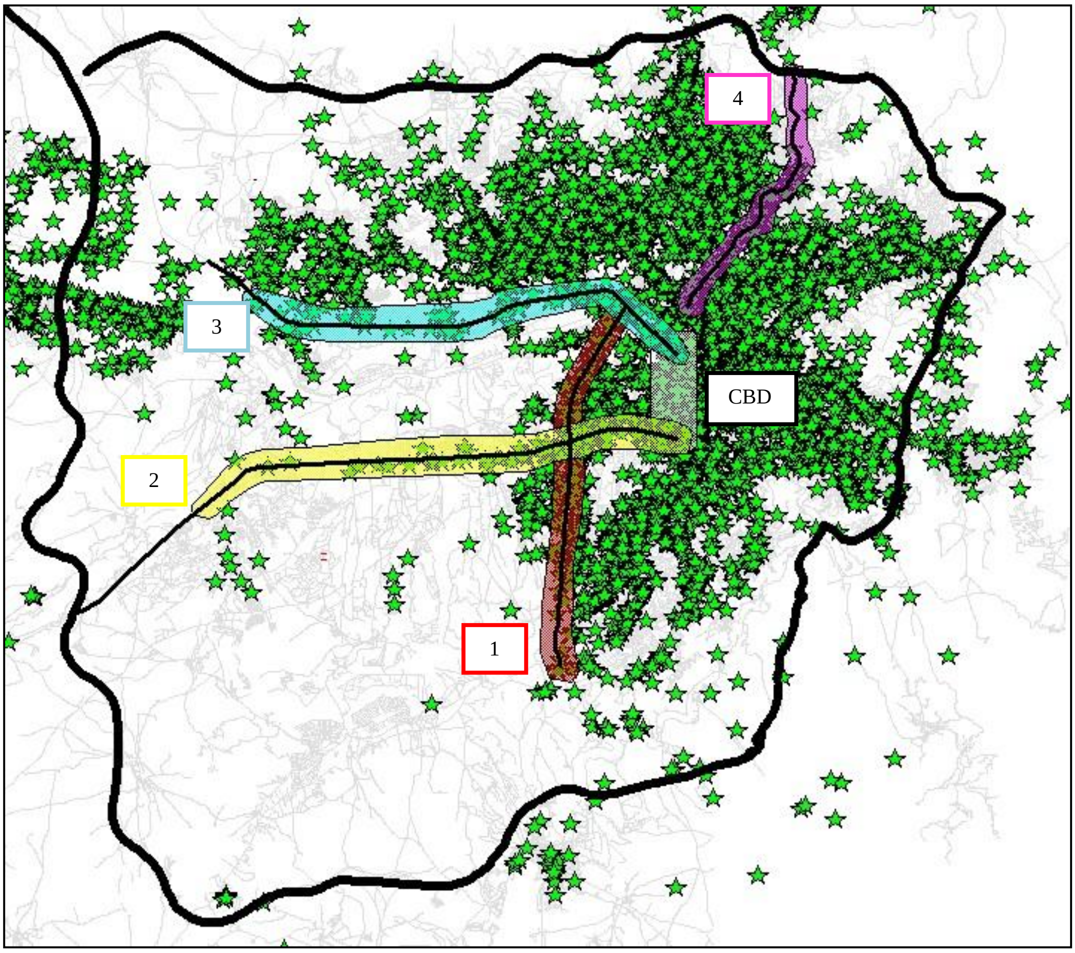


Figure 4.2 Distribution of Pedestrian Related Accidents in Ankara between 2007-2009

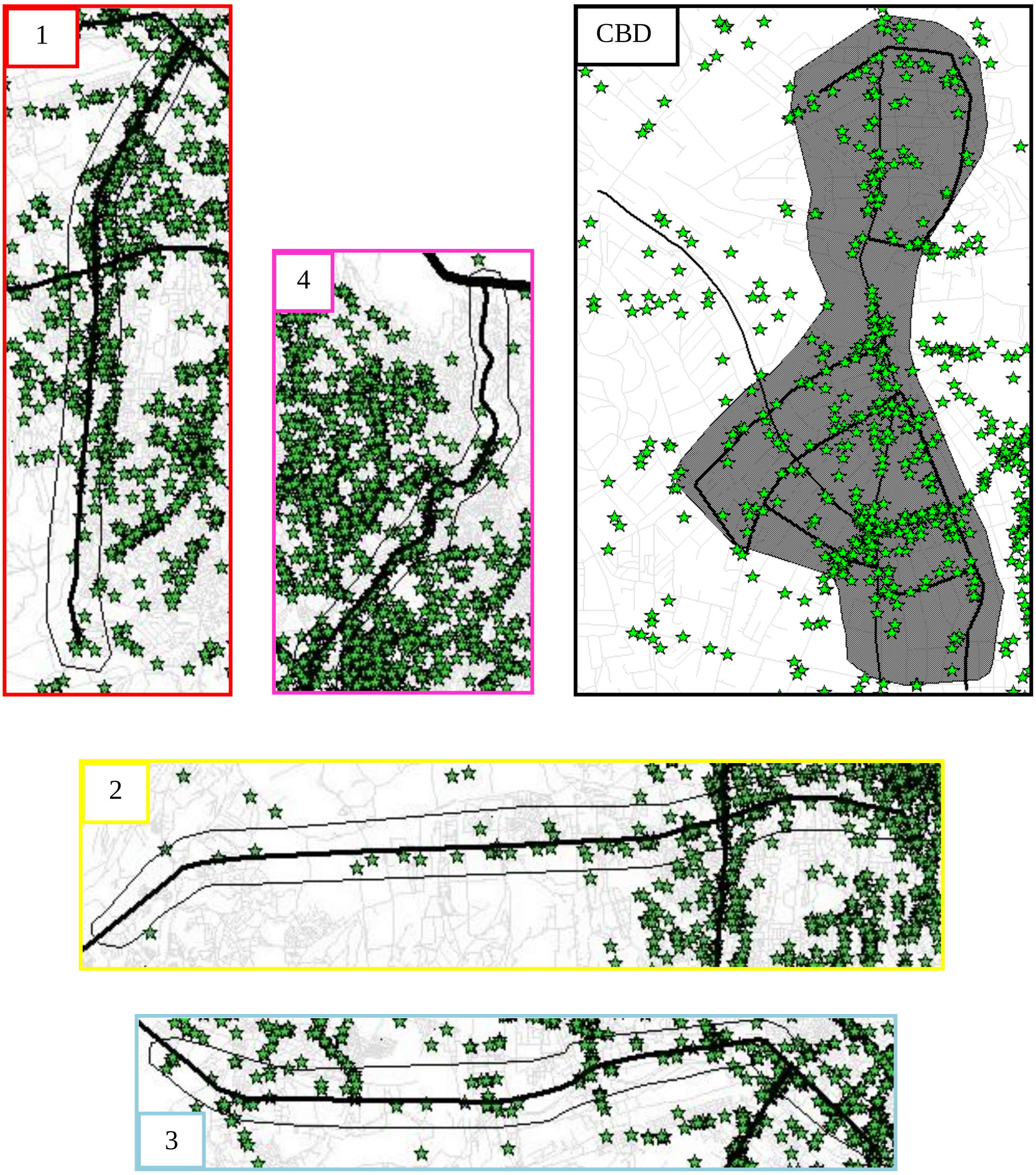


Figure 4.3 Distribution of Pedestrian Related Accidents for UTZ and CBD Corridors

4.1 Description of the Study Area

More detailed description of the four main arterials and CBD corridors were given in Figure 4.4 and Figure 4.5, respectively. Except the CBD region, the common property of other four main corridors is that they provide connection between Ankara and neighboring cities. In addition, these corridors can be categorized as UTZ corridors, while a small portion would have urban characteristics where they join or come closer to the CBD region.

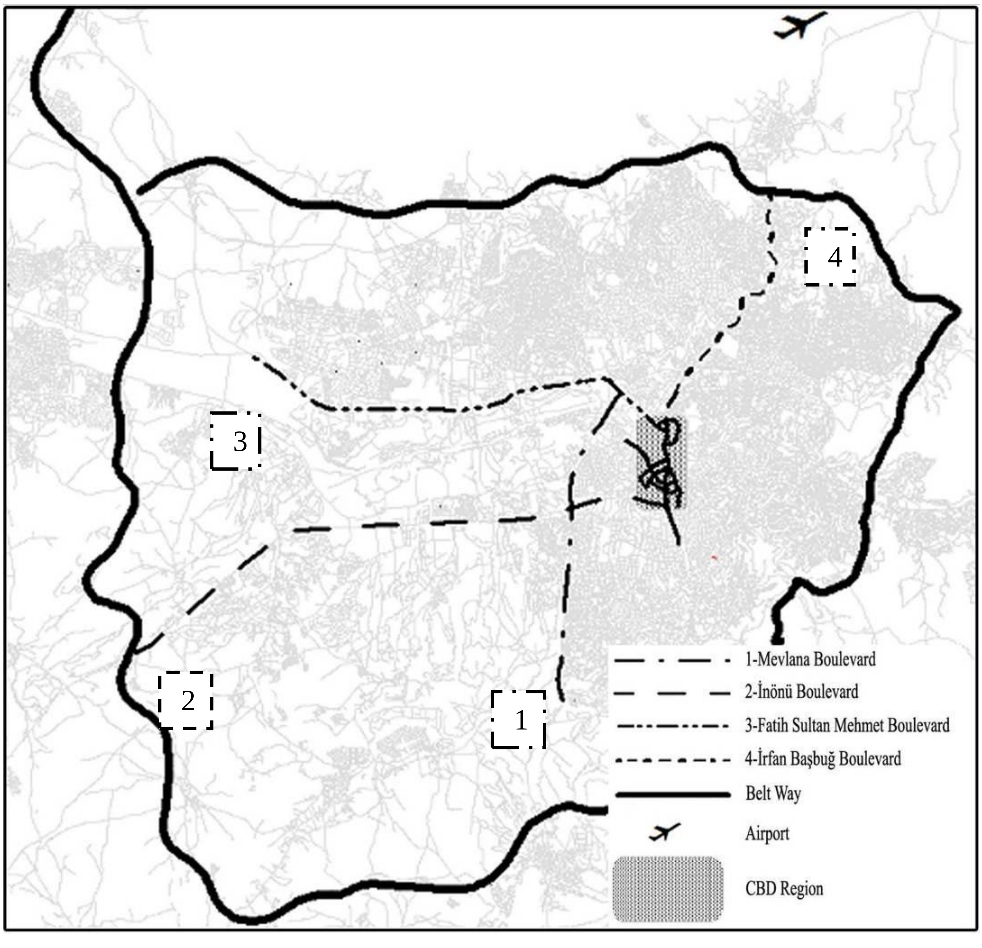


Figure 4.4 The Selected UTZ Corridors in Ankara

To cover the CBD region, two main corridors, Atatürk and Gazi Mustafa Kemal Boulevard were selected firstly. As these two arterials run along the old city center, there are so many small streets crossing them. In the vicinity around these two corridors, there are 12 more small streets crossing them. In addition, high pedestrian accident density on the CBD corridors was the second reason in the selection. Share of pedestrian accidents of CBD roads is 4.72 % of total pedestrian accidents in Ankara occurred (see Table 4.2). Still, it is expectable to have a high number of accidents in the Atatürk Boulevard with the share of 2.40%.

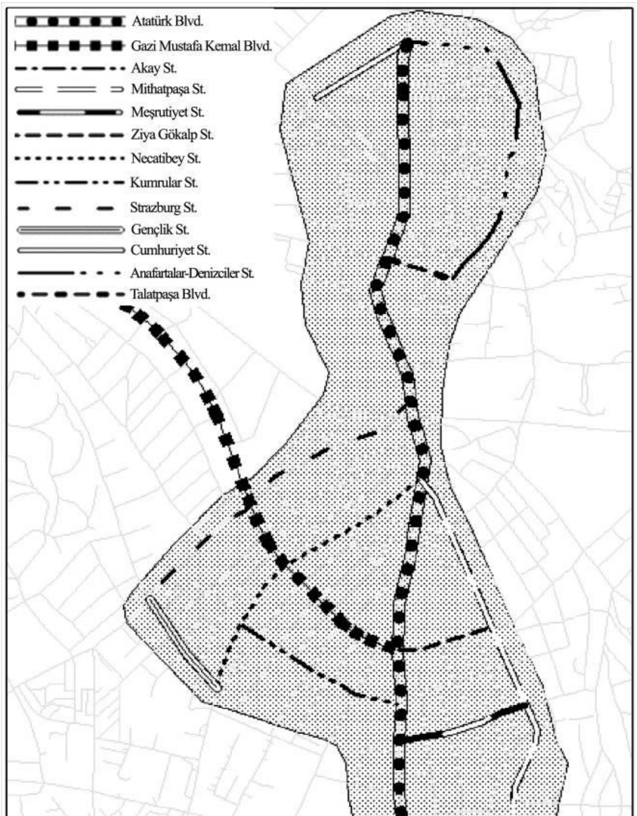


Figure 4.5 The Selected CBD Corridors in Ankara

Table 4.2 Distribution of Pedestrian Accidents on the CBD and UTZ Corridors

Corridor Name	No. of Pedestrian Accidents (NP_A)	Share (%)*
CBD Region	**370**	**4.72**
Atatürk Blvd.	*188*	*2.40*
Ziya Gökalp St.+Gazi Mustafa Kemal Blvd.	*66*	*0.84*
Mithatpaşa St.+Necatibey St.	*42*	*0.54*
Anafartalar St.+Denizciler St.+Talatpaşa Blvd.	*35*	*0.45*
Others	*39*	*0.50*
Mevlana Blvd.	**168**	**2.14**
İnönü Blvd.	**70**	**0.89**
Fatih Sultan Mehmet Blvd.	**88**	**1.12**
İrfan Baştuğ Blvd.	**110**	**1.40**
Total	**806**	**10.29**

* Total number of accident=7833 in Ankara for (2007-2009)

Characteristics of the Corridors:

In investigating the affecting factors behind the pedestrian accidents, some aspects of the accident location have to be collected regarding traffic network; traffic flow and built-environment (see Table 3.10). Among built environment characteristics, a significant one is the land-use, because it generates both pedestrian and motor vehicle traffic. To find the land use characteristics influencing the pedestrian safety, the selected main arterials and the CBD region were studied in GIS environment to create a thematic map showing the land use pattern (Figure 4.6). For the sake of simplicity, only 8 categories were created to display the main usage purpose. Shopping, green, governmental and residential categories are common ones in the literature, while business simply denoted commercial locations, which are mostly work places. Military, health and education zones were created specifically to denote campus like regions that have high but very specific travel demand pattern (limited access).

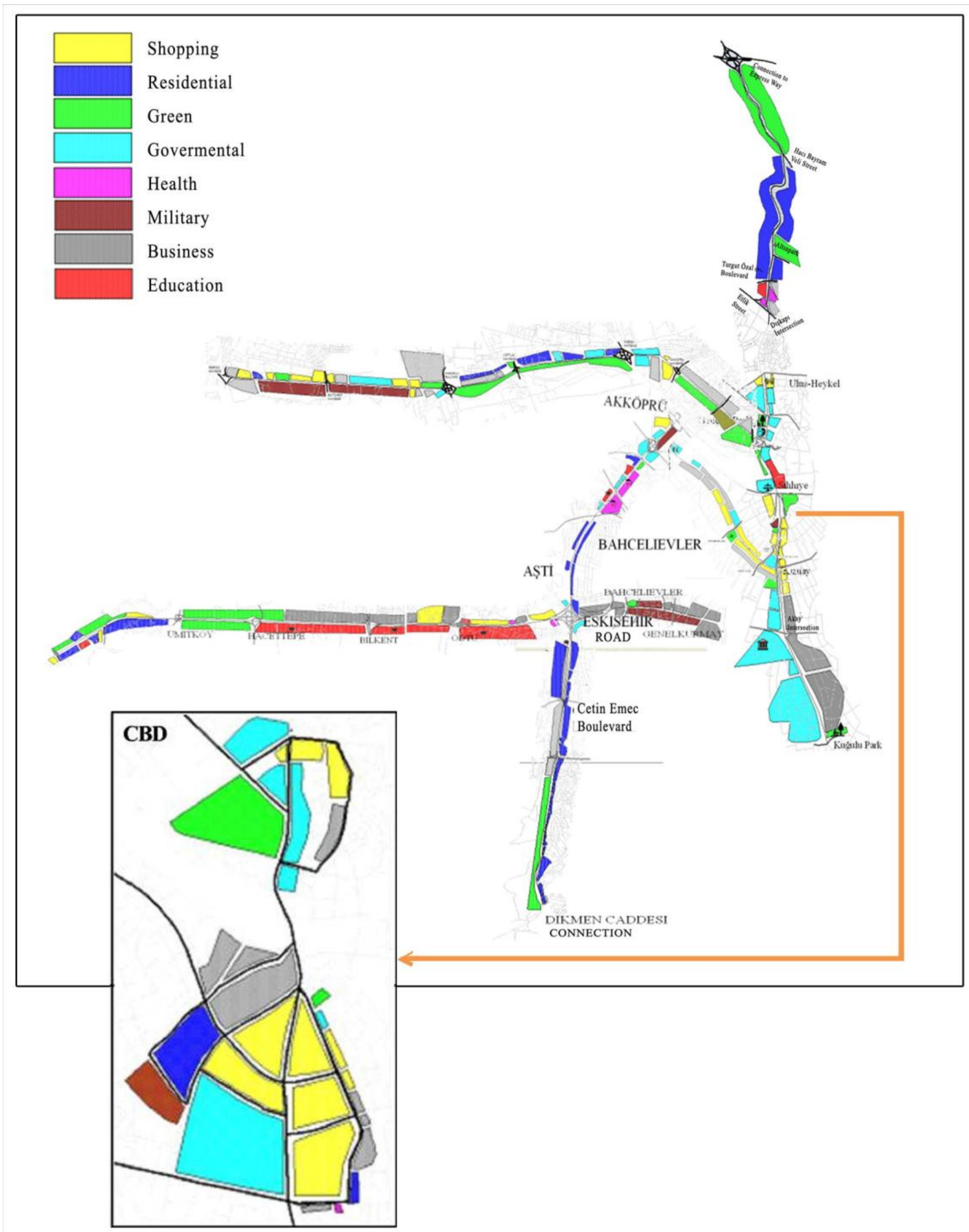


Figure 4.6 Primary Land-use Characteristics along the CBD and UTZ Corridors

For the CBD, it can be easily seen that land-use includes more shopping, business and governmental usage. Major differences among the UTZ arterials, which were the main motivation in choosing them, can also be easily observed: while there are university campuses and governmental offices along the İnönü Boulevard, industrial areas are located more along the Fatih Sultan Mehmet Boulevard. On the other hand, Mevlana Boulevard connects the city center to the one of the biggest residential/summer housing area, Gölbaşı County. In addition, İrfan Baştuğ Boulevard is used to arrive the airport form the city center. This way, a more random sampling of the city arterials is achieved.

The land-use and built environment aspects of the selected corridors were eventually converted to numeric (nominal or ratio) data for regression analysis purposes mainly. Besides the land use characteristics, other built environment characteristics that were discussed in Section 3.4 before like, the existence of underpass, overpass, pedestrian barrier etc. are also added to the statistical models. Some of these were presented in figures in Appendix C. Segments with urban zone assumption is presented in Figure C.8 , while existence of pedestrian barriers is shown in Figure C.6 (photographs of different road elements and designs that were assumed to act as if pedestrian barriers were presented in Figure C.7).

4.2 Segmentation of Main Arterials in Ankara

The reasons and the process of segmentation were discussed before in Section 3.1.2. In this section, the details of unit segmentation while it was applied to the study corridors will be discussed. To determine the accident-prone segments of the corridors, they were divided into the pieces. The traditional and density safety measures were calculated by using a) the original traffic network links and b) a selected "unit" segment definitions, separately. The comparisons of the results from two analyses showed the problems related to existence of very long and very short segment length, and capacity of usage of "unit" segment concept to solve the problems. For instance, while there are short links in CBD, like 0.50 km, in suburbia, they are getting longer; like 3 km.

Number of accidents for each link was calculated by drawing buffer zones with 200 m total width that their radiuses are 100 m around each link and the accidents that fall into the buffer zones were counted and assigned to the corresponding link (Figure 4.7). When this was applied to the original traffic network links, existence of various lengths of links resulted in some problems. One of the biggest problems of this method especially in the short links is that buffer zones of the two short links are overlapping (Figure 4.8). When the buffer zones are overlapped, the accidents that occur in the overlapping zone are counted two times. On the other hand, when the link is long (for example 3 km) and the accidents are clustered in a region on the link, occurrence of all accidents can be perceived throughout the link.

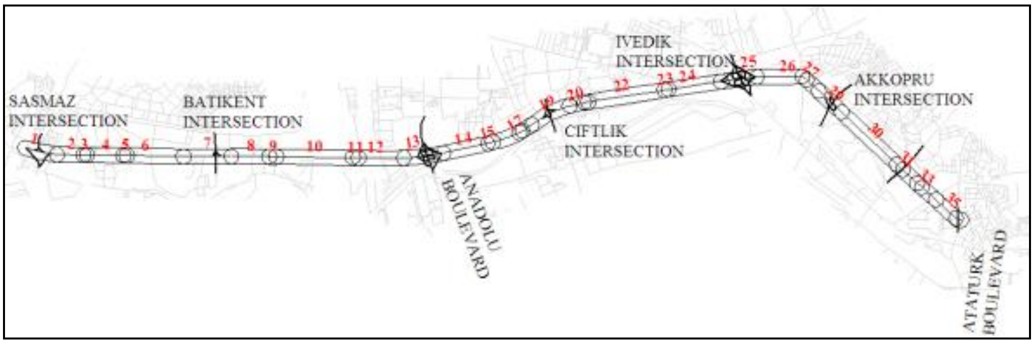


Figure 4.7 The Road Links along Fatih Sultan Mehmet Boulevard

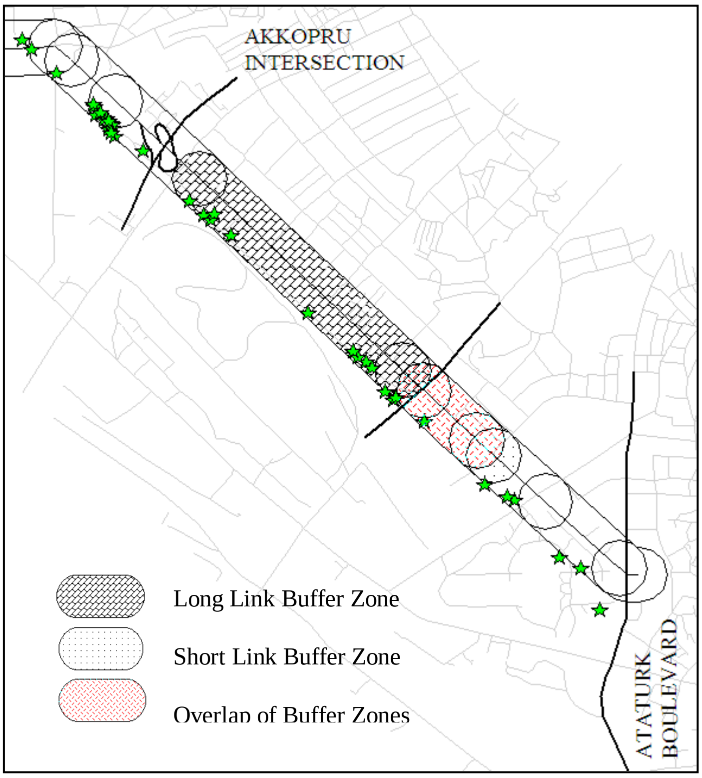
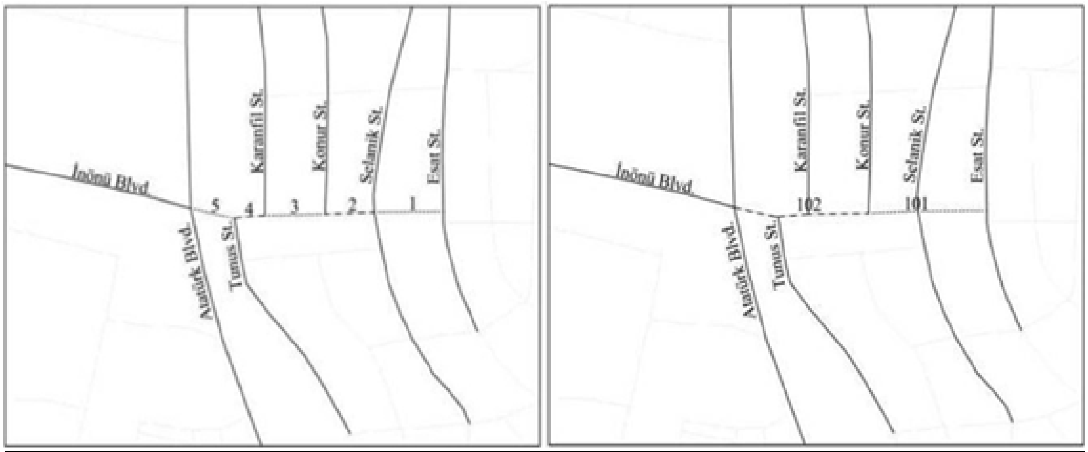


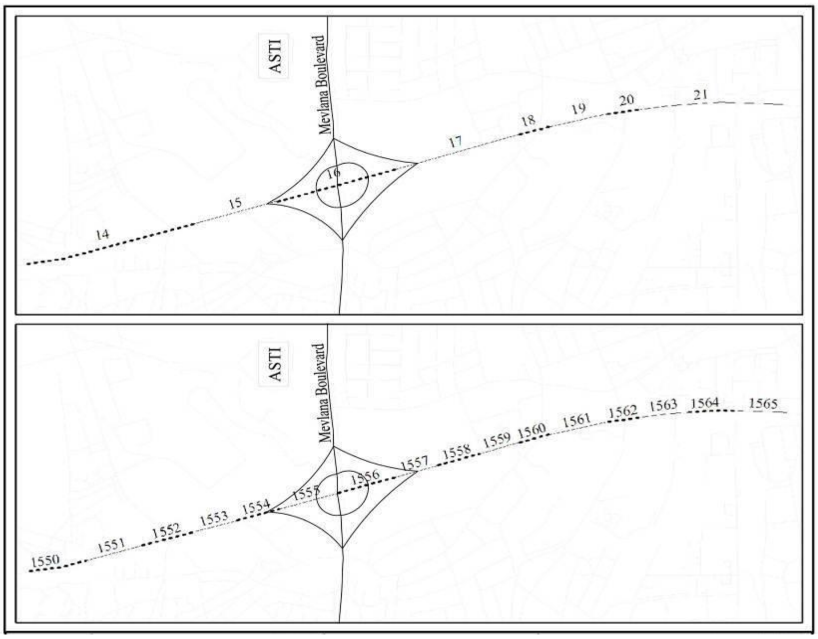
Figure 4.8 An Example of MapInfo Link Buffer Zone Definition and Overlaps

The unit segmentation of links can solve the problem of clustering of accidents in long links however; it is not an exact solution in solving the overlapping of buffer zones of links. For an exact solution of overlapping problem, a GIS development focused study should be conducted that assign each accident only one link. Scope of this study does not cover such a focus. Still, to minimize the overlapping problem, link lengths that were aggregated or disaggregated tried to be selected same to keep multiplication of count of accidents same as well. However, it is not easy to divide each link into same lengths in a city network. Choice of the segment length requires a discussion, as it depends on the link length distribution on a road network. For the City of Ankara, there were small links (around 50 m) representing blocks between small streets in the CBD. On the other hand, there were links up to 3 km on UTZ corridors. To create an almost equal segmentation, this length was determined as 200 m, if possible with a flexibility to change in the range of 0.15 km -0.3 km for Ankara. This required either the aggregation of very small links, or disaggregation of very long ones as shown in Figure 4.9 and Figure 4.10, respectively.



ID	Length (km)	ID	Length (km)
1	0.114	101	0.195
2	0.081	101	0.195
3	0.099	102	0.223
4	0.051	102	0.223
5	0.074	102	0.223

Figure 4.9 An Example of Aggregation of Links during Segmentation



ID	Length (km)	ID	Length (km)
14	0.746	1550	0.249
14	0.746	1551	0.249
14	0.746	1552	0.249
15	0.374	1553	0.187
15	0.374	1554	0.187
16	0.528	1555	0.264
16	0.528	1556	0.264
17	0.564	1557	0.188
17	0.564	1558	0.188
17	0.564	1559	0.188
18	0.132	1560	0.132
19	0.26	1561	0.260
20	0.146	1562	0.146
21	0.637	1563	0.212
21	0.637	1564	0.212
21	0.637	1565	0.212

Figure 4.10 An Example of Disaggregation of Links during Segmentation

At the end of segmentation process, the UTZ corridors and the roads in the CBD region were divided into 295 segments with lengths changing from 0.105 km to 0.300 km (see Table 4.3). The variety in the segment length originated due to both the manual segmentation process and the nature of the original network that can not be divided anytime into equal selected segment length.

Table 4.3 Distribution of the Segment Lengths in Ankara Study Corridors

Segment Length Range (km)	Number of Segments in Corridors		
	All	CBD	UTZ
0.100-0.150	19	11	8
0.150-0.200	55	21	34
0.200-0.250	145	26	119
0.250-0.300	76	14	62

To understand the effect of the unit segmentation process, especially change in the density measures for the original link and segment density measures were calculated. In Table 4.4, the change intervals of density measures and the number of the segments that their density measures change in corresponding interval were identified. As the average number of injuries per accident is very close to 1 for Ankara study region, number of accidents and number of injuries-based calculations produce very similar results and patterns. Thus, from now on, only the results of number of accidents and total accident density values will be given in the text, while the rest will be presented in Appendix.

Table 4.4 Number of Segments with the Density Change Based on Link versus Segment Based Calculations

Density Interval (acc/km) or (fatality/km)	Number of Segments							
	Aggregation				Disaggregation			
	ΔD_A	δD_A	ΔD_F	δD_F *	ΔD_A	δD_A **	ΔD_F	δD_F ***
<-250	1	14	0	0	0	7	0	15
(-201)-(-250)	1	9	0	0	0	3	0	13
(-151)-(-200)	3	8	0	0	0	1	0	11
(-101)-(-150)	3	16	0	0	1	5	0	20
(-51)-(-100)	12	9	0	5	0	8	5	38
(-41)-(-50)	4	6	0	2	0	4	8	16
(-31)-(-40)	5	1	0	0	3	4	5	5
(-21)-(-30)	8	2	1	0	2	5	12	13
(-11)-(-20)	16	3	7	3	11	7	19	9
(-1)-(-10)	23	8	24	7	67	12	107	12
0-9	24	8	23	9	107	12	67	13
10-19	7	2	16	4	19	16	11	8
20-29	1	0	8	2	12	16	2	10
30-39	0	5	5	6	5	26	3	4
40-49	0	2	4	8	8	23	0	5
50-99	0	15	12	47	5	63	0	44
100-149	0	0	3	0	0	0	1	0
>150	0	0	5	0	0	0	0	0

* For 14 segments, % change could not be calculated as the new value is “zero”
** For 28 segments, % change could not be calculated as the new value is “zero”
*** For 4 segments, % change could not be calculated as the new value is “zero”

CHAPTER 5

PEDESTRIAN SAFETY ANALYSES RESULTS

5.1 Corridor Pedestrian Safety Analysis

As a case study, pedestrian accidents in the City of Ankara from 2007-2009 were analyzed for the CBD region and four UTZ arterials. The thematic maps of studied corridors for traditional and density measures were created in GIS environment. The results were visualized separately for CBD and UTZ corridors. In this chapter, only thematic maps of the number of accidents and total accident density will be given. The injury and fatality maps are in the Appendix.

For the number of accidents measure, while there is only one segment that has no accidents assigned to it out of 66 ones, the most critical segments in the CBD region is mostly placed in Atatürk Boulevard and Ziya Gökalp Street. Although, the most critical segments almost remains same, the change can be observed in less critical segments in terms of number of accident become more critical and opposite situation is also valid when the total accident density measure is visualized (Figure 5.1). For example, more segments on the Atatürk Boulevard are critical when it was compared with number of accidents map in the total accident density map. However, all of the segments were tried to be drawn in the same length still there are little differences between the lengths of the segments, these differences can be observed obviously in comparison of number, and density thematic maps

As it was discussed before, the ratio of number of accidents to number of injuries is nearly 1. Thus, the criticality of segments in terms of injury is same as the accident for CBD corridors (Figure B.1). In addition the thematic maps of the number of fatalities were not prepared for CBD corridors because, there was no fatality in the CBD for 3-year period.

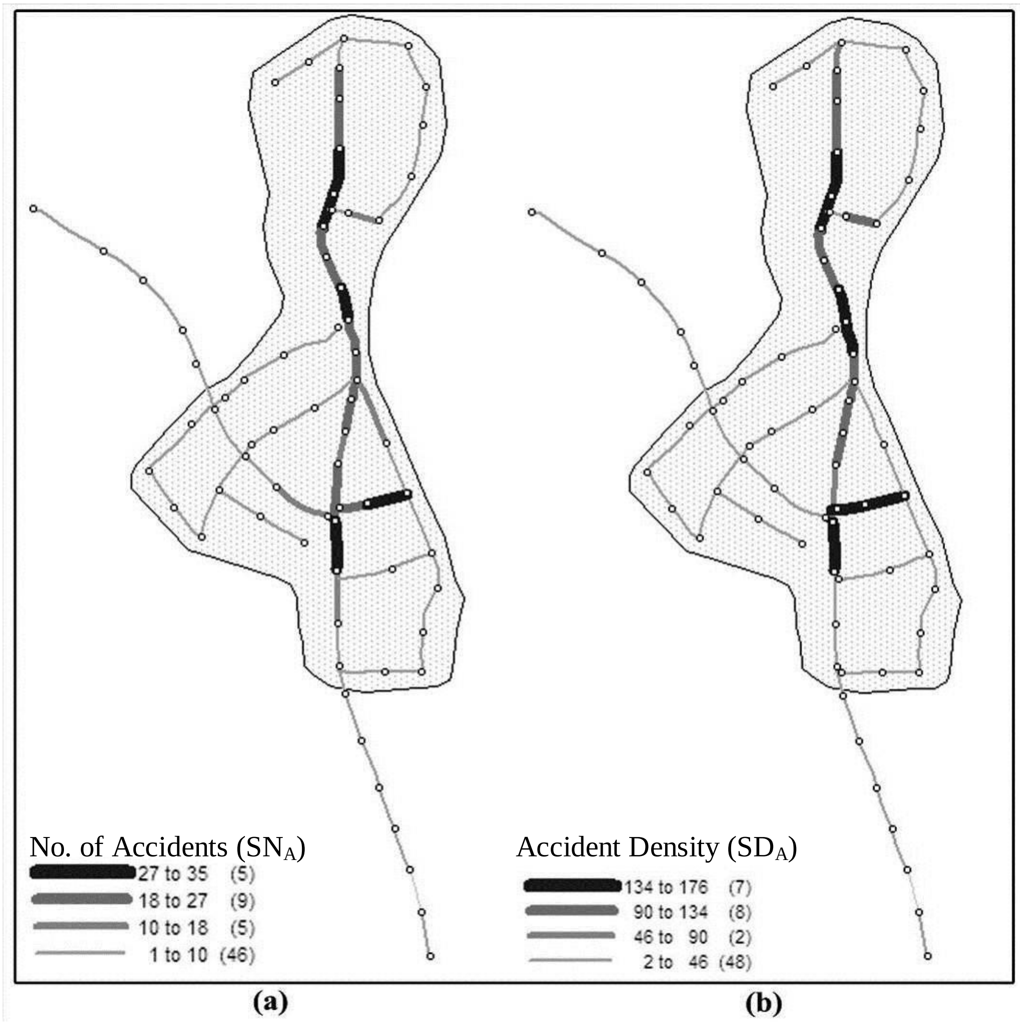


Figure 5.1 CBD Thematic Maps of (a) the Segment Number of Accidents, SN_A, (b) the Segment Accident Density, SD_A (for a 3 Year Period)

Same procedure was conducted for UTZ corridors and accident-prone segments were visualized for number of accidents and accident density values (Figure 5.2 and Figure 5.3). There are no accidents on 55 segments out of 219 ones in the UTZ corridors. The most critical segments are observed in generally where the pedestrian activity is high. The most risky segment is in front of the intercity bus terminal (AŞTİ). One difference of UTZ maps from CBD ones, is that there are segments that number of fatalities change 1 to 3. This is most probably higher speed of UTZ corridors than CBD ones that cause more severe accidents, which will be analyzed in Section 5.4. The thematic maps of number of injuries, injury density, number of fatalities and fatality density of UTZ corridors can be seen in Figure B2, Figure B3, Figure B4 and Figure B5 respectively.

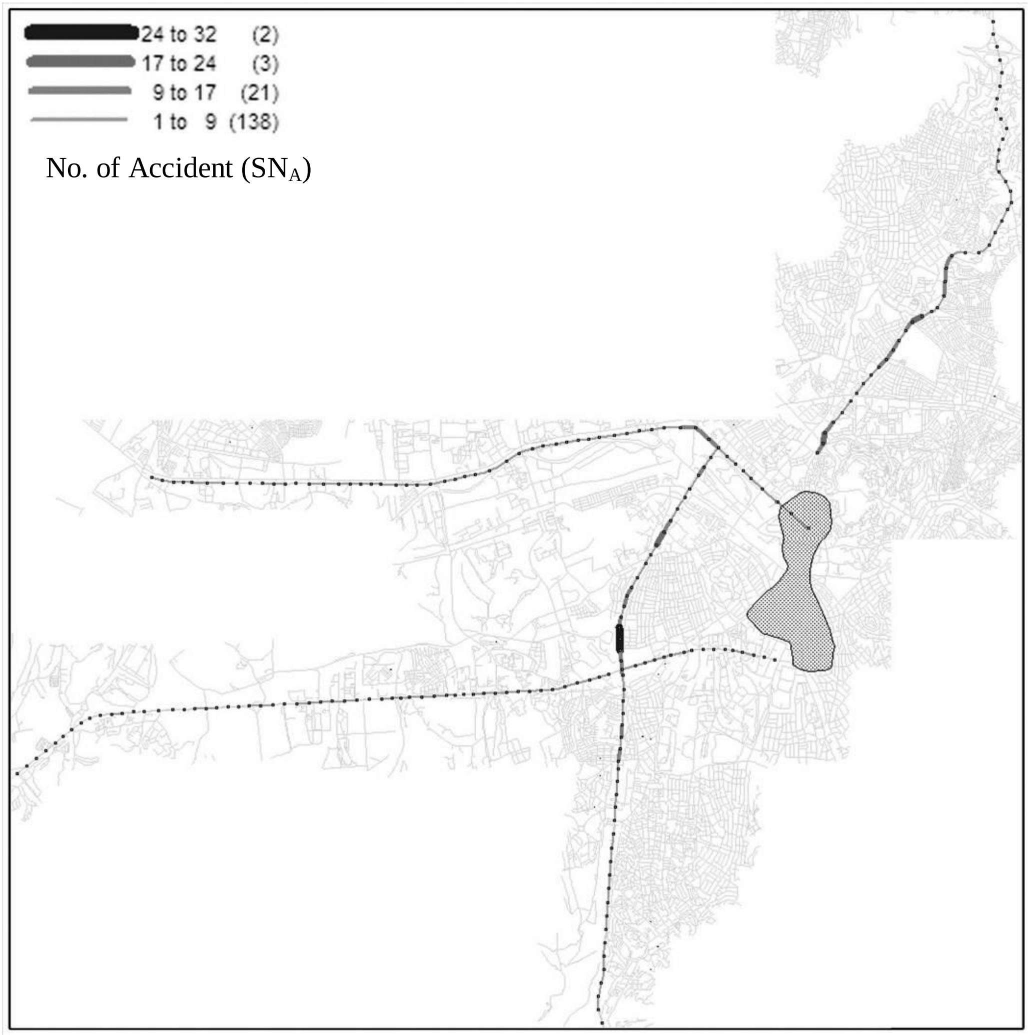


Figure 5.2 UTZ Thematic Map of the Segment Number of Accidents, SN_A (for a 3 Year Period)

Figure 5.3 UTZ Thematic Map of the Segment Accident Density, SD_A (for a 3 Year Period)

5.2 Bus Stop Pedestrian Safety Analysis

As in the corridor pedestrian safety analysis, the aim is to identify the accident-prone bus stops on the CBD and UTZ corridors. Out of 344 bus stops for study region, 174 ones are on the CBD and 170 ones are on the UTZ corridors the distribution of bus stops and pedestrian accidents along the studied corridors can be observed in Figure 5.4 and Figure 5.5. Although the UTZ covers wider area than CBD, the numbers of bus stops are almost equal to each other. Due to distances between bus stops in the CBD are very short, the density of bus stops is higher than the UTZ bus stops that have longer distances between each others. While, the average distance was found 0.056 km for the CBD, this value was 0.412 km for the UTZ. The average distance between bus stops were calculated by measuring the distances between consecutive bus stop and then divided by the number of them. Thus, the average distance did not represent the distance between every consecutive bus stop.

The bus stop pedestrian safety analysis was conducted by drawing the buffer zones around each bus stop. Considering the average distance between bus stops, radiuses of buffer zones were selected as 50 m (especially for CBD bus stops), 100 m and 400 m (especially for UTZ bus stops) for CBD and UTZ corridors. Thus, the results that show the accident-prone bus stops will be given for different buffer zones and for CBD and UTZ corridors separately.

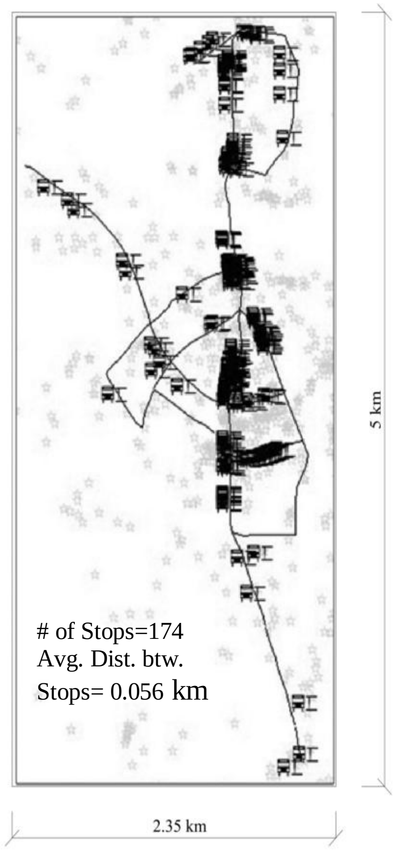


Figure 5.4 Distribution of Bus Stops in the CBD

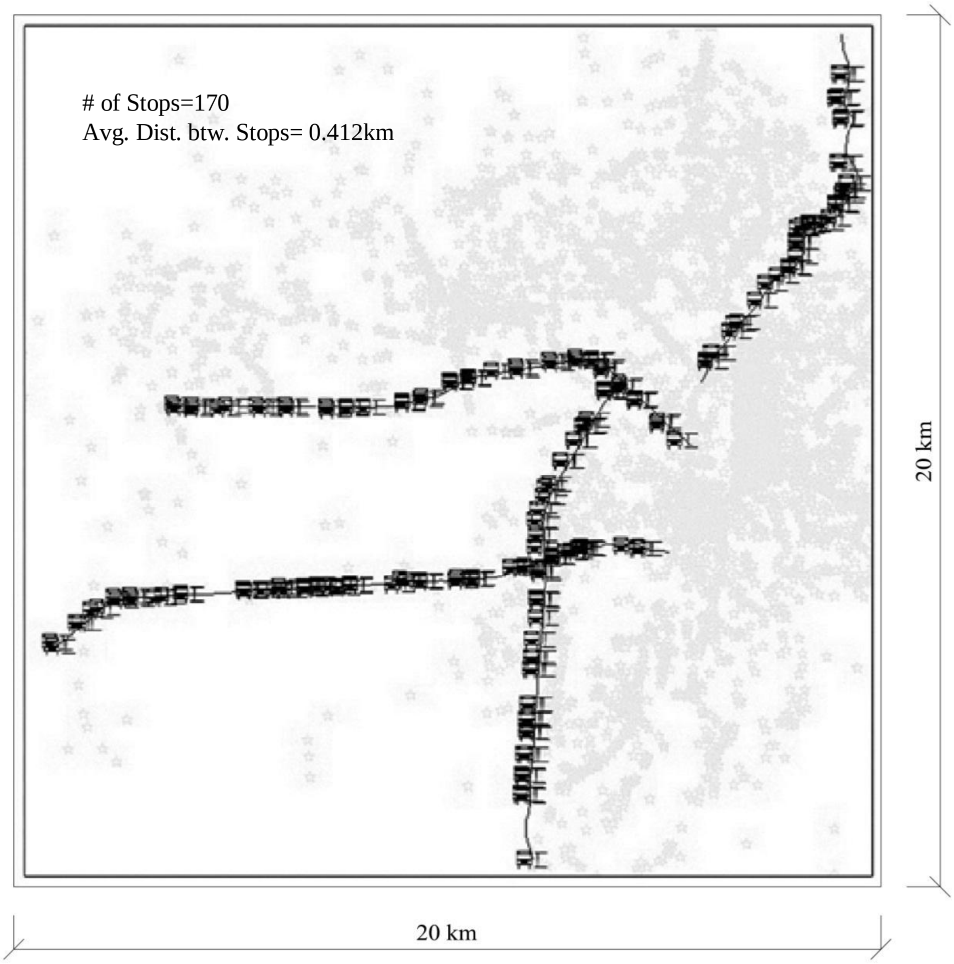


Figure 5.5 Distribution of Bus Stops in the UTZ

Results of Bus Stops on CBD Corridors

The results of bus stop pedestrian safety analysis for CBD corridors show that while, there are no accidents for 46 stops for 50 m buffer zones, this number decreases to 8 stops when the radius of buffer zones increase to 100 m, the number of non-accident bus stops is 0 for 400 m buffer zones. This means that 46 stops are safer for pedestrians when the immediate impacts of bus stops are analyzed. When the radius of buffer zones increase, the number of safer stops decreases however the probability of occurrence of pedestrian crashes due to the bus stops also decreases. For a general visualization of change in the number of pedestrian accidents for 50 m, 100 m and 400 m buffer zones for CBD region, thematic maps were created with accident information of 3-year period (Figure 5.6, Figure 5.7 and Figure 5.8). The thematic maps for other traditional measurements, the total number of injuries, for different buffer zones of bus stops can be seen in the Figure C1, Figure C2 and Figure C3.

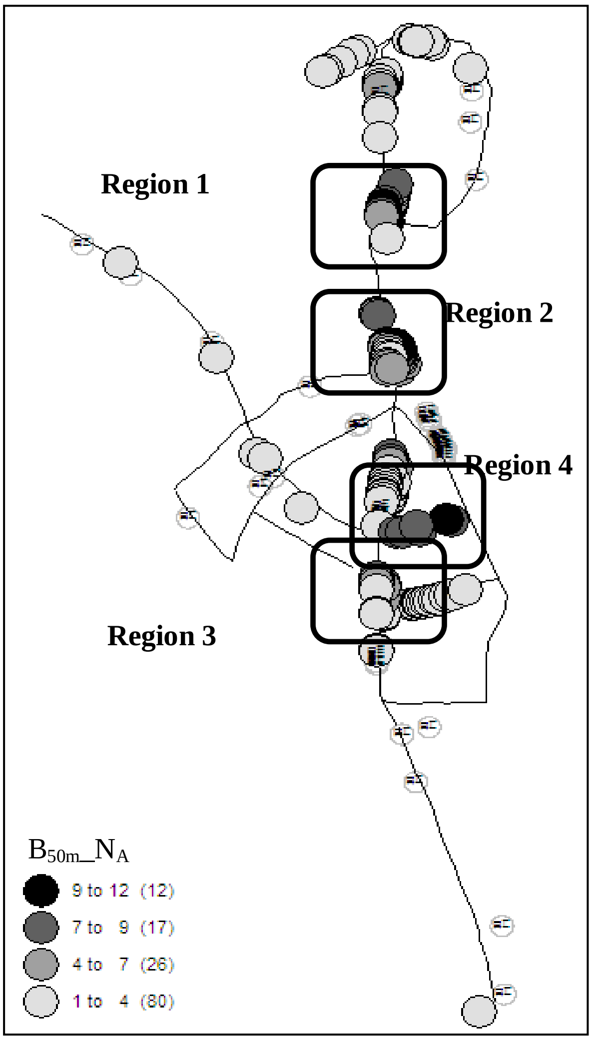


Figure 5.6 CBD Thematic Map for Bus Stop Safety for the 50m Buffer Zone, $B_{50m}\_N_A$ (for a 3 Year Period)

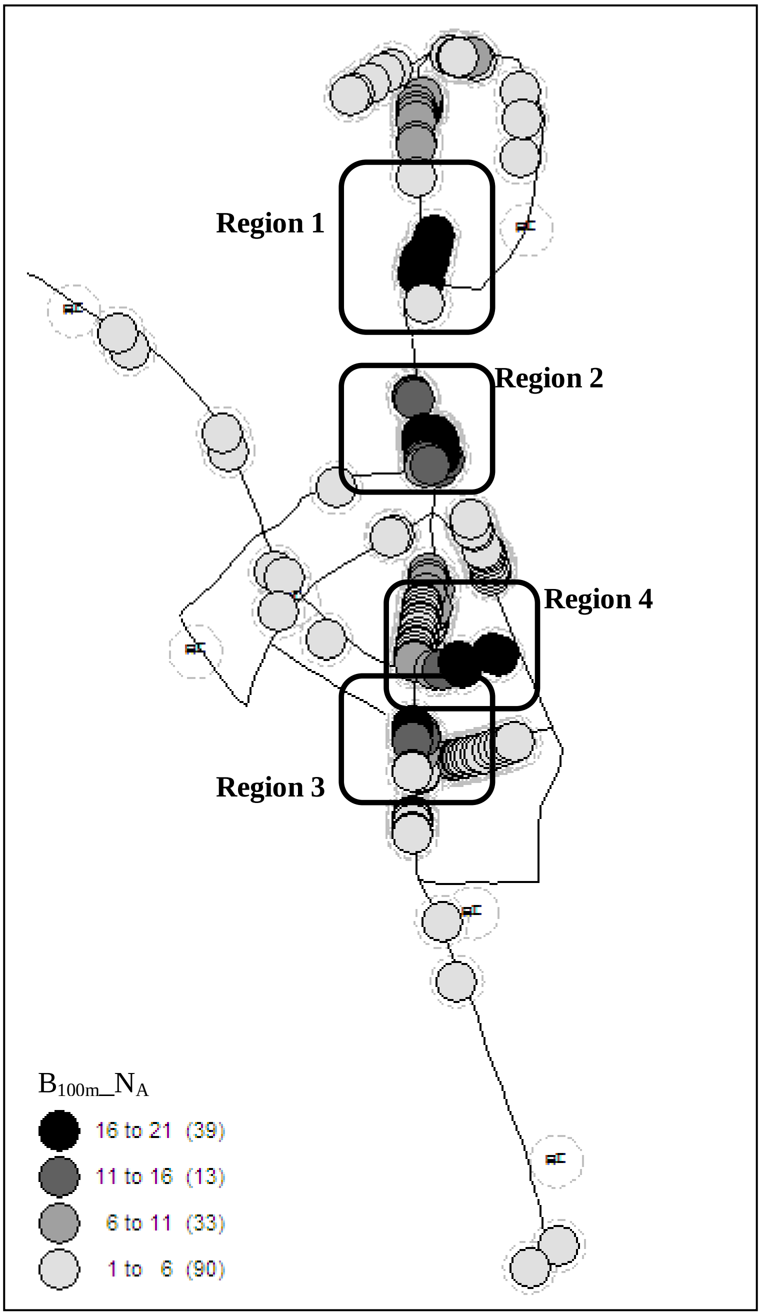


Figure 5.7 CBD Thematic Map for Bus Stop Safety for the 100m Buffer Zone, B_{100m}_N_A (for a 3 Year Period)

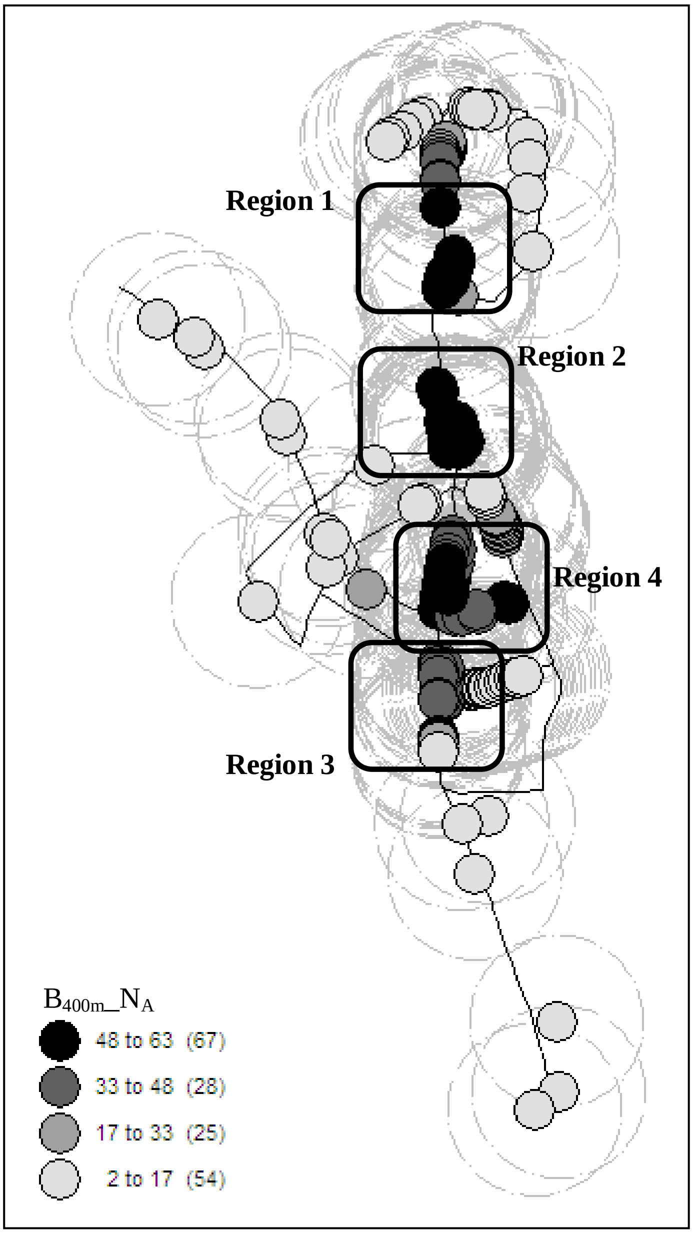
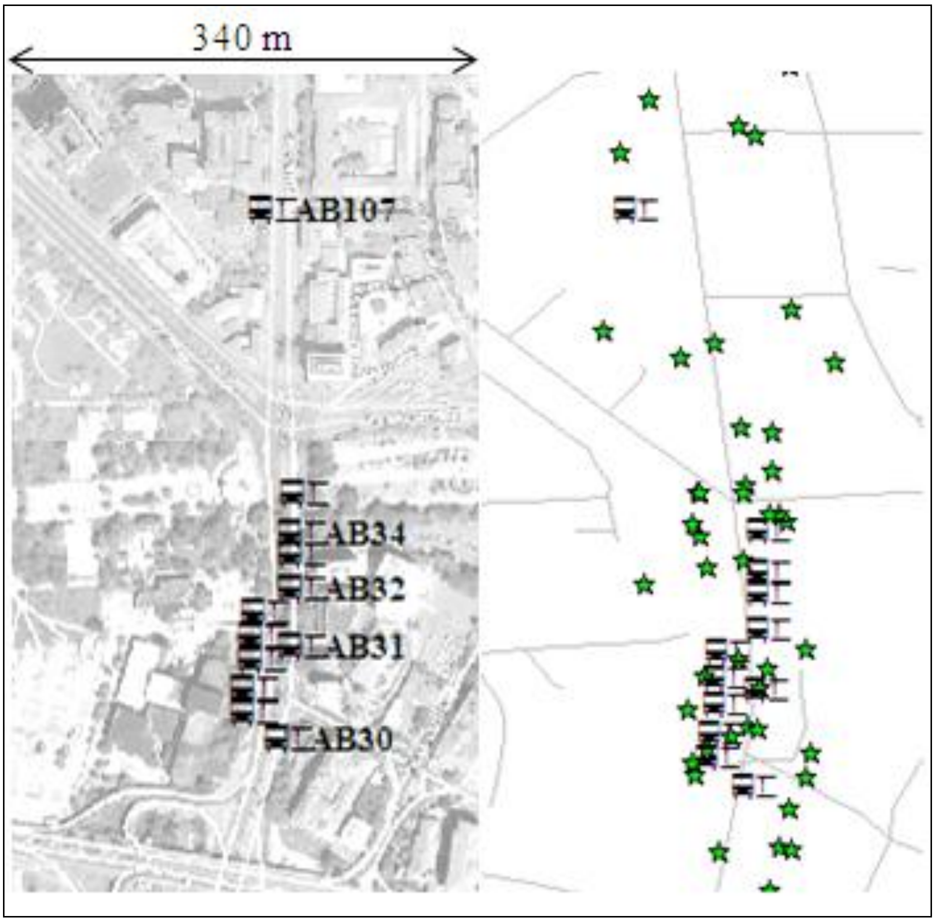


Figure 5.8 CBD Thematic Map for Bus Stop Safety for the 400m Buffer Zone, $B_{400m}\_N_A$ (for a 3 Year Period)

When these thematic maps were analyzed, number of accidents of the bus stops that are in the specific regions are higher than the others. There are “4” critic regions for CBD. Region 1 denotes the bus stops that are between the Gençlik Park and Opera House and bus stops that place opposite side of the road; Region 2, in front of the Court House and Ankara University Faculty of Languages, History and Geography; Region 3, in front of Güvenpark. On the other hand, Region 4 denotes bus stops at two different corridors. First group in Region 4 are in between Güvenpark and Sıhhıye and the second group are at the end of Ziya Gökalp Street, which meets with Atatürk Boulevard. As it was expected, the locations of the critical bus stops in terms of pedestrian safety show parallelism with the accident-prone segments of the same corridors.

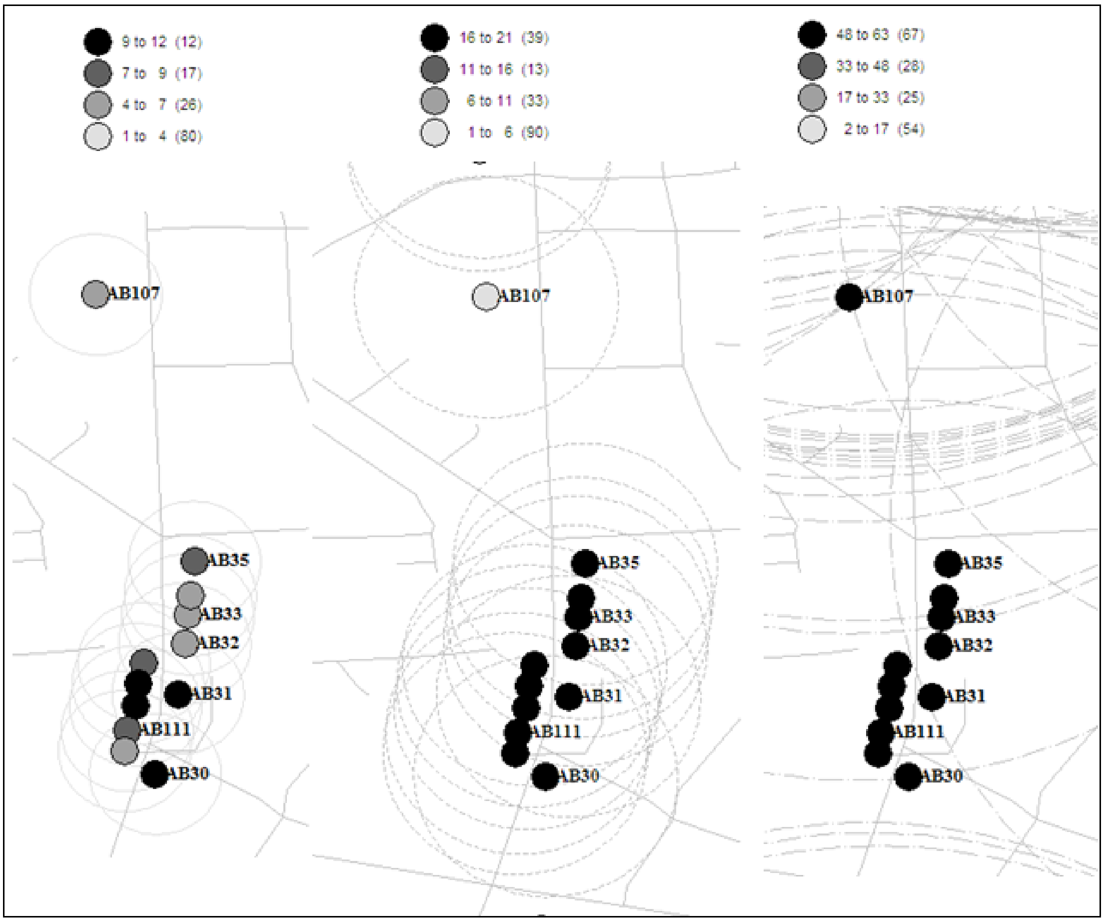
The close up display of critical Region 1, Region2, Region 3 and Region 4 were taken for better understanding in Google Earth. In addition, the distribution of pedestrian accidents with road network visualized in Figure 5.9, Figure 5.12, Figure 5.14 and Figure 5.16. The change in the number of pedestrian accidents for different buffer zones dimensions were represented throughout thematic maps in Figure 5.10, Figure 5.13, Figure 5.15 and Figure 5.17 for critical Region 1, Region2, Region 3 and Region 4 respectively.

Especially in the CBD bus stops, to see the effect of 400 m buffer zones, a close up display was taken to show the change in the number of accidents that fall into the bus stop AB 107. As it can be seen from the thematic maps of Region 1 (Figure 5.10), when the radiuses of buffer zones were changed, the number of the accidents that fall into the buffer zones also changed. The sharpest change can be observed in the bus stop AB107 that when 400 m buffer zone was applied, AB 107 covered the accidents that might not related with the existence of bus stop (see Figure 5.11).



(a) (b)

Figure 5.9 For the Critical Region 1 in the CBD (a) Satellite Picture (b) Road Network, with Bus Stops



(a) (b) (c)

Figure 5.10 For the Critical Region 1 in the CBD, Thematic Maps for Bus Stop Safety for the Buffer Zone (a) 50 m (b) 100 m (c) 400 m (for a 3 Year Period)

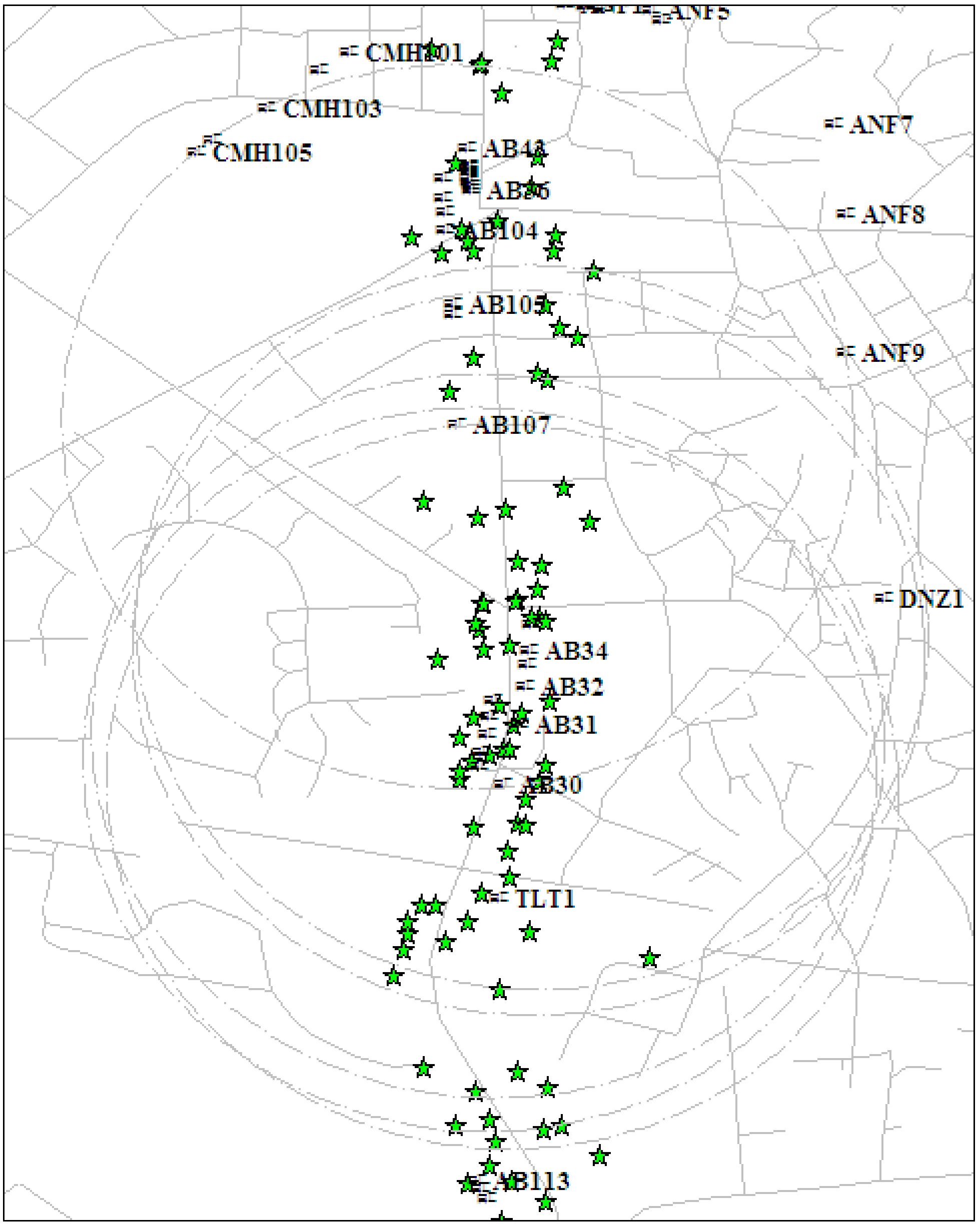
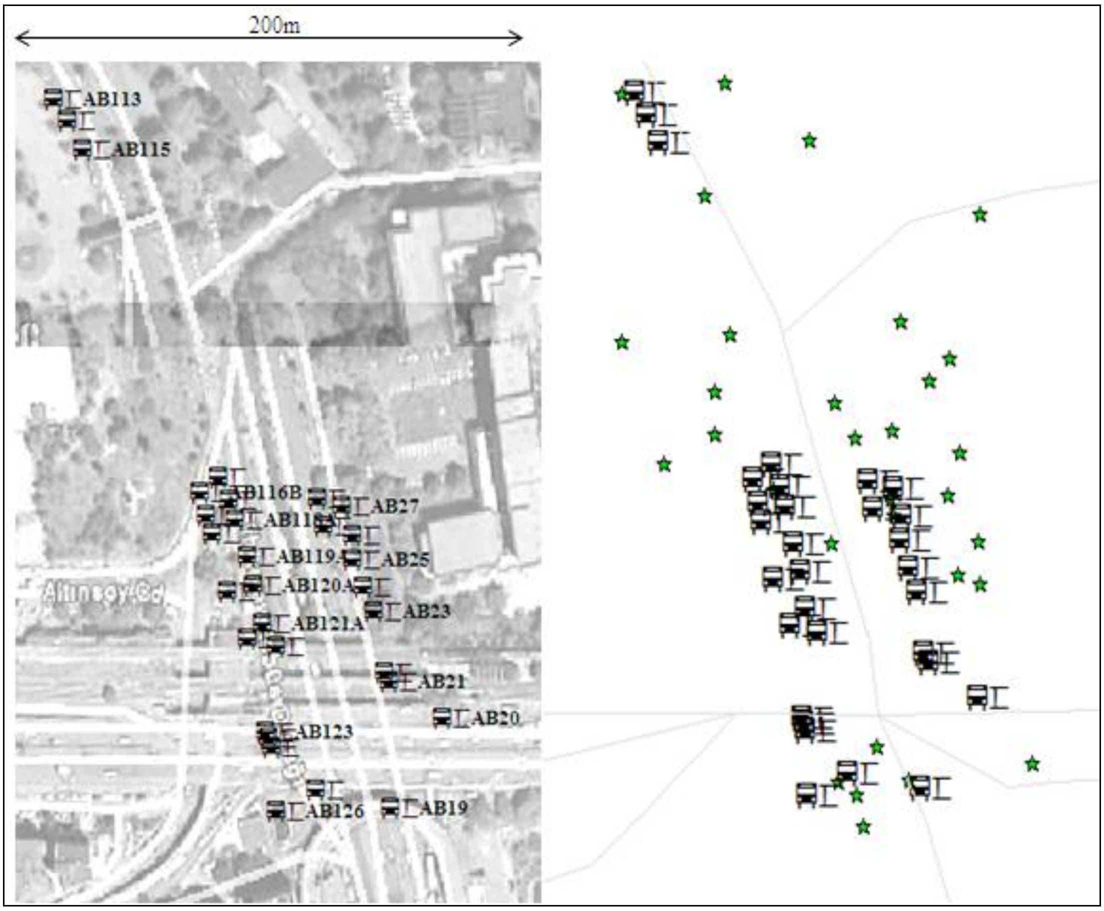
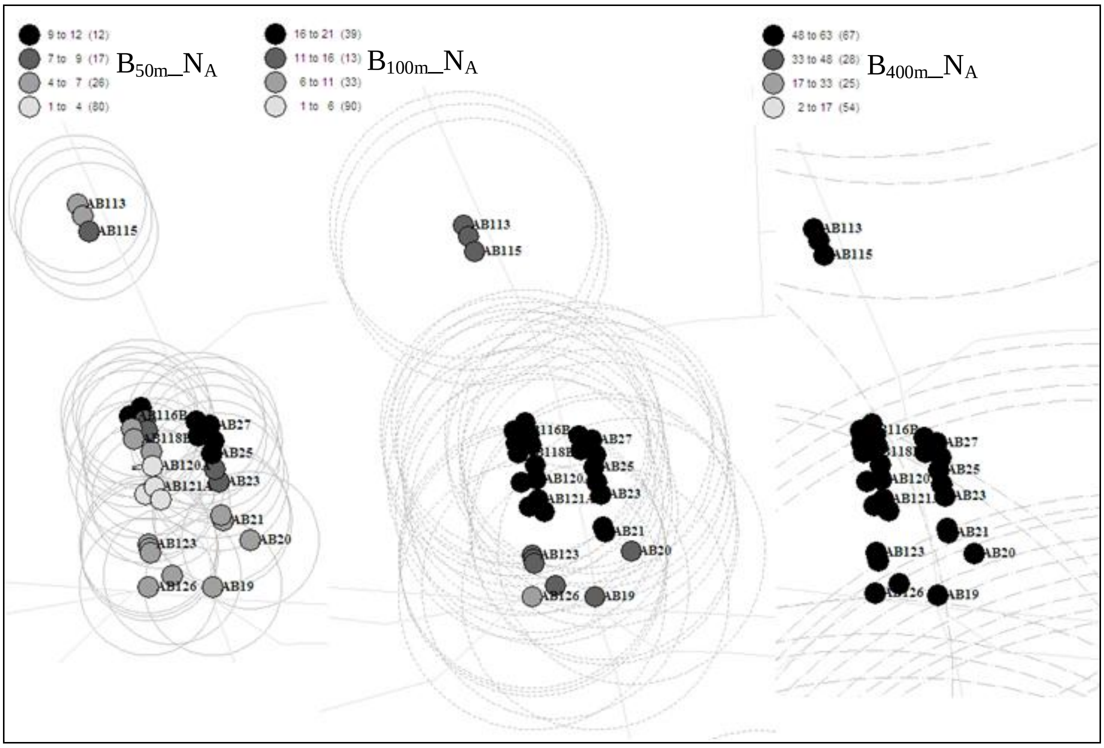


Figure 5.11 For the Critical Region 1 in CBD, Close up Display of the 400m Bus Stops Buffer Zones and Accidents around the Bus Stop AB107



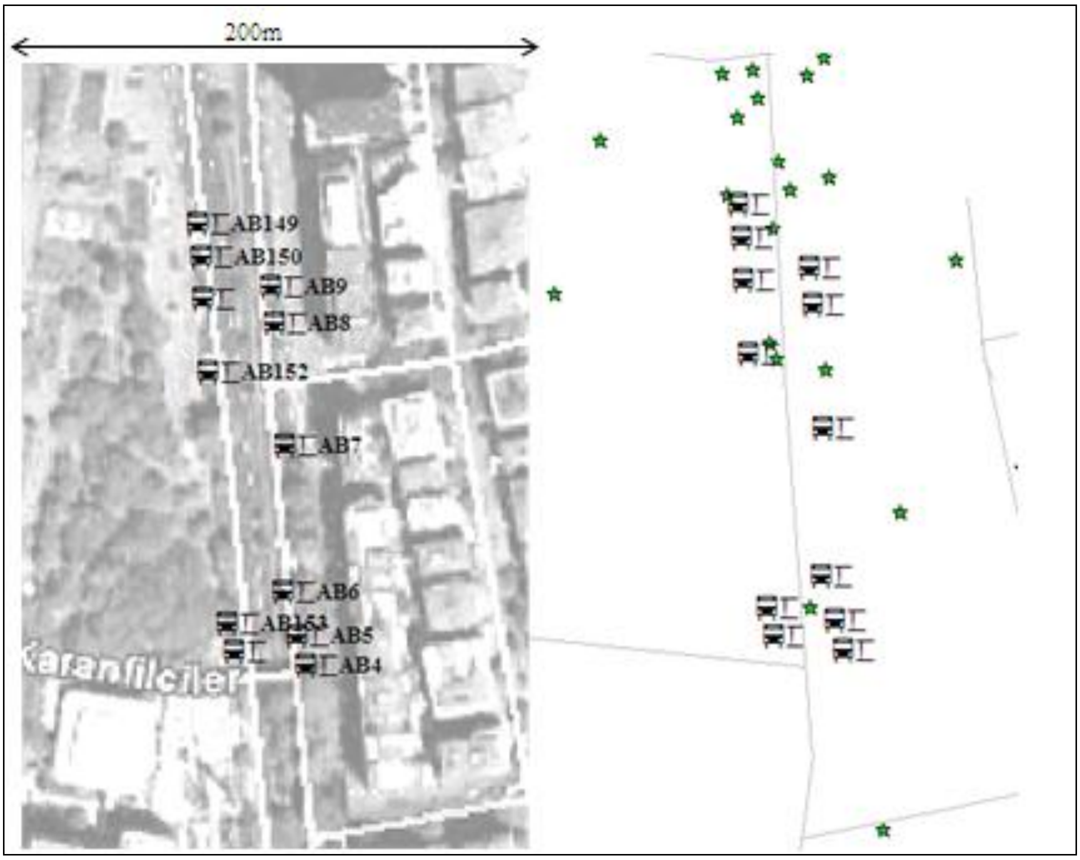
(a) (b)

Figure 5.12 For the Critical Region 2 in the CBD (a) Satellite Picture (b) Road Network, with Bus Stops



(a) (b) (c)

Figure 5.13 For the Critical Region 2 in the CBD, Thematic Maps for Bus Stop Safety for the Buffer Zone (a) 50 m (b) 100 m (c) 400 m (for a 3 Year Period)



(a) (b)

Figure 5.14 For the Critical Region 3 in the CBD (a) Satellite Picture (b) Road Network, with Bus Stops

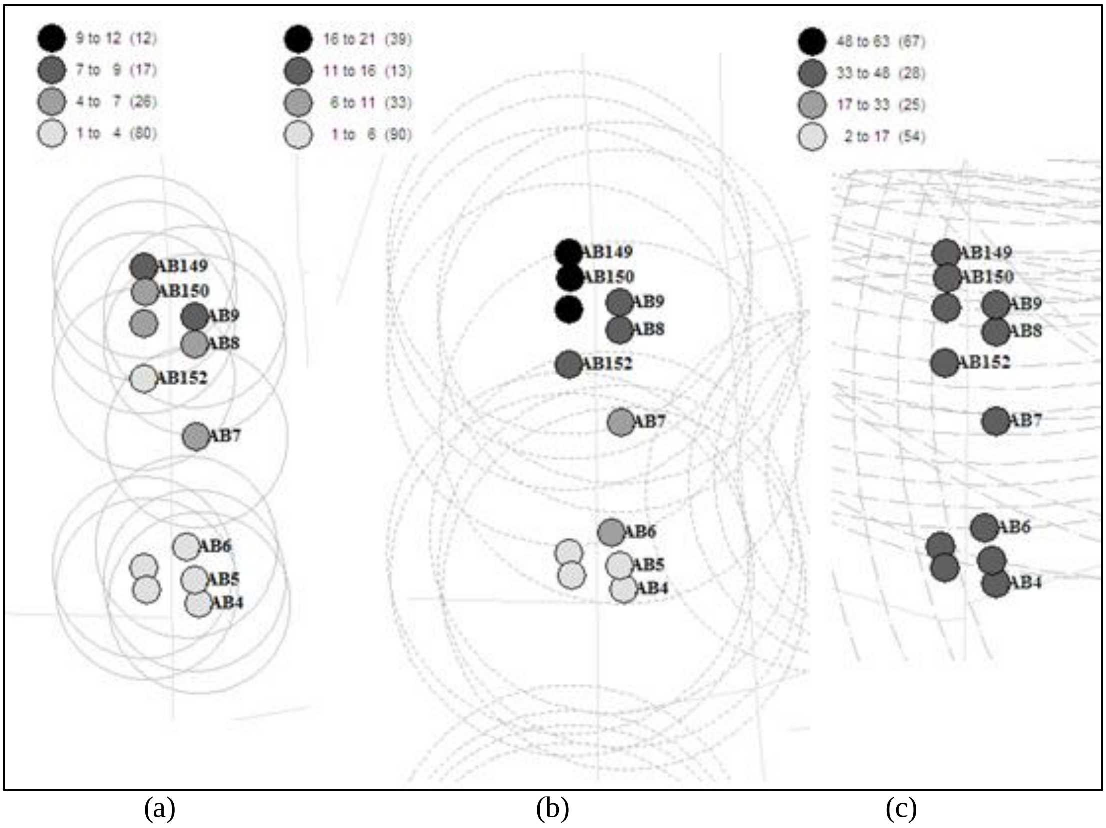
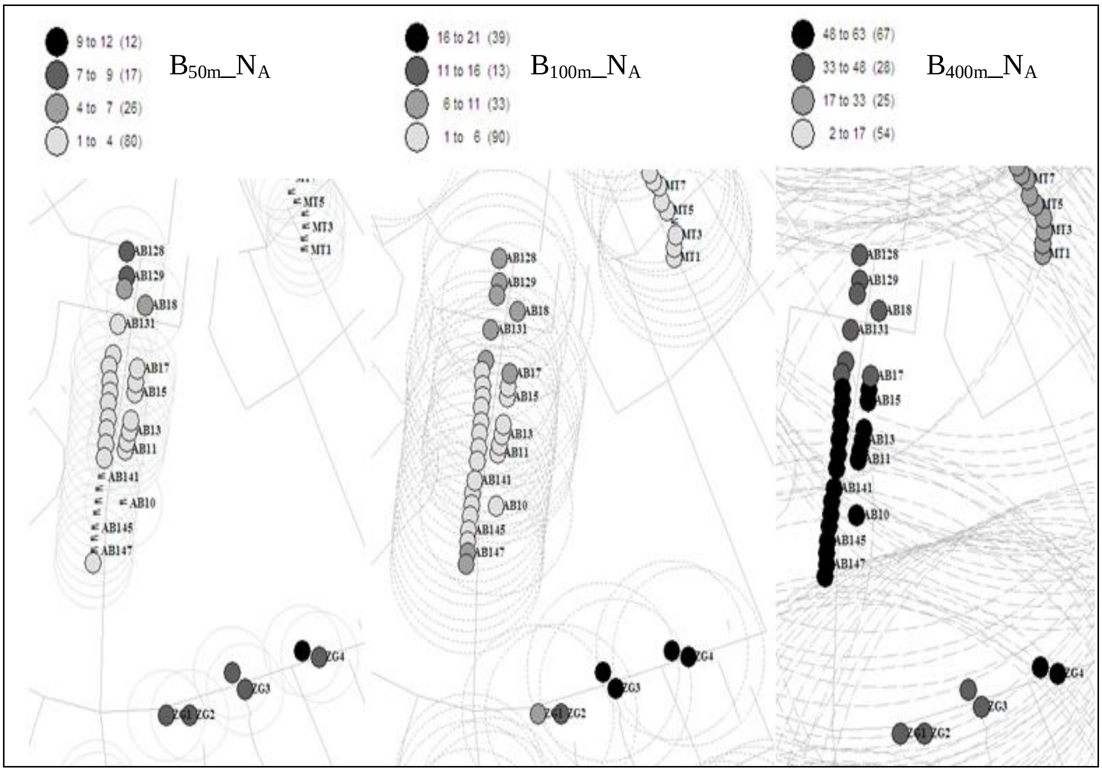


Figure 5.15 For the Critical Region 3 in the CBD, Thematic Maps for Bus Stop Safety for the Buffer Zone (a) 50 m (b) 100 m (c) 400 m (for a 3 Year Period)

(a) (b)

Figure 5.16 For the Critical Region 4 in the CBD (a) Satellite Picture (b) Road Network, with Bus Stops



(a) (b) (c)

Figure 5.17 For the Critical Region 4 in the CBD, Thematic Maps for Bus Stop Safety for the Buffer Zone (a) 50 m (b) 100 m (c) 400 m (for a 3 Year Period)

Bus Stops on UTZ Corridors

The same analysis processes were conducted for bus stops on the UTZ corridors. As a general result, the number of bus stops that have no accidents on their buffer zones decreases while the radiuses of buffer zones increase. While, the number of bus stops that there are no pedestrian accidents on is 100 for 50 m buffer zones, it is 48 for 100 m and 15 for 400 m. On the other hand, number of accidents increases for the bus stops while the radius of buffer zones increases. Thematic maps of whole bus stops on the UTZ corridors for 50 m, 100 m and 400 m buffer zone radiuses were shown in Figure 5.18, Figure 5.19 and Figure 5.20.
Three critical regions that bus stops on these regions have higher number of accidents than others were determined. Region 1 denotes the bus stops in front of intercity bus terminal and

covers very wide area. The bus stops of Region 2 and Region 3 are starting from the beginning of İrfan Baştuğ Boulevard and ends at the middle of it.

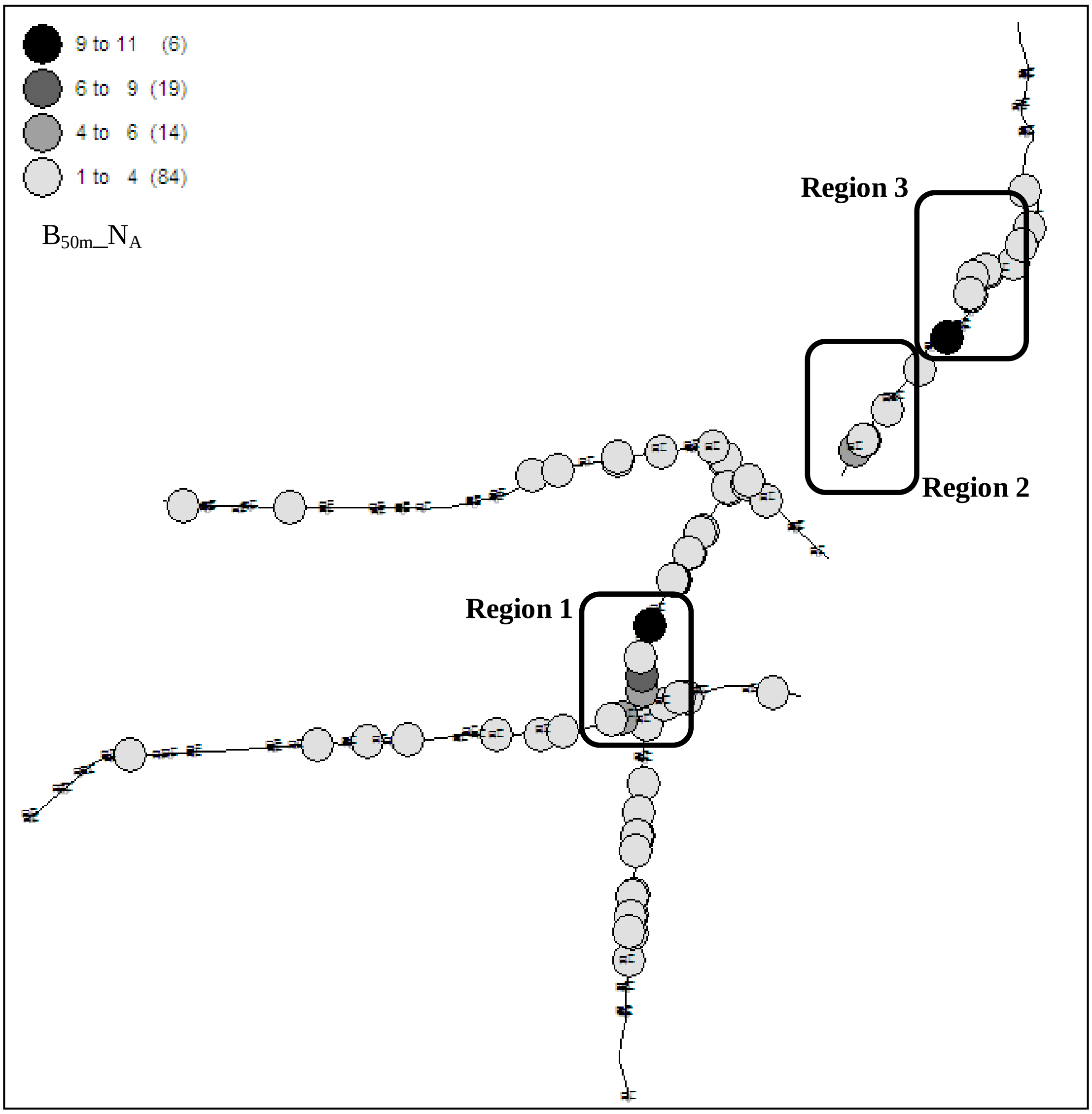


Figure 5.18 UTZ Thematic Map for Bus Stop Safety for the 50m Buffer Zone, B_{50m}_N_A (for a 3 Year Period)

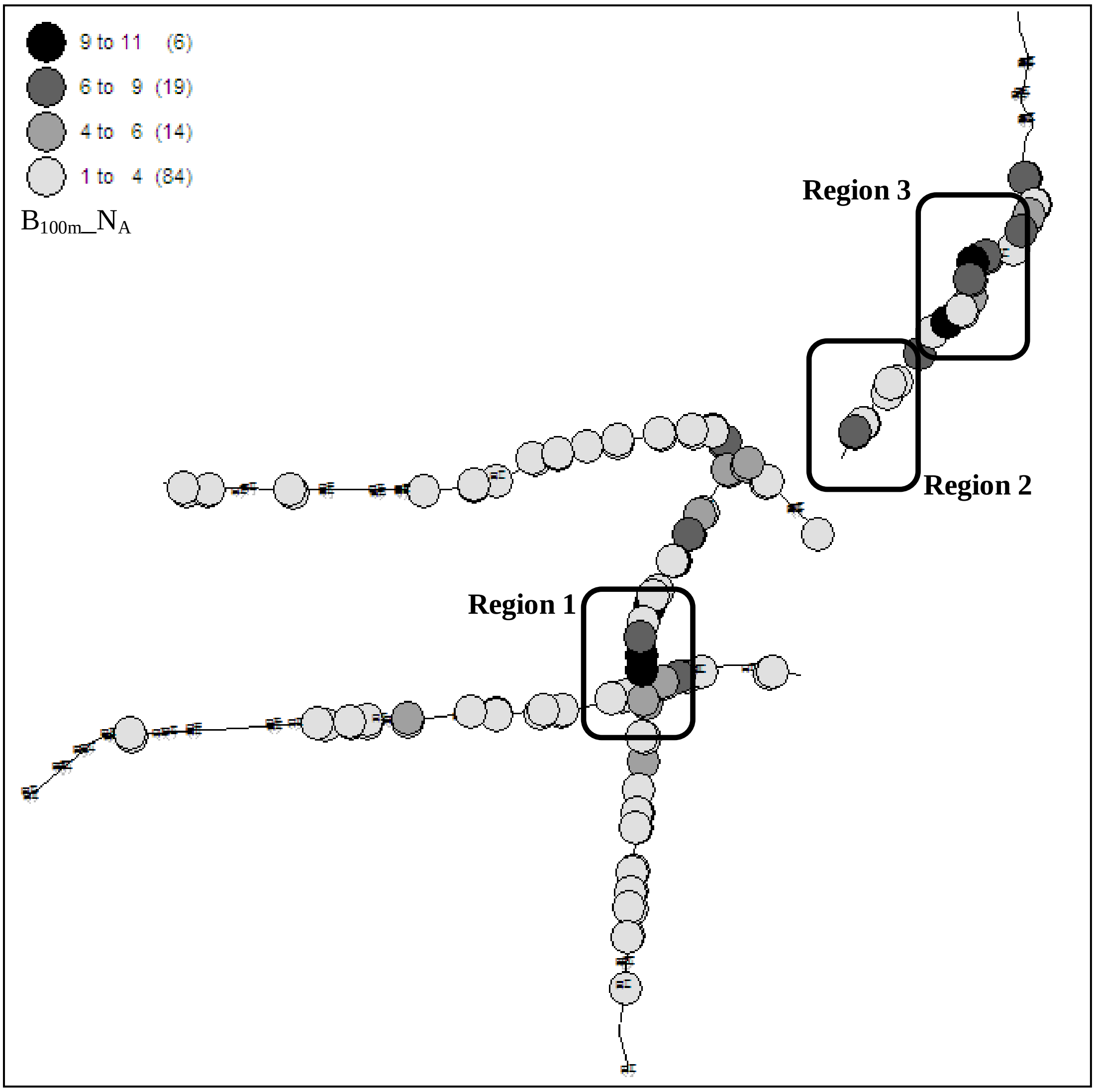


Figure 5.19 UTZ Thematic Map for Bus Stop Safety for the 100m Buffer Zone, $B_{100m}\_N_A$ (for a 3 Year Period)

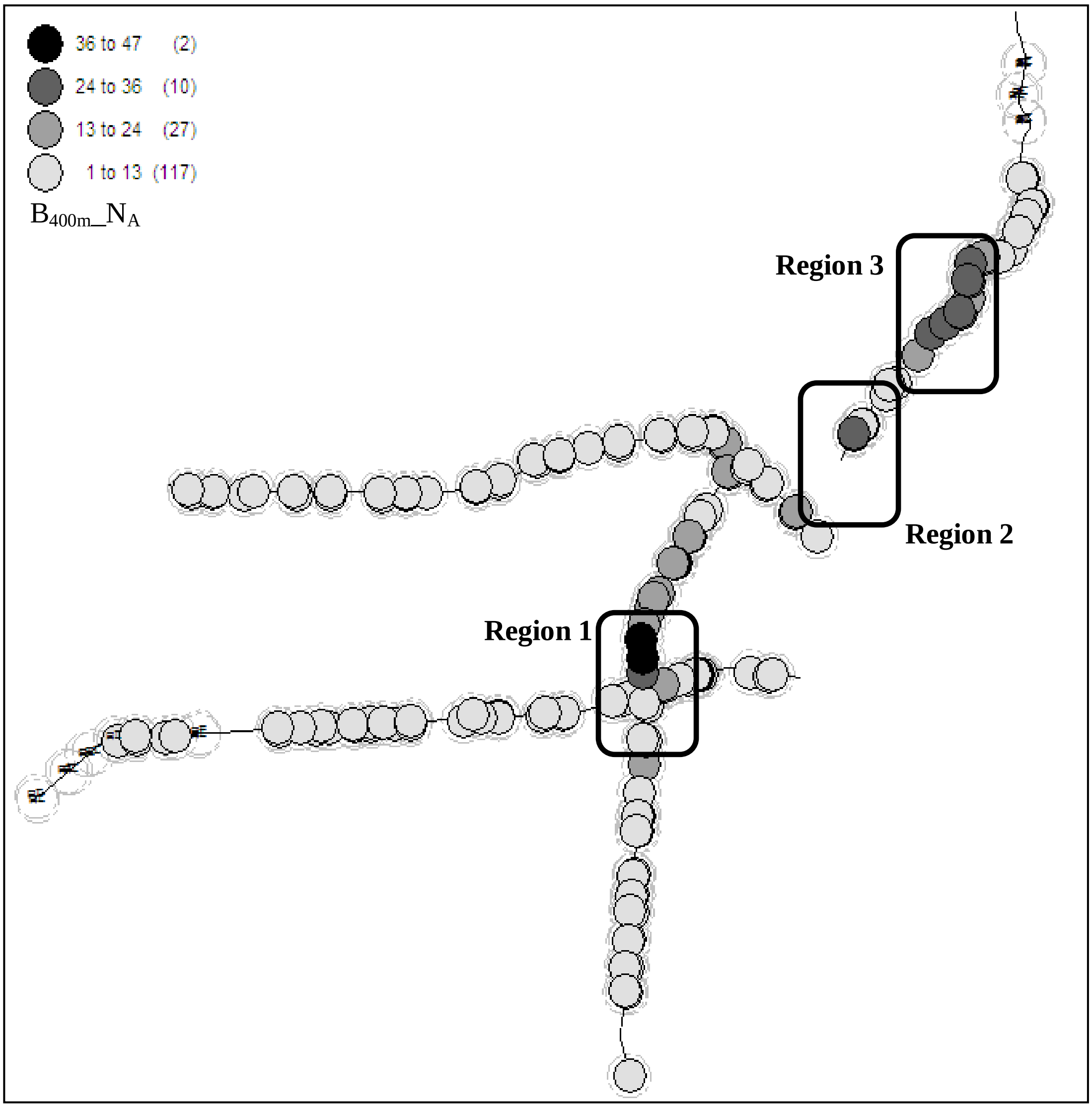
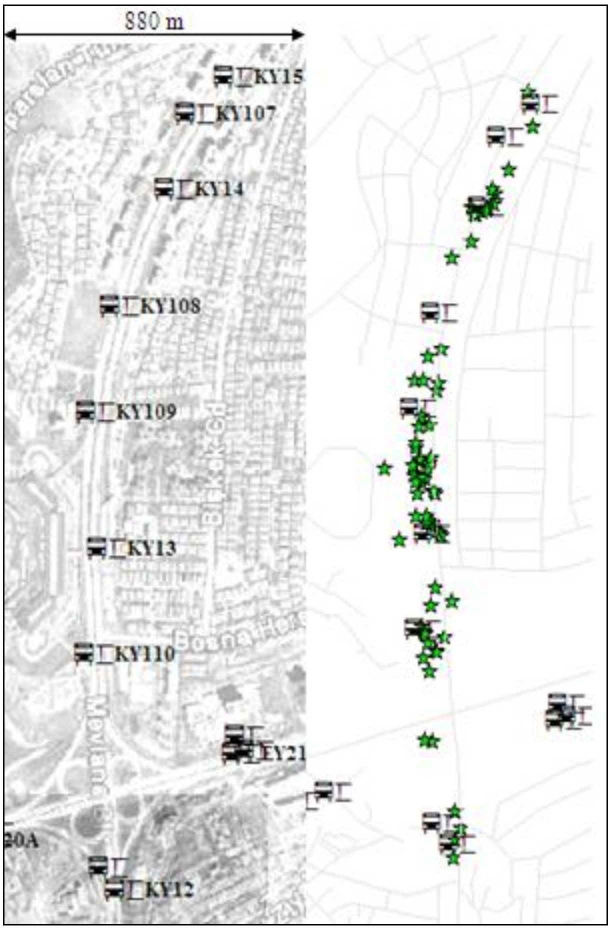


Figure 5.20 UTZ Thematic Map for Bus Stop Safety for the 400m Buffer Zone, $B_{400m}\_N_A$ (for a 3 Year Period)

As in the CBD bus stops, the regions defined as critical covers the segments that were found accident-prone in the corridor pedestrian safety analysis. The more detailed satellite pictures of these regions were taken with the geocoded bus stops on. In addition, the road network and the pedestrian accidents of the critical Region 1, Region 2 and Region 3 of UTZ were shown with the distribution of bus stops in Figure 5.21, Figure 5.23 and Figure 5.25

respectively. The change in the number of accidents in the buffer zones of bus stops was illustrated throughout thematic maps in Figure 5.22, Figure 5.24 and Figure 5.26.



(a) (b)

Figure 5.21 For the Critical Region 1 in the UTZ (a) Satellite Picture (b) Road Network, with Bus Stops

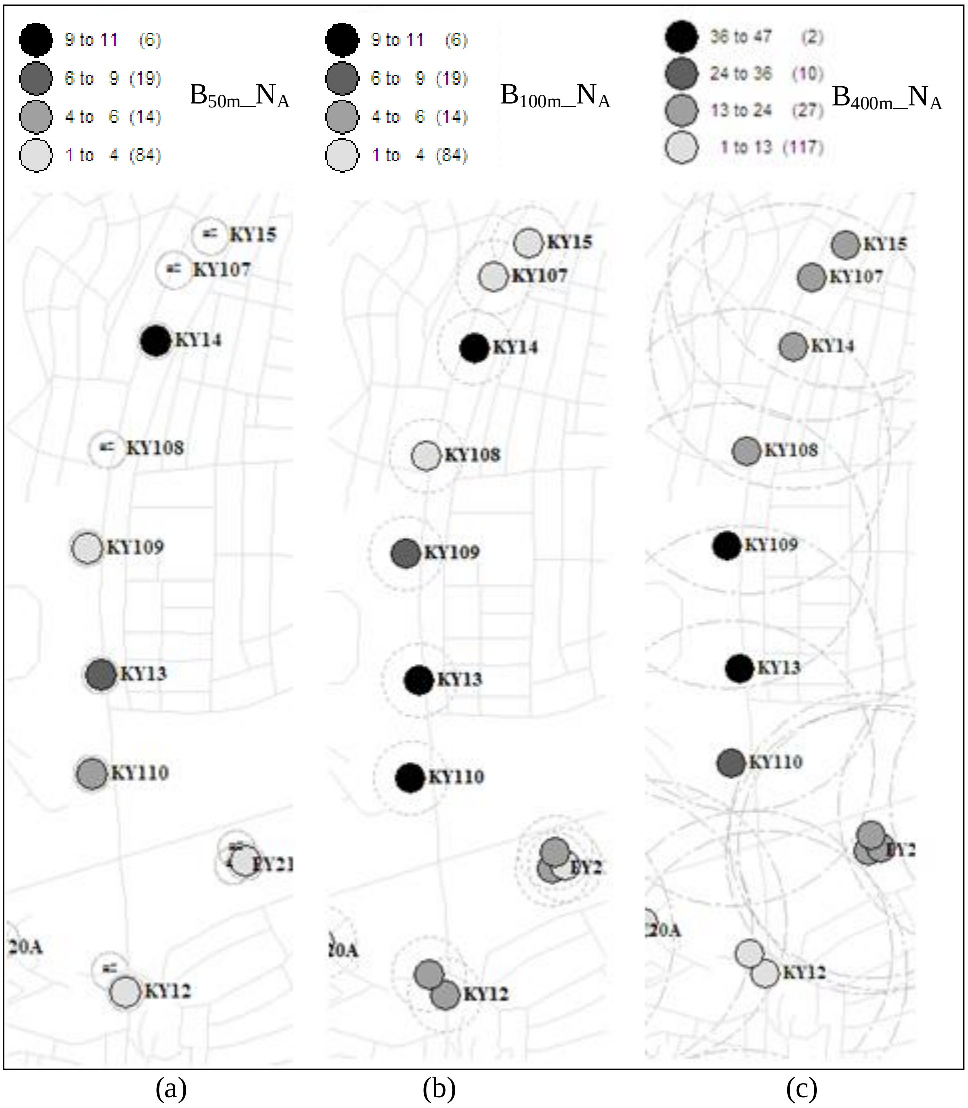
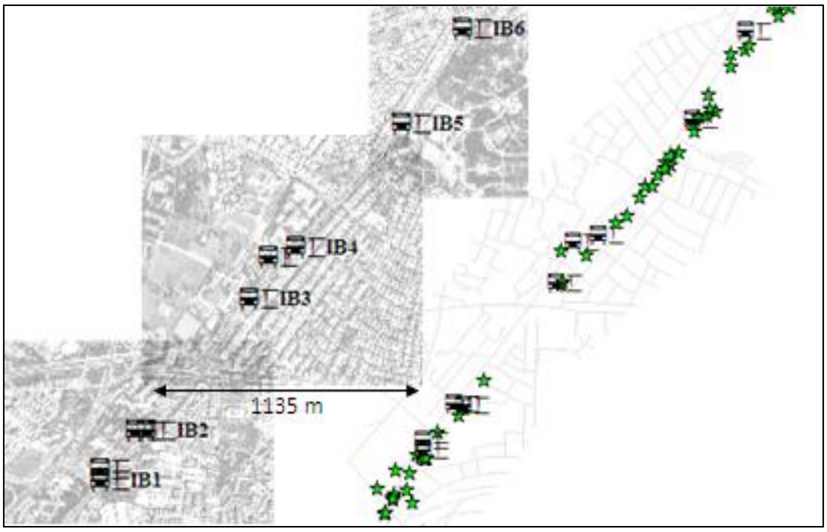
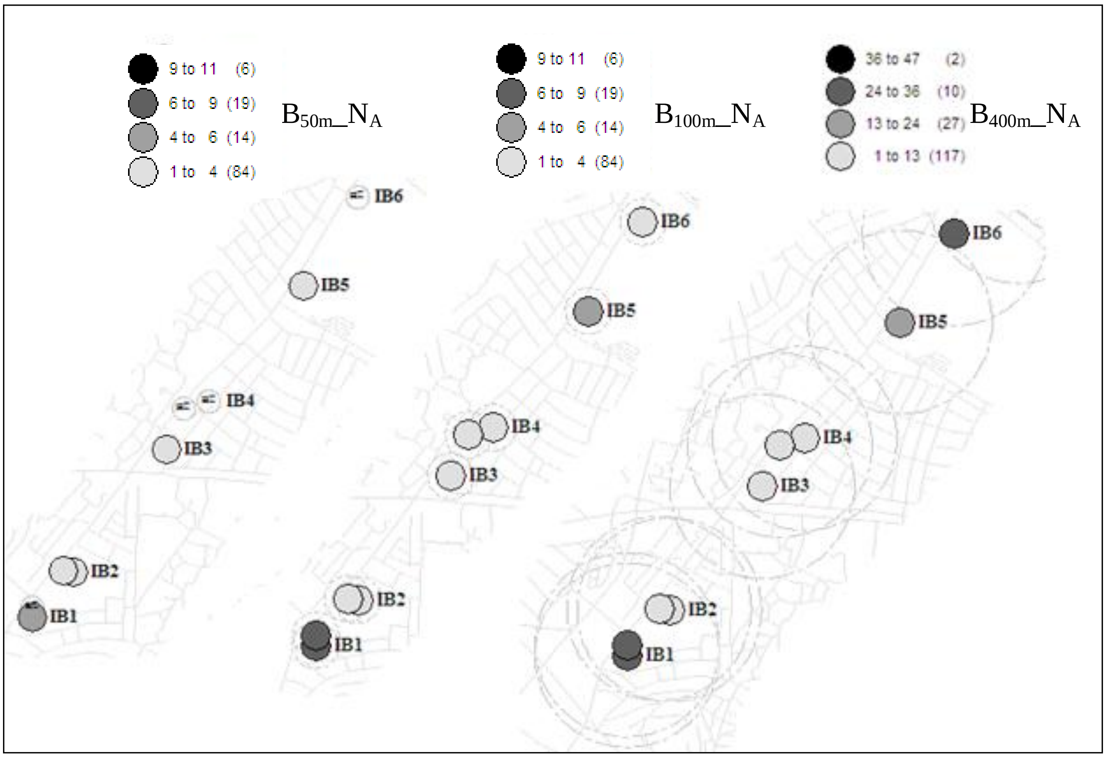


Figure 5.22 For the Critical Region 1 in the UTZ, Thematic Maps for Bus Stop Safety for the Buffer Zone (a) 50 m (b) 100 m (c) 400 m (for a 3 Year Period)



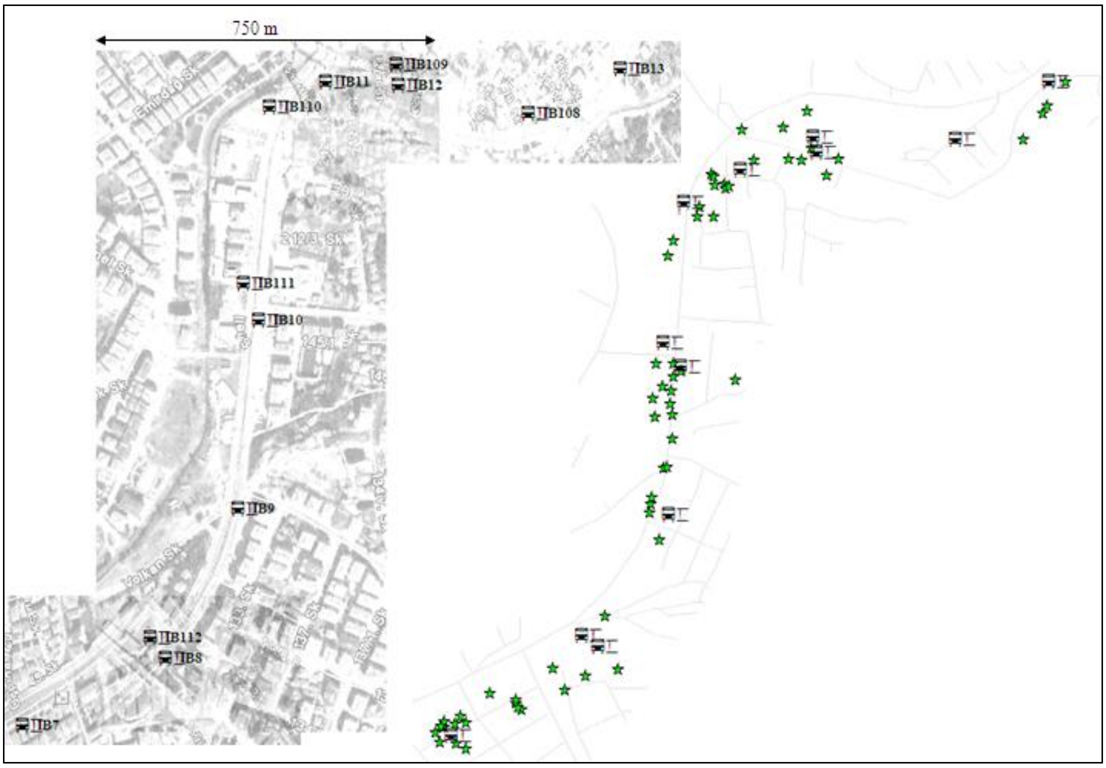
(a) (b)

Figure 5.23 For the Critical Region 2 in the UTZ (a) Satellite Picture (b) Road Network, with Bus Stops



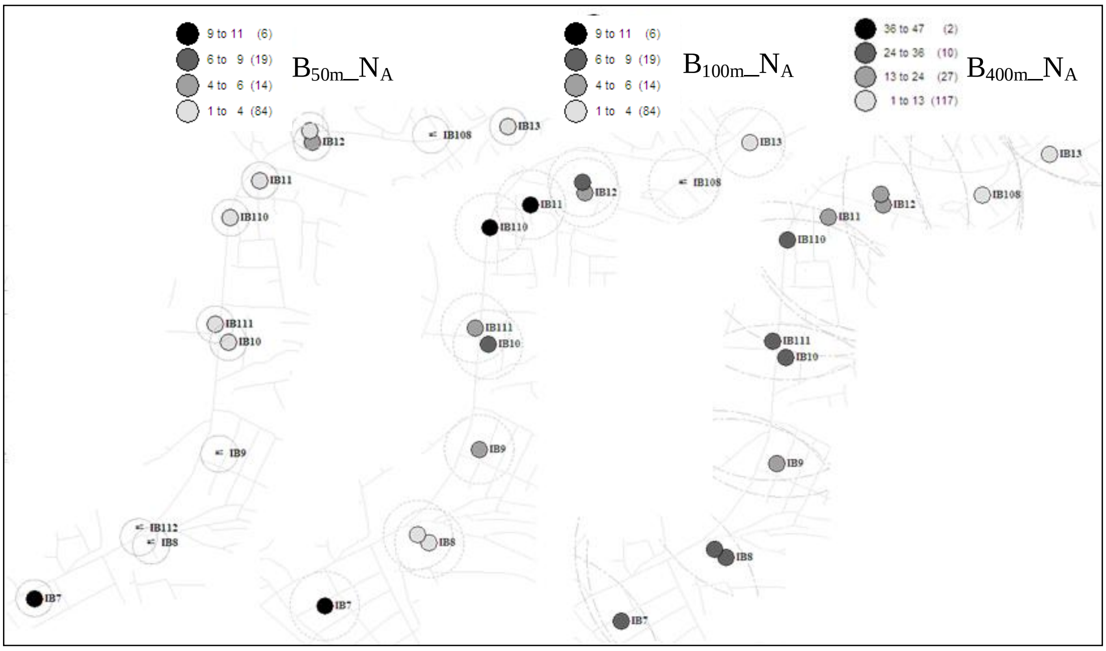
(a) (b) (c)

Figure 5.24 For the Critical Region 2 in the UTZ, Thematic Maps for Bus Stop Safety for the Buffer Zone (a) 50 m (b) 100 m (c) 400 m (for a 3 Year Period)



(a) (b)

Figure 5.25 For the Critical Region 3 in the UTZ (a) Satellite Picture (b) Road Network, with Bus Stops

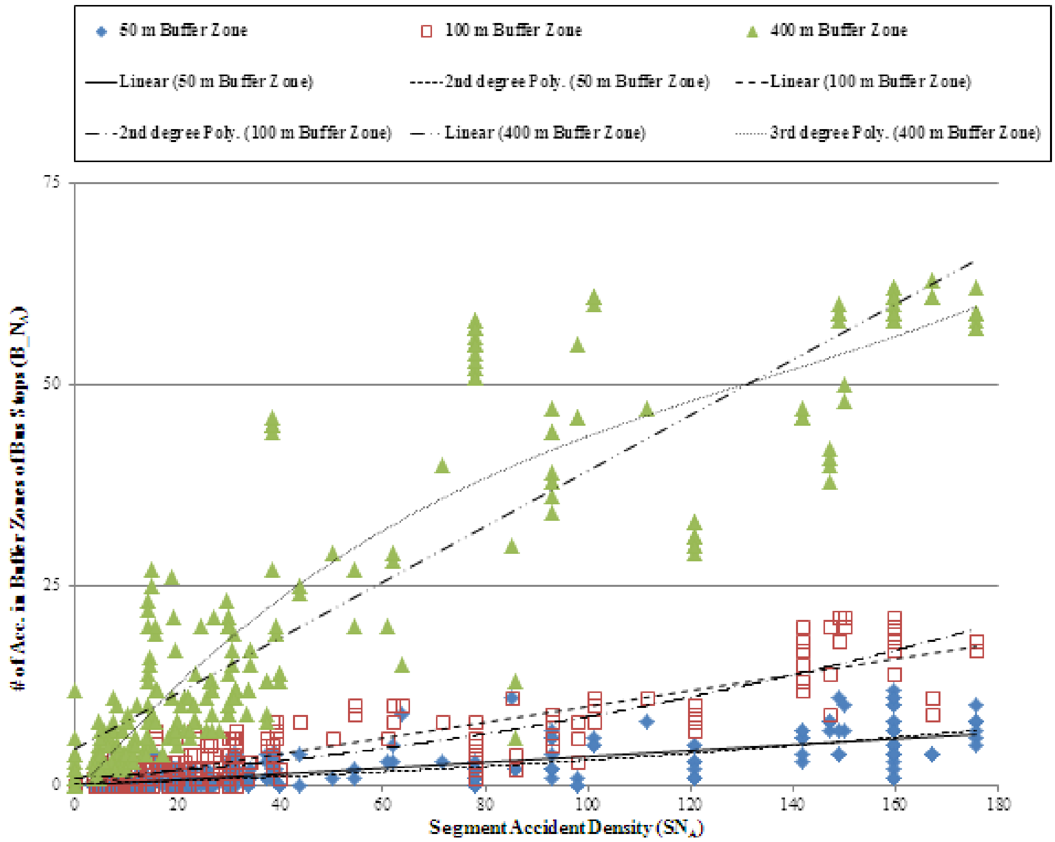


(a) (b) (c)

Figure 5.26 For the Critical Region 3 in the UTZ, Thematic Maps for Bus Stop Safety for the Buffer Zone (a) 50 m (b) 100 m (c) 400 m (for a 3 Year Period)

5.3 Corridor – Bus Stop Pedestrian Safety Correlation Analysis

The reason of analyzing both corridor and bus stop pedestrian safety in evaluating the pedestrian safety around bus stops is to investigate a relationship between the behavior of corridor segments and the bus stops' buffer zones. When the scattered graph of accident density of corridor segments and the number of accidents that fall into different dimensions of buffer zones of bus stops was plotted for all study corridors from the R-values the relation can be observed. The scattered graphs and the formulas that fit the points for all corridors are given in Figure 5.27. The y values represent the number of accidents in the buffer zones of bus stops (B_N_A) and the x values are the accident density of the segments (SN_A). Although the location of bus stops on a segment (at the beginning, middle or end of the segment) can change the relationship, a general statement can be read from the correlation analysis. As the radius of buffer zones of bus stops getting larger, the buffer zones start to behave like a corridor segment until the radiuses of buffer zones exceed the length of the segments. It can be clearly seen that for 50 m buffer zones the R-values are not that much high because the buffer zone covers only a part of a segment from the figures. However, when the radius become 100 m that is diameter is almost equal to average unit segment length, buffer zone almost represent the segment of corridors.

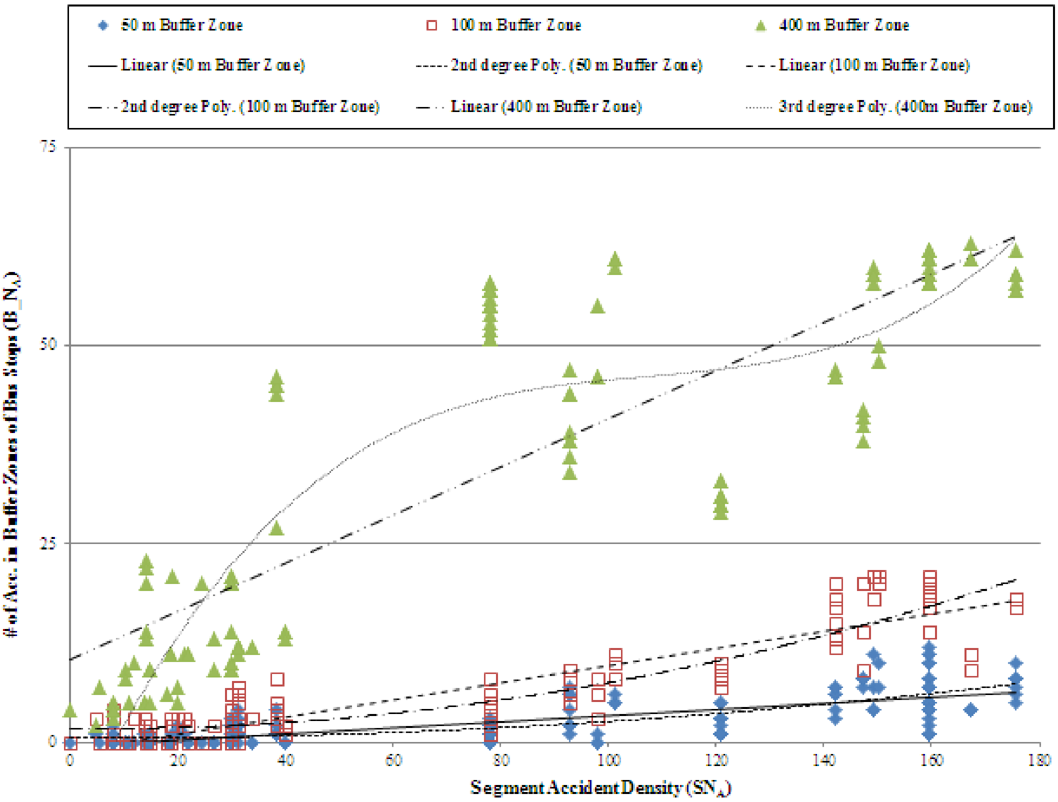


		Linear		Higher Order Polynomial	
		Formula	R	Formula	R
All	50 m	y = 0.0364x - 0.0571	0.764	$y = 0.0001x^2 + 0.0185x + 0.2344$	0.769
	100 m	y = 0.0995x + 0.0001	0.906	$y = 0.0004x^2 + 0.0374x + 1.0101$	0.918
	400 m	y = 0.3473x + 4.6072	0.895	$y = 1E\text{-}05x^3 - 0.0042x^2 + 0.765x - 1.1577$	0.914

Figure 5.27 The Graph of Corridor and Bus Stop Pedestrian Safety Correlation Results for Ankara Study Corridors

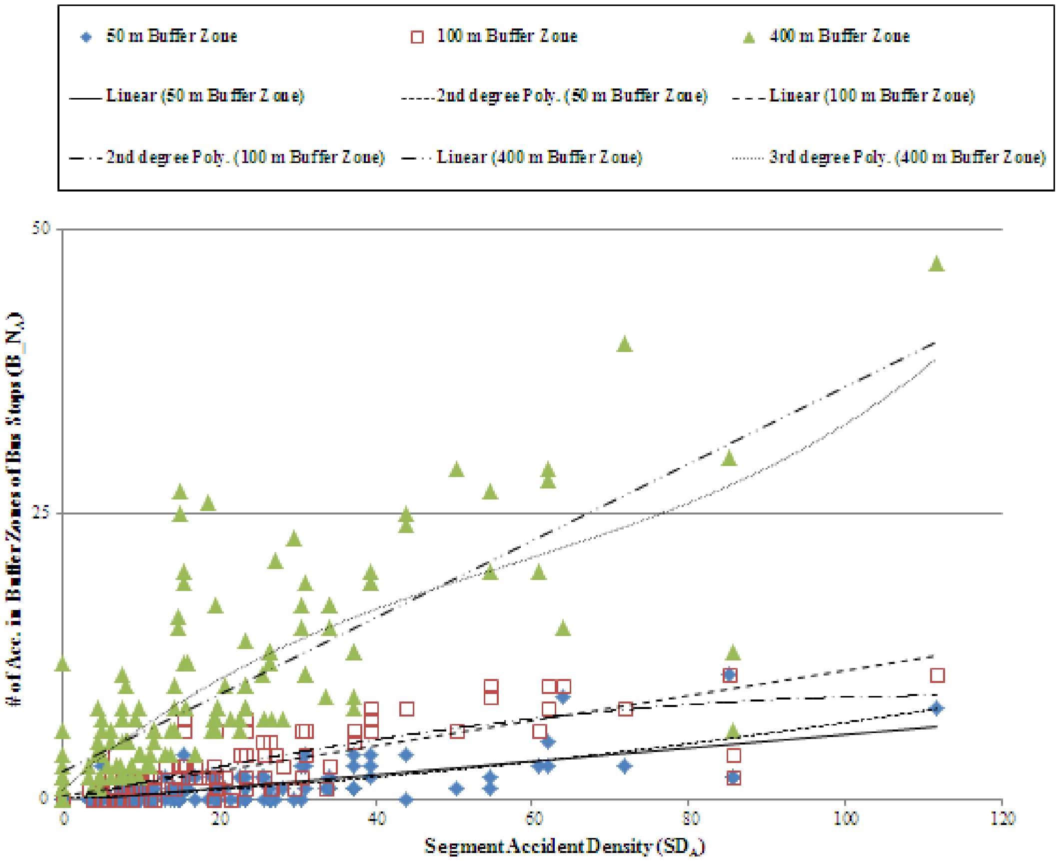
100 m buffer zones, the relationship between the two seems more linear. Nevertheless, the R-value is not exactly “1”. One reason is the location of bus stops on the segments. If a bus stop with a 100m buffer zone is located at the middle of the segment, the bus stop buffer zone will correspond to the segment stripe total; thus will include the same number of accidents. Thus, it is reasonable to see the linearity is lost when the radius of buffer zones turn into 400 m that covers more than one corridor segment. Furthermore, though minor, the length of the segment may also create variability as the correlation graph is plotted as SD_A versus bus stop safety measure, not the SN_A.

Similarly, CBD and UTZ corridor correlations are given in Figure 5.28 and Figure 5.29 respectively. The major differences between the CBD and UTZ bus stop characteristics are also observed in their correlation graphs. CBD bus stops are more linearly correlated with the segment densities for all three buffer zone definitions. While the slopes of the 100-m linear fits were found almost the same (around 0.11), the major difference is in the constant (intercept) term, which was 1.2401 accidents/bus stops in the CBD, and only 0.3381 accidents/bus stops in the UTZ. The non-linear fits provided even R-values of 0.924 for 100-m radius. On the other hand, the highest correlation in the UTZ was seen, again for a 100-m buffer zone, for a non-linear fit of R= 0.817 value.



		Linear		Higher Order Polynomial	
		Formula	R	Formula	R
CBD	50 m	y = 0.0393x - 0.5729	0.754	$y = 0.0002x^2 - 0.0029x + 0.5025$	0.774
	100 m	y = 0.1083x - 1.2401	0.900	$y = 0.0007x^2 - 0.0074x + 1.708$	0.924
	400 m	y = 0.305x + 10.328	0.838	$y = 4E\text{-}05x^3 - 0.0135x^2 + 1.5021x - 11.737$	0.889

Figure 5.28 The Graph of Corridor and Bus Stop Pedestrian Safety Correlation Results for the CBD Corridors



		Linear		Higher Order Polynomial	
		Formula	R	Formula	R
UTZ	50 m	y = 0.0589x - 0.1745	0.708	$y = 0.0003x^2 + 0.0355x + 0.0259$	0.719
	100 m	y = 0.1099x + 0.3381	0.798	$y = -0.0007x^2 + 0.163x - 0.1149$	0.817
	400 m	y = 0.3371x + 2.5125	0.783	$y = 4E\text{-}05x^3 - 0.0067x^2 + 0.6015x + 0.8484$	0.797

Figure 5.29 The Graph of Corridor and Bus Stop Pedestrian Safety Correlation Results for the UTZ Corridors

5.4 Pedestrian Safety Regression Analyses

For pedestrian safety regression analysis, two databases, road segment safety and bus stop pedestrian safety, were created based on the results of the corridor and bus stop pedestrian safety analyses, respectively. The regression analysis was conducted for segments of corridors and bus stops separately. Corridor pedestrian safety regressions produced models relating corridor pedestrian safety measures (number of accidents SN_A or segment accident density, SD_A) to preselected regression parameters described in Section 3.4 and were summarized in Table 3.11. For bus stop pedestrian safety regression models, bus stop accident numbers within a predefined buffer zone ($B_{50m}\_N_A$, $B_{100m}\_N_A$, and $B_{400m}\_N_A$), was used as the dependent variable and related to the similar set of segment characteristics.

All the regression models developed in this section were created using SPSS Version 20 software, more specifically "Linear Regression" module under the "Analyze" tab. The software also calculated the significance levels automatically. In these models, coefficients of independent variables with a significance level of 10% or less are assumed statistically significant. Models with some independent variables statistically not significant were left out of the analysis; thus, the models presented as the "best models" in this study are those that have high R-values and all the included variables are statistically significant. As the total and regression sums of square of errors would change based on the number of independent variables and cases included in the regression, a basic summary of the model performance is always presented besides the parameter coefficients and *t*-values in the model.

As many variables could not be measured (such as pedestrian flow rates, traffic flow rates, average speeds, etc.), most of them were represented qualitatively by dummy variables that simply showed existence of a special situation (such as, high pedestrian volume compared to the rest of the segments). Visual display of these generated dummy variables in the all-region was presented in the figures in Appendix C. However, certain rules of thumbs were followed to keep consistency such as, for the high pedestrian volume variable, all the segments in the CBD and segments with hospitals, shopping malls etc. in the UTZ, were defined as "high pedestrian volume" segments and given the value of "1" (see Figure C.1). Similarly, for the high-speed variable, traffic speed limits on the segments were used to decide whether it was a high-speed segment or not (see Figure C.2). The segments with underpasses or overpasses were presented with the corresponding dummy variables; the underpasses in the study were very few (only at 10 locations), but 7 of them served metro or light rail stations, thus needed to be classified separately (see Figure C.3, Figure C.4, Figure C.5).

Descriptive statistics for all the dependent and independent variables used in the regression models were presented in Appendix D for the all the segments (labeled as “All”), as well as, segments in the CBD and UTZ (see Table D.1, Table D.2, Table D.3) . Furthermore, as there were many dummy variables generated to describe the built-environment, traffic flow and network characteristics, it was necessary to check the correlations between them that may affect the success of regression analysis. The correlation tables, determined using SPSS, were also presented in Table D.4, Table D.5, Table D.6 and Table D.7.

How these variables act in the models (and even their significance levels) vary greatly between models, based on the scope of the model (whether it included all corridors in the study region, or a more specific one, such as a CBD or UTZ model), as well as the type of the analysis (whether it is a corridor or bus stop regression model). Thus, the corridor and bus stop pedestrian safety regression models will be discussed separately from this point on. Though it would be possible to model number of injuries (or the corresponding density measure) instead of number of accidents, the models were not expected to change much, as the average number of injuries per accident was close to 1 for this case study. Furthermore, there were only few segments with fatality accidents on, which was not enough to develop statistical models. Thus, they were not created in this study.

5.4.1 Factors affecting corridor pedestrian safety

To find the major factors affecting pedestrian safety on the selected corridors, 3-year accident numbers (and densities) on segments on all the corridors were modeled by linear regressions (models labeled “All”) with all the independent variables in Table 3.11. Some of these models included variables not statistically significant at $\alpha = 0.10$ level, and were not selected in this study. Models with all the variables statistically significant, were studied further to select those with higher “coefficient of determination”, that is high R-values. Summaries of the three of the good models with high R-values were shown in Figure 5.30. The only difference between the first and the second model is the addition of number of intersection and number of legs variables (*No_Intersections* and *No_Legs*) in the latter. The last model was developed without constant term (the intercept). Another difference between the first two and the last two models is the change in the statistically significant variables.

	All_Model 1		All_Model 2		All_Model 3	
Dependent Variable	SN_A		SN_A		SN_A	
Model Statistics	**R = 0.634**		**R = 0.641**		**R = 0.778**	
Total SSE	*11804.732*		*11804.732*		*18195.000*	
Regression SSE	*4738.534*		*4847.788*		*11013.404*	
Number of cases	*295*		*295*		*295*	
Independent Variables	**Coeff.**	***t* value**	**Coeff.**	***t* value**	**Coeff.**	***t* value**
Constant	-4.244	*-1.934*	-4.773	*-1.989*		
Traffic Flow						
D_High speed	-2.256	*-2.197*	-2.144	*-2.086*	-3.385	*-3.864*
D_ HighPedVol	1.581	*1.801*	1.681	*1.919*		
Traffic Network						
No_Lanes	1.040	*3.701*	1.062	*3.648*	0.638	*4.495*
Seg_L_km						
No_Intersections			-2.425	*-2.044*		
No_Legs			0.752	*1.883*		
Built Environment						
D_Urban zone	3.065	*3.632*	3.285	*3.879*	2.396	*3.447*
D_Shop						
D_Shop_2sides	2.012	*2.352*	1.942	*2.260*	1.812	*2.304*
D_BusComInd						
D_PedBarr						
D_Underpass						
D_Overpass						
D_Metro_Underpass	8.745	*4.337*	8.606	*4.247*	9.324	*4.651*
No_Bus_Stops	0.615	*4.836*	0.607	*4.796*	0.629	*4.932*

Figure 5.30 Best Regressions Models for the All Study Corridors with Segment Number of Accidents

As can be seen from the figure, all three models have close R-values; models with constants (allowing an intercept) did not perform as successfully as the model without constant. The models with the constants have *D_High speed*, *D_ HighPedVol*, *No_Lanes*, *No_Intersections*, *No_Legs*, *D_Urban zone*, *D_Shop_2sides*, *D_Metro_Underpass and No_Bus_Stops* variables as statistically significant.

Even though the t-values of the constant term were statistically significant at 10%, Models with negative constants had a problem; under certain conditions, they would predict a negative accident number on a segment, which would not possible in reality. To display the problem numerically, the formulaic representation of one of the regression model with the constant (All_Model 1) is given in Eqn. 6.

$$SN_A = -4.244 - 2.256 \times D_{Highspeed} + 1.581 \times D_{HighPedVol} + 1.040 \times N_{Lanes} + 3.065 \times D_{Urbanzone} + 2.012 \times D_{Shop\_2sides} + 8.745 \times D_{Metro\_Underpass} + 0.615 \times N_{Bus\_Stops} \qquad \text{(Eqn. 6)}$$

The model in Eqn. 6, will estimate a segment accident number of -5.460 for a "1-lane high speed road which is not in the urban zone and does not have high pedestrian volume; has no shopping region both sides, no metro underpass and no bus stops on it". This weakness suggested that one should choose linear models without constant terms to be on the safe side. (The negative intercept value also provided the first clue for a potential non-linear relationship between the number of accidents on a segment and the selected variables). Thus, all of the models that were presented in this study will be developed without constant term from now on.

The model without constant (All_Model 3) had major change in the coefficients of statistically significant variables in Models 1 and 2. Furthermore, dummy variable for high pedestrian volume variable was not statistically significant anymore in Model 3. Assuming Model 3 as the best, one can write an equation of segments' number of accidents as follows:

$$SN_A = -3.385 \times D_{Highspeed} + 0.638 \times N_{Lanes} + 2.396 \times D_{Urbanzone} + 1.812 \times D_{Shop\_2sides} + 9.324 \times D_{Metro\_Underpass} + 0.629 \times N_{Bus\_Stops} \qquad \text{(Eqn. 7)}$$

The performance of similar four good models for segment densities, SDA's, was summarized in Figure 5.31. The R-values were found very close but smaller than to those in the models with SN_A's, except for the second model (All_Model 2 with SD_A), which had slightly higher R-value. However, the constant term in this model was not statistically significant. Omitting the constant term, better and more reliable models were found as the third and fourth models in the Figure 5.31. Another noticeable result different from SN_A models, was the statistically significance of *Seg_L_km* variable in SD_A models. Most probably small variances in the lengths of the segments were resulted in this situation. In addition, the negative behavior of the *Seg_L_km* variable can be explained by the reason of the fact that it was denominator in the calculation of SD_A values.

	All_Model 1		All_Model 2		All_Model 3		All_Model 4	
Dependent Variable	SD_A		SD_A		SD_A		SD_A	
Model Statistics	**R = 0.630**		**R = 0.644**		**R = 0.765**		**R = 0.774**	
Total SSE	*307048.165*		*307048.165*		*454805.426*		*454805.426*	
Regression SSE	*121899.725*		*127326.575*		*265969.72*		*272394.701*	
Number of cases	*295*		*295*		*295*		*295*	
Independent Variables	**Coeff.**	***t* value**	**Coeff.**	***t* value**	**Coeff.**	***t* value**	**Coeff.**	***t* value**
Constant	-23.936	*-2.131*	-0.169	*-0.012*				
Traffic Flow								
D_High speed	-12.513	*-2.381*	-11.757	*-2.264*	-18.900	*-4.208*	-16.276	*-3.618*
D_HighPedVol	8.962	*1.995*	8.329	*1.876*				
Traffic Network								
No_Lanes	5.631	*3.916*	5.745	*4.047*	3.368	*4.629*	5.877	*5.520*
Seg_L_km			-112.311	*-2.939*			-94.843	*-3.185*
No_Intersections								
No_Legs								
Built Environment								
D_Urban zone	13.956	*3.231*	13.839	*3.247*	10.194	*2.861*	15.012	*3.929*
D_Shop								
D_Shop_2sides	11.274	*2.574*	11.185	*2.588*	10.163	*2.520*	13.354	*3.261*
D_BusComInd								
D_PedBarr								
D_Underpass								
D_Overpass								
D_Metro_Underpass	40.829	*3.956*	39.385	*3.862*	44.093	*4.290*	40.317	*3.957*
No_Bus_Stops	3.152	*4.845*	3.246	*5.050*	3.232	*4.945*	3.289	*5.110*

Figure 5.31 Best Regressions Models for the All Study Corridors with Segment Accident Density

It should be noted that number of bus stops was statistically significant in all SN_A or SD_A based models. Though the impact was found rather small (around 0.6 acc/one bus stop in SN_A models) compared to other variables; considering the average number of bus stops of 1.166; one can say that on average, the impact of bus stops on segment accidents would a total of 0,6*1,166 accidents in total. However, this becomes more critical in CBD region segments where the average number of bus stops is 2.431, adding a total accident value of 0.6*2.431. Although the coefficient of number of bus stops on a segment rather small, from the *t-values* of *No_Bus_Stops* variable, it can be seen that the effect of it definite.

As the characteristics of the CBD and UTZ differ from each other's significantly, the factors that affect the pedestrian safety can change for these regions. Thus, models for segment accident numbers in these two regions were created separately (see Figure 5.32).

	CBD_Model 1		CBD_Model 2		UTZ_Model 1		UTZ_Model 2	
Dependent Variable	SN_A		SN_A		SN_A		SN_A	
Model Statistics	**R = 0.859**		**R = 0.860**		**R = 0.762**		**R = 0.770**	
Total SSE	*11120.000*		*11120.000*		*7075.000*		*7075.000*	
Regression SSE	*8213.692*		*8223.966*		*4105.466*		*4189.396*	
Number of cases	*72*		*72*		*223*		*223*	
	Coeff.	***t* value**	**Coeff.**	***t* value**	**Coeff.**	***t* value**	**Coeff.**	***t* value**
Independent Variables								
Constant								
Traffic Flow								
D_High speed					-2.818	*-4.124*	-2.786	*-4.104*
D_ HighPedVol								
Traffic Network								
No_Lanes	2.713	*6.682*	2.427	*6.643*				
Seg_L_km					15.617	*4.791*	10.726	*2.787*
No_Intersections							-3.158	*-2.264*
No_Legs							1.204	*2.419*
Built Environment								
D_Urban zone					2.583	*4.333*	2.477	*4.150*
D_Shop								
D_Shop_2sides			3.583	*2.080*	3.476	*3.632*	3.433	*3.600*
D_BusComInd	-6.257	*-3.355*	-7.943	*-3.673*	1.344	*2.605*	1.406	*2.741*
D_PedBarr	-4.099	*-2.019*						
D_Underpass								
D_Overpass								
D_Metro_Underpass	14.916	*4.074*	12.603	*3.443*				
No_Bus_Stops	0.616	*3.222*	0.508	*2.672*				

Figure 5.32 Best Regressions Models for the CBD and the UTZ Corridors with Segment Number of Accidents

As it was discussed before, the models without constant term perform better. Thus, the CBD and UTZ models were developed without constant terms also. The results of the models show that there was a more linear relationship between the dependent and independent variables of CBD region than ones of UTZ. In addition, the statistically significant variables of these two regions vary considerably. One of the significant differences can be observed in existence of variable *No_Bus_Stops* in the CBD models as significant but not in UTZ models. Firstly, in the CBD region, segments show more common properties than UTZ segments. These variations among UTZ segments also created not so linearly relationship between selected dependent and independent parameters. In addition, there were more segments with no accidents in the UTZ corridors than the CBD ones. Also, the higher density of bus stops in the CBD make *No_Bus_Stops* variable important.

There SD_A models of segments (Figure D.1) did not show major change from the SN_A models of CBD and UTZ corridors.

Addressing Non-linearity in Pedestrian Regression Models

To improve the R-values and avoid negative intercept cases, a simple logarithmic transformation of the dependent variables (SN_A or SD_A) was performed. The major change in these models was the omission of segments without accidents, as logarithm of "0" was not defined. In another way, models with $ln(SN_A)$ (and also with $ln(SD_A)$ relates safety measures with segment characteristics for those segments that have accident history only, which were 237 out of 295 segments. Best models of all and CBD, UTZ corridors with logarithmic transformation for SN_A and SD_A were presented in Figure 5.33 and Figure 5.34 respectively.

Models with logarithmic transformation performed better in terms of R-values compared to regular variables. Especially for CBD models, R-values were considerably high which means that the real relationship between dependent and independent variables is non-linear. While R-values for SN_A and SD_A models were close to each other, R-values for $ln(SD_A)$ were certainly better than those models with $ln(SN_A)$. The statistically significant variables did not change from the results of SN_A and SD_A models.

	All_Model 1		All_Model 1	
Dependent Variable	ln(SN_A)		ln(SD_A)	
Model Statistics	**R = 0.870**		**R = 0.962**	
Total SSE	*603.444*		*2106.599*	
Regression SSE	*456.262*		*1949.84*	
Number of cases	*237*		*237*	
	Coeff.	***t* value**	**Coeff.**	***t* value**
Independent Variables				
Constant				
Traffic Flow				
D_High speed	-0.452	*-3.024*		
D_ HighPedVol			0.571	*4.417*
Traffic Network				
No_Lanes	0.145	*5.894*	0.275	*20.573*
Seg_L_km				
No_Intersections				
No_Legs				
Built Environment				
D_Urban zone	0.637	*5.225*	0.953	*7.667*
D_Shop				
D_Shop_2sides	0.416	*2.942*	0.497	*3.219*
D_BusComInd				
D_PedBarr				
D_Underpass				
D_Overpass				
D_Metro_Underpass				
No_Bus_Stops	0.074	*3.656*	0.067	*3.218*

Figure 5.33 Best Regression Models for All Study Corridors with Logarithmic Transformation of the Dependent Variables

	CBD_Model 1		CBD_Model 1		UTZ_Model 1		UTZ_Model 1	
Dependent Variable	$\ln(SN_A)$		$\ln(SD_A)$		$\ln(SN_A)$		$\ln(SD_A)$	
Model Statistics	**R = 0.943**		**R = 0.981**		**R = 0.866**		**R = 0.956**	
Total SSE	*260.758*		*807.827*		*342.686*		*1298.771*	
Regression SSE	*231.65*		*777.707*		*257.159*		*1187.398*	
Number of cases	*69*		*69*		*168*		*168*	
	Coeff.	***t* value**	**Coeff.**	***t* value**	**Coeff.**	***t* value**	**Coeff.**	***t* value**
Independent Variables								
Constant								
Traffic Flow								
D_High speed					-0.504	*-3.633*		
D_ HighPedVol							0.591	*3.974*
Traffic Network								
No_Lanes	0.444	*10.006*	0.757	*16.883*			0.277	*20.367*
Seg_L_km					5.777	*7.624*		
No_Intersections								
No_Legs								
Built Environment								
D_Urban zone					0.507	*4.021*	0.994	*7.821*
D_Shop								
D_Shop_2sides	0.618	*3.255*	0.871	*4.629*	1.010	*3.783*	0.637	*2.089*
D_BusComInd	-1.027	*-4.328*	-0.915	*-3.836*				
D_PedBarr	-0.550	*-2.461*	-0.850	*-3.842*	-0.220	*-1.852*		
D_Underpass								
D_Overpass	0.385	*1.775*	0.413	*1.889*				
D_Metro_Underpass	0.821	*2.112*						
No_Bus_Stops	0.060	*2.982*	0.053	*2.677*				

Figure 5.34 Best Regression Models for the CBD and UTZ Corridors with Logarithmic Transformation of the Dependent Variables

As a summary all of the good models created in corridor pedestrian safety regression models for different regions and with different dependent variables were presented in Table 5.1. Although the statistically significant variables vary from model to model, the common properties of the variables are the direction of the impact on the number of accidents (decreasing or increasing). Based on the models that were developed the general behaviors of independent variables on the dependent ones are summarized in Figure 5.35.

Table 5.1 Summary of Corridor Pedestrian Safety Regression Models

Corridor Models	Dependent Variable	Constant	Number of cases	Total SSE	R
All_Model 1	SN_A	Yes	295	11804.732	0.634
All_Model 2	SN_A	Yes	295	11804.732	0.641
All_Model 3	SN_A	No	295	18195.000	0.778
All_Model 1	SD_A	Yes	295	307048.165	0.63
All_Model 2	SD_A	Yes	295	307048.165	0.644
All_Model 3	SD_A	No	295	454805.426	0.765
All_Model 4	SD_A	No	295	454805.426	0.774
CBD_Model 1	SN_A	No	72	11120.000	0.859
CBD_Model 2	SN_A	No	72	11120.000	0.860
UTZ_Model 1	SN_A	No	223	7075.000	0.762
UTZ_Model 2	SN_A	No	223	7075.000	0.770
CBD_Model 1	SD_A	No	72	305565.402	0.850
CBD_Model 2	SD_A	No	72	305565.402	0.865
UTZ_Model 1	SD_A	No	223	149240.024	0.740
UTZ_Model 2	SD_A	No	223	149240.024	0.745
All_Model 1	$ln(SN_A)$	No	237	603.444	0.870
CBD_Model 1	$ln(SN_A)$	No	69	260.758	0.943
UTZ_Model 1	$ln(SN_A)$	No	168	342.686	0.866
All_Model 1	$ln(SD_A)$	No	237	2106.599	0.962
CBD_Model 1	$ln(SD_A)$	No	69	807.827	0.981
UTZ_Model 1	$ln(SD_A)$	No	168	1298.771	0.956

As it can be observed from Figure 5.35, the behaviors of most of the variables were in the same direction for different models. The opposite behavior was observed in for only two variables that are *Seg_L_km* and *D_BusComInd*. While segment length showed positive behavior in SN_A models, it was negative in SD_A models. The reason behind this situation is the existence of segment length as denominator in the calculations of SD_A values. Thus, it was meaningful to behave in opposite directions in SN_A and SD_A models. The second difference was observed in the behavior of *D_BusComInd*. While it had negative affect in the CBD, the affect became positive in the UTZ due to the definition of the variable. As it was discussed before, the land-use definitions cover only the major characteristics of the region. As it can be observed in Figure 4.6, the major characteristic of the CBD was determined as shopping. Although shopping could be assumed as business as well, in the CBD business land use characteristic was defined for the private sector offices where mostly. Although, all of the CBD corridors were defined as high pedestrian volume corridors (Figure C.1), when the distribution of accidents was observed (Figure 4.3), the areas defined as shopping had higher density of pedestrian accidents than areas defined as business. Thus, the occurrence of the accidents in the CBD firstly associated positively with land use characteristic of shopping. On the other hand, business regions in UTZ corridors generally identify the industrial areas where large number of people work. Thus, these areas naturally attraction zones with high pedestrian volume, which resulted in positive association with the number of accidents in UTZ corridors.

	SN_A Models			SD_A Models			$\ln(SN_A)$ Models			$\ln(SD_A)$ Models		
	ALL	CBD	UTZ	ALL	CBD	UTZ	ALL	CBD	UTZ	ALL	CBD	UTZ
Traffic Flow												
D_High speed	↓		↓	↓					↓			
D_HighPedVol	↑			↑		↑	↑			↑		↑
Traffic Network												
No_Lanes	↑	↑		↑	↑	↑	↑	↑		↑	↑	↑
Seg_L_km			↑	↓					↑			
No_Intersections	↓		↓						↓			↓
No_Legs	↑		↑						↑			↑
Built Environment												
D_Urban zone	↑		↑	↑		↑	↑		↑	↑		↑
D_Shop												
D_Shop_2sides	↑	↑	↑	↑	↑	↑	↑	↑	↑	↑	↑	↑
D_BusComInd		↓	↑		↓			↓			↓	
D_PedBarr		↓			↓			↓	↓		↓	
D_Underpass												
D_Overpass								↑			↑	
D_Metro_Underpass	↑	↑		↑	↑			↑				
No_Bus_Stops	↑	↑		↑	↑	↑	↑	↑		↑	↑	

Figure 5.35 Behaviors of Independent Variables in Corridor Pedestrian Safety Models

5.4.2 Factors Affecting Bus Stop Pedestrian Safety

Similar to corridor pedestrian safety, factors affecting pedestrian safety can be sought after by regressing bus stop pedestrian safety measures over the selected independent variables. As bus stop pedestrian safety was defined for different buffer zone sizes, and the relationship of these bus stop variables had different levels of correlation with their link densities (see Section 5.3), a series of bus stop pedestrian safety regression models using linear relationship was developed as shown in Figure 5.36. The regression models developed for all, CBD and UTZ bus stops separately as in the regression analyses of the segments due to the different characteristics of the corridors. The models were developed with dependent variable B_N_A and also $ln(B\_N_A)$ for different sizes of the bus stops' buffer zones. The results of the bus stops regression models for 50, 100 and 400 m buffer zones' radiuses of all bus stops were presented for in Figure 5.36, Figure 5.37 and Figure 5.38 respectively.

Firstly, the relationship between the number of accidents in the buffer zones and accident densities of the corridor segments that the bus stops take place was discovered throughout the regression analysis. The R-value of these models (labeled as All_Model 1 with dependent variables of $B_{50m}\_N_A$, $B_{100m}\_N_A$ and $B_{400m}\_N_A$) were considerably high. Substantially, the same results were found in Section 5.3, in the correlation analysis and the proof was conducted with regression analysis. Thus, if the data of accident densities of segments of corridors exists, one can predict the number of accidents that could occur at the vicinity of bus stops on these segments.

Secondly, models were developed including other variables of segments excluding the accident densities of them (labeled as All_Model 2 with dependent variables of $B_{50m}\_N_A$, $B_{100m}\_N_A$ and $B_{400m}\_N_A$). The observations from the results showed that the R-values and the coefficients of statistically significant variables were getting higher when the radiuses of buffer zones changed from 50 m to 400 m. While the variables of *D_High speed, No_Lanes, D_Urban zone, D_Shop_2sides, D_BusComInd* were found as statistically significant in 50 m and 100 m models, *D_ HighPedVol, D_Overpass and D_Metro_Underpass* variables added as significant in 400 m models. The formula for predicting the numbers of accidents for buffer zones with 50 m, 100 m and 400 m were presented in Eqn. 8, Eqn. 9 and Eqn. 10 respectively.

$$B_{50m}\_N_A = -1.719 \times D_{Highspeed} + 0.364 \times N_{Lanes} + 0.721 \times D_{Urbanzone} + 1.255 \times D_{Shop\_2sides} - 1.020 \times D_{BusComInd} \quad \text{(Eqn. 8)}$$

$$B_{100m}\_N_A = -4.326 \times D_{Highspeed} + 0.904 \times N_{Lanes} + 2.095 \times D_{Urbanzone} + 2.859 \times D_{Shop\_2sides} - 1.971 \times D_{BusComInd} \quad \text{(Eqn. 9)}$$

$$B_{400m}\_N_A = -15.705 \times D_{Highspeed} + 4.562 \times D_{HighPedVol} + 3.220 \times N_{Lanes} + 6.334 \times D_{Urbanzone} + 11.699 \times D_{Shop\_2sides} - 4.801 \times D_{BusComInd} - 8.008 \times D_{BusComInd} + 15.071 \times D_{Metro\_Underpass} \quad \text{(Eqn. 10)}$$

Thirdly, models developed by combining All_Model 1 and All_Model 2 for different buffer zones (labeled as All_Model 3 with dependent variables of $B_{50m}\_N_A$, $B_{100m}\_N_A$ and $B_{400m}\_N_A$). Although, the R-values were considerably high in these models, due to the high association of the segment accident density variable with the number of accidents in the buffer zones, the number of statistically significant variables decreased in these models. In addition, the general behaviors of the statistically significant variables, which were expected and found in the corridor pedestrian safety regression analysis changed in the opposite direction. Substantially, most of the relationship was explained by segment accident density variable, thus, when the models were developed with this one and other variable together, most of the other statistically significant variables were capturing the residual effects. Except 400 m buffer zones' model, both the number of statistically significant variables decreased and the behaviors changed in the opposite direction. Thus, for reliable results, the models should be developed for segment accident density and other variables separately.

The very same models were developed for CBD and UTZ bus stops also which are in Appendix D. The results were found for CBD and UTZ bus stops showed parallelism with all bus stops were analyzed together. The power of segment accident density was valid for CBD and UTZ bus stops too. In addition, when the models were developed both segment accident density variable and other variables, the results were still problematic for CBD and UTZ corridors

	All_Model 1		All_Model 2		All_Model 3	
Dependent Variable	B_{50m}_N_A		B_{50m}_N_A		B_{50m}_N_A	
Model Statistics	**R = 0.848**		**R = 0.693**		**R = 0.853**	
Total SSE	*3389.000*		*3389.000*		*3389.000*	
Regression SSE	*2435.673*		*1628.723*		*2463.749*	
Number of cases	*344*		*344*		*344*	
Independent Variables	**Coeff.**	***t* value**	**Coeff.**	***t* value**	**Coeff.**	***t* value**
Constant						
Segment Accidents						
Seg_Acc_Den	0.036	*29.603*			0.038	*28.272*
Traffic Flow						
D_High speed			-1.719	*-4.304*		
D_ HighPedVol						
Traffic Network						
No_Lanes			0.364	*5.446*		
Seg_L_km						
No_Intersections						
No_Legs						
Built Environment						
D_Urban zone			0.721	*1.950*		
D_Shop						
D_Shop_2sides			1.225	*4.122*		
D_BusComInd			-1.020	*-3.184*		
D_PedBarr						
D_Underpass						
D_Overpass						
D_Metro_Underpass					-0.961	*-3.221*

Figure 5.36 Best Regression Models for Bus Stop Pedestrian Safety Using Number of Accidents for 50 m Buffer Zones (All Corridors)

	All_Model 1		All_Model 2		All_Model 3	
Dependent Variable	B_{100m}_N_A		B_{100m}_N_A		B_{100m}_N_A	
Model Statistics	**R = 0.948**		**R = 0.762**		**R = 0.957**	
Total SSE	*20778.000*		*20778.000*		*20778.000*	
Regression SSE	*18660.979*		*12066.394*		*19038.918*	
Number of cases	*344*		*344*		*344*	
Independent Variables	**Coeff.**	***t* value**	**Coeff.**	***t* value**	**Coeff.**	***t* value**
Constant						
Segment Accidents						
Seg_Acc_Den	0.100	*54.986*			0.109	*41.171*
Traffic Flow						
D_High speed			-4.326	*-4.868*	0.420	*2.067*
D_ HighPedVol					-0.394	*-1.646*
Traffic Network						
No_Lanes			0.904	*6.087*		
Seg_L_km						
No_Intersections						
No_Legs						
Built Environment						
D_Urban zone			2.095	*2.545*		
D_Shop						
D_Shop_2sides			2.859	*4.325*		
D_BusComInd			-1.971	*-2.765*		
D_PedBarr						
D_Underpass						
D_Overpass						
D_Metro_Underpass					-3.278	*-7.953*

Figure 5.37 Best Regression Models for Bus Stop Pedestrian Safety Using Number of Accidents for 100 m Buffer Zones (All Corridors)

	All_Model 1		All_Model 2		All_Model 3	
Dependent Variable	$B_{400m}\_N_A$		$B_{400m}\_N_A$		$B_{400m}\_N_A$	
Model Statistics	**R = 0.946**		**R = 0.894**		**R = 0.967**	
Total SSE	*319681.000*		*319681.000*		*319681.000*	
Regression SSE	*285892.975*		*255642.073*		*298841.638*	
Number of cases	*344*		*344*		*344*	
Independent Variables	**Coeff.**	***t* value**	**Coeff.**	***t* value**	**Coeff.**	***t* value**
Constant						
Segment Accidents						
Seg_Acc_Den	0.390	*53.873*			0.277	*26.916*
Traffic Flow						
D_High speed			-15.705	*-5.410*	-4.414	*-2.970*
D_ HighPedVol			4.562	*1.886*		
Traffic Network						
No_Lanes			3.220	*6.926*	0.880	*3.559*
Seg_L_km						
No_Intersections						
No_Legs						
Built Environment						
D_Urban zone			6.334	*2.784*	3.026	*2.504*
D_Shop						
D_Shop_2sides			11.699	*5.930*	4.525	*4.364*
D_BusComInd			-4.801	*-2.445*		
D_PedBarr						
D_Underpass						
D_Overpass			-8.008	*-4.284*	-3.221	*-3.015*
D_Metro_Underpass			15.071	*5.729*	10.049	*6.676*

Figure 5.38 Best Regression Models for Bus Stop Pedestrian Safety Using Number of Accidents for 400 m Buffer Zones (All Corridors)

Although the R-values of models, which were developed with dependent variable of B_N_A, were considerably high, to stick to the linear regression procedure, the logarithmic transformation was applied to the B_N_A values and results were illustrated for 50 m, 100 m and 400 m in Figure 5.39, Figure 5.40 and Figure 5.41 respectively. Models were labeled as All_Model 1, All_Model 2, and All_Model 3 with dependent variables of $ln(B_{50m}\_N_A)$, $ln(B_{100m}\_N_A)$ and $ln(B_{400m}\_N_A)$. Generally, most of the R-values increased when logarithmic transformation was applied to the dependent variable. Although, exactly the same independent variables were modeled with B_N_A models, in some models' R-values decreased because the number of statistically significant variables also decreased when the dependent variable was changed. Another noticeable result from these models was the any of the variables were statistically significant except segment accident density one when the segment accident density variable were modeled with other characteristics of the segment. In addition, this logarithmic transformation was applied to the models of CBD and UTZ bus stops and they can be observed in Appendix D

	All_Model 1		All_Model 2	
Dependent Variable	$\ln(B_{50m}\_N_A)$		$\ln(B_{50m}\_N_A)$	
Model Statistics	**R = 0.863**		**R = 0.774**	
Total SSE	*257.383*		*257.383*	
Regression SSE	*191.604*		*154.324*	
Number of cases	*198*		*198*	
Independent Variables	**Coeff.**	***t* value**	**Coeff.**	***t* value**
Constant				
Segment Accidents				
Seg_Acc_Den	0.010	*23.955*		
Traffic Flow				
D_High speed			-0.620	*-4.180*
D_ HighPedVol				
Traffic Network				
No_Lanes			0.140	*8.906*
Seg_L_km				
No_Intersections				
No_Legs				
Built Environment				
D_Urban zone				
D_Shop				
D_Shop_2sides			0.283	*2.518*
D_BusComInd				
D_PedBarr				
D_Underpass				
D_Overpass				
D_Metro_Underpass				

Figure 5.39 Best Regression Models for Bus Stop Pedestrian Safety Using Logarithmic Transformation of Number of Accidents for 50 m Buffer Zones (All Corridors)

	All_Model 1		All_Model 2		All_Model 3	
Dependent Variable	$\ln(B_{100m}\_N_A)$		$\ln(B_{100m}\_N_A)$		$\ln(B_{100m}\_N_A)$	
Model Statistics	**R = 0.929**		**R = 0.858**		**R = 0.946**	
Total SSE	*784.017*		*784.017*		*784.017*	
Regression SSE	*676.648*		*577.298*		*701.743*	
Number of cases	*288*		*288*		*288*	
Independent Variables	**Coeff.**	***t* value**	**Coeff.**	***t* value**	**Coeff.**	***t* value**
Constant						
Segment Accidents						
Seg_Acc_Den	0.019	*42.529*			0.015	*24.550*
Traffic Flow						
D_High speed			-0.617	*-3.786*	0.236	*3.814*
D_ HighPedVol						
Traffic Network						
No_Lanes			0.220	*6.775*		
Seg_L_km						
No_Intersections						
No_Legs						
Built Environment						
D_Urban zone			0.471	*2.887*	0.398	*7.063*
D_Shop						
D_Shop_2sides			0.623	*4.551*		
D_BusComInd			-0.403	*-2.902*		
D_PedBarr			-0.257	*-2.221*		
D_Underpass						
D_Overpass						
D_Metro_Underpass						

Figure 5.40 Best Regression Models for Bus Stop Pedestrian Safety Using Logarithmic Transformation of Number of Accidents for 100 m Buffer Zones (All Corridors)

	All_Model 1		All_Model 2		All_Model 3	
Dependent Variable	ln(B_{400m}_N_A)		ln(B_{400m}_N_A)		ln(B_{400m}_N_A)	
Model Statistics	**R = 0.869**		**R = 0.969**		**R = 0.982**	
Total SSE	*2671.099*		*2671.099*		*2671.099*	
Regression SSE	*2017.08*		*2506.672*		*2574.125*	
Number of cases	*329*		*329*		*329*	
Independent Variables	**Coeff.**	***t* value**	**Coeff.**	***t* value**	**Coeff.**	***t* value**
Constant						
Segment Accidents						
Seg_Acc_Den	0.033	*31.806*			0.011	*16.116*
Traffic Flow						
D_High speed			-0.290	*-1.887*		
D_ HighPedVol			0.497	*3.904*	0.260	*3.129*
Traffic Network						
No_Lanes			0.216	*8.922*	0.132	*14.583*
Seg_L_km						
No_Intersections						
No_Legs						
Built Environment						
D_Urban zone			1.320	*11.111*	1.145	*13.817*
D_Shop						
D_Shop_2sides			0.674	*6.055*	0.277	*3.451*
D_BusComInd			-0.322	*-3.015*		
D_PedBarr						
D_Underpass						
D_Overpass			-0.211	*-2.105*		
D_Metro_Underpass			0.461	*3.353*	0.278	*2.654*

Figure 5.41 Best Regression Models for Bus Stop Pedestrian Safety Using Logarithmic Transformation of Number of Accidents for 400 m Buffer Zones (All Corridors)

A table summarizing the model fit statistics was presented in Table 5.2 to identify the performance of the models. Further, the behaviors of the independent variables in the models were given in Figure 5.42.

Table 5.2 Summary of Bus Stop Pedestrian Safety Regression Models

Bus Stop Models	Dependent Variable	Constant	Number of cases	Total SSE	R
All_Model 1	$B_{50m}\_N_A$	No	344	3389.000	0.848
All_Model 2	$B_{50m}\_N_A$	No	344	3389.000	0.693
All_Model 3	$B_{50m}\_N_A$	No	344	3389.000	0.853
All_Model 1	$B_{100m}\_N_A$	No	344	20778.000	0.948
All_Model 2	$B_{100m}\_N_A$	No	344	20778.000	0.762
All_Model 3	$B_{100m}\_N_A$	No	344	20778.000	0.957
All_Model 1	$B_{400m}\_N_A$	No	344	319681.000	0.946
All_Model 2	$B_{400m}\_N_A$	No	344	319681.000	0.894
All_Model 3	$B_{400m}\_N_A$	No	344	319681.000	0.967
All_Model 1	$ln(B_{50m}\_N_A)$	No	198	257.383	0.863
All_Model 2	$ln(B_{50m}\_N_A)$	No	198	257.383	0.774
All_Model 1	$ln(B_{100m}\_N_A)$	No	288	784.017	0.929
All_Model 2	$ln(B_{100m}\_N_A)$	No	288	784.017	0.858
All_Model 3	$ln(B_{100m}\_N_A)$	No	288	784.017	0.946
All_Model 1	$ln(B_{400m}\_N_A)$	No	329	2671.099	0.869
All_Model 2	$ln(B_{400m}\_N_A)$	No	329	2671.099	0.969
All_Model 3	$ln(B_{400m}\_N_A)$	No	329	2671.099	0.982
CBD_Model 1	$B_{50m}\_N_A$	No	174	2845.000	0.868
CBD_Model 2	$B_{50m}\_N_A$	No	174	2845.000	0.805
CBD_Model 3	$B_{50m}\_N_A$	No	174	2845.000	0.88
CBD_Model 1	$B_{100m}\_N_A$	No	174	18761.000	0.955
CBD_Model 2	$B_{100m}\_N_A$	No	174	18761.000	0.851
CBD_Model 3	$B_{100m}\_N_A$	No	174	18761.000	0.968
CBD_Model 1	$B_{400m}\_N_A$	No	174	296413.000	0.95
CBD_Model 2	$B_{400m}\_N_A$	No	174	296413.000	0.952
CBD_Model 3	$B_{400m}\_N_A$	No	174	296413.000	0.977
CBD_Model 1	$ln(B_{50m}\_N_A)$	No	128	212.939	0.889
CBD_Model 2	$ln(B_{50m}\_N_A)$	No	128	212.939	0.851
CBD_Model 3	$ln(B_{50m}\_N_A)$	No	128	212.939	0.899
CBD_Model 1	$ln(B_{100m}\_N_A)$	No	166	625.059	0.964
CBD_Model 2	$ln(B_{100m}\_N_A)$	No	166	625.059	0.911
CBD_Model 3	$ln(B_{100m}\_N_A)$	No	166	625.059	0.969
CBD_Model 1	$ln(B_{400m}\_N_A)$	No	174	2011.281	0.91
CBD_Model 2	$ln(B_{400m}\_N_A)$	No	174	2011.281	0.987
CBD_Model 3	$ln(B_{400m}\_N_A)$	No	174	2011.281	0.991
UTZ_Model 1	$B_{50m}\_N_A$	No	170	544.000	0.782
UTZ_Model 2	$B_{50m}\_N_A$	No	170	544.000	0.577
UTZ_Model 1	$B_{100m}\_N_A$	No	170	2017.000	0.888
UTZ_Model 2	$B_{100m}\_N_A$	No	170	2017.000	0.743
UTZ_Model 3	$B_{100m}\_N_A$	No	170	2017.000	0.897
UTZ_Model 1	$B_{400m}\_N_A$	No	170	23268.000	0.888
UTZ_Model 2	$B_{400m}\_N_A$	No	170	23268.000	0.840
UTZ_Model 3	$B_{400m}\_N_A$	No	170	23268.000	0.921
UTZ_Model 1	$ln(B_{50m}\_N_A)$	No	70	44.444	0.794
UTZ_Model 2	$ln(B_{50m}\_N_A)$	No	70	44.444	0.718
UTZ_Model 1	$ln(B_{100m}\_N_A)$	No	122	158.958	0.855
UTZ_Model 2	$ln(B_{100m}\_N_A)$	No	122	158.958	0.792
UTZ_Model 3	$ln(B_{100m}\_N_A)$	No	122	158.958	0.867
UTZ_Model 1	$ln(B_{400m}\_N_A)$	No	155	659.818	0.852
UTZ_Model 2	$ln(B_{400m}\_N_A)$	No	155	659.818	0.936
UTZ_Model 3	$ln(B_{400m}\_N_A)$	No	155	659.818	0.958

	$B_{50m}\_N_A$ Models			$B_{100m}\_N_A$ Models			$B_{400m}\_N_A$ Models			$\ln(B_{50m}\_N_A)$ Models			$\ln(B_{100m}\_N_A)$ Models			$\ln(B_{400m}\_N_A)$ Models		
	ALL	CBD	UTZ	ALL	CBD	UTZ	ALL	CBD	UTZ	ALL	CBD	UTZ	ALL	CBD	UTZ	ALL	CBD	UTZ
Segment Accidents																		
Seg_Acc_Den	↑	↑	↑	↑	↑	↑	↑	↑	↑	↑	↑	↑	↑	↑	↑	↑	↑	↑
Traffic Flow																		
D_High speed	↓		↓	↓			↓			↓			↓			↓		
D_HighPedVol						↑	↑		↑						↑	↑		↑
Traffic Network																		
No_Lanes	↑	↑	↑	↑	↑	↑	↑	↑	↑	↑	↑		↑	↑	↑	↑	↑	↑
Seg_L_km																		
No_Intersections																		
No_Legs																		
Built Environment																		
D_Urban zone	↑		↑	↑		↑	↑		↑			↑	↑		↑	↑		↑
D_Shop																		
D_Shop_2sides	↑	↑	↑	↑	↑		↑	↑		↑	↑		↑	↑		↑	↑	
D_BusComInd	↓	↓		↓	↓		↓	↓			↓		↓	↓		↓	↓	
D_PedBarr									↓		↓		↓	↓			↓	
D_Underpass																		
D_Overpass		↑					↓									↓		
D_Metro_Underpass							↑	↑	↑			↑				↑	↑	
No_Bus_Stops																		

Figure 5.42 Behaviors of Independent Variables in Bus Stop Pedestrian Safety Models

5.4.3 Predicting Pedestrian Safety

Once a regression model is developed, it is also possible to use it to determine future safety measure values (or predict), if any of the segment's characteristics would be changed (such as adding a new bus stop on the segment). For the former, it would be acceptable to work with 3-year period total values, as an accident at any time in the period would contribute to determine the major factors. On the other hand, prediction is generally done with more uniform time step such as "per year" and thus, requires working with "annual average" accident numbers or densities per segment. If the regression models were linear, this difference would not create any problems; however, non-linear models developed with 3-year data versus "annual average" values may not have the same success in explaining the relationship between the dependent and independent variables.

The power of regression analysis is formula of the dependent variable in terms of independent ones is obtained at the end of the process. If any change in the independent variables is conducted, the effect of this change on the dependent variable can be calculated from these formulas. In previous sections of this chapter, the formulas of the number of accidents in corridor segments and bus stops were given in Eqn. 6, Eqn. 7, Eqn. 8, Eqn. 9 and Eqn. 10. However, the regression analyses were conducted to the 3 years total accident data, which can not be used in the yearly predictions.

If linear models want to be used in yearly predictions of the pedestrian safety, using an annual average (AA) accident values of the dependent variables ($AAN_A = N_A/3$ or $AAD_A = D_A/3$) is more suitable. The structure of the models developed for $AASN_A$, $AASD_A$ variables were exactly same with the models of SN_A and SD_A variables (see Figure 5.43). Only difference was that the coefficients of variables of the AA models were scaled down by 1/3 from the corresponding 3-year models. However, the same ratio was not valid between the coefficients of logarithmic transformations of annual average and 3-year accident values that created better models in explaining pedestrian safety for 3-year accident data in corridor segments and around bus stops, as discussed in Sections 5.4.1 and 5.4.2.

To develop a prediction model using AA safety measures in a non-linear regression requires further care. A straightforward logarithmic transformation of the annual average segment accident values ($ln(AASN_A)$) resulted in a major drop in the R-values, while models with $ln(AASD_A)$ did not suffer such change (see Figure 5.45). In addition, still there was a decrease in the R-value of $ln(AASD_A)$ model due to the decrease in the number of cases. The logarithmic transformation eliminated the segments without accidents. Thus, 237 cases out of 295 were analyzed in regression models after the logarithmic transformation process. Further analysis of the data revealed that a simple averaging of accident numbers over a 3-year period created $AASN_A$ values smaller than 1 on many segments, which created negative values under logarithmic transformation. On the other hand, this was not observed in the $AASD_A$ values, as the SD_A values were in order of 10s, due to definition of accident/km. Elimination of the segments with relatively smaller values of number of accidents ($SN_A < 4$) was an attempt to improve the desired prediction model. This simply means searching a relationship for segments that have more frequently accidents. After eliminating the segments that had number of accidents smaller than 4, resulted in high R-values again for dependent variable of $ln(AASN_A)$ although the number of cases decreased (see Figure 5.45).

	All_Model 1		All_Model 1		All_Model 2	
Dependent Variable	$AASN_A$		$AASD_A$		$AASD_A$	
Model Statistics	**R = 0.778**		**R = 0.765**		**R = 0.774**	
Total SSE	*2021.667*		*50533.936*		*50533.936*	
Regression SSE	*1223.712*		*29552.191*		*30266.078*	
Number of cases	*295*		*295*		*295*	
Independent Variables	**Coeff.**	***t* value**	**Coeff.**	***t* value**	**Coeff.**	***t* value**
Constant						
Traffic Flow						
D_High speed	-1.128	*-3.864*	-6.300	*-4.208*	-5.425	*-3.618*
D_ HighPedVol						
Traffic Network						
No_Lanes	0.213	*4.495*	1.123	*4.629*	1.959	*5.520*
Seg_L_km					-31.614	*-3.185*
No_Intersections						
No_Legs						
Built Environment						
D_Urban zone	0.799	*3.447*	3.398	*2.861*	5.004	*3.929*
D_Shop						
D_Shop_2sides	0.604	*2.304*	3.388	*2.520*	4.451	*3.261*
D_BusComInd						
D_PedBarr						
D_Underpass						
D_Overpass						
D_Metro_Underpass	3.108	*4.651*	14.698	*4.290*	13.439	*3.957*
No_Bus_Stops	0.210	*4.932*	1.077	*4.945*	1.096	*5.110*

Figure 5.43 Best Regressions Models for the All Study Corridors with Annual Average Segment Number of Accidents

	All_Model 1		All_Model 1	
Dependent Variable	ln($AASN_A$)		ln($AASD_A$)	
Model Statistics	**R = 0.565**		**R = 0.916**	
Total SSE	*222.651*		*924.63*	
Regression SSE	*71.043*		*776.111*	
Number of cases	*237*		*237*	
Independent Variables	**Coeff.**	***t* value**	**Coeff.**	***t* value**
Constant				
Traffic Flow				
D_High speed	-0.452	*-5.683*	-0.319	*-1.889*
D_ HighPedVol			0.287	*2.025*
Traffic Network				
No_Lanes			0.178	*6.528*
Seg_L_km				
No_Intersections				
No_Legs				
Built Environment				
D_Urban zone	0.390	*4.769*	0.693	*5.426*
D_Shop				
D_Shop_2sides	0.246	*1.723*	0.414	*2.745*
D_BusComInd				
D_PedBarr				
D_Underpass				
D_Overpass				
D_Metro_Underpass				
No_Bus_Stops	0.079	*3.959*	0.069	*3.388*

Figure 5.44 Best Regressions Models for the All Study Corridors with Logarithmic Transformation of Annual Average Segment Number of Accidents

	All_Model 1		All_Model 1	
Dependent Variable	ln($AASN_A$>1)		ln($AASD_A$>1)	
Model Statistics	**R = 0.908**		**R = 0.919**	
Total SSE	*156.277*		*924.585*	
Regression SSE	*128.802*		*781.43*	
Number of cases	*116*		*234*	
Independent Variables	**Coeff.**	***t* value**	**Coeff.**	***t* value**
Constant				
Traffic Flow				
D_High speed	-0.338	*-2.699*	-0.380	*-2.258*
D_ HighPedVol			0.259	*1.844*
Traffic Network				
No_Lanes	0.128	*5.242*	0.197	*7.101*
Seg_L_km				
No_Intersections				
No_Legs				
Built Environment				
D_Urban zone	0.326	*2.612*	0.623	*4.845*
D_Shop				
D_Shop_2sides			0.429	*2.879*
D_BusComInd				
D_PedBarr				
D_Underpass				
D_Overpass				
D_Metro_Underpass	0.860	*3.560*		
No_Bus_Stops	0.030	*2.202*	0.068	*3.341*

Figure 5.45 Best Regressions Models for the All Study Corridors for Segments with Non-negative *ln($AASN_A$)* Values

Prediction of the safety of the bus stops was conducted with the same procedure of prediction of corridor pedestrian safety. The regression models were developed firstly for annual average number of accidents (AAN_A) values for 50 m, 100 m and 400 m buffer zones (see Figure 5.46). The results showed that the coefficients were 1/3 of 3-year total accidents as in the prediction models of corridor segments.

	All_Model 1		All_Model 1		All_Model 1	
Dependent Variable	B_{50m}_AAN_A		B_{100m}_AAN_A		B_{400m}_AAN_A	
Model Statistics	**R = 0.693**		**R = 0.762**		**R = 0.894**	
Total SSE	*376.556*		*2308.667*		*35520.111*	
Regression SSE	*180.969*		*1340.71*		*28404.675*	
Number of cases	*344*		*344*		*344*	
Independent Variables	**Coeff.**	***t* value**	**Coeff.**	***t* value**	**Coeff.**	***t* value**
Constant						
Segment Accidents						
Seg_Acc_Den						
Traffic Flow						
D_High speed	-0.573	*-4.304*	-1.442	*-4.868*	-5.235	*-5.410*
D_HighPedVol					1.521	*1.886*
Traffic Network						
No_Lanes	0.121	*5.446*	0.301	*6.087*	1.073	*6.926*
Seg_L_km						
No_Intersections						
No_Legs						
Built Environment						
D_Urban zone	0.240	*1.950*	0.698	*2.545*	2.111	*2.784*
D_Shop						
D_Shop_2sides	0.408	*4.122*	0.953	*4.325*	3.900	*5.930*
D_BusComInd	-0.340	*-3.184*	-0.657	*-2.765*	-1.600	*-2.445*
D_PedBarr						
D_Underpass						
D_Overpass					-2.669	*-4.284*
D_Metro_Underpass					5.024	*5.729*

Figure 5.46 Best Regression Models for Bus Stop Pedestrian Safety Using Annual Average Number of Accidents for Buffer Zones (All Corridors)

With logarithmic transformation of AAN_A values of bus stops safety measures, the R-values were considerably low – except for 400 m – when compared with 3-years total accident results (see Figure 5.47). Reason of this decrease was same with the prediction models of corridor pedestrian safety. To avoid minus values after logarithmic transformation of annual average accident values, the bus stops that had accidents lower than “4” were eliminated and models were developed again (see Figure 5.48). Although the number of cases decreased when the elimination process was conducted, the association created higher R-values, which was expected and still there were statistically significant variables that explained the relationship between dependent and independent variables.

	All_Model 1		All_Model 1		All_Model 1	
Dependent Variable	$\ln(B_{50m}\_AAN_A)$		$\ln(B_{100m}\_AAN_A)$		$\ln(B_{400m}\_AAN_A)$	
Model Statistics	**R = 0.480**		**R = 0.508**		**R = 0.904**	
Total SSE	*138.824*		*296.498*		*1175.404*	
Regression SSE	*32.031*		*76.48*		*959.69*	
Number of cases	*198*		*288*		*329*	
Independent Variables	**Coeff.**	***t* value**	**Coeff.**	***t* value**	**Coeff.**	***t* value**
Constant						
Segment Accidents						
Seg_Acc_Den						
Traffic Flow						
D_High speed	-0.562	*-4.360*	-0.817	*-5.729*	-1.168	*-7.291*
D_HighPedVol					0.337	*2.444*
Traffic Network						
No_Lanes			0.114	*4.867*	0.232	*10.377*
Seg_L_km						
No_Intersections						
No_Legs						
Built Environment						
D_Urban zone			0.471	*2.887*		
D_Shop						
D_Shop_2sides	0.332	*2.372*	0.516	*3.698*	0.727	*6.687*
D_BusComInd	-0.371	*-3.093*	-0.379	*-2.743*		
D_PedBarr			-0.216	*-1.847*		
D_Underpass						
D_Overpass						
D_Metro_Underpass						

Figure 5.47 Best Regression Models for Bus Stop Pedestrian Safety Using Logarithmic Transformation of Annual Average Number of Accidents for Buffer Zones (All Corridors)

	All_Model 1		All_Model 1		All_Model 1	
Dependent Variable	$\ln(B_{50m}\_AAN_A>1)$		$\ln(B_{100m}\_AAN_A>1)$		$\ln(B_{400m}\_AAN_A>1)$	
Model Statistics	**R = 0.921**		**R = 0.938**		**R = 0.950**	
Total SSE	*39.478*		*213.447*		*1161.582*	
Regression SSE	*33.515*		*187.66*		*1048.993*	
Number of cases	*65*		*133*		*290*	
Independent Variables	**Coeff.**	***t* value**	**Coeff.**	***t* value**	**Coeff.**	***t* value**
Constant						
Segment Accidents						
Seg_Acc_Den						
Traffic Flow						
D_High speed			-0.326	*-2.474*	-0.521	*-3.227*
D_HighPedVol					0.291	*2.005*
Traffic Network						
No_Lanes	0.145	*10.225*	0.256	*14.522*	0.225	*7.495*
Seg_L_km						
No_Intersections						
No_Legs						
Built Environment						
D_Urban zone					0.448	*3.425*
D_Shop						
D_Shop_2sides	0.507	*3.283*	0.727	*5.400*	0.687	*6.356*
D_BusComInd	-0.468	*-3.077*	-0.478	*-3.604*	-0.278	*-2.578*
D_PedBarr	-0.225	*-2.078*	-0.536	*-5.149*	-0.181	*-2.017*
D_Underpass						
D_Overpass					-0.178	*-1.833*
D_Metro_Underpass					0.488	*3.944*

Figure 5.48 Best Regression Models for Bus Stop Pedestrian Safety with Non-negative Logarithmic Transformation of Bus Stop Safety Measures

5.5 Case Study Summary

Studying pedestrian safety along selected Ankara corridors was very helpful in many perspectives. However, it is very difficult to generalize the numeric outcomes to other locations or networks, several insights about pedestrian safety in urban locations were gained at different steps of the proposed methodology, which are listed below:

Data Collection:

- Data collection is a massive operation requiring man power, unless it is available. If the accident data are not geocoded, this man power requirement may increase substantially, as the accident records must be geocoded manually using the details in the records. Similarly, the collection of bus stop data may be overwhelming, if it is not performed before.
- Data quality is very important. If there are reliability problems regarding the geocoding, a more flexible matching has to be done between the location (bus stop or accident) and the road network to compensate for the problem.
- More data should be collected about the characteristics of the vehicular and pedestrian flows in the study region, and via quantitative measurements as opposed to qualitative assessments of the aspects.

Corridor Pedestrian Safety Analysis

- As the road centerlines would not be enough to represent the pedestrian safety on the road segments or sidewalks, a buffer zone has to be defined along the centerlines to represent the study corridor.
- Furthermore, the traditional link definitions between road network nodes are not appropriate for urban studies, as they show a great variation from very short distances (50 m) to very long ones (3 km) based on the location and type of road. Thus, road network data has to be processed to create "almost equal" unit segments for the study corridor in urban region. This brings more accuracy to detection of accident-prone locations. This step may require aggregation or disaggregation of road network links. For Ankara network, unit segment length is chosen as 0.2 km, whenever possible; with a possibility to vary in the range of 0.1 km - 0.3 km, if needed.
- The results of the corridor pedestrian safety analysis revealed that SN_A, the number of accidents per segment (0.2 km on average), could go up to 35 accidents, with an average value of 4.65 accidents (the averages for CBD and UTZ were found as 8.39 accidents and 3.45 accidents, respectively).
- When the segment densities SD_A were examined, the maximum value observed was

found as 176 accidents/km, with an average value of 22.38 accidents/km. These values include multiple counting of the same accident, if it was located in the overlapping regions of the segment buffer zones.

- GIS-based maps revealed the location of 5 critical road segments in the CBD region that have 27-35 accidents, while there were only 2 segments in the UTZ that have number of accidents between 24-32. They can be described as “black” segments in terms of pedestrian safety in the Ankara study region.
- As the average number of injuries per accidents is very close to one, study of number of injuries or injury density revealed similar patterns to the number of accidents and accident densities. There were only 41 out of 295 segments with fatalities, and had a maximum number of fatalities is 3. These were only in the UTZ region, and no fatality accident in the CBD.
- The average number of bus stops in the road segments (on both sides of the segment) was found as 1.16 stops in overall region, which was 2.43 stops in average in the CBD and 0.758 in the UTZ. The maximum number was observed as 25 stops in a segment in the Atatürk Boulevard in the CBD.

Bus Stop Pedestrian Safety Analysis

- To associate the pedestrian accidents with a bus stop, a buffer zone had to be defined. The size of this buffer zone was selected as 50 m, 100 m and 400 m for the bus stops in the study region.
- Accident-prone bus stops were mostly observed on or close to accident-prone road segments. The highest number of accidents at a stop was found as 12 (within 50 m buffer zone) and observed in the CBD. This number was found as 21 accidents for a 100 m buffer zone and found in the CBD. For 400 m buffer zone, the maximum accident was found as 63 accidents to be associated with the bus stops in the CBD.
- There was 1 stop with 2 fatalities within a 50 m and 6 stops with a single fatality within the buffer zone. Within 100m definition, there were 30 stops associated with fatalities. There were at most 3 fatalities at 2 stops. For 400 m buffer zone, there were 59 bus stops that had fatalities up to 5.

Regression Analysis

- Corridor pedestrian safety regression models developed for all study segments reached an R-value on average 0.7. The models were developed for CBD and UTZ corridors gave R-values that are more realistic. The logarithmic transformation of dependent variables resulted in better R-values around 0.9, which proposed a significant non-linear relationship between the selected safety variable (number of accidents or accident densities of segments) and independent variables.
- Statistically significant variables were generally found as *D_High speed, D_High speed, No_Lanes, D_Urban zone, D_Shop_2sides, D_Metro_Underpass* and *No_Bus_Stops* although some variety could be observed in the developed models. The coefficients and the direction of impact (negative or positive) of these variables changed from overall model to region-specific (CBD only and UTZ only) models.
- In general, the regression of CBD must be carried separate from UTZ zones to capture the true dynamics of the relationship.
- Number of bus stops on a segment was always positively related with number of accidents or accident density, and mostly statistically significant. However, the impact of marginal increase in the bus stop number on a segment was relatively small compared with the other factors.
- Bus stop pedestrian safety regression models were developed for different sizes of buffer zones and also CBD and UTZ bus stops separately. The results of models developed with segment accident density as independent variable, most of the relationship was explained as in the correlation analyses between segment accident density and number of accidents in the buffer zones of the bus stops. The R-values were found around 0.9 when models were developed only segment accident density values.
- When the regression analyses were conducted for other independent variables that are characteristics of corridor segments the R-values were found as 0.693, 0.762 and 0.894 for 50 m, 100 m and 400 m respectively. The constant increase in the R values explained by the fact that more local factors would be captured in the smaller buffer zones compared to capture of impact of more regional and overall aspects in the larger buffer zones. However, unrealistically large buffer zones would associate many accidents with many bus stops at the same time, which may cause a decrease

in the success of linear regression.

- The statistically significant variables were found parallel with the corridor pedestrian safety regression models as; *Seg_Acc_Den, D_High speed, D_ HighPedVol, No_Lanes, D_Urban zone, D_Shop_2sides, D_BusComInd, D_Overpass* and *D_Metro_Underpass*.
- The same bus stop models with logarithmic transformation of dependent variables gave better R-values that show the relationship was non-linear.
- The results of the 3-year total regression analyses except the models with logarithmic transformation can be used in the prediction of pedestrian safety by dividing the constants with 3. The logarithmic transformation models should be carried out again after eliminating the segments or bus stops that have small number of accidents, which cause negative values after calculation of annual average accident value in logarithmic transformation.

CHAPTER 6

CONCLUSIONS AND FURTHER RECOMMENDATIONS

This study aimed at analyzing pedestrian safety around bus stops, which was brought up to the attention of Highway Traffic Safety Commission (HTSC), under which a subcommittee was formed and was the main motivation behind this study. As the accident data structure available in Turkey allows use of only the coordinates of the accidents and do not provide anything further on the direct or indirect involvement of bus or bus stops, the proposed approach focused on developing a systematic method to assess the potential relation between number and distribution of pedestrian accidents and bus stops, and the built-environment around them. Due to the limited resources on data collection and time allocated by the subcommittee, the study was designed to investigate pedestrian safety around bus stops in the CBD and four main arterials in the City of Ankara.

The study included two levels of analysis: GIS-based and regression analyses. In addition, these analyses were performed for road segments as well as bus stops. As a results, the former produced accident-prone road segment, which can be called "black spots", while the latter produced accident-prone bus stops. The regression models searched for statistically significant relations between the selected pedestrian safety measures and many parameters describing the traffic flow, road network and built-environment, similar to those mentioned in the literature. Finally, the developed methodology can be briefly summarized with the following steps:

- Step 1: Data Collection
- Step 2: GIS-based Corridor Pedestrian Safety Analysis
  - Outcome: Accident-prone corridor segments
- Step 3: GIS-based Bus Stop Pedestrian Safety Analysis
  - Outcome: Accident-prone bus stops segments

- Step 4: Bus Stop-Corridor Pedestrian Safety Correlation Analysis
- Step 5: Regression Analysis
    - Outcome: Corridor pedestrian safety regression models,
    - Outcome: Bus stop pedestrian safety regression models

The major findings and conclusions will be made based on the experience gained from the study of pedestrian safety along the selected corridors in Ankara. While the study was developed for the major corridors of Ankara, it can be easily generalized for the whole corridors, if the required data is available. Secondly, the methodology proposed here can be easily applied to another location to produce the targeted outcomes.

6.1 General Findings about Corridor Pedestrian Safety

- Pedestrian safety problem in a region is not an easy one to analyze. First, pedestrian accidents are not uniformly distributed spatially. Secondly, pedestrian movements and volumes are not constant or easily predictable throughout a region without detailed data and measurements. Finally, factors have non-linear combined effects; thus, they do not have the same at every location in the region in terms of pedestrian safety.
- To analyze pedestrian safety in a region, it is important to geocode pedestrian accidents to the road network; furthermore, to study pedestrian accidents at bus stops, it is important to have information on bus stop locations.
- GIS based representation of the selected pedestrian safety measure is crucial to see accident-prone locations visually.
- In thematic maps, using road network links developed for purposes other than accident analysis may be misleading; the longer the link length is, more likely it would have higher number of accidents. Alternatively, it is possible to work with "density" based safety measures, such as, accident density, injury density, fatality density etc.
- Introduction of "density" measures has one disadvantage, which is the underestimation of safety measures in very long links. To overcome this shortness, it is important to have more unified length definition for the spatial analyses, which can be achieved with "unit segment" creations.

- The length choice for "unit segmentation" in an urban region is very challenging; there can be very small links (close to 50 m) mostly in CBD regions, and very long ones (up to 3 km) in UTZ corridors. The links longer than the unit segment length have to be "disaggregate" into multiple segments, while links shorter than the unit segment have to be "aggregated". Even so, it may not be possible to enforce a constant unit segment length, but an average unit segment length with relatively a small variability should be aimed.
- An analysis segment length of 200 m is chosen for Ankara study, which resulted in creation of segments within a length range of [105m, 300m]. Nevertheless, it is important to look at the distribution of link lengths of a region, before making a unit segment length and range choices.
- For Ankara study, the impact of the studied variables in the corridor pedestrian safety for different dependent variable choice is summarized in Figure 5.35.
- In the absence of measurements, variables for high pedestrian volume and high speed were introduced as "dummy" variables, and variables such as "sidewalk capacity" could not be included, at all. As built environment characteristics were also represented by dummy variables, the regression models developed included more dummy variables than measured ones.
- In the pedestrian safety assessment of a segment, built environment and traffic network parameters were found significant; however, some of them were more influential in CBD region, while other were in UTZ. It is important to study pedestrian safety in these two regions separately, as much as possible.
- Although the effect of bus stops existence on a segment safety is smaller in magnitude compared to other independent variables, it was found statistically significant in CBD models, (thus also in all corridor models).
- In linear regression models, using number of segment accidents (SN_A) versus segment accident densities (SD_A) did not show a big difference, as the latter is very close to the former due to the usage of "almost uniform" segment length.
- Regression models revealed a non-linear relationship, where a logarithmic transformation of the dependent variable (SN_A or SD_A) improved the R-value of the regressions, significantly. However, it should be noted that since *ln(0)* could not be defined, these non-linear models were fit to a subset of segments with non-zero accident (density) values. The models with logarithm of segment accident densities,

ln(SD_A), produced better fits than that of accident numbers, *ln(SN_A)*. This also suggested the use of density based safety measures, as opposed to traditional ones.

- If the independent variables do not change much over the years, prediction can be done based on the regression models with the total study period, creating "annual average" values for the safety parameters. In prediction, the linear regression models using the annual average values would not be affected much. However, non-linear models with logarithmic transformation of the new annual average values may lose their power to fit the data. The reason behind this loss was found to be the negative values created by the logarithmic transformation of $AASN_A$ (or $AASD_A$) values less or equal to "1". This should be noted down as a weakness of logarithmic transformation of segment safety measures with few or no accidents on them.

6.2 General Findings about Pedestrian Safety around Bus Stops

- Pedestrian safety around bus stops does not have a constant pattern at every bus stop in a region. This means, pedestrians do not have a constant risk of accident at every bus stop.
- Choice of bus stop buffer zone size affects the determination of "accident-prone" bus stops and factors that affect the safety around the bus stops a lot. Thus, it is important to try different sizes of buffer zones in a study.
- The smaller the bus stop buffer zone size is, the pedestrian safety in more immediate vicinity is reflected by the proposed bus stop safety measure.
- As the approach in this study is a corridor based one, bus stop safety is found highly correlated with the corridor segment safety parameters.
- Furthermore, size of the bus stop buffer zone relative to unit segment length in the study is crucial. For bus stop buffer zone sizes close to unit segment length a more linear correlation is observed compared to sizes smaller or much larger than the unit segment length.
- Bus stop scattering and layout structure and pedestrian levels around them may differ significantly between the CBD and UTZ regions, as well. Pedestrian safety around bus stops in CBD must be evaluated separately from the UTZ.
- The impact of the studied variables in the corridor pedestrian safety for different

dependent variable choice is summarized in Figure 5.42.

- Bus stop safety regression models had non-linearity and weaknesses similar to those of the corridor safety models.

6.3 Proposed Pedestrian Safety Policies

- For a city or an urban zone, to assess the pedestrian safety around bus stops, both corridor and bus stop safety analyses should be performed. The locations and bus stops with higher density of pedestrian accidents can be selected for immediate improvements.
- Regression models in general revealed an obvious fact that there are more pedestrian accidents, where there are more pedestrian activities, such as locations with shopping on both sides and/or in CBD regions, metro stations, etc. Pedestrian movements in such areas have to be studied more carefully. In such locations, barriers that discourage pedestrian crossing may reduce the number of accidents and should be used properly.
- Locations and bus stops on high-speed traffic corridors have more fatalities, thus, bus stops at these locations has to be avoided, if possible, or designed properly.
- Before adding a new bus stop or changing the location of an existing one, it is very important to study the accident history of the location and possible/existing pedestrian flows at the location.
- Bus stops closely located each other create a “bus stop zone” (which is generally observed in the CBD) rather behaving separately and result in higher pedestrian volumes around them increasing the risk; thus bus stop zones must be designed very carefully.

6.4 Further Recommendations

- The pedestrian accident data should include as much detail as possible, such as time of the accident, weather, pedestrian characteristics, etc. accessibility to accident records for the potential mentioning of intent or use of bus or bus stop can further clarify the direct and indirect involvement of bus public transit system in the pedestrian safety.

- Developed regression models should be verified on some corridors in the regions outside the study area.
- Prediction models can be tested, if accident data from other years for the same locations can be obtained
- Dummy variables are easy to create and can capture major relationships. The correlations between the dummy variables may be stronger (in a negative or positive way) than the correlations between dummy variables and the dependent variable. This may result in some of them being insignificant when used together in a model, or even change of direction in their impact. Idealistically, there should be as few dummy variables as possible.
- Other variables as sidewalk, bus stop waiting area, intersection traffic control (signalized or not), pedestrian volumes, speed, personal characteristics, accident characteristics, weather, bus stop service statistics (ridership) should have been included in the regression analyses.
- In bus stop analysis, the number of lines and services at the bus stop must be included in the regressions, which may reflect the number of travelers indirectly, if available ridership data at the bus stops can be included.
- To develop regression models for prediction purposes, the regression models could be fit to yearly data. This however requires multiplication of the dependent and independent variables for every year, which further requires yearly information for all the independent variables. One disadvantage of this approach is the potential risk of not having enough accidents on many of the segments in a yearly base approach. In addition, in such studies, it is preferred to have a longer duration data (up to 10 years), where each year would be assumed statistically independent of others.
- If a regional analysis would be preferred as opposed to corridor based, the regression models have to be developed separately as pedestrian activities on the roads crossing the corridors may also have influence on the corridor and bus stop safety.

REFERENCES

af Wahlberg, A.E., (2004). "Characteristics of low speed accidents with buses in public transport: part II", *Accident Analysis and Prevention*, 36, 63-71.

Amoros, E., Martin, J.L., Laumon, B., (2003). "Comparison of road crashes incidence and severity between some French counties", *Accident Analysis and Prevention*, 35, 537-547.

Ankara Büyükşehir Belediyesi EGO Genel Müdürlüğü, (2000). Brifingler Dosyası, Brifing CD'si, Ankara.

Ardic E., Z. (2008). An approach to investigate relationship between speed and safety on urban arterials. (MS thesis). Ankara: METU.

Bernhoft, I.M., Carstensen, G., (2008). "Preferences and behaviour of pedestrians and cyclists by age and gender", *Transportation Research Part F*, 11, 83-95.

Brenac, T., Clabaux, N., (2005). "The indirect involvement of buses in traffic accident process", *Safety Science*, 43, 835-843.

Cameron, I.C., Harris, N.J., Kehoe, N.J.S., (2001). "Tram-relatedinjuries in Sheffield", *Injury International Journal of the Care of the Injured*, 32 (4), 275-277.

Clifton, K.J., Burnier, C.V., Akar, G., (2009). "Severity of injury resulting from pedestrian-vehicle crashes: What can we learn from examining the built environment?", *Transportation Research Part D*, 14, 425-436.

Clifton, K.J., Kreamer-Fults, K., (2007). "An examination of the environmental attributes associated with pedestrian−vehicular crashes near public schools", *Accident Analysis and Prevention*, 39 (4), 708-715.

Çubuk M.K., Türkmen M., (2003). "Ankara'da Raylı Ulaşım", *Gazi Üniviversitesi Mühendislik Mimarlık Fakültesi Dergisi*, Cilt 18, No 1.

EGO Genel Müdürlüğü –EGO, (2012a), [online]. Last accessed from: http://www.ego.gov.tr/parts/egootobushizmetleri.asp, Last accessed on: 12.08.2012.

EGO Genel Müdürlüğü –EGO, (2012b), [online]. Last accessed from: www.ego.gov.tr/filelib/kurumsalistatistikler/hareket_noktasi_hat_durak_sayilari.pdf, Last accessed on: 12.08.2012.

Erdogan, S., Yilmaz, I., Baybura, T., Gullu, M., (2008). "Geographical information systems aided traffic accident analysis system case study: city of Afyonkarahisar", *Accident Analysis and Prevention*, 40, 174-181.

Federal Highway Administration- FHWA. (2012), [online]. Last accessed from: http://safety.fhwa.dot.gov/ Last accessed on: 03.08.2012.

General Directorate of Security Traffic Services Department-GDS, (2012), [online]. Last accessed from: www.trafik.gov.tr Last accessed on: 15.01.2012.

Golob, T.F., Recker, W.W., (2004). "A method for relating type of crash to traffic flow characteristics on urban freeways", *Transportation Research Part A*, 38, 53-80.

Golob, T.F., Recker, W.W., Alvarez, V.M., (2004). "Freeway safety analysis as a function of traffic flow", *Accident Analysis and Prevention*, 36, 933-946.

Hijar, M., Trostle, J., Bronfman, M., (2003). "Pedestrian injuries in Mexico: a multi-method approach", *Social Science & Medicine*, 57, 2149-2159.

Iowa DOT. (1989). *Accident and Related Data for Rural and Municipal Intersections in Iowa*. Ames, Iowa: Iowa Department of Transportation. Iowa.

Iowa State University Center for Transportation Research and Education- CTRE. (2012), [online]. Last accessed from: www.ctre.iastate.edu/pubs/traffichandbook/5CrashAnalysis.pdf Last accessed on: 03.08.2012.

Keskin, F., Yenilmez, F., Çolak, M., Yavuzer, İ., Düzgün, H.Ş., (2011)."Analysis of traffic incidents in METU campus", *Procedia - Social and Behavioral Sciences*, 19, 61-70.

Lee, C., Abdel-Aty, M., (2005). "Comprehensiveanalysis of vehicle–pedestriancrashes at intersections in Florida", *Accident Analysis and Prevention*, 37 (4), 775-786.

Levine, N., Kim, K., Nitz, L., (1995). "Spatial Analysis of Honolulu motor vehicle crashes: I. Spatial Patterns", *Accident Analysis and Prevention*, 27 (5), 663-674.

Loukaitou-Sideris, A., Liggett, R., Sung, H., (2007). "Death on the crosswalk: A study of pedestrian-automobile collisons in Los Angeles", *Journal of Planning Education and Research*, 26, 338-351.

Map of Ankara City Center, (2012), [online]. Last accessed from: <www.turkcebilgi.com/harita/ankara-sehir-merkezi/ Last accessed on: 01.09.2012.

Martinez, K.L.H., Porter, B.E., (2004). "The likelihood of becoming a pedestrian fatality and drivers' knowledge of pedestrian rights and responsibilities in the Commonwealth of Virginia", *Transportation Research Part F*, 7, 43-58.

Miranda-Moreno, L.F., Morency, P., El-Geneidy, A.M., (2011). "The link between built environment, pedestrian activity and pedestrian-vehicle collision occurrence at signalized intersections", *Accident Analysis and Prevention*, 43 (5), 1624-1634.

National Highway Traffic Safety Administration, (2001). *Traffic Safety Facts 2000: A Compilation of Motor Vehicle Crash Data from the Fatality Analysis Reporting System and the General Estimate System*, US Government Printig Office, Washington, DC.

O'Sullivan, D., Unwin, D.J., (2010). *Geographic Information Analysis*. New Jersey: John Wiley & Sons Inc.

Ossenbruggen, P.J., Pendharkar, J., Ivan J., (2001). "Roadway safety in rural and small urbanized areas", *Accident Analysis and Prevention*, 33, 485-498.

Planning Commission Transit-Oriented Committee, (2012), [online]. Last accessed from: http://www.fairfaxcounty.gov/planning/tod_docs/walking_distance_abstracts.pdf Last accessed on: 03.09.2012.

Retting, R.A., Ferguson, S.A., McCartt, A.T., (2003). "A review of evidence-based traffic engineering measures designed to reduce pedestrian–motor vehicle crashes", *American Journal of Public Health*, 93 (9), 1456-1463.

Sjolinder, K., Ek, H., (2001). *Black Spot Manual*, Ankara: General Directorate of Highways

Thill, J.C., (2000). "Geographic information systems for transportation in perspective." *Transportation Research Part C: Emerging Technologies*, 8, 3-12.

Tiwari, G., Bangdiwala, S., Saraswat, A., Gaurav, S., (2007). "Survival analysis: Pedestrian risk exposure at signalized intersections", *Transportation Research Part F*, 10, 77-89.

Türkiye İstatistik Kurumu- TUİK, (2012a), [online]. Last accessed from: www.tuik.gov.tr/PreHaberBultenleri.do?id=10736, Last accessed on: 12.08.2012.

Türkiye İstatistik Kurumu- TUİK, (2012b), [online]. Last accessed from: http://rapor.tuik.gov.tr/reports/rwservlet?ulastirmadb2=&report=tablo23.RDF&p_yil1=2012&p_ar1=1&p_ay1=5&p_tur=2&p_duz1=TR51&desformat=html&ENVID=ulastirmadb2Env, Last accessed on: 12.08.2012.

Unger, R., Eder, C., Mayr, J.M., Wernig, J., (2002). "Child pedestrian injuries at tram and bus stops", *Injury International Journal of the Care of the Injured*, 33, 485-488.

Valverde, J.A., Jovanis, P.P., (2006). "Spatial analysis of fatal and injury crashes in Pennsylvania", *Accident Analysis and Prevention*, 38 (3), 618-625.

APPENDICES

A. THEMATIC MAPS FOR CORRIDOR PEDESTRIAN SAFETY ANALYSIS FOR CBD AND UTZ

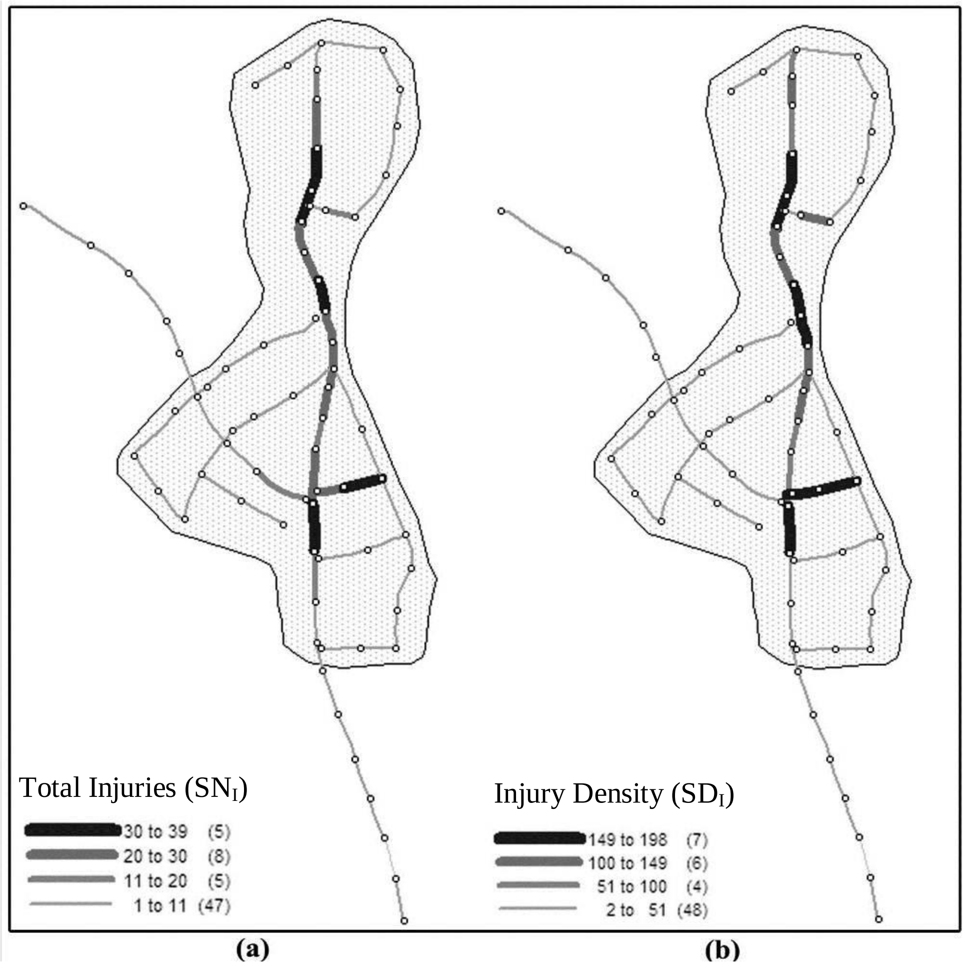


Figure A.1 CBD Thematic Maps of (a) the Segment Number of Injuries, SN_I, (b) the Segment Injury Density, SD_I (for a 3 Year Period)

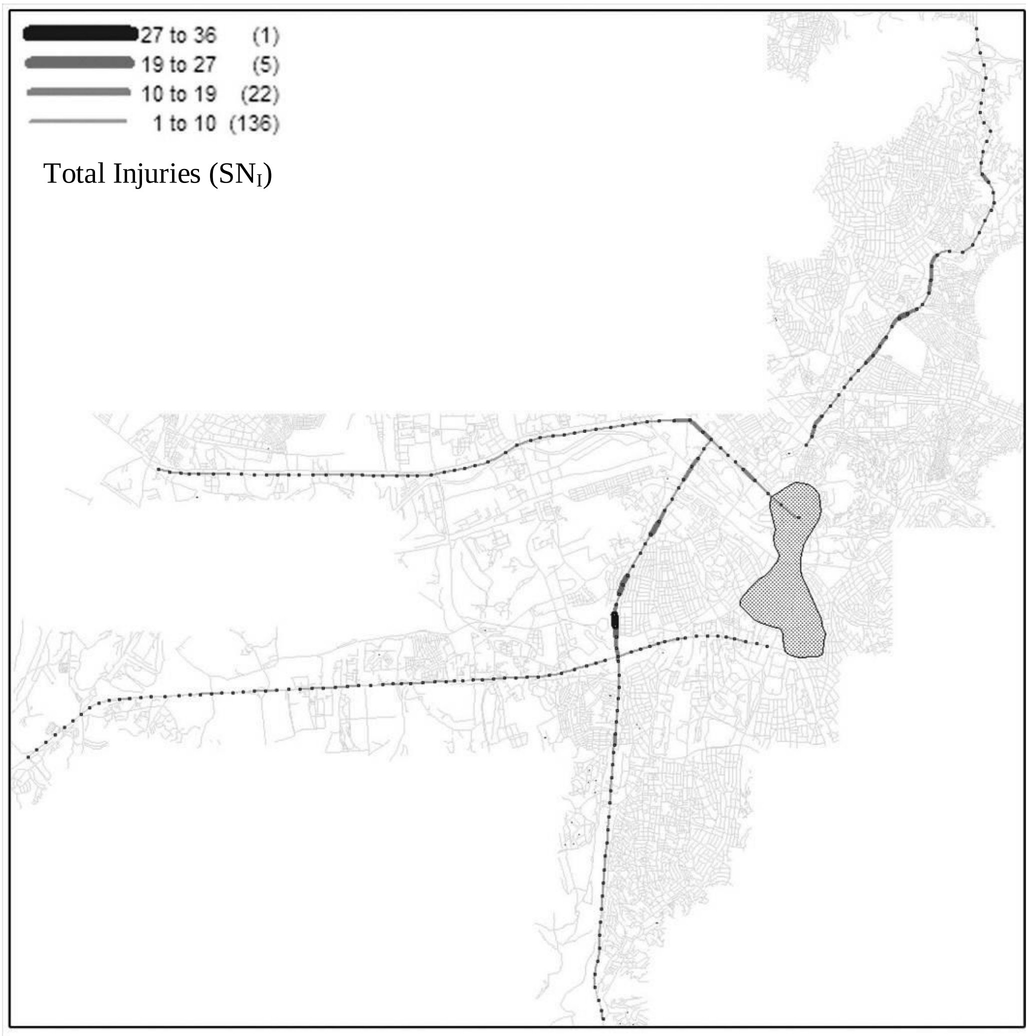


Figure A.2 UTZ Thematic Map of the Segment Number of Injuries, SN_I (for a 3 Year Period)

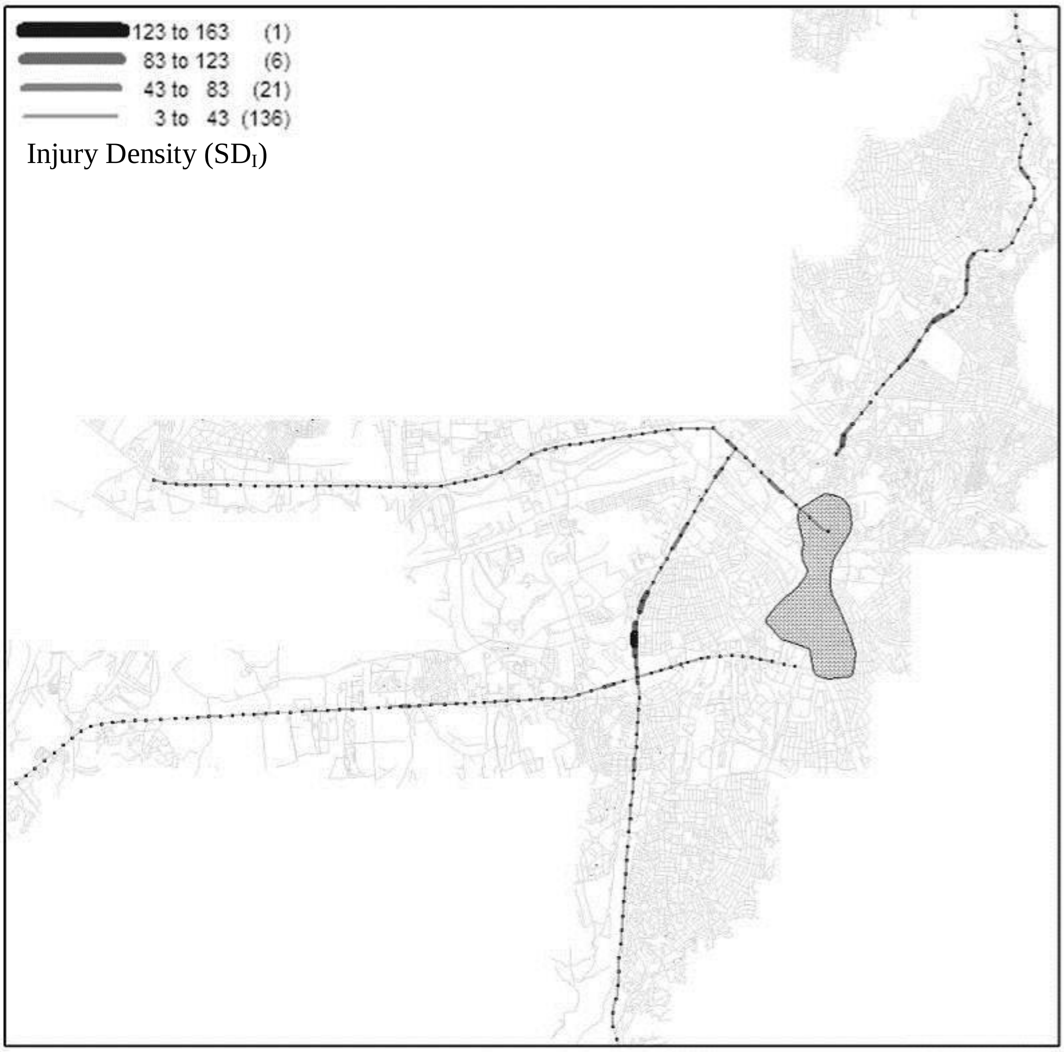


Figure A.3 UTZ Thematic Map of the Segment Injury Density, SD_I (for a 3 Year Period)

Figure A.4 UTZ Thematic Map of the Segment Number of Fatalities, SN_F (for a 3 Year Period)

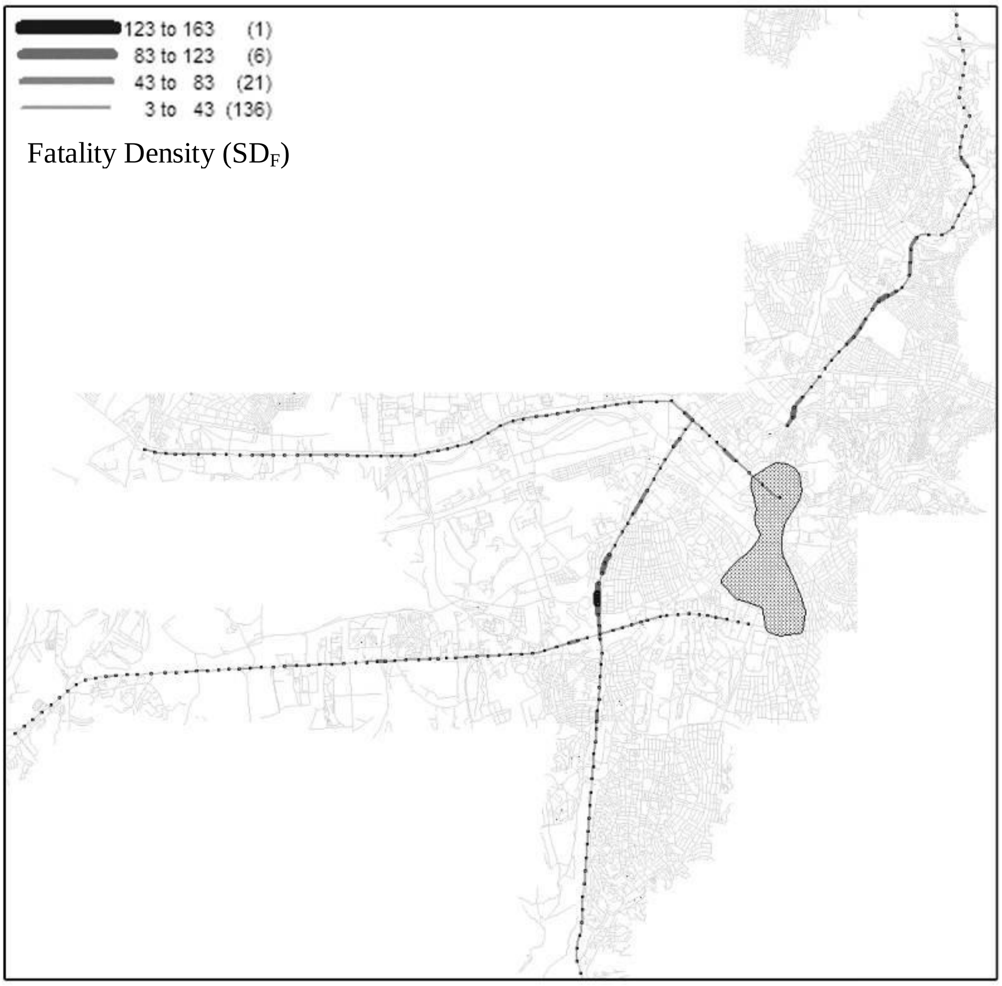


Figure A.5 UTZ Thematic Map of the Segment Fatality Density, SD_F (for a 3 Year Period)

B. THEMATIC MAPS FOR BUS STOP PEDESTRIAN SAFETY ANALYSIS FOR CBD AND UTZ

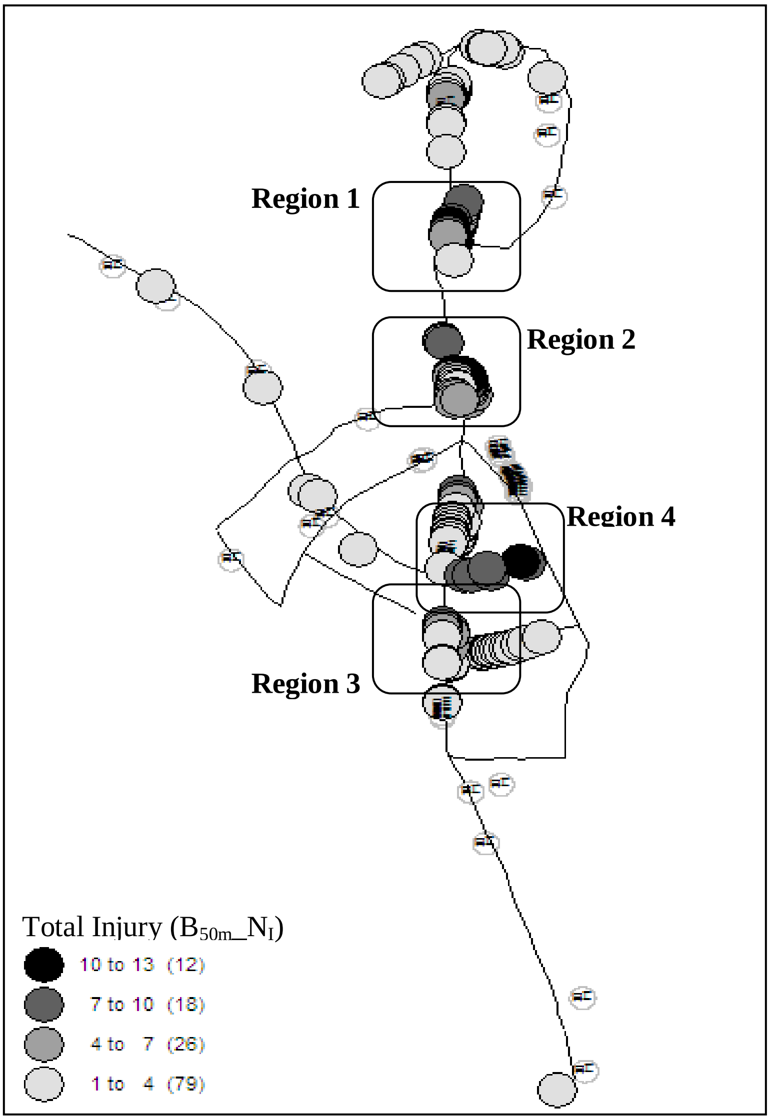


Figure B.1 CBD Thematic Map for Bus Stop Safety for the 50m Buffer Zone with the Number of Injuries, B_{50m}_N_I, (for a 3 Year Period)

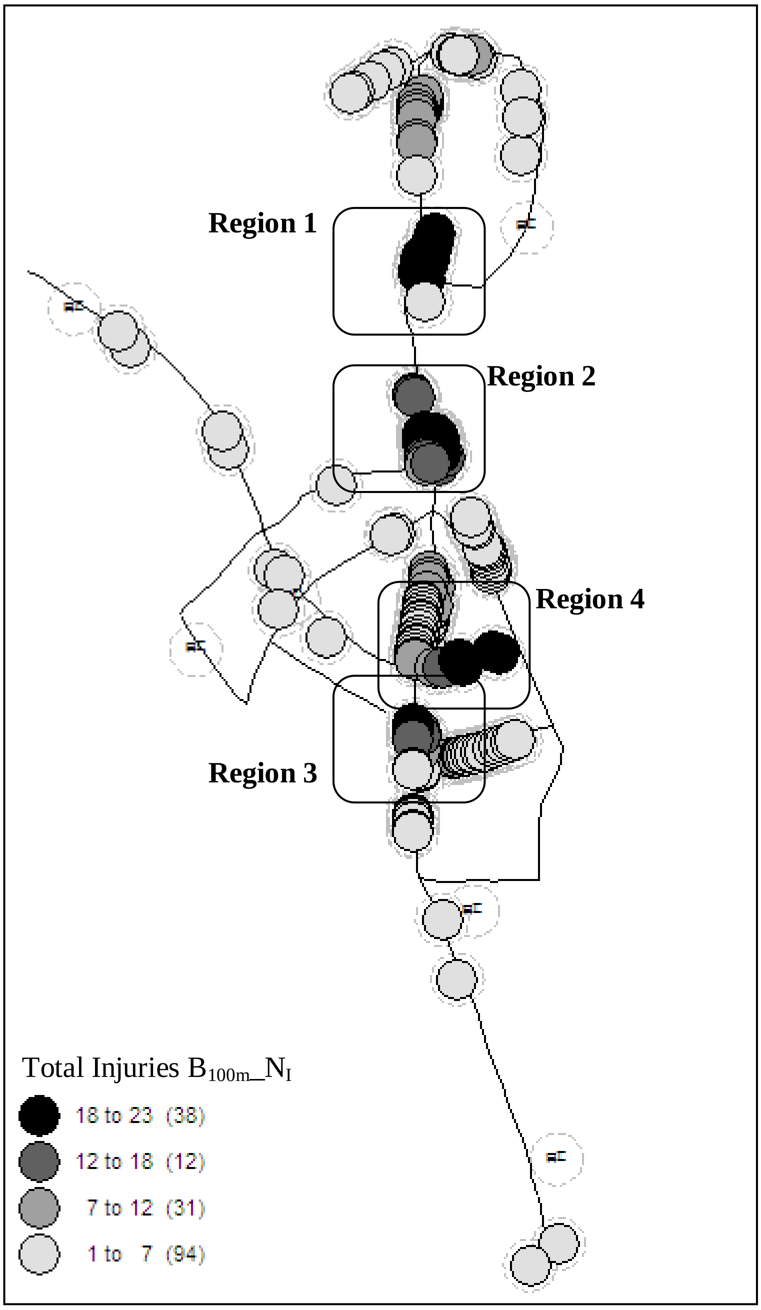


Figure B.2 CBD Thematic Map for Bus Stop Safety for the 100m Buffer Zone with the Number of Injuries, $B_{100m}\_N_I$, (for a 3 Year Period)

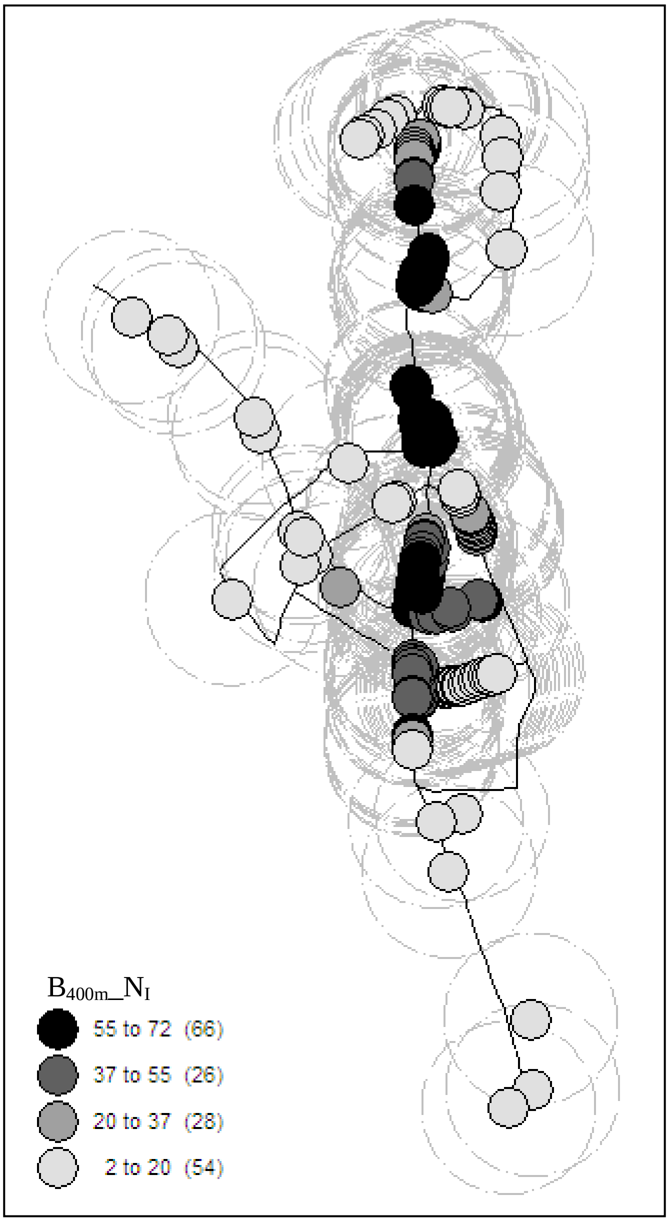


Figure B.3 CBD Thematic Map for Bus Stop Safety for the 400m Buffer Zone with the Number of Injuries, $B_{400m}\_N_I$, (for a 3 Year Period)

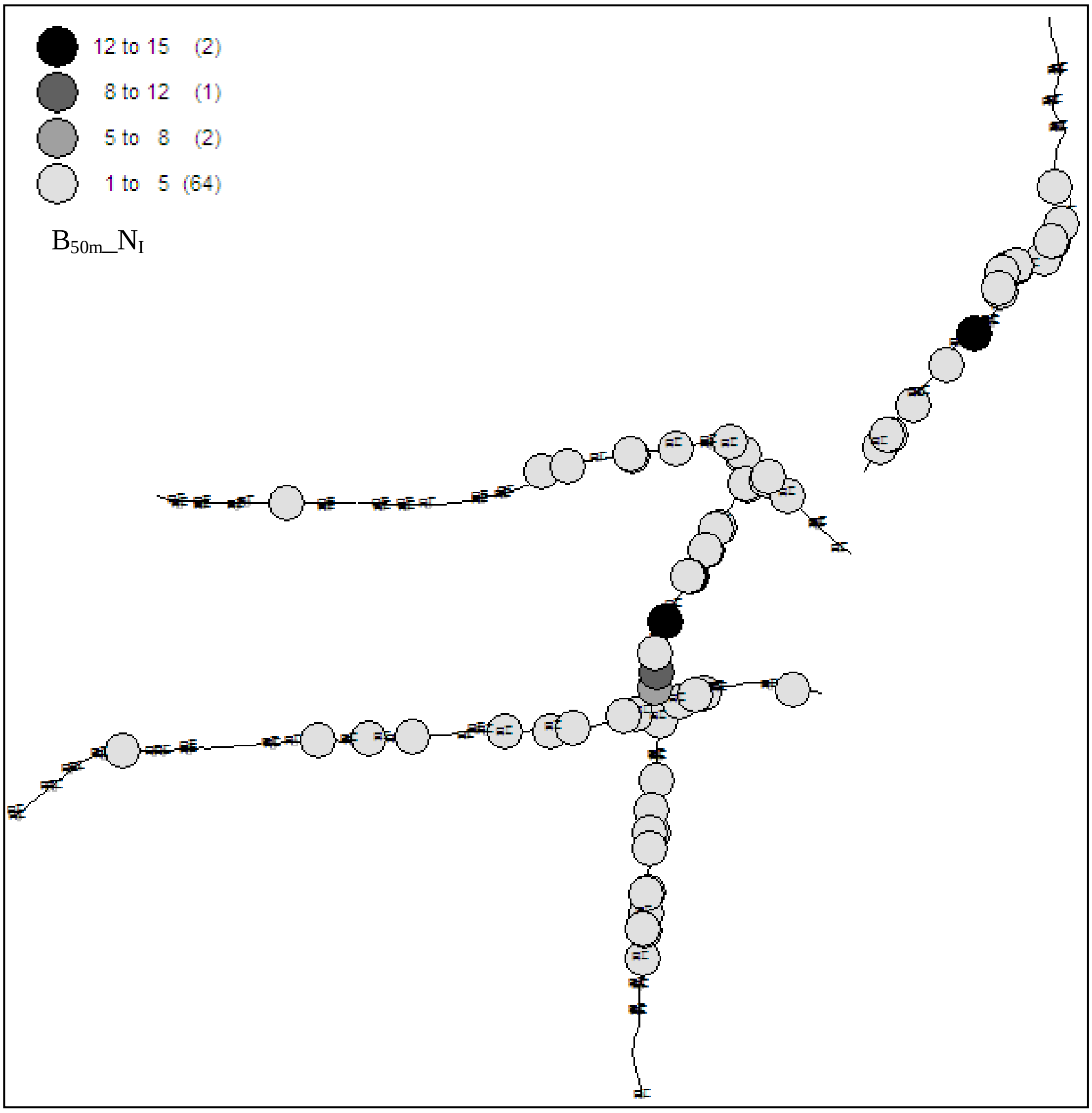


Figure B.4 UTZ Thematic Map for Bus Stop Safety for the 50m Buffer Zone with the Number of Injuries, $B_{50m}\_N_I$, (for a 3 Year Period)

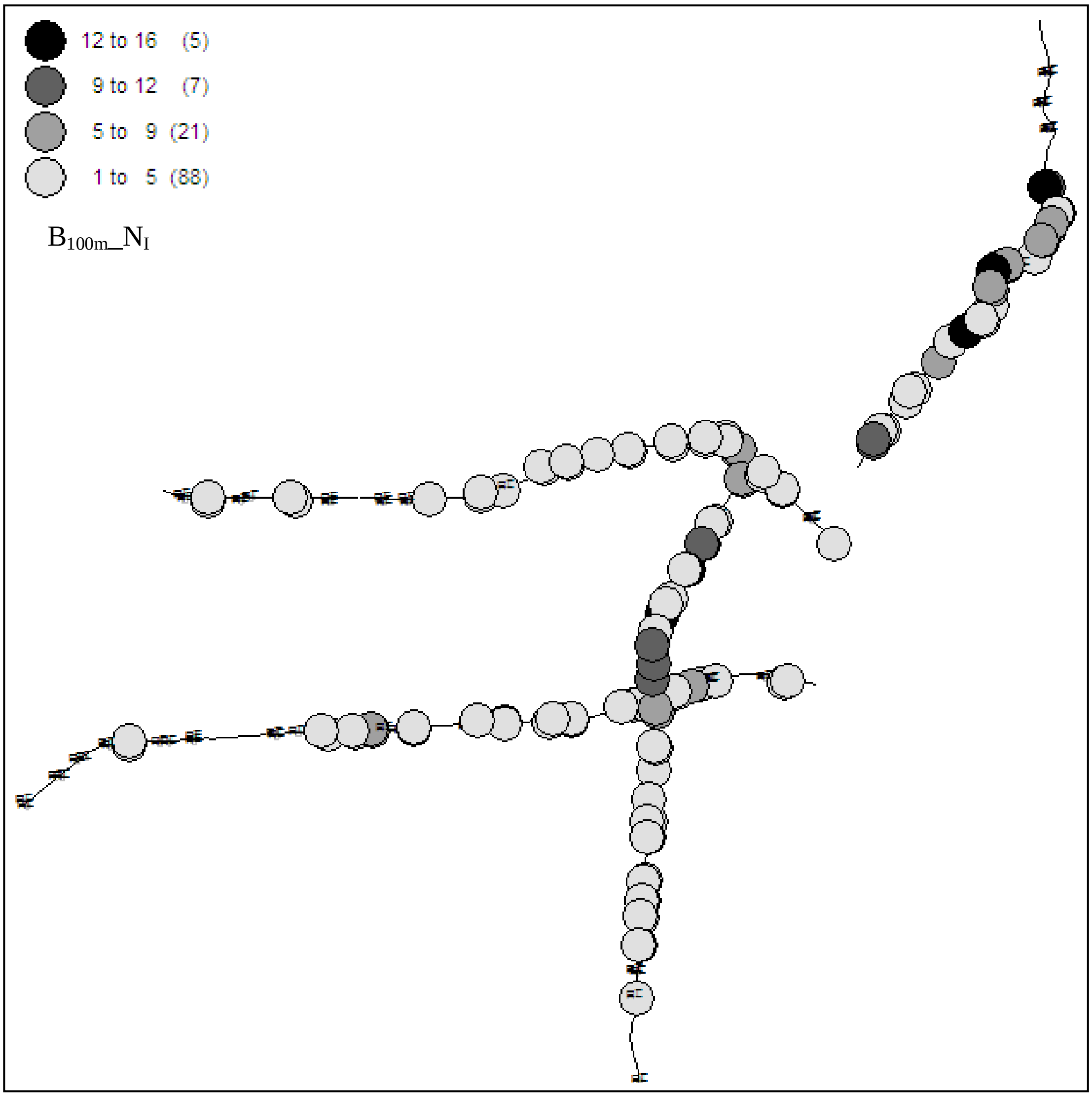


Figure B.5 UTZ Thematic Map for Bus Stop Safety for the 100m Buffer Zone with the Number of Injuries, $B_{100m}\_N_I$, (for a 3 Year Period)

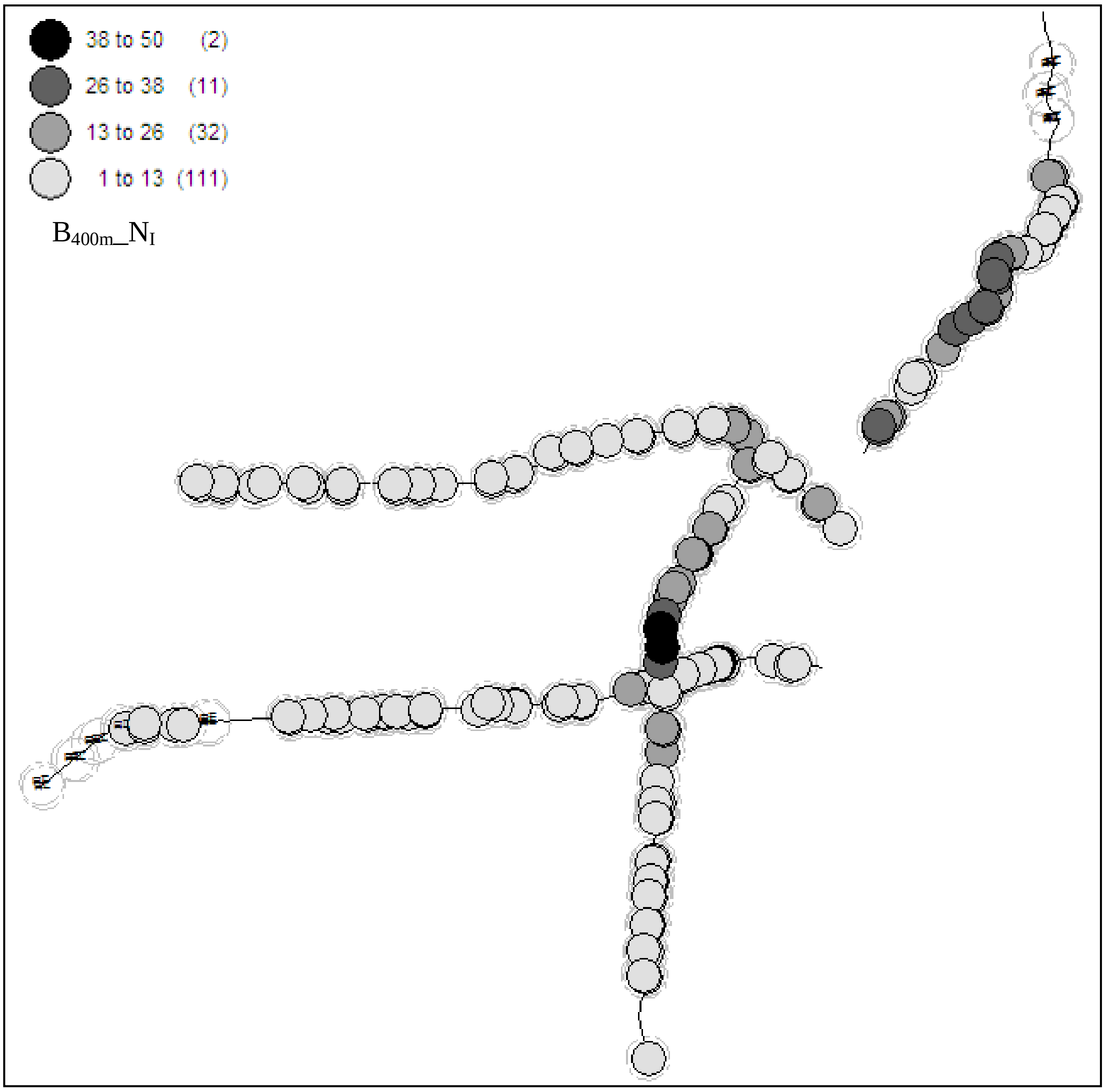


Figure B.6 UTZ Thematic Map for Bus Stop Safety for the 400m Buffer Zone with the Number of Injuries, $B_{400m}\_N_I$, (for a 3 Year Period)

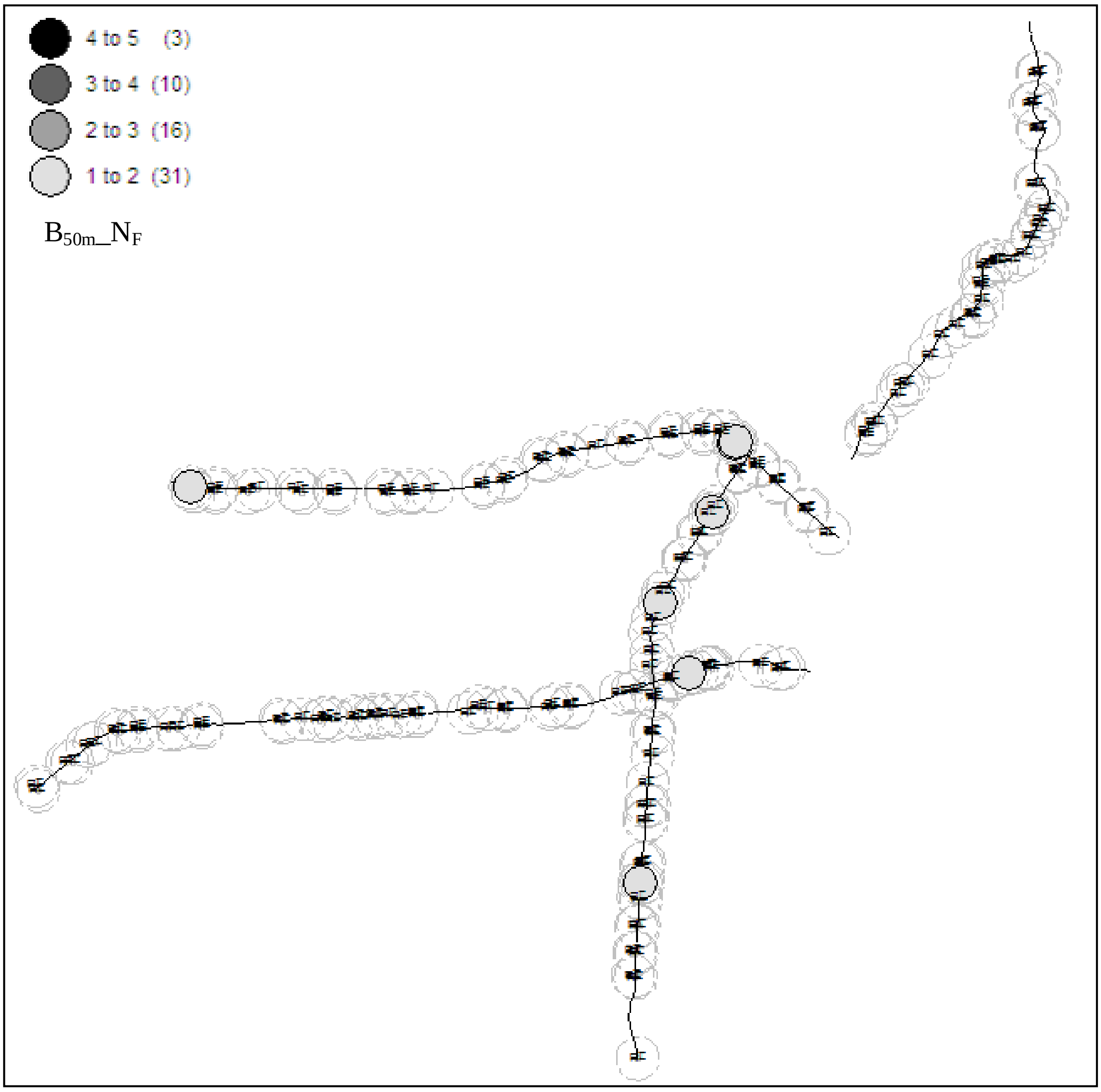


Figure B.7 UTZ Thematic Map for Bus Stop Safety for the 50m Buffer Zone with the Number of Fatalities, $B_{50m}\_N_F$, (for a 3 Year Period)

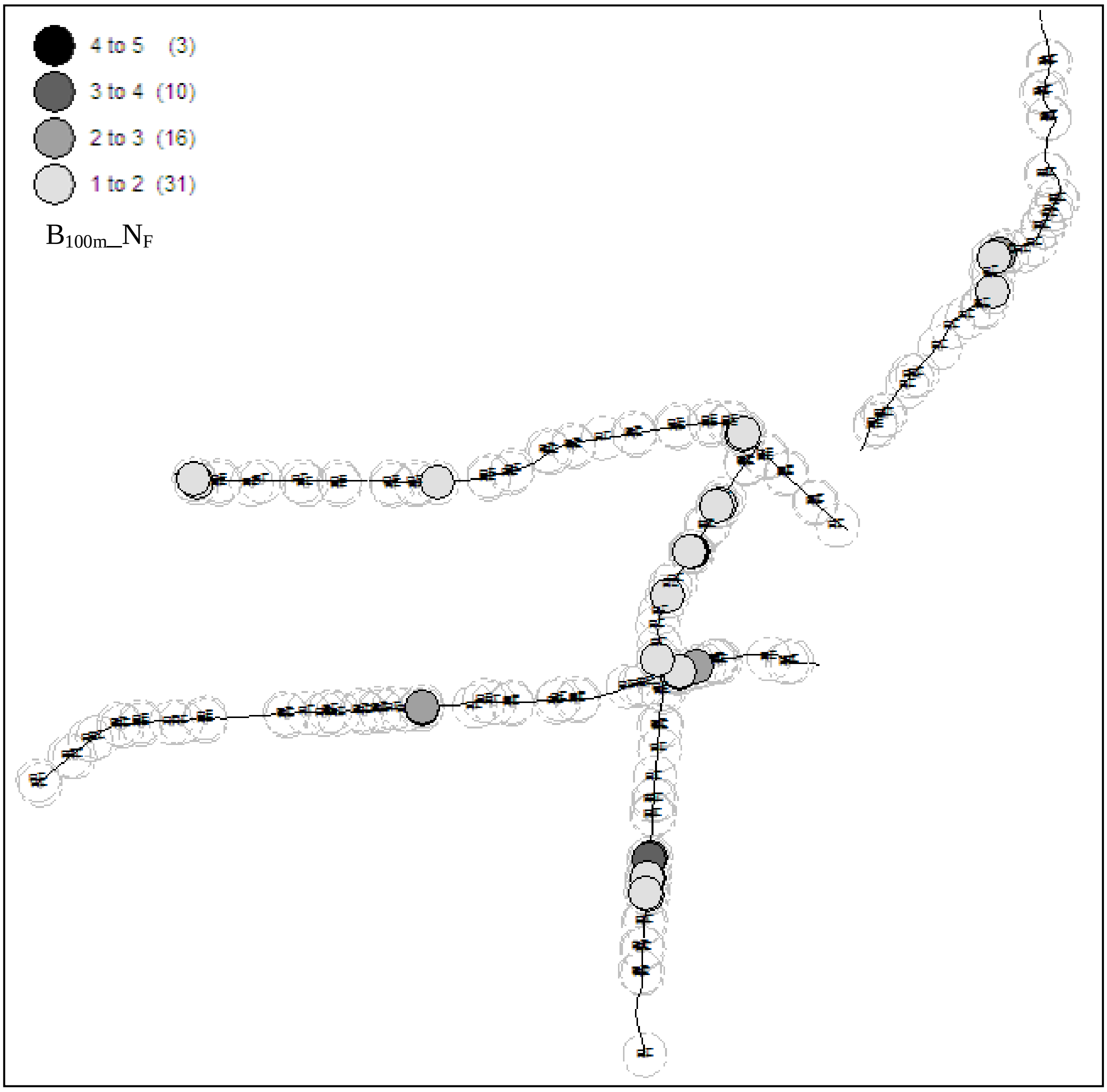


Figure B.8 UTZ Thematic Map for Bus Stop Safety for the 100m Buffer Zone with the Number of Fatalities, $B_{100m}\_N_F$, (for a 3 Year Period)

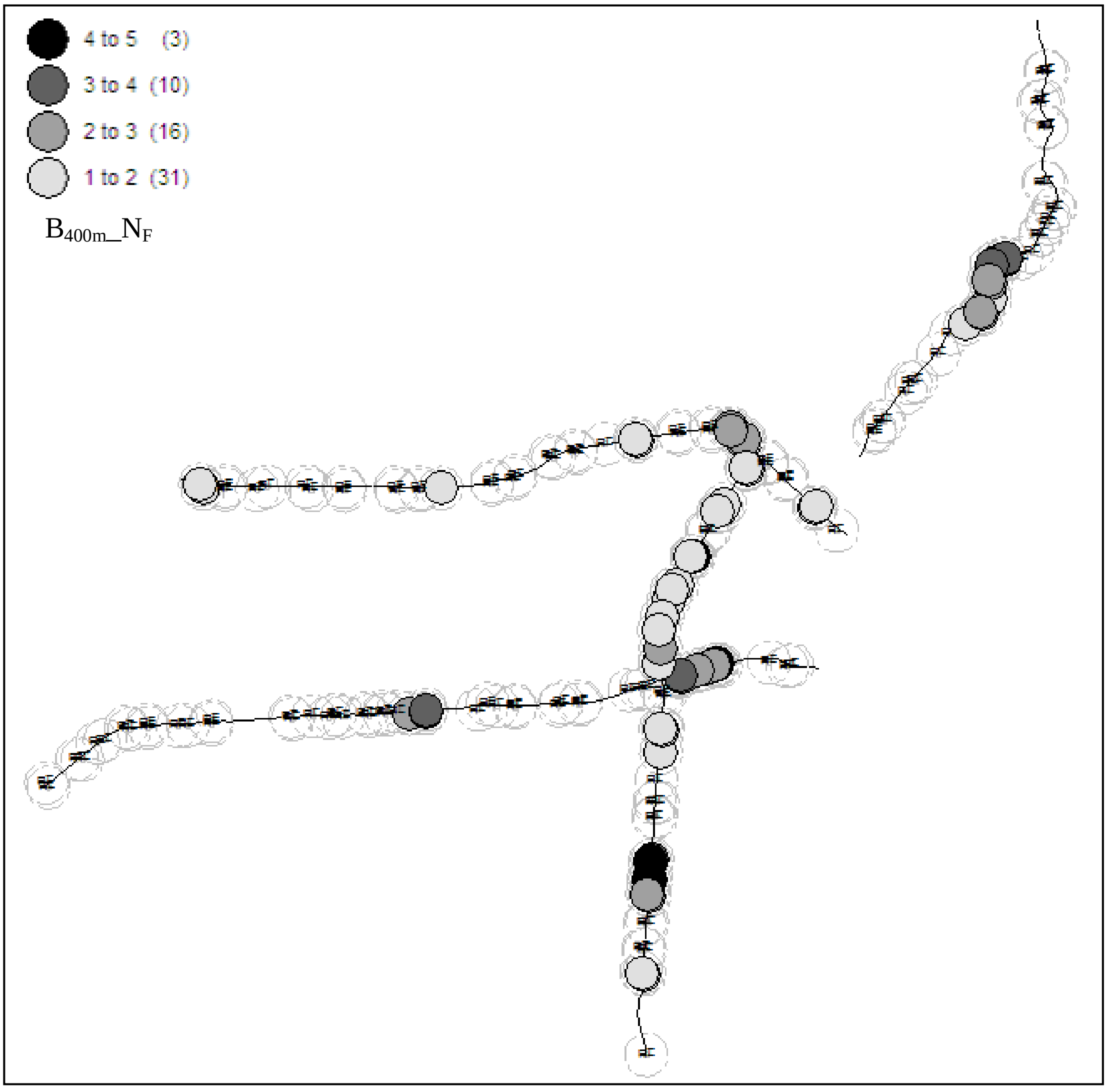


Figure B.9 UTZ Thematic Map for Bus Stop Safety for the 400m Buffer Zone with the Number of Fatalities, $B_{400m}\_N_F$, (for a 3 Year Period)

C. SPATIAL DISPLAY OF INDEPENDENT VARIABLES IN REGRESSION ANALYSIS

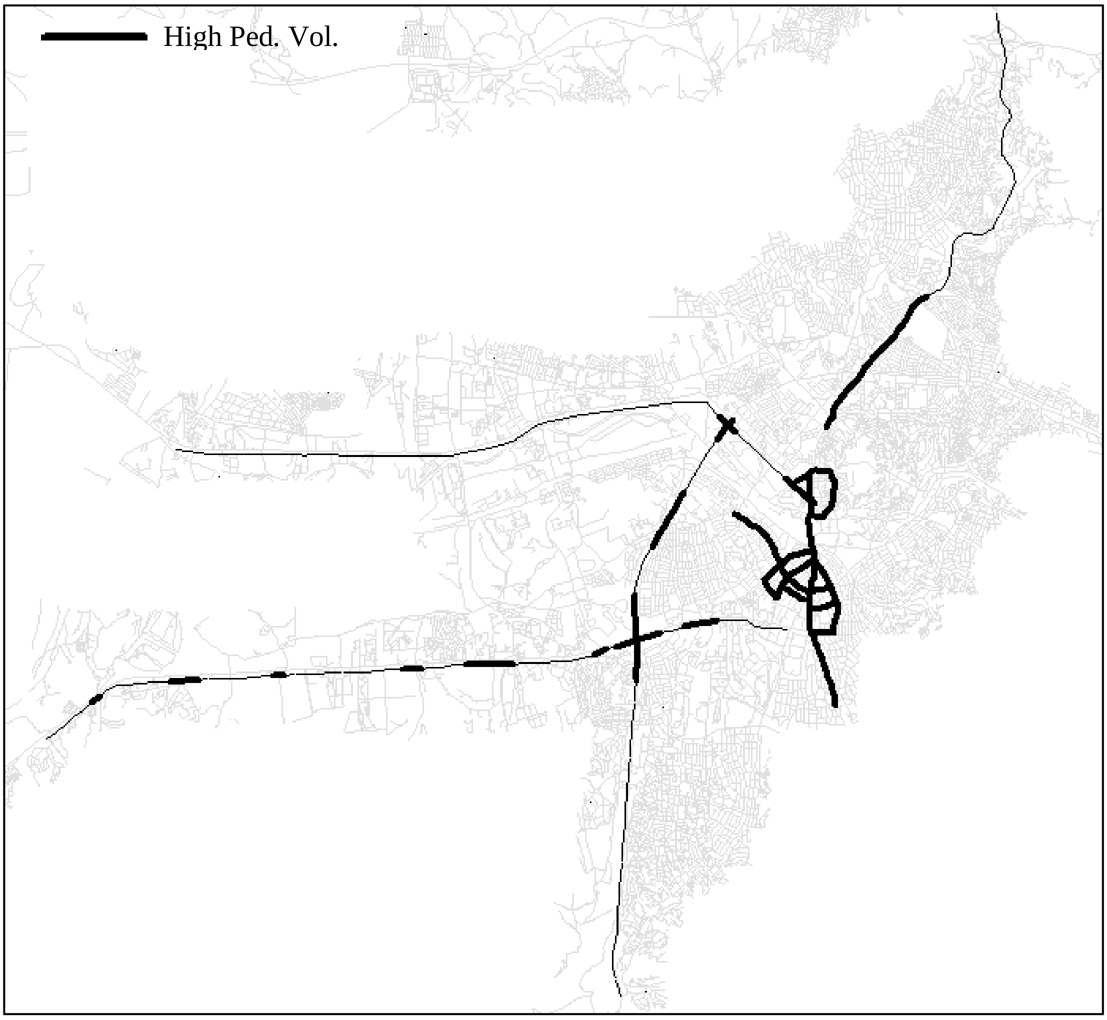


Figure C.1 Locations of Segments with High Pedestrian Volume (*D_HighPedVol* = 1)

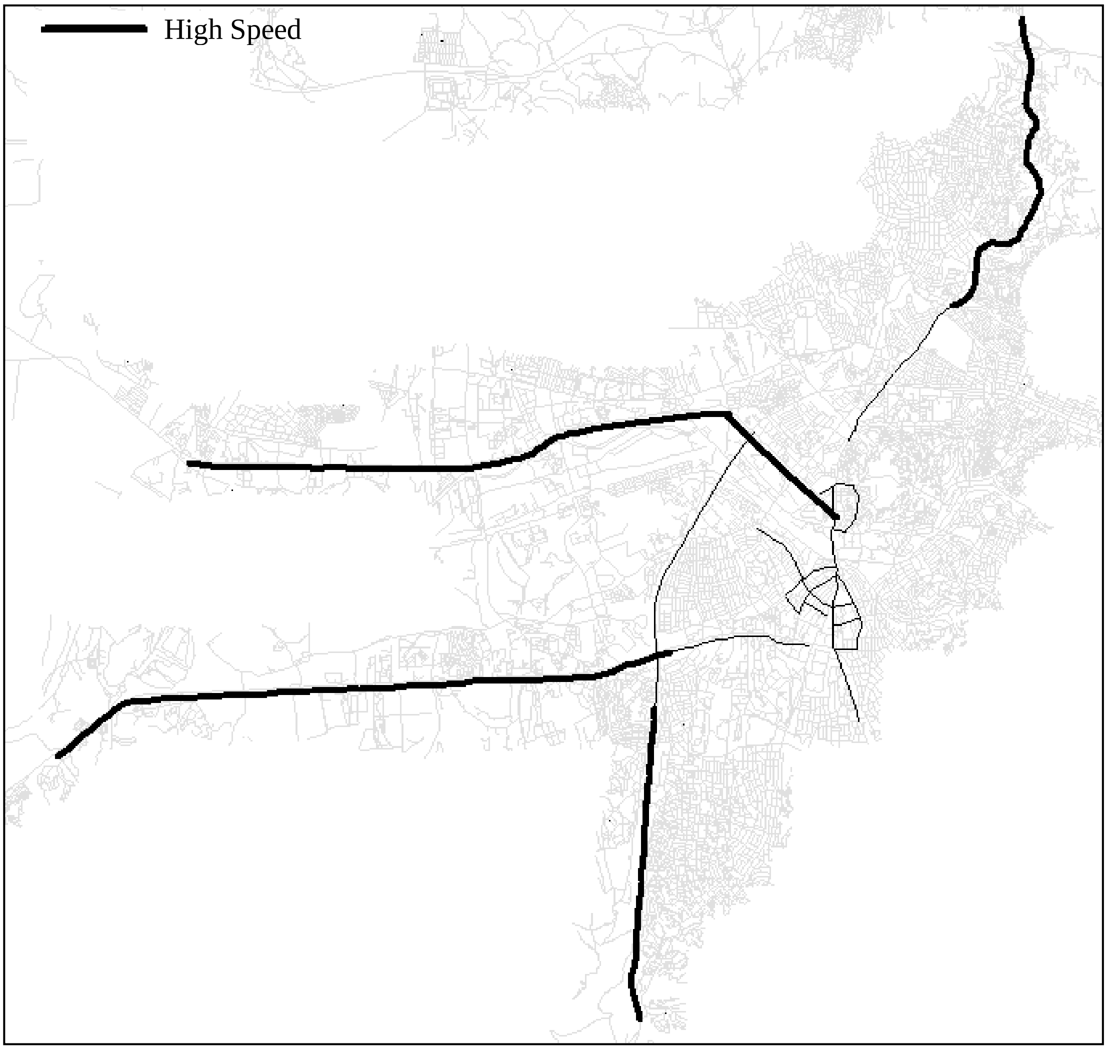


Figure C.2 Locations of Segments with High Speed (*D_High speed*= 1)

Figure C.3 Segments with Underpass (*D_Underpass* = 1)

Figure C.4 Segments with Metro Underpass (*D_Metro_Underpass* = 1)

Figure C.5 Segments with Overpass (*D_Overpass* = 1)

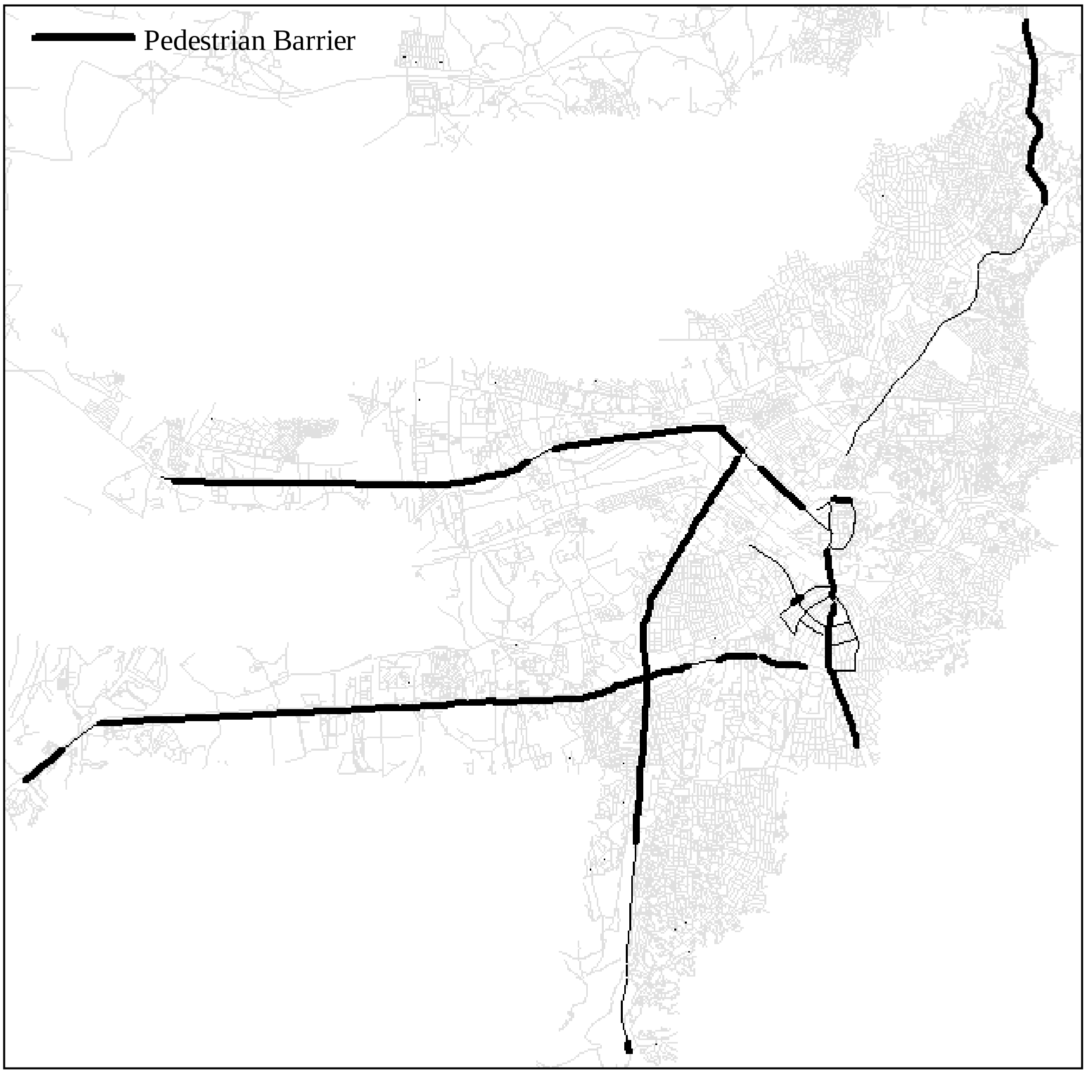


Figure C.6 Segments with Pedestrian Barrier (*D_PedBarr* = 1)

Figure C.7 Examples of Road Designs Constituting a Pedestrian Barrier in this Study

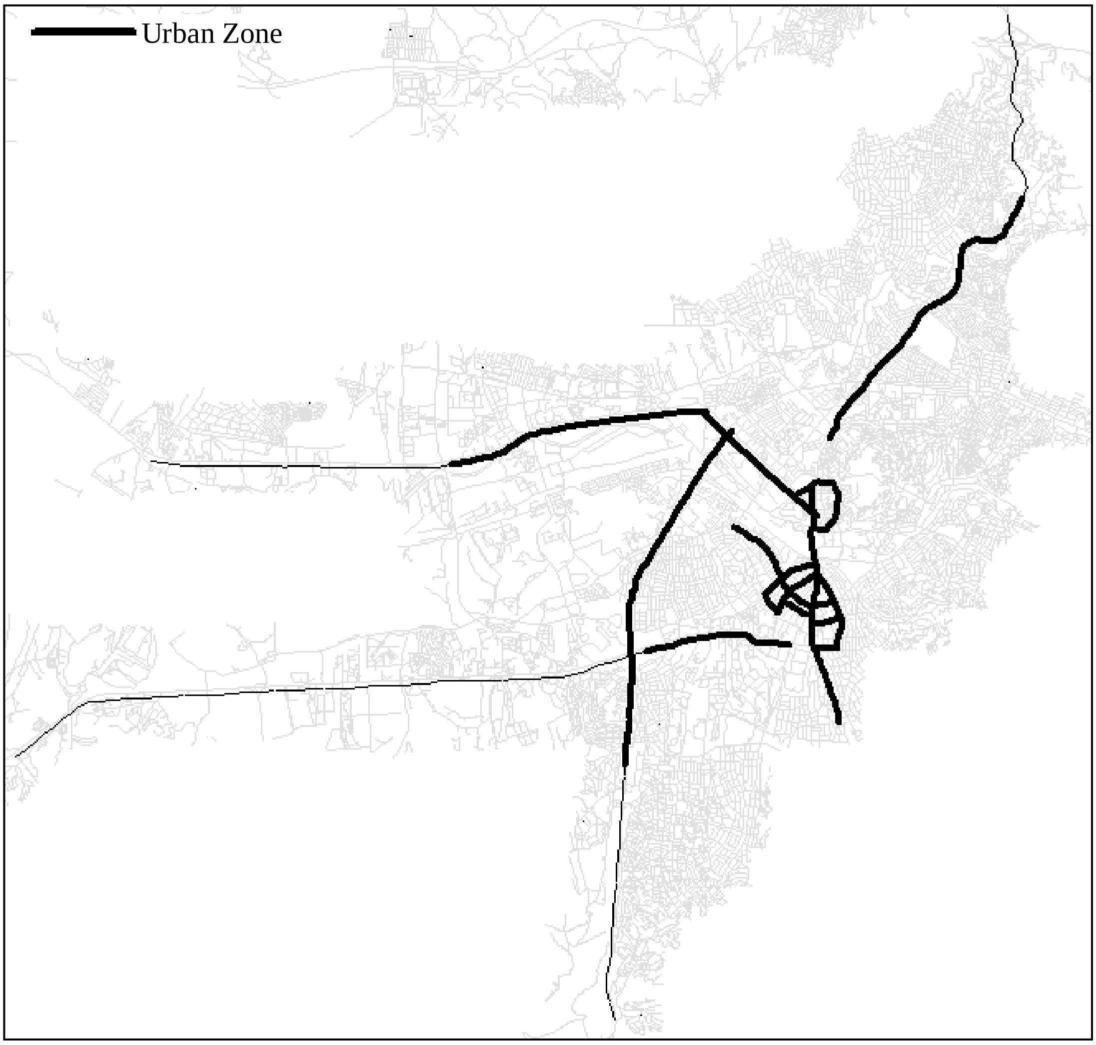


Figure C.8 Segments Defined as Urban Zone (*D_Urban zone* = 1)

D. PEDESTRIAN SAFETY REGRESSION ANALYSIS FOR CBD AND UTZ

Table D.1 Descriptive Statistics of the Variables Used in the Regressions (All Corridors)

	N	Minimum	Maximum	Mean	Std. Deviation
Dependent Variables					
SN_A	295	0	35	4.65	6.337
SD_A	295	0	176	22.38	32.316
$B_{50m}\_N_A$	344	0	12	1.811	2.567
$B_{100m}\_N_A$	344	0	21	5.052	5.914
$B_{400m}\_N_A$	344	0	63	22.137	20.989
Traffic Flow					
D_High speed	295	0	1	0.573	0.496
D_ HighPedVol	295	0	1	0.427	0.496
Traffic Network					
No_Lanes	295	2	10	6.119	1.354
Seg_L_km	295	0.105	0.300	0.219	0.039
No_Intersections	295	0	8	1.478	1.327
No_Legs	295	1	26	5.000	4.010
Built Environment					
D_Urban zone	295	0	1	0.603	0.490
D_Shop	295	0	1	0.383	0.487
D_Shop_2sides	295	0	1	0.190	0.393
D_BusComInd	295	0	1	0.559	0.497
D_PedBarr	295	0	1	0.617	0.487
D_Underpass	295	0	1	0.010	0.101
D_Overpass	295	0	1	0.153	0.360
D_Metro_Underpass	295	0	1	0.024	0.152
No_Bus_Stops	295	0	25	1.166	2.460

Table D.2 Descriptive Statistics of the Variables Used in the Regressions (CBD)

	N	Minimum	Maximum	Mean	Std. Deviation
Dependent Variables					
SN_A	72	0	35	8.389	9.233
SD_A	72	0	176	43.099	49.194
$B_{50m}\_N_A$	174	0	12	2.741	2.981
$B_{100m}\_N_A$	174	0	21	7.776	6.902
$B_{400m}\_N_A$	174	2	63	35.534	21.056
Traffic Flow					
D_High speed	72	0	0	0	0
D_ HighPedVol	72	1	1	1	0
Traffic Network					
No_Lanes	72	2	6	4.708	1.156
Seg_L_km	72	0.105	0.300	0.208	0.049
No_Intersections	72	0	8	2.472	1.610
No_Legs	72	1	26	8.222	5.144
Built Environment					
D_Urban zone	72	1	1	1	0
D_Shop	72	0	1	0.750	0.436
D_Shop_2sides	72	0	1	0.556	0.500
D_BusComInd	72	0	1	0.861	0.348
D_PedBarr	72	0	1	0.278	0.451
D_Underpass	72	0	1	0.014	0.118
D_Overpass	72	0	1	0.181	0.387
D_Metro_Underpass	72	0	1	0.056	0.231
No_Bus_Stops	72	0	25	2.431	4.506

Table D.3 Descriptive Statistics of the Variables Used in the Regressions (UTZ)

	N	Minimum	Maximum	Mean	Std. Deviation
Dependent Variables					
SN_A	223	0	32	3.448	4.464
SD_A	223	0	149	15.691	20.614
$B_{50m}\_N_A$	170	0	11	0.859	1.574
$B_{100m}\_N_A$	170	0	11	2.265	2.603
$B_{400m}\_N_A$	170	0	47	8.424	8.143
Traffic Flow					
D_High speed	223	0	1	0.758	0.429
D_ HighPedVol	223	0	1	0.242	0.429
Traffic Network					
No_Lanes	223	6	10	6.574	1.071
Seg_L_km	223	0.130	0.300	0.223	0.035
No_Intersections	223	0	6	1.157	1.039
No_Legs	223	1	18	3.960	2.895
Built Environment					
D_Urban zone	223	0	1	0.475	0.501
D_Shop	223	0	1	0.265	0.442
D_Shop_2sides	223	0	1	0.072	0.259
D_BusComInd	223	0	1	0.462	0.500
D_PedBarr	223	0	1	0.726	0.447
D_Underpass	223	0	1	0.009	0.094
D_Overpass	223	0	1	0.143	0.351
D_Metro_Underpass	223	0	1	0.013	0.115
No_Bus_Stops	223	0	5	0.758	0.913

Table D.4 Correlation of Independent Variables

		$D_{High\ speed}$	$D_{HighPedVol}$	No_{Lanes}	Seg_L_{km}	$No_{Intersections}$	No_{Legs}	$D_{Urban\ zone}$
$D_{High\ speed}$	*All*	1	-.709**	.547**	.161**	-.377**	-.394**	-.700**
	CBD	1	.[a]	.[a]	.[a]	.[a]	.[a]	.[a]
	UTZ	1	-.487**	.304**	.095	-.187**	-.196**	-.594**
$D_{HighPedVol}$	*All*	-.709**	1	-.547**	-.157**	.356**	.370**	.490**
	CBD	.[a]	1	.[a]	.[a]	.[a]	.[a]	.[a]
	UTZ	-.487**	1	-.304**	-.088	.147*	.146*	.279**
No_{Lanes}	*All*	.547**	-.547**	1	.129*	-.444**	-.472**	-.472**
	CBD	.[a]	.[a]	1	-.296*	-.568**	-.560**	.[a]
	UTZ	.304**	-.304**	1	.206**	-.106	-.123	-.326**
Seg_L_{km}	*All*	.161**	-.157**	.129*	1	.172**	.170**	-.125*
	CBD	.[a]	.[a]	-.296*	1	.479**	.479**	.[a]
	UTZ	.095	-.088	.206**	1	.125	.119	-.074
$No_{Intersections}$	*All*	-.377**	.356**	-.444**	.172**	1	.983**	.329**
	CBD	.[a]	.[a]	-.568**	.479**	1	.976**	.[a]
	UTZ	-.187**	.147*	-.106	.125	1	.986**	.220**
No_{Legs}	*All*	-.394**	.370**	-.472**	.170**	.983**	1	.329**
	CBD	.[a]	.[a]	-.560**	.479**	.976**	1	.[a]
	UTZ	-.196**	.146*	-.123	.119	.986**	1	.212**
$D_{Urban\ zone}$	*All*	-.700**	.490**	-.472**	-.125*	.329**	.329**	1
	CBD	.[a]	.[a]	.[a]	.[a]	.[a]	.[a]	1
	UTZ	-.594**	.279**	-.326**	-.074	.220**	.212**	1

**. Correlation is significant at the 0.01 level (2-tailed).
*. Correlation is significant at the 0.05 level (2-tailed).
a. Cannot be computed because at least one of the variables is constant.

Table D.4 Correlation of Independent Variables (cont'd)

		D_{Shop}	D_{Shop_2sides}	$D_{BusComInd}$	$D_{PedBarr}$	$D_{Underpass}$	$D_{Overpass}$	$D_{Metro\_Underpass}$	No_{Bus_Stops}
$D_{High\ speed}$	*All*	-.236**	-.333**	-.228**	.335**	.019	-.072	-.091	-.204**
	CBD	.[a]	.[a]	.[a]	.[a]	.[a]	.[a]	.[a]	.[a]
	UTZ	.078	.035	-.001	.123	.054	-.067	-.025	-.047
$D_{HighPedVol}$	*All*	.377**	.386**	.366**	-.391**	.049	.091	.136*	.201**
	CBD	.[a]	.[a]	.[a]	.[a]	.[a]	.[a]	.[a]	.[a]
	UTZ	.159*	.086	.211**	-.217**	.057	.097	.116	.035
No_{Lanes}	*All*	-.198**	-.401**	-.089	.358**	-.009	-.009	-.014	-.057
	CBD	-.258*	-.252*	-.137	.671**	.134	-.038	.273*	.316**
	UTZ	.192**	-.052	.226**	-.009	-.051	.043	-.063	-.014
Seg_L_{km}	*All*	-.063	-.078	.047	.027	-.118*	.018	-.057	.001
	CBD	.082	.074	.296*	-.055	-.156	-.002	.034	.057
	UTZ	-.026	-.048	.060	-.036	-.093	.040	-.105	.069
$No_{Intersections}$	*All*	.147*	.256**	.243**	-.321**	.014	.082	-.073	.036
	CBD	.151	.194	.169	-.455**	-.109	-.093	-.223	-.211
	UTZ	-.140*	-.109	.103	-.053	.077	.160*	-.055	.112
No_{Legs}	*All*	.181**	.281**	.242**	-.333**	.017	.087	-.072	.044
	CBD	.176	.192	.167	-.446**	-.098	-.042	-.212	-.188
	UTZ	-.129	-.104	.088	-.047	.084	.152*	-.066	.085
$D_{Urban\ zone}$	*All*	.012	.233**	.327**	-.339**	-.056	.093	.126*	.156**
	CBD	.[a]	.[a]	.[a]	.[a]	.[a]	.[a]	.[a]	.[a]
	UTZ	-.266**	-.021	.217**	-.222**	-.091	.097	.123	.076

**. Correlation is significant at the 0.01 level (2-tailed).
*. Correlation is significant at the 0.05 level (2-tailed).
a. Cannot be computed because at least one of the variables is constant.

Table D.4 Correlation of Independent Variables (cont'd)

		$D_{High\ speed}$	$D_{HighPedVol}$	No_{Lanes}	Seg_L_{km}	$No_{Intersections}$	No_{Legs}	$D_{Urban\ zone}$
D_{Shop}	*All*	-.236**	.377**	-.198**	-.063	.147*	.181**	.012
	CBD	.[a]	.[a]	-.258*	.082	.151	.176	.[a]
	UTZ	.078	.159*	.192**	-.026	-.140*	-.129	-.266**
D_{Shop_2sides}	*All*	-.333**	.386**	-.401**	-.078	.256**	.281**	.233**
	CBD	.[a]	.[a]	-.252*	.074	.194	.192	.[a]
	UTZ	.035	.086	-.052	-.048	-.109	-.104	-.021
$D_{BusComInd}$	*All*	-.228**	.366**	-.089	.047	.243**	.242**	.327**
	CBD	.[a]	.[a]	-.137	.296*	.169	.167	.[a]
	UTZ	-.001	.211**	.226**	.060	.103	.088	.217**
$D_{PedBarr}$	*All*	.335**	-.391**	.358**	.027	-.321**	-.333**	-.339**
	CBD	.[a]	.[a]	.671**	-.055	-.455**	-.446**	.[a]
	UTZ	.123	-.217**	-.009	-.036	-.053	-.047	-.222**
$D_{Underpass}$	*All*	.019	.049	-.009	-.118*	.014	.017	-.056
	CBD	.[a]	.[a]	.134	-.156	-.109	-.098	.[a]
	UTZ	.054	.057	-.051	-.093	.077	.084	-.091
$D_{Overpass}$	*All*	-.072	.091	-.009	.018	.082	.087	.093
	CBD	.[a]	.[a]	-.038	-.002	-.093	-.042	.[a]
	UTZ	-.067	.097	.043	.040	.160*	.152*	.097
$D_{Metro\_Underpass}$	*All*	-.091	.136*	-.014	-.057	-.073	-.072	.126*
	CBD	.[a]	.[a]	.273*	.034	-.223	-.212	.[a]
	UTZ	-.025	.116	-.063	-.105	-.055	-.066	.123
No_{Bus_Stops}	*All*	-.204**	.201**	-.057	.001	.036	.044	.156**
	CBD	.[a]	.[a]	.316**	.057	-.211	-.188	.[a]
	UTZ	-.047	.035	-.014	.069	.112	.085	.076

**. Correlation is significant at the 0.01 level (2-tailed).
*. Correlation is significant at the 0.05 level (2-tailed).
a. Cannot be computed because at least one of the variables is constant.

Table D.4 Correlation of Independent Variables (cont'd)

		D_{Shop}	D_{Shop_2sides}	$D_{BusComInd}$	$D_{PedBarr}$	$D_{Underpass}$	$D_{Overpass}$	$D_{Metro\_Underpass}$	No_{Bus_Stops}
D_{Shop}	*All*	1	.614**	.362**	-.096	-.010	.112	.106	.180**
	CBD	1	.645**	.232	-.143	.069	.271*	.140	.134
	UTZ	1	.464**	.260**	.163*	-.057	.044	.018	.003
D_{Shop_2sides}	*All*	.614**	1	.273**	-.259**	-.049	.083	.208**	.256**
	CBD	.645**	1	.368**	-.132	-.133	.202	.217	.180
	UTZ	.464**	1	.021	-.024	-.026	-.015	.118	-.002
$D_{BusComInd}$	*All*	.362**	.273**	1	-.208**	-.046	.054	.138*	.146*
	CBD	.232	.368**	1	.070	.048	.189	.097	.101
	UTZ	.260**	.021	1	-.118	-.088	.006	.126	.049
$D_{PedBarr}$	*All*	-.096	-.259**	-.208**	1	.080	.043	.031	-.012
	CBD	-.143	-.132	.070	1	.191	.031	.256*	.294*
	UTZ	.163*	-.024	-.118	1	.058	.079	-.016	-.064
$D_{Underpass}$	*All*	-.010	-.049	-.046	.080	1	.051	-.016	-.048
	CBD	.069	-.133	.048	.191	1	.253*	-.029	-.064
	UTZ	-.057	-.026	-.088	.058	1	-.039	-.011	-.079
$D_{Overpass}$	*All*	.112	.083	.054	.043	.051	1	-.004	.067
	CBD	.271*	.202	.189	.031	.253*	1	.044	.027
	UTZ	.044	-.015	.006	.079	-.039	1	-.048	.151*
$D_{Metro\_Underpass}$	*All*	.106	.208**	.138*	.031	-.016	-.004	1	.271**
	CBD	.140	.217	.097	.256*	-.029	.044	1	.343**
	UTZ	.018	.118	.126	-.016	-.011	-.048	1	.031
No_{Bus_Stops}	*All*	.180**	.256**	.146*	-.012	-.048	.067	.271**	1
	CBD	.134	.180	.101	.294*	-.064	.027	.343**	1
	UTZ	.003	-.002	.049	-.064	-.079	.151*	.031	1

**. Correlation is significant at the 0.01 level (2-tailed).
*. Correlation is significant at the 0.05 level (2-tailed).
a. Cannot be computed because at least one of the variables is constant.

	CBD_Model 1		CBD_Model 2		UTZ_Model 1		UTZ_Model 2	
Dependent Variable	SD_A		SD_A		SD_A		SD_A	
Model Statistics	**R = 0.850**		**R = 0.865**		**R = 0.740**		**R = 0.745**	
Total SSE	*305565.402*		*305565.402*		*149240.024*		*149240.024*	
Regression SSE	*220836.056*		*228515.543*		*81763.500*		*82838.989*	
Number of cases	*72*		*72*		*223*		*223*	
	Coeff.	***t* value**	**Coeff.**	***t* value**	**Coeff.**	***t* value**	**Coeff.**	***t* value**
Independent Variables								
Constant								
Traffic Flow								
D_High speed								
D_ HighPedVol					10.053	*3.552*	9.503	*3.366*
Traffic Network								
No_Lanes	13.836	*7.001*	19.006	*6.743*			0.602	*2.556*
Seg_L_km			-139.010	*-2.023*				
No_Intersections								
No_Legs								
Built Environment								
D_Urban zone					19.198	*8.908*	17.384	*7.415*
D_Shop								
D_Shop_2sides	19.249	*2.066*	17.115	*1.845*	17.565	*3.892*	16.200	*3.576*
D_BusComInd	-47.864	*-4.092*	-34.988	*-2.652*				
D_PedBarr			-23.145	*-2.042*				
D_Underpass								
D_Overpass								
D_Metro_Underpass	57.134	*2.886*	60.240	*3.081*				
No_Bus_Stops	2.525	*2.456*	2.794	*2.76*	1.963	*1.714*		

Figure D.1 Best Regressions Models for the CBD and the UTZ Corridors with Segment Accident Density

	CBD_Model 1		CBD_Model 2		CBD_Model 3	
Dependent Variable	$B_{50m}\_N_A$		$B_{50m}\_N_A$		$B_{50m}\_N_A$	
Model Statistics	**R = 0.868**		**R = 0.805**		**R = 0.880**	
Total SSE	*2845.000*		*2845.000*		*2845.000*	
Regression SSE	*2143.23*		*1842.310*		*2204.787*	
Number of cases	*174*		*174*		*174*	
Independent Variables	**Coeff.**	***t* value**	**Coeff.**	***t* value**	**Coeff.**	***t* value**
Constant						
Segment Accidents						
Seg_Acc_Den	0.035	*22.868*			0.039	17.271
Traffic Flow						
D_High speed						
D_ HighPedVol						
Traffic Network						
No_Lanes			1.142	*11.357*		
Seg_L_km						
No_Intersections						
No_Legs						
Built Environment						
D_Urban zone						
D_Shop						
D_Shop_2sides			2.310	*4.988*	0.684	*1.823*
D_BusComInd			-5.630	*-7.850*	-1.275	*-3.598*
D_PedBarr						
D_Underpass						
D_Overpass			0.892	*1.856*	0.971	*2.540*
D_Metro_Underpass						

Figure D.2 Best Regression Models for Bus Stop Pedestrian Safety Using Number of Accidents for 50 m Buffer Zones (CBD Corridors)

	CBD_Model 1		CBD_Model 2		CBD_Model 3	
Dependent Variable	B_{100m}_N_A		B_{100m}_N_A		B_{100m}_N_A	
Model Statistics	**R = 0.955**		**R = 0.851**		**R = 0.968**	
Total SSE	*18761.000*		*18761.000*		*18761.000*	
Regression SSE	*17113.141*		*13574.847*		*17589.694*	
Number of cases	*174*		*174*		*174*	
Independent Variables	**Coeff.**	***t* value**	**Coeff.**	***t* value**	**Coeff.**	***t* value**
Constant						
Segment Accidents						
Seg_Acc_Den	0.098	*42.387*			0.109	*36.025*
Traffic Flow						
D_High speed						
D_ HighPedVol						
Traffic Network						
No_Lanes			2.853	*12.765*		
Seg_L_km						
No_Intersections						
No_Legs						
Built Environment						
D_Urban zone						
D_Shop						
D_Shop_2sides			5.136	*4.906*	1.831	*3.514*
D_BusComInd			-12.132	*-7.706*	-1.909	*-4.068*
D_PedBarr						
D_Underpass						
D_Overpass						
D_Metro_Underpass					-3.870	*-7.302*

Figure D.3 Best Regression Models for Bus Stop Pedestrian Safety Using Number of Accidents for 100 m Buffer Zones (CBD Corridors)

	CBD_Model 1		CBD_Model 2		CBD_Model 3	
Dependent Variable	B_{400m}_N_A		B_{400m}_N_A		B_{400m}_N_A	
Model Statistics	**R = 0.950**		**R = 0.952**		**R = 0.977**	
Total SSE	*296413.000*		*296413.000*		*296413.000*	
Regression SSE	*267551.061*		*268636.103*		*282751.949*	
Number of cases	*174*		*174*		*174*	
Independent Variables	**Coeff.**	***t* value**	**Coeff.**	***t* value**	**Coeff.**	***t* value**
Constant						
Segment Accidents						
Seg_Acc_Den	0.398	*40.046*			0.215	*13.215*
Traffic Flow						
D_High speed						
D_ HighPedVol						
Traffic Network						
No_Lanes			10.291	*19.391*	4.342	*7.426*
Seg_L_km						
No_Intersections						
No_Legs						
Built Environment						
D_Urban zone						
D_Shop						
D_Shop_2sides			20.470	*8.020*	11.674	*6.097*
D_BusComInd			-38.747	*-10.445*	-16.722	*-5.401*
D_PedBarr						
D_Underpass						
D_Overpass						
D_Metro_Underpass			8.790	*3.375*	8.804	*4.805*

Figure D.4 Best Regression Models for Bus Stop Pedestrian Safety Using Number of Accidents for 400 m Buffer Zones (CBD Corridors)

	UTZ_Model 1		UTZ_Model 2	
Dependent Variable	B_{50m}_N_A		B_{50m}_N_A	
Model Statistics	**R = 0.782**		**R = 0.577**	
Total SSE	*544.000*		*544.000*	
Regression SSE	*332.79*		*180.830*	
Number of cases	*170*		*170*	
Independent Variables	**Coeff.**	***t* value**	**Coeff.**	***t* value**
Constant				
Segment Accidents				
Seg_Acc_Den	0.054	*16.318*		
Traffic Flow				
D_High speed			-0.552	*-1.777*
D_ HighPedVol				
Traffic Network				
No_Lanes			0.111	*2.227*
Seg_L_km				
No_Intersections				
No_Legs				
Built Environment				
D_Urban zone			0.885	*3.611*
D_Shop				
D_Shop_2sides			0.843	*1.896*
D_BusComInd				
D_PedBarr				
D_Underpass				
D_Overpass				
D_Metro_Underpass				

Figure D.5 Best Regression Models for Bus Stop Pedestrian Safety Using Number of Accidents for 50 m Buffer Zones (UTZ Corridors)

	UTZ_Model 1		UTZ_Model 2		UTZ_Model 3	
Dependent Variable	$B_{100m}\_N_A$		$B_{100m}\_N_A$		$B_{100m}\_N_A$	
Model Statistics	**R = 0.888**		**R = 0.743**		**R = 0.897**	
Total SSE	*2017.000*		*2017.000*		*2017.000*	
Regression SSE	*1592.151*		*1114.930*		*1622.279*	
Number of cases	*170*		*170*		*170*	
Independent Variables	**Coeff.**	***t* value**	**Coeff.**	***t* value**	**Coeff.**	***t* value**
Constant						
Segment Accidents						
Seg_Acc_Den	0.119	*25.166*			0.114	*15.725*
Traffic Flow						
D_High speed						
D_ HighPedVol			0.778	*1.888*		
Traffic Network						
No_Lanes			0.120	*3.237*		
Seg_L_km						
No_Intersections						
No_Legs						
Built Environment						
D_Urban zone			2.392	*6.841*	0.443	*1.786*
D_Shop						
D_Shop_2sides					-1.302	*-2.751*
D_BusComInd						
D_PedBarr						
D_Underpass						
D_Overpass						
D_Metro_Underpass						

Figure D.6 Best Regression Models for Bus Stop Pedestrian Safety Using Number of Accidents for 100 m Buffer Zones (UTZ Corridors)

	UTZ_Model 1		UTZ_Model 2		UTZ_Model 3	
Dependent Variable	B_{400m}_N_A		B_{400m}_N_A		B_{400m}_N_A	
Model Statistics	**R = 0.888**		**R = 0.840**		**R = 0.921**	
Total SSE	*23268.000*		*23268.000*		*23268.000*	
Regression SSE	*18364.397*		*16437.313*		*19724.320*	
Number of cases	*170*		*170*		*170*	
Independent Variables	**Coeff.**	***t* value**	**Coeff.**	***t* value**	**Coeff.**	***t* value**
Constant						
Segment Accidents						
Seg_Acc_Den	0.404	*25.158*			0.284	13.161
Traffic Flow						
D_High speed						
D_ HighPedVol			2.828	*2.479*		
Traffic Network						
No_Lanes			0.611	*3.943*	0.372	*3.330*
Seg_L_km						
No_Intersections						
No_Legs						
Built Environment						
D_Urban zone			9.259	*9.404*	4.523	*5.668*
D_Shop						
D_Shop_2sides						
D_BusComInd						
D_PedBarr			-2.011	*-1.899*	-2.156	*-2.855*
D_Underpass						
D_Overpass						
D_Metro_Underpass			7.476	*1.962*		

Figure D.7 Best Regression Models for Bus Stop Pedestrian Safety Using Number of Accidents for 400 m Buffer Zones (UTZ Corridors)

	CBD_Model 1		CBD_Model 2		CBD_Model 3	
Dependent Variable	$\ln(B_{50m}\_N_A)$		$\ln(B_{50m}\_N_A)$		$\ln(B_{50m}\_N_A)$	
Model Statistics	**R = 0.889**		**R = 0.851**		**R = 0.899**	
Total SSE	*212.939*		*212.939*		*212.939*	
Regression SSE	*168.207*		*154.232*		*172.285*	
Number of cases	*128*		*128*		*128*	
Independent Variables	**Coeff.**	***t* value**	**Coeff.**	***t* value**	**Coeff.**	***t* value**
Constant						
Segment Accidents						
Seg_Acc_Den	0.010	*21.853*			0.010	*14.324*
Traffic Flow						
D_High speed						
D_ HighPedVol						
Traffic Network						
No_Lanes			0.323	*10.049*		
Seg_L_km						
No_Intersections						
No_Legs						
Built Environment						
D_Urban zone						
D_Shop						
D_Shop_2sides			1.039	*5.131*	0.480	*3.472*
D_BusComInd			-1.496	*-6.281*	-0.430	*-3.151*
D_PedBarr			-0.377	*-2.253*		
D_Underpass						
D_Overpass						
D_Metro_Underpass						

Figure D.8 Best Regression Models for Bus Stop Pedestrian Safety Using Logarithmic Transformation of Number of Accidents for 50 m Buffer Zones (CBD Corridors)

	CBD_Model 1		CBD_Model 2		CBD_Model 3	
Dependent Variable	$\ln(B_{100m}\_N_A)$		$\ln(B_{100m}\_N_A)$		$\ln(B_{100m}\_N_A)$	
Model Statistics	**R = 0.964**		**R = 0.911**		**R = 0.969**	
Total SSE	*625.059*		*625.059*		*625.059*	
Regression SSE	*581.442*		*518.403*		*586.676*	
Number of cases	*166*		*166*		*166*	
Independent Variables	**Coeff.**	***t* value**	**Coeff.**	***t* value**	**Coeff.**	***t* value**
Constant						
Segment Accidents						
Seg_Acc_Den	0.018	*46.899*			0.015	*20.377*
Traffic Flow						
D_High speed						
D_ HighPedVol						
Traffic Network						
No_Lanes			0.501	*13.499*	0.033	*1.901*
Seg_L_km						
No_Intersections						
No_Legs						
Built Environment						
D_Urban zone						
D_Shop						
D_Shop_2sides			0.913	*5.224*	0.184	*2.394*
D_BusComInd			-1.621	*-6.545*		
D_PedBarr			-0.399	*-2.608*		
D_Underpass						
D_Overpass						
D_Metro_Underpass						

Figure D.9 Best Regression Models for Bus Stop Pedestrian Safety Using Logarithmic Transformation of Number of Accidents for 100 m Buffer Zones (CBD Corridors)

	CBD_Model 1		CBD_Model 2		CBD_Model 3	
Dependent Variable	$\ln(B_{400m}\_N_A)$		$\ln(B_{400m}\_N_A)$		$\ln(B_{400m}\_N_A)$	
Model Statistics	**R = 0.910**		**R = 0.987**		**R = 0.991**	
Total SSE	*2011.281*		*2011.281*		*2011.281*	
Regression SSE	*1664.122*		*1959.661*		*1976.2*	
Number of cases	*174*		*174*		*174*	
Independent Variables	**Coeff.**	***t* value**	**Coeff.**	***t* value**	**Coeff.**	***t* value**
Constant						
Segment Accidents						
Seg_Acc_Den	0.031	*28.797*			0.008	*11.564*
Traffic Flow						
D_High speed						
D_ HighPedVol						
Traffic Network						
No_Lanes			0.664	*26.692*	0.415	*25.136*
Seg_L_km						
No_Intersections						
No_Legs						
Built Environment						
D_Urban zone						
D_Shop						
D_Shop_2sides			1.109	*9.650*	0.687	*9.226*
D_BusComInd			-0.982	*-6.061*		
D_PedBarr			-0.410	*-3.959*	-0.288	*-3.408*
D_Underpass						
D_Overpass						
D_Metro_Underpass			0.203	*1.759*	0.190	*2.000*

Figure D.10 Best Regression Models for Bus Stop Pedestrian Safety Using Logarithmic Transformation of Number of Accidents for 400 m Buffer Zones (CBD Corridors)

	UTZ_Model 1		UTZ_Model 2	
Dependent Variable	$\ln(B_{50m}\_N_A)$		$\ln(B_{50m}\_N_A)$	
Model Statistics	**R = 0.794**		**R = 0.718**	
Total SSE	*44.444*		*44.444*	
Regression SSE	*27.986*		*22.902*	
Number of cases	*70*		*70*	
Independent Variables	**Coeff.**	***t* value**	**Coeff.**	***t* value**
Constant				
Segment Accidents				
Seg_Acc_Den	0.017	*10.832*		
Traffic Flow				
D_High speed				
D_ HighPedVol				
Traffic Network				
No_Lanes				
Seg_L_km				
No_Intersections				
No_Legs				
Built Environment				
D_Urban zone			0.629	*7.658*
D_Shop				
D_Shop_2sides				
D_BusComInd				
D_PedBarr				
D_Underpass				
D_Overpass				
D_Metro_Underpass			1.451	*2.551*

Figure D.11 Best Regression Models for Bus Stop Pedestrian Safety Using Logarithmic Transformation of Number of Accidents for 50 m Buffer Zones (UTZ Corridors)

	UTZ_Model 1		UTZ_Model 2		UTZ_Model 3	
Dependent Variable	ln(B_{100m}_N_A)		ln(B_{100m}_N_A)		ln(B_{100m}_N_A)	
Model Statistics	**R = 0.855**		**R = 0.792**		**R = 0.867**	
Total SSE	*158.958*		*158.958*		*158.958*	
Regression SSE	*116.311*		*99.675*		*119.367*	
Number of cases	*122*		*122*		*122*	
Independent Variables	**Coeff.**	***t* value**	**Coeff.**	***t* value**	**Coeff.**	***t* value**
Constant						
Segment Accidents						
Seg_Acc_Den	0.032	*18.166*			0.026	*9.760*
Traffic Flow						
D_High speed						
D_ HighPedVol			0.395	*2.728*		
Traffic Network						
No_Lanes			0.053	*3.537*		
Seg_L_km						
No_Intersections						
No_Legs						
Built Environment						
D_Urban zone			0.599	*4.727*	0.307	*3.043*
D_Shop						
D_Shop_2sides						
D_BusComInd						
D_PedBarr						
D_Underpass						
D_Overpass						
D_Metro_Underpass						

Figure D.12 Best Regression Models for Bus Stop Pedestrian Safety Using Logarithmic Transformation of Number of Accidents for 100 m Buffer Zones (UTZ Corridors)

	UTZ_Model 1		UTZ_Model 2		UTZ_Model 3	
Dependent Variable	$\ln(B_{400m}\_N_A)$		$\ln(B_{400m}\_N_A)$		$\ln(B_{400m}\_N_A)$	
Model Statistics	**R = 0.852**		**R = 0.936**		**R = 0.958**	
Total SSE	*659.818*		*659.818*		*659.818*	
Regression SSE	*479.214*		*578.658*		*605.724*	
Number of cases	*155*		*155*		*155*	
Independent Variables	**Coeff.**	***t* value**	**Coeff.**	***t* value**	**Coeff.**	***t* value**
Constant						
Segment Accidents						
Seg_Acc_Den	0.065	*20.214*			0.026	*9.121*
Traffic Flow						
D_High speed						
D_HighPedVol			0.321	*2.440*		
Traffic Network						
No_Lanes			0.147	*11.777*	0.140	*9.395*
Seg_L_km						
No_Intersections						
No_Legs						
Built Environment						
D_Urban zone			1.339	*11.825*	0.917	*8.785*
D_Shop						
D_Shop_2sides						
D_BusComInd						
D_PedBarr					-0.174	*-1.717*
D_Underpass						
D_Overpass						
D_Metro_Underpass						

Figure D.13 Best Regression Models for Bus Stop Pedestrian Safety Using Logarithmic Transformation of Number of Accidents for 400 m Buffer Zones (UTZ Corridors)