EFFECTS OF STREET PATTERN ON FREQUENCY OF TRAFFIC CRASH: A CASE STUDY OF GAINESVILLE, FLORIDA

By

DIXUE LI

A THESIS PRESENTED TO THE GRADUATE SCHOOL
OF THE UNIVERSITY OF FLORIDA IN PARTIAL FULFILLMENT
OF THE REQUIREMENTS FOR THE DEGREE OF
MASTER OF ARTS IN URBAN AND REGIONAL PLANNING

UNIVERSITY OF FLORIDA

2011

To my beloved parents

ACKNOWLEDGMENTS

I would like to thank all my committee members, Dr. Blanco and Dr. Zwick, for their mentoring, keen assistance, and generous support. I would also like to thank my parents for their constant support and loving encouragement, which motivated me to complete this milestone. Finally, I would like to thank all my friends for helping me through the rough times and never giving up on me.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS

AIC	Akaike's Information Criterion
FGDL	Florida Geographic Data Library
GIS	Geographic Information Software
GWR	Geographically Weighted Regression
NMVCCS	National Motor Vehicle Crash Causation Statistic
NRI	National Resources Inventory
OLS	Ordinary Least Squares Regression
VMT	Vehicle Miles Traveled

Abstract of Thesis Presented to the Graduate School
of the University of Florida in Partial Fulfillment of the
Requirements for the Degree of Master of Arts in Urban and Regional Planning

EFFECTS OF STREET PATTERN ON FREQUENCY OF TRAFFIC CRASH:
GAINESVILLE AS A CASE STUDY

By

Dixue Li

December 2011

Chair: Andres Blanco
Cochair: Paul Zwick
Major: Urban and Regional Planning

Since the vehicles have been widely used in our daily life, the occurrence of the traffic accidents has never been stopped. During the six years from 2003 to 2008, there were about 6,000,000 crashes happened annually in The United States. Street pattern as one of the factors for the road traffic accident can influence travel behavior of residents. Furthermore, different socioeconomic and demographic factors in same street pattern can contribute different safety issues. In the previous study, only a few research mentioned about effects of street network design on traffic accidents. Therefore, this thesis is aimed to explore the relationship between the street pattern design and crash incidence.

In this study, the crash data came from City of Gainesville. It used the block group as the unit of analysis, which helped the regression model to examine the effects of street network design on crash. The results from regression models provide planners and policy makers with visual map based spatial distribution on the type of street pattern and transportation factors, which would reduce road safety problems in the face of urban sprawl. The research found that grid pattern was associated with more traffic

accidents. Compared with grid, loop and parallel pattern play a significant role in reducing the frequency of traffic crashes by controlling other contributing factors. In this paper, two main statistical models were used to analyze the effect of street pattern and other relevant factors on the frequency of traffic accidents, including the Poisson Regression and Multiple Linear Regression. Besides, this research investigated the predicting map, which generated from Geographically Weighted Regression (GWR) in ArcGIS to examine the hypothesis coming from the literature review.

Based on the analysis and results, we could discern that the loop and parallel pattern were highly recommended for improving transportation safety in southeastern Gainesville. More population density and Vehicle Miles Traveled (VMT) would contribute high frequency of crashes in the western area. Less household density would lead to the higher traffic incident.

CHAPTER 1
INTRODUCTION

Background

Road network constitutes the frame of the city. It is also the carrier of most social-economic activities, especially passenger and freight transportation. Urban transportation is the proverbial “blood” of the city; however, it also the reason that city roads face grimmer challenges today than ever before. Ever since motor vehicles appeared on the road over a hundred years ago, traffic accidents have never ceased to be commonplace events. Although the number of accidents per vehicle tends to decrease with the increasing quality of vehicles (Carlsson and Hedman, 1990), the rates are still too high. In the U.S., for instance, the total number of accidents has been steady around 6,000,000 per year; this statistic comes from the police-report of crash severity from 2003 to 2008 (Table 1-1). The reasons for traffic accidents can be divided into two types: the subjective factor, such as driver behavior, and the objective factor, such as vehicle condition and road or traffic environment. Based on the National Motor Vehicle Crash Causation Statistic (NMVCCS) data, about 73.6 percent of the estimated annual crashes featured critical reasons attributed to drivers, while the vehicle-or environment-attributed critical reasons were assigned to less than 17 percent of these crashes (NHTSA, 2008). This means that the road environment has been usually ignored in previous considerations of vehicle crash causation. In fact, with the exception of some accidents obviously attributable to careless driving and other subjective causes, a number of accidents are caused by improper operation and difficult driving conditions. And difficult driving conditions are directly related to road network design and maintenance. Therefore, the tremendous importance of road network design on traffic security cannot be ignored.

In the past, engineers and builders did not generally have a clear sense of street pattern. The main function of the road was to help people transport materials from one place to another with high efficiency. Irregular street networks, largely of medieval origin, were shaped by the location of pre-existing structures. With growing populations and the proliferation of large urban centers, the traditional grid pattern evolved from such original street layouts. People began to divide the new area into blocks, which were flanked by a grid of straight roads; this linear block plan was not prevalent until World War II (Marshall and Garrick, 2010). Due to the number of automobiles used by residents, other varieties of street patterns appeared for accommodating vehicle transportation such as parallel, loops, lollipops and cul-de-sac (Grammenos 2002).

Different street patterns can influence the travel behavior of residents. James M. (1979) mentioned that Perry (1929) proposed the concept of the neighborhood unit plan which promoted familiarity and interaction among residents. Perry suggested that the use of non-linear and cul-de-sac streets should be widely generalized. This concept was adopted by designers, who put it into practice, and researchers have since found that these types of street pattern can strongly influence neighboring behavior. They provide comparatively traffic-friendly neighborhoods, which encourage residents to comfortably ride bicycles and walk on streets and sidewalks. Such environments greatly reduced the chance of driving cars, generally seen as a positive development by traffic engineers and city planners desirous of stemming the frequency of traffic accidents. However, for the city with old infrastructures, where buildings and roads cannot be uprooted, it is not feasibility to change the basic network, especially in the downtown area. Even if the old city districts need to be reconstructed, many complicated limitations influence the re-configuring of road networks.

Nevertheless, it does not follow that this idea is meaningless. With the development of modern cities, new communities have mushroomed. The survey of National Resources Inventory (NRI) showed that "about 8,900 square kilometers (2.2 million acres) of land in the United States were developed between 1992 and 2002"(Lubowski, Marlow, Shawn, Alba and Michael, 2006). A series of numbers mean that there were plenty of new communities, industrial districts and commercial districts which appeared in the process of rural development. This phenomenon is more common in the developing world. In China, most cities are seeking a wider developing space to accommodate enormous populations. For example, "Pudong New Area" in Shanghai and "Qianjiang CBD" in Hangzhou are the representatives of this trend of massive expansion; they cover an area of more than 1000 kms^2.

Before claiming a new undeveloped area, planners always design the road network first. Under such circumstances, they need not be concerned with the existing and historical structures which are common to older cities; the question of what street patterns are ideal for traffic safety are paramount considerations for road engineers and planners designing new districts.

The aim of this study is to explore the relationship between street patterns and the occurrence of traffic accidents in Gainesville, Florida, and, furthermore, determine whether some types of street patterns incline towards a lower frequency of such accidents. If a certain relationship exists, this study seeks to gauge its significance and determine what the spatial distribution looks like in study area, based on the statistical analysis and tools in Geographic Information Software (GIS). The description has a significant impact on the analysis of street pattern spatial distribution and transportation safety.

Gainesville, Florida as Study Area

As a study area, Gainesville is used as a case to verify the effects of street pattern on the frequency of traffic accidents. Based on the analysis of the relationship, I will not only describe the significance of street patterns on traffic wrecks with the help of a statistic model, but also explain the socio-economic factors which contribute to safety issues. Thus, Gainesville is an ideal case for these particular reasons.

Firstly, Gainesville is, by far, the largest city in the Alachua County. There are multiple types of street patterns in the city layout such as grid, parallel, loops and cul-de-sac. It meets the basic demand of the main existing types of pattern. In order to analyze the relationship between street pattern and traffic wreck more effectively, all the data within city limits are valid. So the larger area and more various the types of street pattern, the better for this study.

Secondly, Gainesville is a rapidly expanding city. It is experiencing such dramatic growth because of the population boom in education and business. Based on U.S. Census data from 1990 to 2010, the population in 1990 totaled 84,770; in 2000, the population rose by 12.6%. By 2010, it kept growing by 30.3% annually (U.S. Census, 2010). If the number continues rising by this rate, within the next twenty years, the total population will exceed one million. In the meanwhile, more communities were built. Therefore, it is necessary to provide evidence for the ideal types of road network to accommodate development of the neighborhoods.

Thirdly, the existence of abundant traffic accident data and other information resources such as GIS shape files, social-economic and demographic data are also main advantages to using Gainesville as a case. Moreover, the greater familiarization with the area greatly assists the author in the course of the study; and, in many cases, the more adequate reason can be explained by local observations.

Overview of the Thesis

The main purpose of this thesis is to explore the relationship between different pattern of street network and frequency of crashes. Using the data I got from Florida Geographic Data Library (FGDL) and Census website, I will also try to explain the socioeconomic or demographic factors which contribute safety issues. So the regression model will be conducted to test these hypotheses:

Due to the grid pattern consists of the linear road which is made drivers speed up easily and contributed to the proliferation of accidents. The curving streets and limited connectivity usually contributed the increasing the attention of drivers. Also the limited interconnection will reduce the usage of automobile. The grid pattern should have positive correlation with crash incidence.

Besides being affected by the street pattern, the other social-economic and demographic factors also tend to change the frequency of crash. The appropriate model takes into account all of the factors that contribute to crash incidence to measure the effects of different predictors.

The thesis consists of an introduction and other four main parts as follows. Chapter 1 (Introduction) briefly describes the background of the research, study objectives and significance. Chapter 2 (Literature Review) introduces the classification of different types of street pattern and their evolution. Then it reviews all the theories and statistical analysis used by other researchers, with an emphasis on the relationship between street patterns in neighborhoods and transportation safety. Chapter 3 (Data and Methodology) introduces the sources of the data, study area, statistical model and analysis procedures. Based on the literature review, the proper statistical models in both SPSS and GIS software are discussed. Chapter 4 (Analysis and Results Discussions) demonstrates that the frequency of traffic wrecks occurring on the grid pattern is much higher than that in the other two patterns. And this chapter also discusses the influence of relative social-economic factors on traffic safety. Chapter 5 (Conclusions and Future Research) illustrates that the study results can contribute to the objective of improving road safety and help planners and decision makers design better neighborhoods in the light of urban planning. Meanwhile, some issues for this research and what problems need to be explored in the future are also mentioned.

Table.1-1. Crashes number by crash severity：2003 to 2008

Item	2003	2004	2005	2006	2007	2008
Crashes(1,000)	6,394	6,181	6,159	5,973	6,024	5,811
Fatal	38.5	38.4	39.3	38.6	37.4	34.0
Nonfatal Injury	1,925	1,862	1,816	1,746	1,711	1,630
Property Damage Only	4,365	4,281	4,304	4,189	4,275	4,146

Source: U.S. National Highway Safety Traffic Administration, Traffic Safety Facts, annual. (2008). http://www-nrd.nhtsa.dot.gov/CATS/index.aspx. Last accessed October, 2011.edited by author.

CHAPTER 2
LITERATURE REVIEW

This chapter presents the classification of street patterns, such as grid-iron, parallel, loop and lollipops. Following that, the second part discusses the main risk factors in traffic safety issues. The last part reviews all the regression models that are used by relevant researches.

Type of Street Pattern

The street patterns can be classified by different principles. Some types of patterns were usually described based on typologies of network (Lynch, 1990). Ray Brindle clarified the grid and the tributary as two broad types of layout (Brindle, 1996). Stephen Marshal defined four types of street pattern in the urban scope. These were "Altstadt, Bilateral, Conjoint and Distributory" (Marshall, 2005). Altstadt usually indicated "the irregular and fine scale angular streets, which mostly short or crooked, varying in width and going in all directions". Bilateral mainly consisted of grid with crossroads. Conjoint was a type of street morphologically situated between regularity and irregularity, with curved or rectilinear formations. Distributory consists of curlinear or rectilinear formations with many tree-like configurations in it (Marshall, 2005). Other scholars divided route structure by its heterogeneity into the irregular, regular, recursive and characteristic pattern (DTLR, 2001). Among these researchers, Southworth and Owens demonstrated the detailed classification of street pattern by discussing design characteristics in neighborhoods. They classified the pattern into five typical categories (Figure 2-1). They were Grid-iron, Fragmented Parallel, Warped Parallel, Loops and Lollipops and Lollipops on a stick (Southworth, 2003). Based on the scheme developed by Southworth, Rifaat reclassified the street pattern into four major types in order to fit his research. He summarized the road patterns in Calgary as Grid-iron, Warped Parallel, Loops and lollipops and Mixed (Riffat, 2011). The scheme developed by Southworth was widely used in practical applications. The following sections provide brief descriptions of each street pattern.

Advantages of Grid-iron Street Pattern

The grid-iron street layout is a traditional pattern which was conceived in the Laws of the Indies (1573), introduced by King Philip II of Spain. The city consists of central plaza, public buildings and residential houses (Garr, 1991).

Grid-iron street pattern recurs in some settlements, which are easily associated with each other. Since the grid was automatically adopted by these regions, no contemporary opposition challenges such an entrenched and idiosyncratic system. Obviously, it has both advantages and disadvantages. In its favor, this pattern helps a compact settlement to use space efficiently. It is also easy for the military to control, which indicates colonial status (Dan, 1946). That's why the first batch of European cities established on the American continent put the grid pattern to use during the heyday of Spanish colonial influence. Since 1573, according to the guidance of the laws which originated from the Indies, numerous grid plan settlements spread throughout the Americas. These early cities include Philadelphia, New York, Washington, D.C. and Savannah. Then it rapidly expanded to the west, largely due to the enchantment over the rapid speculation of land such as Salt Lake City and San Francisco (Moudon and Untermann, 1987). From ancient to modern society, the grid-iron pattern mainly consists of linear roads. This type of road is easy to survey: clear directions which make shortcuts from one site to one's destination simple. Simplicity is obviously one of the reasons to explain why the gird-iron pattern is so popular among most road engineers and planners.

Disadvantages of Grid Street Pattern

Since the Second Industrial Revolution, the emergence of motor vehicles put a great strain on cities and suburbs, which widely used grid-iron scheme. On the basis of observation, the rate of traffic accidents is on the rise. The unbroken monotony of the linear road, the most widely used in grid-iron pattern, made drivers speed up easily and contributed to the proliferation of accidents. Aware of these disadvantages, most cities take street hierarchy, curvilinear design, and disconnected networks into consideration (Wolfe 1987). Initially, as Wolfe said, the grid pattern was rejected because of social and economic issues. The increasing population densities brought dangers by the advent of the automobile in the grid neighborhoods. The other problem was the gridded cities usually lacked of public and semipublic parks and open spaces. The monotony of the grid was also ruled out by aesthetic viewpoints. It lacked a sense of natural contours and increased the costs of construction through more civil engineering (Wolfe, 1987).

Fragmented Parallels

Fragmented Parallel, a variant of the regular grid, has been widely used since 1950s. Most crossings turn into T intersections or L corners, which shaped blocks into the irregular mixture of rectangular. This pattern reduces the number of access points, the choices of routes through a neighborhood and the degree of interconnection. Although, this pattern has a similar street length as the grid-iron, it limits the number of blocks and traffic flow (Southworth, 1993). The reduced number of access points also influence accessibility within the block.

Warped Parallels

Based on the planning literature, neighborhoods built after World-War II tend to have a warped parallel than grid (King, 2005). Warped parallel has similarities with fragmented parallel in spatial shape. It features long and narrow blocks, fewer route choices, limited crossing intersections and a degree of interconnection, but is curvilinear rather than straight. The curving street usually creates a rural impression and shortened visual length. The topography was used to generate this pattern. However, in a flat site, the occurrence of a curving street was a response to the spaces which were filled in by occasional cul-de-sacs. The curvilinear street confuses the orientation of users in the block (Southworth, 1993). As a whole, compared with fragmented parallel, it is not an auto-friendly block with the warped parallel pattern.

Loops and Lollipops

The street design changed from warped parallel to loops and lollipops since 1970 (King, 2005). This street pattern consists of many loops and cul-de-sacs. The direction of the street is multiple. The limited connectivity contributes to a sense of privacy and quietude. However, the greater number of short streets, which could bring more relative safety to children, increases traffic congestion on the existing arterials which connect with it. The limitation of access is an even more serious issue than other non-loops patterns. Thus, both automobiles and pedestrians are not interested in passing through communities with this pattern (Southworth, 1993).

Lollipops on a Stick

The lollipops on a stick are also named cul-de-sacs or “dead-end street”. It has the minimum interconnection, route choices and access points (Southworth, 1993). Compared with the grid pattern, it enhances the pressure on traffic, but reduces the occupation of land. Though it cannot handle too much traffic, most communities prefer to maximize privacy by using “lollipops on a stick” layout (Asabere, 1990).

Summary of Street pattern

From the above discussion, Southworth classified street pattern based on their characteristic on interconnection, blocks, access points and their morphologic differences. The research about the effect of street pattern on traffic accidents mostly followed his classification scheme, for example, Ben-Joseph (1995) and Riffat (2009).

Transportation Safety Issues and Street Pattern

The large number of factors relevant to traffic safety can be summarized into three categories: driver behavior, vehicles and road condition. Driver behavior played the dominant role in the traffic accident. More than 90% of traffic accidents were related to road users. In the US study, the vehicle was identified as the sole factor in 2% of crashes, while the interaction between the vehicle and road user was indicated in 6% of the crashes. 7% of the crashes not linked to the road users (Evans, 1991). Most researchers took the non-road user's wrecks into consideration and found main components relevant to traffic safety.

Population Density and Vehicle Miles Traveled

Clark and Cushing (2004) used multiple linear regression to estimate the effect of population density and Vehicle Miles Traveled (VMT) on the rates of traffic fatalities in rural and urban areas. The number of motor vehicle collisions in rural areas is much higher than in urban areas. The variation of mortality rates in rural areas was significantly affected by population density and VMT. However, the variation in urban mortality rates was not affected by population density. The result demonstrated that increased mortality in rural areas could be partly attributed to the increased VMT per capita (Clark and Cushing, 2004). Their model was oversimplified because it only has two main risk factors. So Dumbaugh and Rae "revisited the relationship between community design and traffic safety". They added more risk factors such as median household income, population age and land use. They observed that the reductions in VMT contributed to the decreased crash incidence. However, they found that increased population density was associated with significantly fewer wrecks (Dumbaugh and Rae, 2009). The same result was obtained by Ewing, Scheiber and Zegeer. They explained that people living in denser and more compact blocks drive less (Ewing, Schieber and Zegeer, 2003).

Speed Limit

Speed limit as one of the risk factors for traffic accidents; many researchers used different research designs to explain the relationship between it and crash incidence. Gallaher et al. (1989) compared the rate of fatalities 5 years before raising limits with actual rate 1 year after on rural interstates. They found that the rate of fatalities was more than twice than before. At the same year, NHTSA (1989) claimed the fatalities on 65 mph roads were much higher than those on the roads below 65 mph. The same result was gained by Wagenaar et al. (1990), McCarthy (1991), Lave and Elias (1994) and Steven (1995). In contrast with the above studies on interstate roads or highways, Dumbaugh and Rae (2009) focused on the effects of the mean speed limits on the traffic safety of a community. They claimed that a safe community had a positive correlation with lower speed.

Median Household Income and Household Density

Median household income and household density as demographic and socioeconomic characteristics were discussed by Ewing et al. (2003). They claimed that the lower density of household area and larger blocks had a higher crash incidence than denser communities (Ewing, Schieber & Zegeer, 2003).

Based on Rifaat's analysis, the high proportion of low-income people in a community would increase the number of wrecks. The author suspected that economically advantaged communities may have a better road infrastructure as well as better vehicles. Besides, lower-income persons easily became the vulnerable road users. They usually have less education and pay less attention to road safety (Riffat, 2011). The same correlation was concluded by Baker et al. (1992) and Graham (2005).

Land Use

Dumbaugh et al. (2009) discussed the relationship between plan land use and traffic safety. When the residential, commercial and retail uses were clustered together, the total wreck rate would increase by 1.3%. Because the high density of population and large number of pedestrians usually clustered in such areas, there were increasing risks between people and automobiles. Riffat (2009) found that residential and commercial lands have significantly increased wreck risk. And more retail uses and commercial zones are often considered to have a high economic condition, which was positively associated with greater frequency of collisions (Graham, 2005).

Street Pattern

Marks (1957), as the originator of the street pattern analysis, divided street layout into two general categories: "limited access" and "grid-iron". There were 86 plots in his study area. He studied five-year crash data to compare the total accident rate in each type of street pattern. In order to minimize the error, the bordering of the plots, which were the same level as the major streets, was not considered in the analysis. Finally, Marks (1957) found that the total crashes for gridiron pattern plots were almost eight times of the crashes for the same area of limited-access plots.

Ben-Joseph (1995) improved the research of Marks, and studied the effect of street pattern on the frequency of traffic accidents. He used three types of street pattern: grids, loops and cul-de-sacs. Compared to Marks' study, he optimized the crash data by crash rates and added other relevant factors such as street length, household density and average daily traffic. He claimed that grid communities have more accidents than loops and Cul-de-sac communities. Furthermore, the grid iron also had the most intersection accidents, 62% of the total intersection collisions. However, he ignored other risk factors, such as population density, income, land use and speed limit, which are equally important to consideration.

In above literature analysis, the conclusions in previous study about effect of the each factor on frequency of crashes are mostly the same as the hypotheses mentioned before. However, some researches indicated that a change in study area or circumstances would cause significant changes in the relationship between population density, VMT, household density and crash incidence. So this research will focus on the factors such as street pattern, population density, VMT, household density, speed limit, median income and land use to test the correlation between all predictors and crashes incidence in specific study area.

Regression Models of the Crash Incidence

In the past, multiple linear regression models were widely used in many traffic accident studies (Miaou and Lum, 1993). However, when the road section was taken as a unit to report the occurrences of crashes, researchers had to deal with a greater number of zero incidences during the research. Therefore multiple linear regression models were not deemed reasonable to describe the dispersed data on road sections (Miaou and Lum, 1993).

To improve the linear regression models, the Poisson regression model was found to analyze count data such as the number of crashes. Miaou and Lum (1993) designed two linear regression models and two Poisson regression models to discuss the relationship between traffic crashes and highway geometric design. The results demonstrated that Poisson regression models obtain more desirable statistical properties than linear regression models (Miaou & Lum, 1993). However, since the Poisson regression model is non-linear, it is harder to explain the quantitative effects of the independent factors than multiple linear regression models (Rifaat, 2011). Miaou et al. (1992) used the Poisson regression model to explain the relationship between truck crashes and highway geometrics. At the end of research, he mentioned that a limitation exists in the Poisson model. It required that the mean and the variance of the dependent variable should be equal. However, in practice, the variance of dependent variable exceeds the mean, and the data would be over dispersed (Radwan, 2000). As Miaou et al. (1992) have stated, the problem caused by over dispersion would not change the conclusion about the relationship between contributing factors and truck accidents.

In order to overcome the problem of over dispersion in the Poisson regression model, Miaou (1994) did the same research using negative binomial regression and maximum likelihood to find the proper model. The value of the likelihood illustrated that the Poisson regression model should be used to establish the relationship between highway geometric and collisions. The Negative binomial regression can be explored when the data was over dispersed (Miaou, 1994).

Based on the previous studies, although the multiple linear regression models could not in probabilistic statements interpret traffic wrecks on the road, it still could describe the general correlation between contributing factors and random traffic wreck events. Thus, that was an overall effective way to explore the research in GIS and obtain a visual correlation between contributing factors and dependent variables in spatial perspective.

Figure 2-1. Classification of street pattern (Source*: Streets and the Shaping of Towns and Cities.* Southworth and Ben-Joseph, 2003. Edited by Author)

CHAPTER 3
DATA AND METHODOLOGY

This thesis uses traffic crash data, census data, geographical information, Poisson Regression model and GIS analysis to explain the relationship between different street patterns and the frequency of road accidents for improving traffic safety.

Description of Study Area

Gainesville (Figure 3-3), the seat of Alachua County, Florida has an area of 127.2 km^2. It is chosen as a case study due to its increasing car accidents yearly, which constitutes more than 50% of the total crashes in the county (Alachua County Crashes, 2011). Figure 3-1 shows the trends of traffic accident for City of Gainesville between 2000 and 2009.

Moreover, the city is experiencing rapid growth of population and urban sprawl. The University established itself in this city in 1905. More and more students come from various other places and significantly impact this relatively small city. Citizens were eager to enlarge the city limits so as to encompass sufficient resources to provide for a growing city. Since 1900, the city population has grown 30 times over; the land resources would reflect a scarcity in city limits in the future. Decision makers encourage suburban style student apartments to sprawl around it. It is necessary to consider traffic safety by building adequate road networks in new neighborhoods to reduce the frequency of traffic. Furthermore, I focused on the Gainesville because of the available data and typical street types. Based on the literature review of street patterns, I used the classification of street pattern by Southworth and Ben-Joseph (2003) for reference. The street pattern consists of "grid-iron, fragmented parallel, wrapped parallel, loops and lollipops and lollipops on a stick" (Rifaat, 2010). However, while discriminating among these patterns, only typical street layouts can be easy to recognize. The difference between two types of parallel and lollipops will be easily confused.

After observing the street network map of Gainesville in each block group, it is easy to discern “gridiron”, “wrapped parallel” and “loops and lollipops”. The type of “fragmented parallel” usually mixes with “gridiron”. And the “lollipops on a stick” always coexists with “loops and lollipops”. In this case, I classify the street pattern into three main categories, shown in Figure 3-2. The “fragmented parallel” pattern has been merged with “gridiron” pattern. It is defined as a grid. At the same time, both patterns with lollipops as the loop pattern will be used. The “Wrapped Parallel” will be considered a simplex category named parallel.

For the unit of the research objects, it will be defined by the block group according to the data in U.S. census. The block group boundary is mostly extracted from the major roads in which crash data is clustered. When the number of cashes was calculated by block group, some crashes on the boundary would be recalculated by adjacent two block groups. Actually, if these roads were contained within the block groups, the result would be more significant in the statistical model. Alternatively, another better method can be used to improve the representativeness of data. The entire city limits can be divided into several unit areas by drawing grids. For example, we can utilize a square of 400 acres as unit grid. It will avoid the boundary of grid sharing the same line with the street. However, during the analysis, the social economic and demographic data such as population, households and total income contained by census data and Florida Department of Revenue are needed; the block groups provided by the census are the best unit for this research so far. So the study area boundary is not equal to the boundary of city limit.

In this research, all the GIS maps come from Florida Geographic Data Library (FGDL). According to the street map and census layer of study area in GIS, the sample size of block groups is 72. Of those areas, 26 are classified as grid, 24 are parallel, and 22 are loops.

Description of Data

In this research, the crash data was gathered from the governmental database maintained by city of Gainesville, Department of Public Works. The crash records contained all the traffic accidents such as motor-motor and motor-pedestrian from 2005 to 2007, and are used in this study. Note that not all the crash records existing in the database will be used in this study. Due to the fact that Interstate 75 highway is not at the same level of hierarchy compared with other streets through the city limits, traffic accidents occurring along it are rejected from the study as invalid data. The other valid data will be processed by the GIS selection tool. Because the boundary of the block group share the same line with the street, some crash records will be dismissed, especially using the spatial selection method of “Target layer features, which are completely within the Source layer feature”. Nevertheless, the expected result is only concerned with the relationship between the street pattern and frequency of crash; so the error is still acceptable. Besides, because of different areas in each block group, the error will be generated by calculating the total number of wrecks. In order to minimize these effects, the density of wrecks will be used as the dependent variable.

The information for each block group is obtained from the 2010 census data. Note that only 2000 and 2010 census data are available, so the year 2010 has been chosen because its proximity to the 2005-2007 crash data. According to the literature review, there are many factors affecting the occurrences of crashes. However, due to the limitation of data resources, only population density, household, total income, number of entrances for each block group, land use and Vehicles Mile Travel (VMT) regarded as covariate will be used.

Traffic Crash Regression Model

Poisson and Negative Binomial Regression Model

In this research, the number of occurrences of crashes in each block group is count data regarded as the dependent variable. The regression models which can provide appropriate analysis for count data are Poisson regression and Negative binomial regression model. Poisson regression is an ideal model for modeling count data. However, it has two assumptions:

The occurrences of crashes should be distributed as Poisson distribution, which is discrete probability distribution.

The value of occurrences of crashes must be non-negative integers.

Therefore, the crash data has to be optimized before using Poisson. Varied as the areas of the block groups are, the frequencies of crashes density are defined as using total crash number in each unit divided by the area of the block group, often named as crashes per 1 km^2. Based on the second requirement of the Poisson distribution, only the integrate part of outcomes are remained. For example, if the number of crash density was more than 1.6 in a block group, the integral number will be approximated to 2. If the number of crash density was less than 3.5 in a block group, the integral number will be approximated to 3.

Negative binomial regression is the variant of Poisson regression. Theoretically, it should be more accurate in over-dispersed count data (Hilbe, 2007). Compared with the Poisson regression, it has an additional parameter to express over-dispersion (Coxe, West, and Aiken, 2009). However, the count variable does not violate the assumption of Poisson regression so that there is, of course, no direct confirmation to guarantee that the negative binomial regression can provide more optimal results. Therefore, one of them will be used as the better model. These two models will be evaluated by the value of log likelihood and deviance. “Large values of the log likelihood statistic indicate poorly fitting statistical models, because the larger the value of the log likelihood, the more unexplained observations there are” (Gigliotti, 2007). And the better model also has a large value of deviance. Another way to evaluate the model is to check the omnibus test. The statistically significant test means that most variables present a significant effect on the dependent variable (D’Agostino, 1971).

Ordinary Least Squares Regression (OLS) and Geographical Weighted Regression (GWR)

In order to show the relationship between the occurrences of crashes and street pattern visually, I will put all the variables into GIS software to generate the spatial distribution map.

Though the Ordinary Least Squares Regression (OLS) is the lack of capacity to analyze the dispersion, it still can re-check the correlation between count data and some independent variables, especially the variable for which the Poisson regression coefficients are not significant. The coefficients in Poisson regression demonstrate the correlation between the independent variables and exponential of dependent variables. The result sometimes is hard to explain. If the correlation of the two research objects is not extremely significant, the coefficient in exponential Poisson model will be easily approaching to zero. As long as the model has a strong explanatory power for regressors, it is also reasonable to explain the correlation between the occurrences of crashes and other explanatory variables that cannot be interpreted well through the Poisson regression model. The R square and Akaike's Information Criterion (AIC) will be used as the rule of selecting appropriate models.

If the OLS shows that the independent variables have a consistent relationship to the occurrences of crashes in data space, and the variations in spatial processes are non-stationarity, this model is good for Geographical Weighted Regression (GWR) analysis. To be noticed, there is a hypothesis to explain why the GWR model still needs to be used if the OLS is sufficient enough. The hypothesis is thus: due to the fact that the occurrences of crashes are affected by many factors, is there any possibility that the independent variable has a positive influence on crashes density in some places but a negative influence in others? Because the OLS is only a global model for variables, we still need the local model to check the direction of the influence of the variables. In additional, by using GWR analysis, I can also explain why the correlation is more significant in some block groups via a geographical map. Even if the two models have the same direction for the influence of those variables, GWR will also show that local coefficients of influence for each variable vary considerably over Gainesville (Charlton, 2005).

Variables in the Regression Model

In the regression model, the different types of street patterns and land use as independent variables are represented as the dummy variable (Table 3-1). For land use variable, the total block groups are classified into two categories. "1" stands for residential and commercial types, and “0” stands for all the other industrial, governmental and agricultural types. Part of the reason is the limitation of samples. There are not too many block groups with industrial, governmental and agricultural land use. Also the common among these types of land use is that they usually have less traffic volume and population density. For the residential and commercial, the population and traffic will be clustered in those areas. So the land use was classified based on the high density of population and traffic. To estimate the effects of street pattern on traffic accident, the grid pattern is used as reference. The correlation coefficients of the other two patterns are explained as comparative to the grid.

All the regression models which analyze the occurrences of crash density from 2005 to 2007 depend on 8 independent variables, which are listed in Table 3-2. From the outcome of each model, the effects of each explanatory variable, especially different street patterns on the collision, will be found.

According to the theoretical discussion in the literature review, due to the grid pattern consists of the linear road which is made drivers speed up easily and contributed to the proliferation of accidents. So my hypothesis is that the grid pattern should have positive correlation with crash incidence.

The hypotheses of other factors as followed:

From the observation of Clark and Cushing (2004), higher population density, and vehicle miles traveled (VMT) should increase the possibilities of occurrences of traffic crashes in this thesis. According to the conclusion which was summarized by Ewing et al. (2003), household density should have the same correlated result as the population density and VMT variables in this study area.

Based on the researches from previous study, most scholars indicated safe community should have a positive correlation with lower speed. This phenomenon is also expected by my research.

The hypothesis about the high proportion of low-income people in a community would increase the number of crashes is considered to be reasonable. Low-income people usually have less chance to take education about transportation safety and do not take much serious about the potential risk when their children are walking or playing within the community.

The high traffic volume area such as commercial and retail used block and high population density area such as residential land use block should be positively associated with greater frequency of collisions. This hypothesis is also discussed by Dumbaugh et al.(2009) and Graham (2005).

Table 3-1. Street pattern variable as dummy variable

Street type	Value	Value	Value
Grid	0	0	0
Parallel	1	0	0
Loops	0	1	0

Source: Edited by author.

Table 3-2. List of variables in regression model

Variable name	Data attribute
Three-year Crash Density (per square kilometers)	Interval (count data)
Street Pattern	Ordinal
Grid	(dummy data)
Parallel	(dummy data)
Loop	(dummy data)
Population density	Interval
Total Income density (per square kilometers)	Interval
Household density (per square kilometers)	Interval
Number of entrances density in a block group (per square kilometers)	Interval
VMT density (per square kilometers)	Interval
Mean speed limit	Interval
Land use	Ordinal
Residential and commercial	(dummy data)
Industrial, agricultural and governmental	(dummy data)

Source: Edited by author.

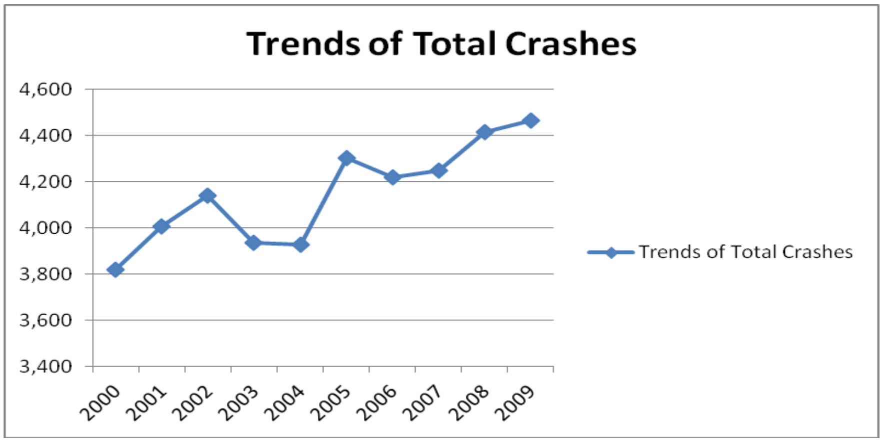

Figure 3-1. Trends of total crashes in Gainesville, FL. (Source: Gainesville Police Department. Edit by author)

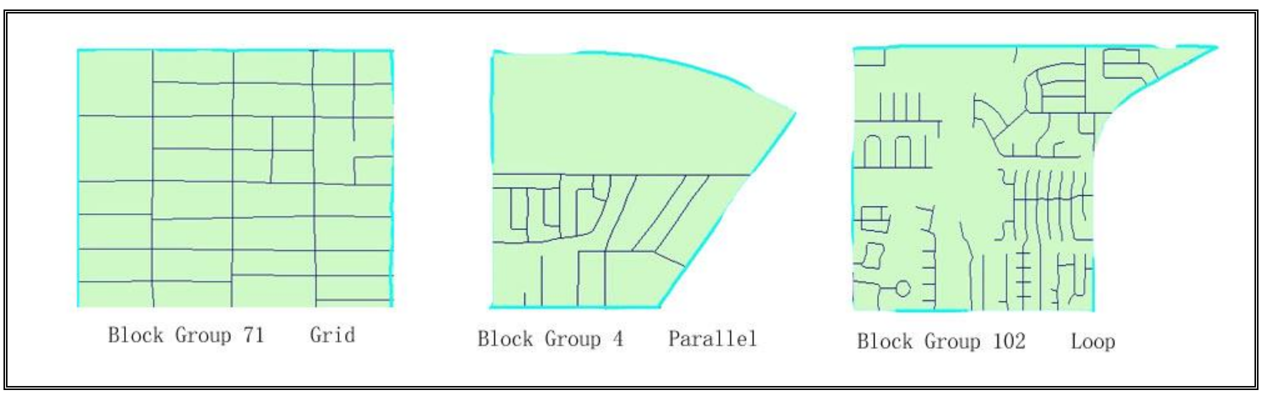

Figure 3-2. Examples of street pattern in Gainesville. Edited by author.

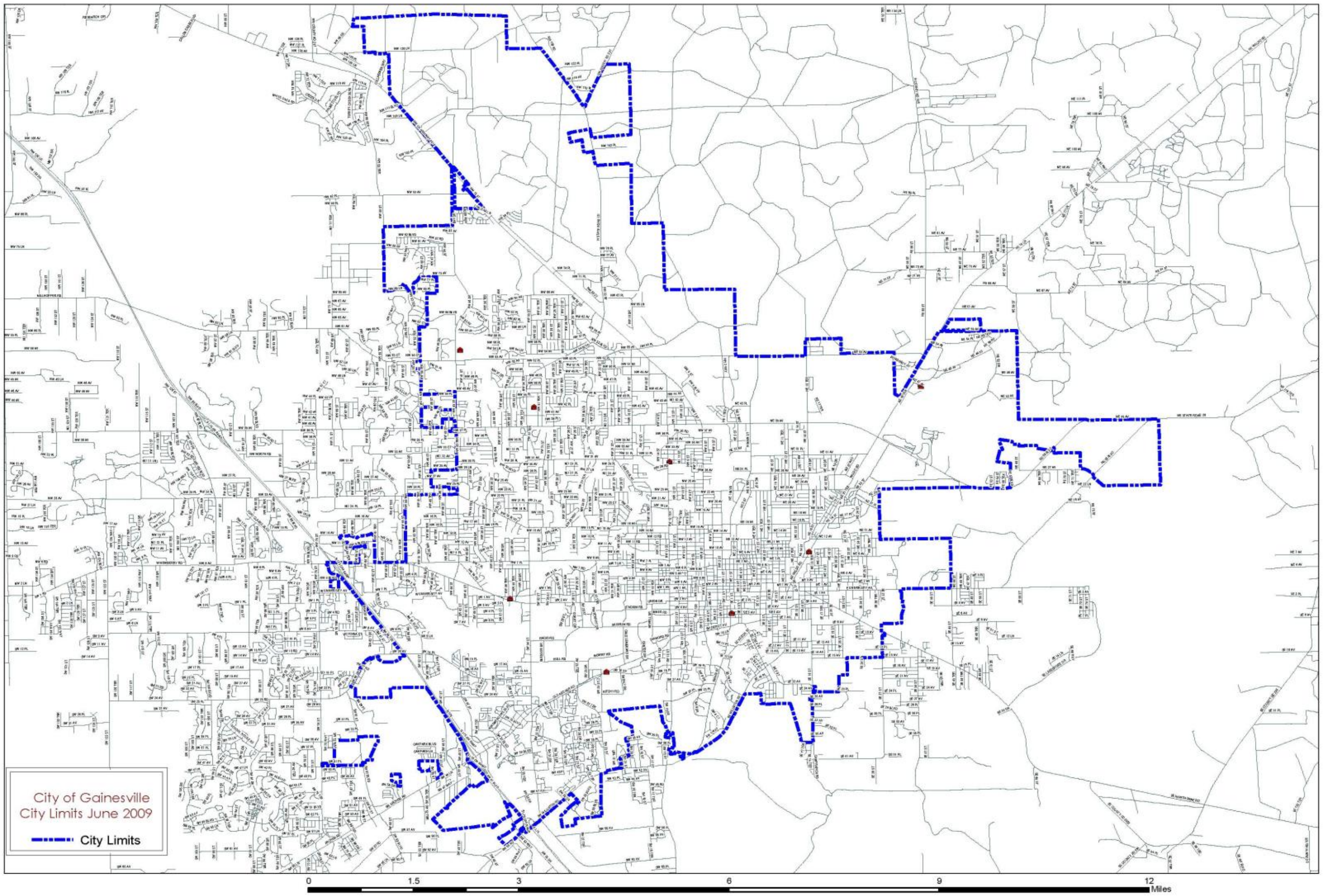

Figure 3-3. City of Gainesville -- City limits. (Source: City of Gainesville)

CHAPTER 4
ANALYSIS AND RESULTS DISCUSSIONS

There are three parts in this chapter. Firstly, the study compares the explanatory power in the Poisson regression model with those in the Negative binomial regression. The more powerful the model's explanatory value, the more accurately the model interprets the effects of the street pattern on the occurrences of crashes. Secondly, this part provides explanations for the results of the relative optimal model. Moreover, the third part will represent the visual correlation between the occurrences of crashes density and other explanatory variables in Gainesville via the Ordinary Least Squares (OLS) and Geographical Weighted Regression (GWR) model in GIS.

The Relative Optimal Regression Model

In order to find a more accurate statistical method from Generalized Linear model's family to measure the effects of eight independent variables on the frequency of crashes, the first step is to test which independent variables are statistically significant to the occurrences of crashes. The second step is to model the correlated variables in both Poisson and Negative binomial regression. At last, based on the outcome from these two models, the relative optimal regression model to measure the frequency of crashes will be interpreted.

Correlation Coefficient Analysis

As a non-parametric correlation, the Spearman Correlation Coefficient can reflect the magnitude and direction of correlation between two variables. It is suitable for the data, which is not assumed to be within normal distribution. All the variables in this study are not assumed to be normal. The correlation coefficient will show if each independent variable is statistically significant to the density of crashes (Table 4-1). If the p-value is less than 0.05, there is a relationship between independent variables and density of crash.

From the above table, we can see that all the independent variables are statistically significant to the three-year crash density. For example, though the significance of the coefficients of parallel street pattern is 0.049, it is still less than 0.05, which means that we can still discern the relationship between parallel and three-year crashes density. The correlation coefficient of parallel, loop and non-residential land use variables illustrates that as the value of one variable increases, the value of crash density decreases. In addition, the coefficient also explains the magnitude of each independent variable; population density, total income density and number of entrances density have a strong effect on crash density. However, “whether a correlation of a given magnitude is substantively or practically significant depends greatly on the phenomenon being studied” (Voelker, Orton and Adams, 2011). So the regression models are used as a better approach to predict the variation of dependent variable.

Poisson regression model compares with Negative binomial regression model

Because all the independent variables are statistical significant to the crash density, they can be measured as factors or covariates in both models. The output of goodness of fit and the omnibus test (Table 4-2) from these two models reveals the relative optimal model with smaller value of Log Likelihood and the larger value of deviance. The Poisson model has a significantly greater value of deviance than the Negative model. Besides, the log likelihood in Poisson is significantly smaller than it is in the Negative. Though both models have the statistical significance of the Omnibus, the Poisson Regression is selected as the relative optimal model to explore the correlation between independent variables and crash density.

Occurrences of Crash in Poisson Regression Model

As mentioned in Chapter 3, all the variables will be run in the Poisson model. Therefore the model consists of one intercept and 9 independent variables. We still want to know how the street pattern will affect the variance of other variables. The models are:

- Log (Three-Year Crash Density) = Intercept + β_1 (Parallel Dummy) + β_2 (Loop Dummy) + β_3 (Population Density) + β_4 (Total Income Density) β_5 (Household Density) + β_6 (Number of Entrances Density in a Block Group) + β_7 (VMT Density) + β_8 (Mean Speed Limits) +β_9 (Land Use Dummy)

Table 4-3 shows the estimated Poisson regression coefficients for the two models. As demonstrated in the first model, the coefficients of population density, household density and VMT density are zero, which indicate that there is no relationship among these three variables and log of expected crash density in this model. Besides, the total income density is not statistically significant due to the value of "Sig." being more than 0.05. And the parallel and loop have a significantly negative effect on the crash density, which is compared with the grid pattern. The coefficients less than zero will indicate that a city with more parallel and loop pattern instead of grid has a lower crash density. In addition, the crash density in the city can, to some extent, be increased by increasing one unit of residential and commercial land use and number of entrances density in a block group. So the final Poisson regression model is:

- Log (Three-Year Crash Density) = 3.679 – 0.627 (Parallel Dummy) – 0.589 (Loop Dummy) + 0.009 (Number of Entrances Density in a Block Group) – 0.004 (Mean Speed Limits) + 0.556 (Land Use Dummy)
- Log (Three-Year Crash Density) = 3.679 + 0.024 (Number of Entrances Density in a Block Group) – 0.005 (Mean Speed Limits) + 0.649 (Land Use Dummy) – 0.001 (Household Density)

The intercept of 3.679 is estimated when other variables are equal to zero. For the grid, when the variable parallel and loop are evaluated at zero with zero numbers of entrances density, mean speed limit and land use, the log of the expected crash density is 3.679. For the parallel, comparing with the grid, when other variables are constant, the difference in the log of expected crash density is expected to be 0.627 units less. The coefficient of residential and commercial land use is 0.556. It illustrates that if other variables hold a constant, we would expect other types of land use with the less crash density than the residential and commercial area. Compared with the second model, all the statistical significant variables have more or less changed without restriction of street pattern variable.

From the result in Table 4-3, the correlation between the control variables and dependent variable is partly consistent with my hypothesis. For example, compared with grid pattern, parallel and loop are significant negative influence the crash incidence. More types of non-grid pattern should contribute to the less frequency of crashes. Also the lower median income would slightly affect crash density, which might cause directly increasing number of traffic accidents. Besides, increasing number of entrances and residential and commercial clustered communities would lead to the highly occurrences of crashes.

From the first models, we can fathom the significant effect of different street patterns on a log of expected crash density. However, in order to show the variance on map, the OLS model in ArcGIS will be used instead of Poisson regression method in GIS.

Ordinary Least Squares (OLS) and Geographical Weighted Regression (GWR) Model

According to the above model, the population density, household density and VMT density didn't generate any effect on the log of expected crash density. However, it was no mean that there wasn't any correlation between them and crash density. The OLS as the global regression model will mainly explain the effect of these variables and street pattern on crash density. The model is:

- Crash Density = Constant + β_1 (Parallel) + β_2 (Loop) + β_3 (Population Density) + β_4 (Household Density) + β_5 (VMT Density) + (α i)

(α i, which is optional, can be multiplied by any of other variables with their coefficient β)

The summary of the diagnostics for different models is shown in Table 4-4. The R Square of the original model is 0.835, which means this model can interpret approximately 83.5% of the variation in the dependent variable. All five models have a closed R Square value and the Joint F-statistic p-value of all models smaller than 0.05 indicates a statistically significant model. As for the Akaike's Information Criterion (AIC), lower values indicate a better fit (Kenny, 2003) and so the Mean Speed Limits model with the lowest AIC is the best-fitting model.

If the OLS shows that the independent variables have a consistent relationship to the occurrences of crashes in data space, and the variations in spatial processes are non-stationarity, this model is acceptable for GWR analysis. So the Koenker (BP) Statistic P-value should be less than 0.05; the variables in this model will likely be useful in the GWR model. The land use and number of entrances variables will be removed from the list of useful variables in the GWR model. Taking into account both the high value of R Square and the low value of AIC, the basic model with a mean speed limit will be selected in the GWR. Besides, as the VIF of all the variables in this model is less than 7.5, there is no explanatory variable redundancy. And the only mean speed limit is not statistically significant. However, the P value at 0.053 is still significant of 10%. It still can explain the positive correlation between speed and crash incidence. The summary of OLS result is shown in Table 4-5. So the OLS model is:

- Crash Density = 347.392 – 24.41 (Parallel) – 19.79 (Loop) + 0.12 (Population Density) – 0.49 (Household Density) + 0.28 (VMT Density)

In order to illustrate that the OLS results can be trusted and avoid the statistically significant clustering of residuals, the regression residuals are run in spatial autocorrelation to ensure that they are spatially random. The result shown in Figure 4-1 indicates that the residuals are spatially random. Because the p-value is more than 0.05, we reject the hypothesis.

From the model above, compared with the grid pattern, if increasing the number of parallel and loop pattern block, the crash density will decrease. And the parallel has a more significant effect on decreasing crash density. The population density and VMT density variables have a positive effect on crash density.

From the above Table 4-4, the Koenker test p-value of this model is 0.43 (<0.05), which indicates that the model is non-stationary, which is appropriate for GWR analysis. In data space, the result of the OLS model shows that the relationship between crash density and every independent variable is stable. However, in geographic space, the changes of magnitudes will be presented in the study area by using the GWR. So the GWR is suitable for predicting the effect of each explanatory variable except for mean speed limits on the locations with grater variation of crash density.

In GWR model, there is a high R square at 0.841, which could explain about 84.1% of the crash density phenomenon. By comparing the coefficient of all the variables with those in the OLS, the GWR has the same direction of the influence of these variables as the OLS. In geographic space, the street pattern and other variables have a changeable correlation to the crash density in different places. We can see that from Figure 4-2 to Figure 4-6. These figures show the variation in the coefficient estimates for two types of street patterns, population density, household density and VMT density, mapped in Gainesville City Limits. For example, Figure 4-2 is the predicting map, which describes the variation of the coefficient for the loop type. The predicted value for OLS is -24.42. The map for the local coefficients of the loop indicates the impact of the loop on crash density over Gainesville, with a reinforced relationship from northwest to southeast area. The range of the GWR coefficient is from -29.23 in the southeastern block groups to -54.05 in the northwestern block groups. So the changes of loop pattern in the northwest of Gainesville will be highly significant for decreasing crash density. The same explanation can be applied in parallel street type (Figur 4-3) and household density (Figure 4-6). The significant effect of parallel on city crash density is aggregated mostly in the eastern area. As to household density, the variation influences the crash density in the western parts of the city. Population and VMT density has a positive influence on the crash density. Then, if we increase one unit of population density, the range from 0.123 to 0.129 units of crash density will also be increased in the northern part of the city. The VMT density pays much attention to the southwest corner of the map. The higher vehicle density will lead more crashes in this region.

Compared with my hypotheses mentioned before, the results from GWR model are consistent with them in some extent. From the map of GWR model, the high frequency of the crash in grid pattern communities has been convinced by the comparison with parallel and loop pattern communities. If we control the other variables and increase the parallel and loop in southeastern of Gainesville, the occurrence of crashes will be decreased. It is probably because of both the block groups of grid pattern and crash densities are clustered in that part of study area via observation of the map. If explore one unit of parallel will obtain significant decreasing correlation with crashes.

Moreover, the result from population density and VMT density GWR map successfully test the hypotheses in the previous chapter, the larger value of variables contributed to the higher degree of accident risk by controlling other variables. The larger value of coefficient in the northern part because of the small number of population density clustered with relative high ratio of crash density. So changing a little bit scale of population density will lead to the high frequency of crash.

The reason for large value of VMT density coefficient clustered in western region probably due to the wide lane of intercity road service for out of the town and significant correlated a plenty of crashes.

For the negative household density coefficient, according to the previous literature, if the city or neighborhoods are walkable or pedestrian friendly, the high density of household would like to choose walking instead of driving. So even increasing the number of household, it can still improve the community safety. Or there are other social economic reasons to explain this phenomenon, due to the limitation of this research, the predicted map only show the negative correlation between household density and crash density.

Table 4-1. Correlations between density of crash and independent variables

Independent Variables	Correlation coefficient	Sig. (2-tailed)
Street pattern (Dummy)		
Grid	.477	.000
Parallel	-.233	.049
Loop	-.258	.028
Population density	.595**	.000
Total Income density (per square kilometers)	.596**	.000
Household density (per square kilometers)	.295	.012
Number of entrances density in a block group (per square kilometers)	.574**	.000
VMT density (per square kilometers)	.549**	.000
Mean speed limit	.532**	.000
Land use (Dummy)		
Residential and commercial ("1")	.493**	.000
Industrial, agricultural and governmental ("0")	-.439**	.000

**. Correlation is significant at the 0.01 level (2-tailed). Source: Generated by SPSS, edited by author

Table 4-2. Goodness of Fit from two models

Criteria	Poisson Regression model	Negative Binomial Regression model
Deviance	4243.159	52.033
Log Likelihood	-2336.246	-396.007
Omnibus Test (Sig.)	.000	.000

Source: Generated by SPSS, edited by author.

Table 4-3. Coefficient of variables

Independent variables	Coefficient (Model 1)	Sig. (Model 1)	Coefficient (Model 2)	Sig. (Model 2)
Intercept	3.679	.000	4.468	.000
Street pattern (Grid as reference)				
Parallel	-.627	.000		
Loop	-.589	.000		
Population density	.000	.000	.000	.000
Total Income density (per square kilometers)	-3.616E-5	.450	.000	.000
Household density (per square kilometers)	.000	.000	-.001	.000
Number of entrances density in a block group (per square kilometers)	.009	.000	.024	.000
VMT density (per square kilometers)	.000	.000	.000	.000
Mean speed limits	-.004	.000	-.005	.000
Land use (Industrial, agricultural and governmental as reference)				
Residential and commercial	.556	.000	.649	.000

Source: Generated by SPSS, edited by author.

Table 4-4. Summary of Ordinary Least Squares Regression (OLS) diagnostics

Model	R Square	Akaike's Information Criterion (AICc)	Koenker (BP) Statistic P-Value	Joint F-Statistic P-Value
α (0)	.835	879.991	.036	.000
α (Mean Speed Limits)	.850	875.860	.043	.000
α (Land Use)	.840	880.442	.056	.000
α (Total Income Density)	.841	879.772	.019	.000
α (Number of Entrances Density in a Block Group)	.838	881.147	.132	.000

Source: Generated by ArcMap, edited by author.

Table 4-5. Summary of OLS result

Variable	Coefficient	P-Value	VIF
Intercept	347.392	.013	---
Parallel	-24.42	.035	1.73
Loop	-19.79	.014	1.47
Population density	.12	.000	1.23
VMT density	.28	.002	3.54
Household density	-.49	.003	3.28
Mean speed limits	11.63	.053	1.34

Source: Generated by ArcMap, Edited by author.

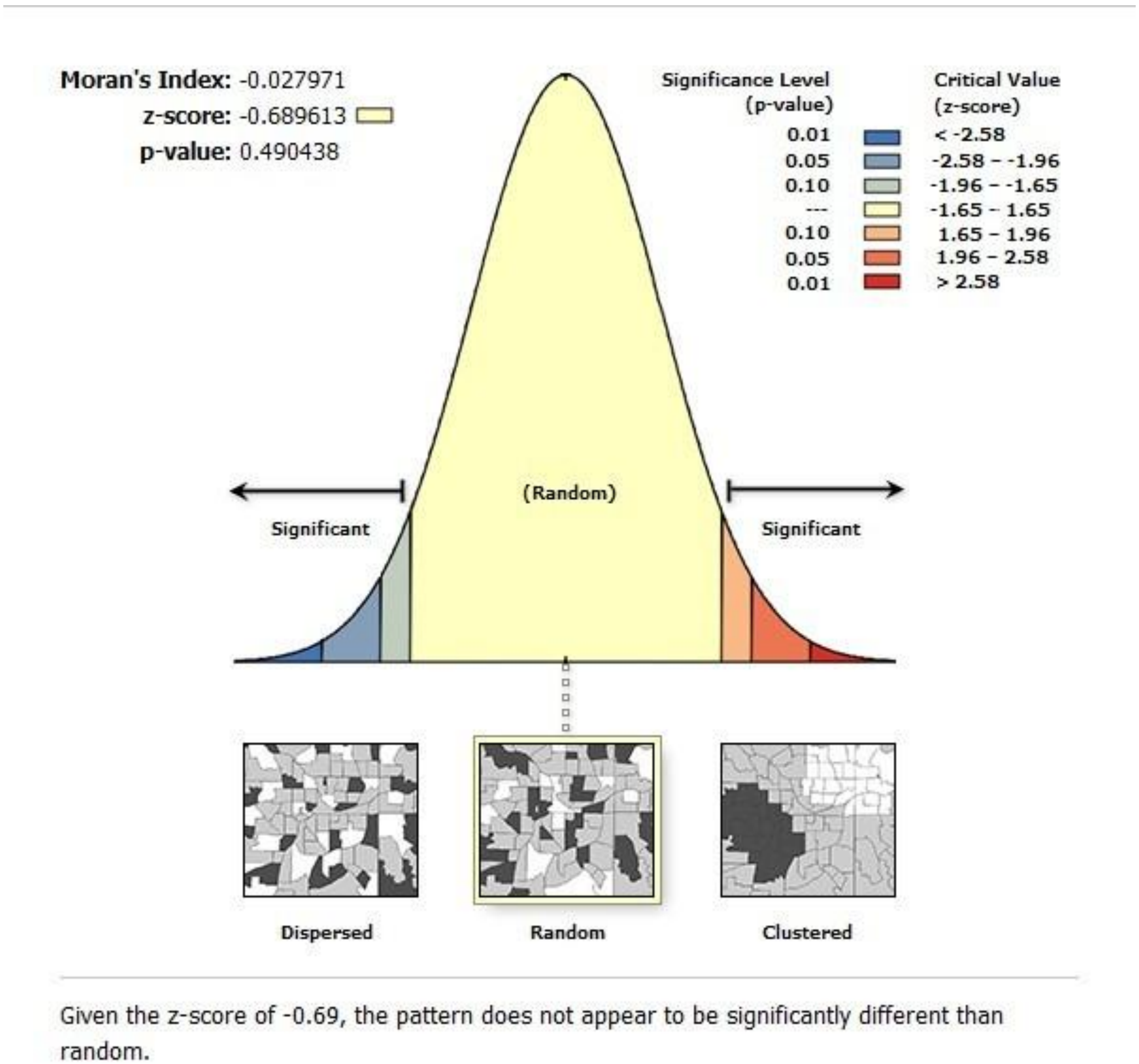

Figure 4-1. Spatial Autocorrelation (Morans I) report of mean speed limits from Ordinary Least Squares Regression (OLS) model. (Source: Generated by ArcMap, Spatial Autocorrelation Tool)

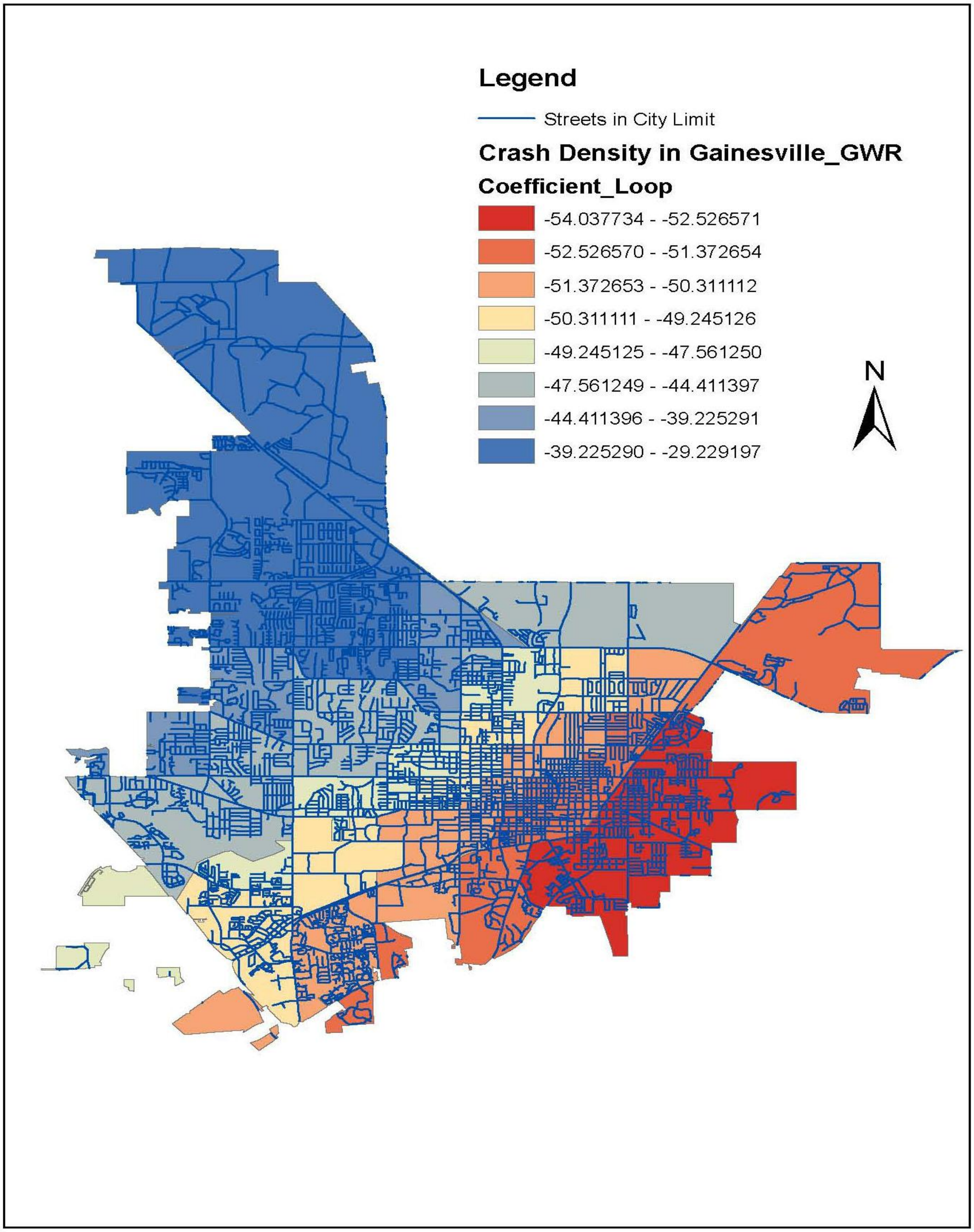

Figure 4-2. Estimated parameter of loop in Geographical Weighted Regression (GWR) model. (Source: Generated by ArcMap, Geographically Weighted Regression Tool, edited by author)

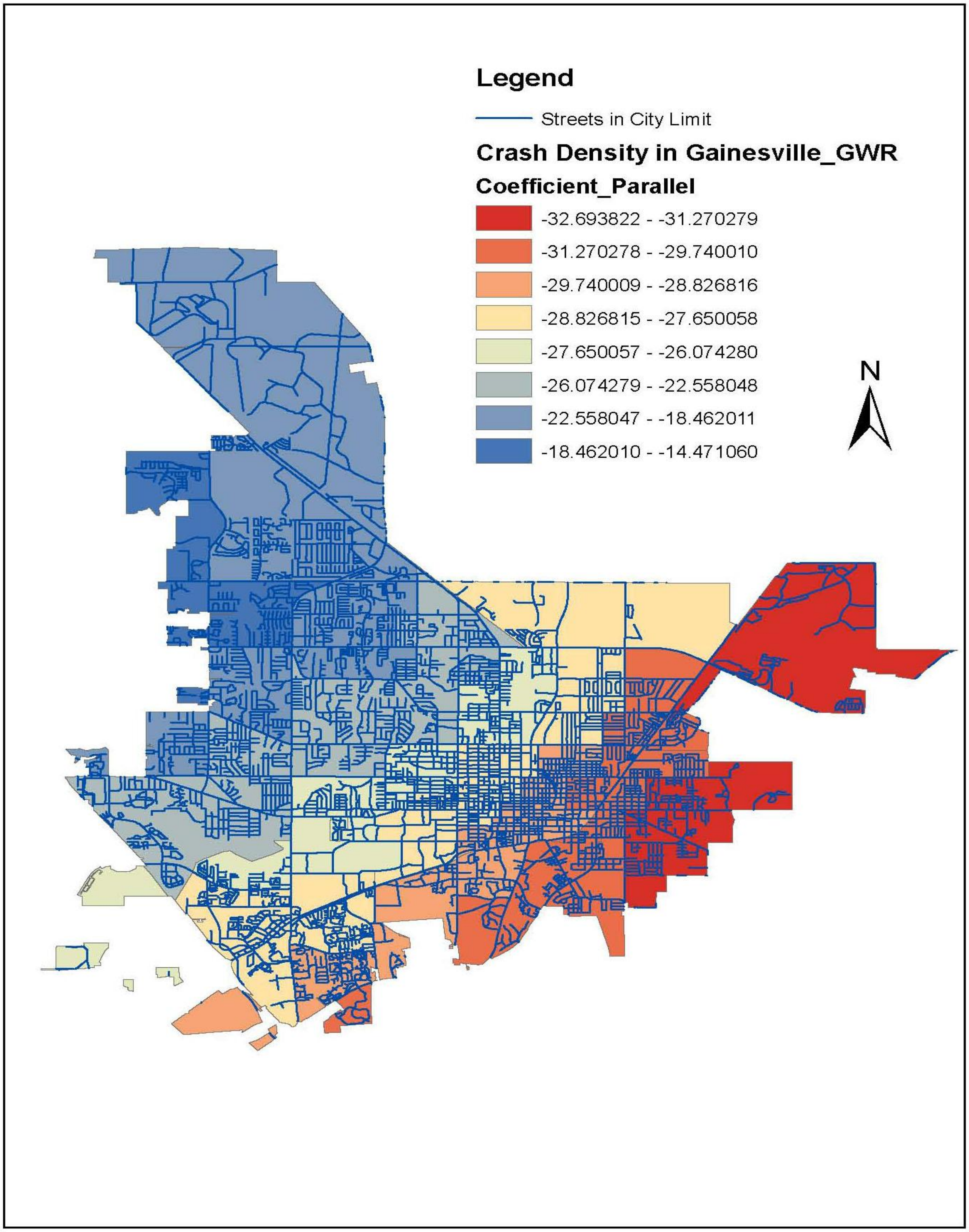

Figure 4-3. Estimated parameter of parallel in GWR model. (Source: Generated by ArcMap, Geographically Weighted Regression Tool, edited by author)

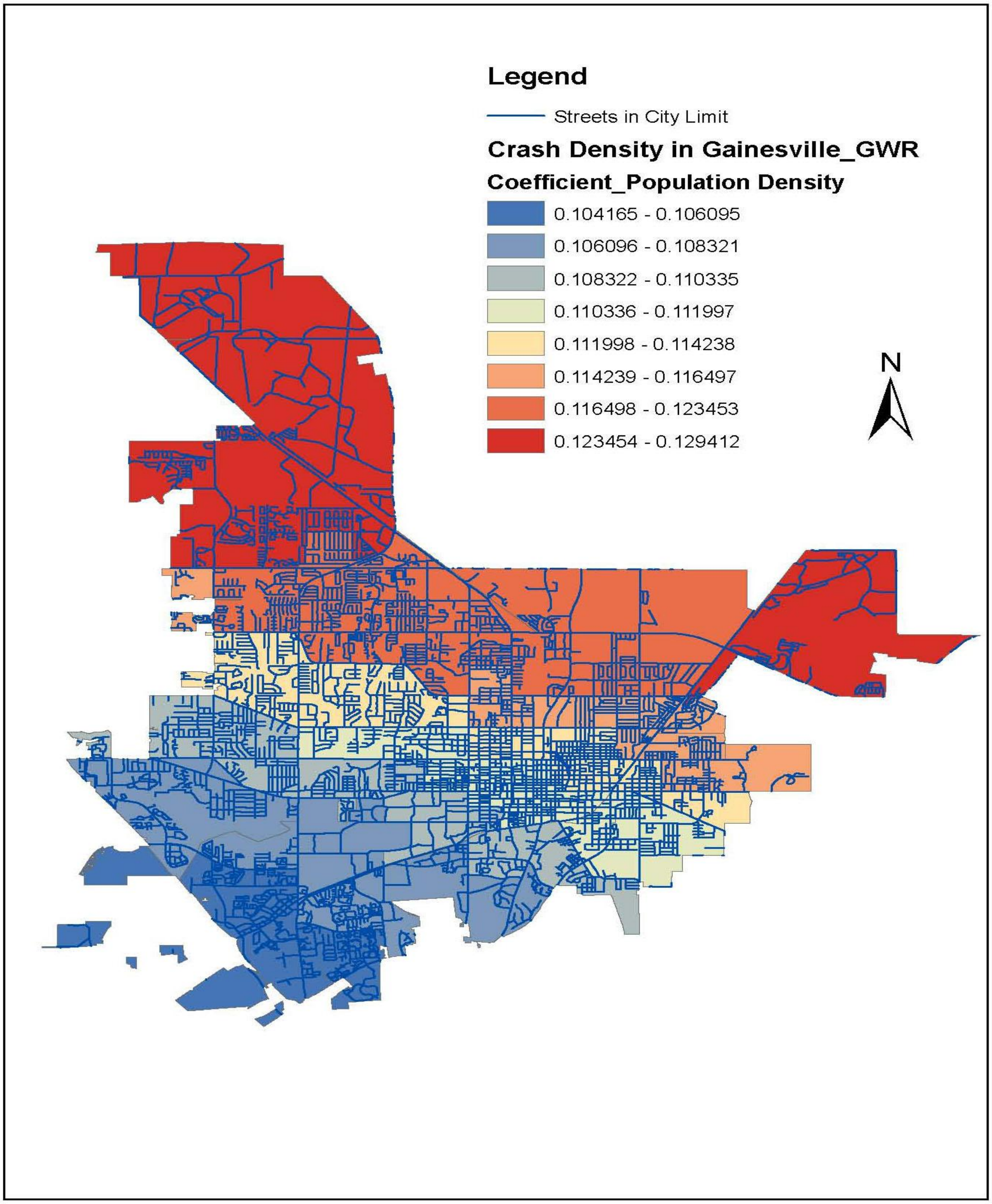

Figure 4-4. Estimated parameter of population density in GWR model. (Source: Generated by ArcMap, Geographically Weighted Regression Tool, edited by author)

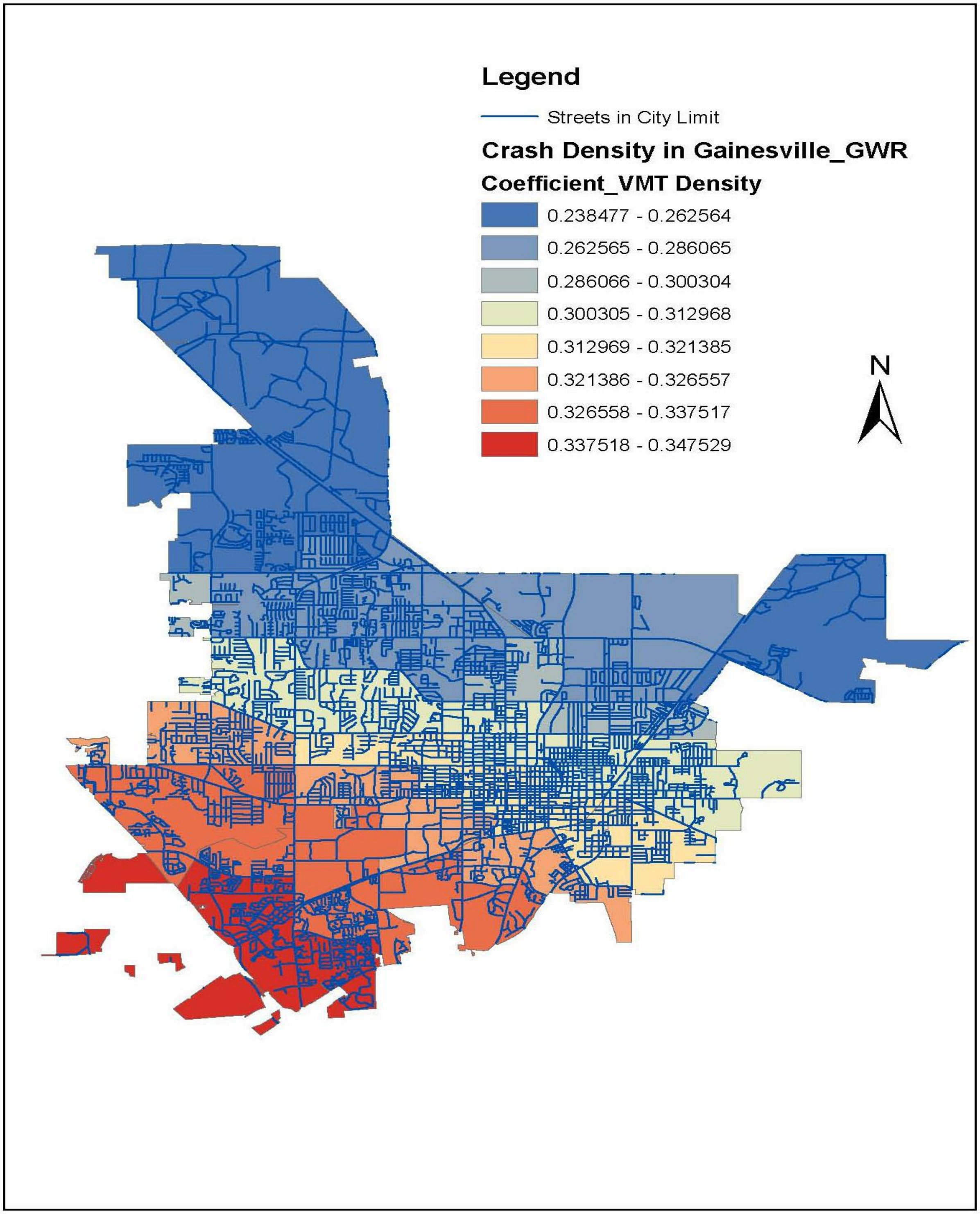

Figure 4-5. Estimated parameter of VMT density in GWR model. (Source: Generated by ArcMap, Geographically Weighted Regression Tool, edited by author)

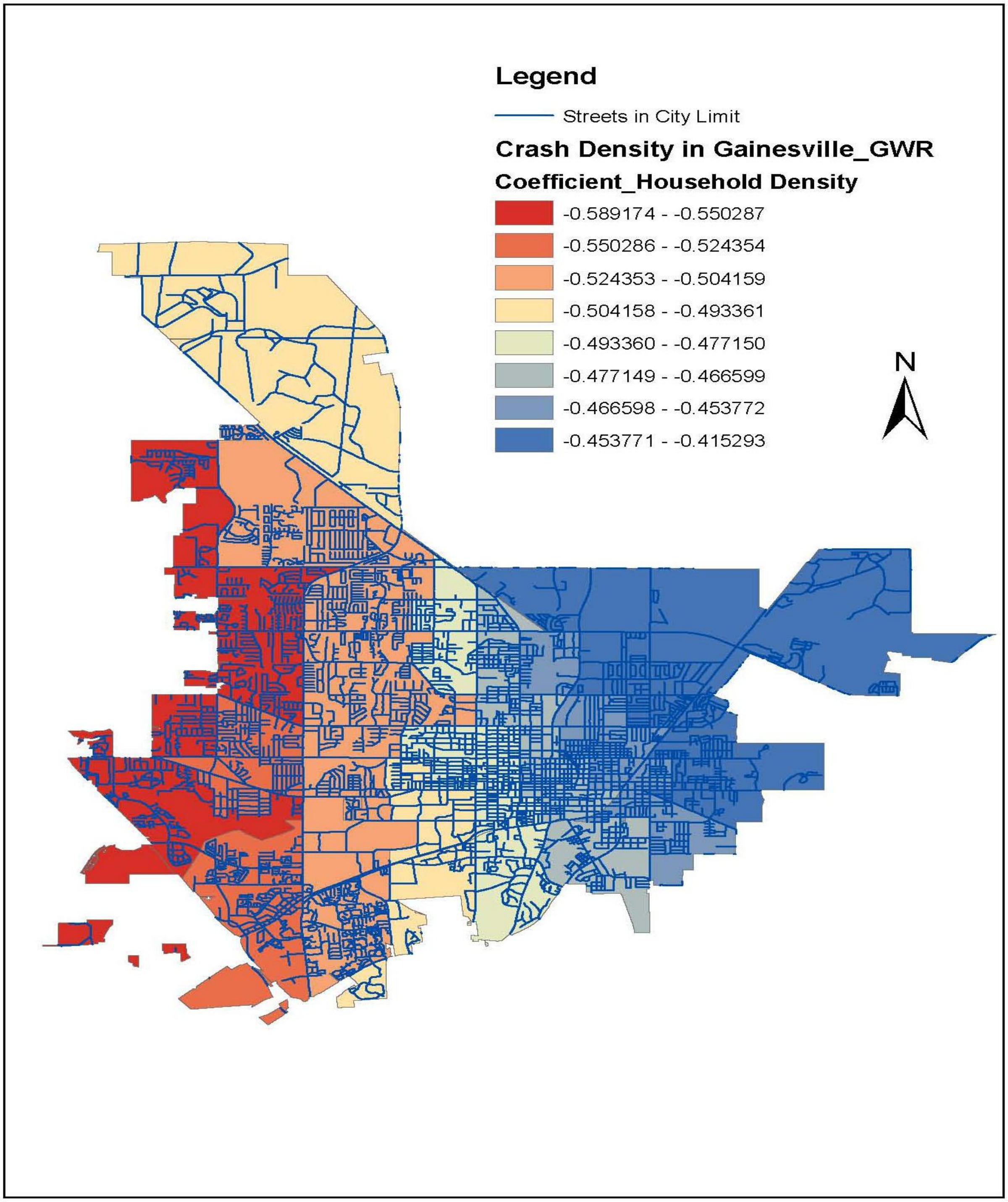

Figure 4-6. Estimated parameter of household density in GWR model. (Source: Generated by ArcMap, Geographically Weighted Regression Tool, edited by author)

CHAPTER 5
CONCLUSIONS

Recommendations

According to the U.S. Census data, the total population of Gainesville, as of 2010, has exceeded 100 thousand. Moreover, from 2000 to 2010, the population has continued growing by 30%. This phenomenon of vast growth in such a short time is mostly due to a large influx of people attracted to a major university; this demographic comprises students, faculty, and staff as well as various related jobs. Therefore, if the population of Gainesville is increasing by the rate of more than 10%, a burgeoning population will put pressure on the existing urban limits, which leads to the expansion of the city to the periphery. This thesis is aimed at exploring the relationship between different street patterns and the frequency of crashes. From the perspective of urban planning, this research can guide planners and policy makers to reduce road safety problems in the face of urban sprawl.

In this paper, different statistical models were used to analyze the effect of street patterns and other relevant factors on the frequency of traffic accidents. During the process, we chose the grid street pattern as the baseline, compared with loop and parallel types by using Poisson regression and multiple linear regression analysis. Both models came to the same correlation results. The results showed that, compared with the grid pattern, loop and parallel have a significant role in reducing the frequency of traffic crashes.

From the multivariate Poisson regression analysis, if the model for the crash density is exp (β), so the impact of increasing the number of entrances density by 1 for the block group with constant of other variables is an increase in crash density by a factor about 0.9%. Similarly, the impact of switching from residential and commercial land use to the other functional land characteristics for a block group decreases crash density by 74%. Therefore, in regards to city planning process, mixed land use should be encouraged. Besides, reasonable residential and commercial sites distributed to avoid spatial aggregation will greatly reduce the occurrence of traffic accidents. As to the number of entrance density, only 0.9% will affect the reduction of crash density; in other words, an insignificant statistic.

Maps generated by GWR model in ArcGIS software demonstrated the variation of different factors affecting the future development of the city. To remedy the expansion caused by urban development in the southeast of Gainesville, city planners need to arrange more blocks with loop and parallel patterns to significantly avoid the high frequency of accidents. As to the northwest area of Gainesville, because of the loop and parallel patterns already established, no significant influence can be exerted on these places. The changes of other factors need to be considered for the community's safety. In the future, planners should control population density in the northeastern and western regions. There is a large traffic flow on the main road in southeast of Gainesville. In order to reduce the frequency of traffic accidents, planners and decision makers can guide the travel influx to other roads by further planning. Reducing traffic density will greatly lower the amount of accidents. As for east of Gainesville, the creation of new communities is suitable for building a multi-family environment.

In a word, compared with the traditional grid pattern, if the loop and parallel patterns are adopted for these new communities, special concern should be taken for improving transportation safety in southeastern Gainesville.

Future Research

Admittedly, this research has its limitations. It was noticed that the limitation of data resulted in an imperfect model. Some results are not statistically significant and hard to explain. Besides, we only calculate the crashes which exist within the block group instead of all accidents within the whole city limit, such as on arterial roads and highways due to the boundary problems. Furthermore, the independent factors used in this thesis are limited by the data resource and census block group. Transportation safety is quite a complicated issue. There are many factors, unmentioned by this paper, which have an effect on traffic accidents.

In summation, this research can be explored by several ways. In the future, researchers can extend the volume of data. Several cities can be combined to constitute a single case study so that the results can enhance the accuracy of correlation. As for the research unit, the study area can be divided into several regular parcels with same areas as units. However, this requires relevant spatial information for each unit. Furthermore, more features of social-economics, roadway and vehicles can be analyzed to determine the effect of these variables on crashes, i.e, road width, age of drivers, median age in each parcels and road condition. Last but not Least, other variables can also be selected by resident's themselves; after all, city planning should be an interactive process, involving all members of the community.

LIST OF REFERENCES

Abdel-Aty, M. A., & Radwan, A. E. (2000). Modeling traffic accident occurrence and involvement. *Accident Analysis & Prevention, 32*(5), 633-642.

Asabere, P. K. (1990). The value of a neighborhood street with reference to the cul-de-sac. *The Journal of Real Estate Finance and Economics, 3*(2), 185-193.

Baker, S. P. (1992). *The injury fact book.* New York: Oxford University Press, USA.

Ben-Joseph, E. (1995). *Livability and safety of suburban street patterns: A comparative study* University of California at Berkeley, Institute of Urban and Regional Development.

Brindle, R. (1996). Road hierarchy and functional classification. *Traffic Engineering and Management.Melbourne: Institute of Transport Studies, Department of Civil Engineering, Monash University,*

Carlsson, G. et al. (1990). *A systematic approach to road safety in developing countries.* World Bank.

Charlton, M., & Fotheringham, A. S. (2009). Geographically weighted regression, white paper. *National Centre for Geocomputation, National University of Ireland Maynooth Maynooth, Co Kildare, Ireland,* 1–17.

Charlton, M. et al. (2005). Geographically weighted regression. *ESRC National Centre for Research Methods NCRM Methods Review Papers NCRM/006,*

Clark, D. E., & Cushing, B. M. (2004). Rural and urban traffic fatalities, vehicle miles, and population density. *Accident Analysis & Prevention, 36*(6), 967-972.

Coxe, S. et al. (2009). The analysis of count data: A gentle introduction to Poisson regression and its alternatives. *Journal of Personality Assessment,*

D'Agostino, R. B. (1971). An omnibus test of normality for moderate and large size samples. *Biometrika, 58*(2), 341.

DTLR, C. (2001). By design: Better places to live; A companion guide to PPG3.

Dumbaugh, E., & Rae, R. (2009). Safe urban form: Revisiting the relationship between community design and traffic safety. *Journal of the American Planning Association, 75*(3), 309-329.

ESRI. (2010). ArcGIS Resource Center. Help. Desktop 10. Assessed Sep 23,2011. http://help.arcgis.com/en/arcgisdesktop/10.0/help/

Evans, L. (1991). *Traffic safety and the driver.* New York: Van Nostrand Reinhold Company.

Ewing, R. & Zegeer, C. V. (2003). Urban sprawl as a risk factor in motor vehicle occupant and pedestrian fatalities. *American Journal of Public Health, 93*(9), 1541.

FROME, E. L. & CHECKOWAY, H. (1985). Use of poisson regression models in estimating incidence rates and ratios. *American Journal of Epidemiology, 121*(2), 309.

Gallaher, M. M. et al. (1989). Effects of the 65-mph speed limit on rural interstate fatalities in New Mexico. *JAMA: The Journal of the American Medical Association, 262*(16), 2243.

Garr, D. J. (1991). *Hispanic urban planning in North America.* London: Routledge press.

Gigliotti, E. (2007). Discovering statistics using SPSS. *Journal of Advanced Nursing, 58*(3), 303-303.

Graham, D. & Anderson, R. (2005). The effects of area deprivation on the incidence of child and adult pedestrian casualties in England. *Accident Analysis & Prevention, 37*(1), 125-135.

Gramenos, F. (2002). Residential street pattern design. *Zell/Lurie Center Working Papers,*

Híjar, M. et al. (2000). Risk factors in highway traffic accidents: A case control study. *Accident Analysis & Prevention, 32*(5), 703-709. doi:10.1016/S0001-4575(99)00116-5

Hilbe, J. (2007). *Negative binomial regression / joseph M. hilbe*. Cambridge: Cambridge University Press. Retrieved from http://www.loc.gov/catdir/enhancements/fy0805/2007282160-t.html

Kenny, D. A. (2003). Measuring model fit. *Retrieved July, 24*, 2005.

King, W. C. et al. (2005). Objective measures of neighborhood environment and physical activity in older women. *American Journal of Preventive Medicine, 28*(5), 461-469.

Lave, C., & Elias, P. (1994). Did the 65 mph speed limit save lives? *Accident Analysis & Prevention, 26*(1), 49-62.

Lubowski, R. N. (2006). *Major uses of land in the United States, 2002* United States, Dept. of Agriculture, Economic Research Service.

Lynch, K. (1992). *The image of the city.* Cambridge: MIT press.

Lynch, K., Banerjee, T., & Southworth, M. (1995). *City sense and city design: Writings and projects of kevin lynch.* Cambridge: The MIT press.

Marks, H. (1957). Subdividing for traffic safety. *Traffic Quarterly, 11*(3), 308-325.

Marshall, S. (2005). *Streets & patterns.* London: Routledge press.

Marshall, W. E., & Garrick, N. W. (2010). Street network types and road safety: A study of 24 California cities. *Urban Design International, 15*(3), 133-147.

Marshall, W. E. et al. (2008). Reassessing on-street parking. *Transportation Research Record: Journal of the Transportation Research Board, 2046*(-1), 45-52.

Mayo, J. M. (1979). Effects of street forms on suburban neighboring behavior. *Environment and Behavior, 11*(3), 375.

McCarthy, P. S. (1991). Highway safety and the 65-mph speed limit. *Contemporary Economic Policy, 9*(4), 82-92.

Miaou, S. P. (1994). The relationship between truck accidents and geometric design of road sections: Poisson versus negative binomial regressions. *Accident Analysis & Prevention, 26*(4), 471-482.

Miaou, S. P. et al. (1992). Relationship between truck accidents and highway geometric design: A poisson regression approach. *Transportation Research Record,* (1376)

Miaou, S. P. & Lum, H. (1993). Modeling vehicle accidents and highway geometric design relationships. *Accident Analysis & Prevention, 25*(6), 689-709.

Moudon, A. V. (1987). *Public streets for public use*. New York: Van Nostrand Reinhold.

NHTSA. (2008). National Motor Vehicle Crash Causation Survey: Report to Congress. Washington, DC: National Highway Traffic Safety Administration. Assessed July, 2008. http://wwwnrd.nhtsa.dot.gov/Pubs/811059.PDF.

Osgood, D. W. (2000). Poisson-based regression analysis of aggregate crime rates. *Journal of Quantitative Criminology, 16*(1), 21-43.

Perry, C. (1929). The neighborhood unit: Regional survey of new york and its environs. *Vol.VII.New York,*

Rifaat, S. M. (2011). *Street Pattern and Traffic Safety.* Ph.D. dissertation, University of Galgary, Canada—Alberta. Retrieved June 20, 2011, from Dissertations & Theses: Full Text. (Publication No. AAT NR69516)

Rifaat, S. M., & Tay, R. (2009). Effects of street patterns on injury risks in two-vehicle crashes. *Transportation Research Record: Journal of the Transportation Research Board, 2102*(-1), 61-67.

Southworth, M., & Ben-Joseph, E. (2003). *Streets and the shaping of towns and cities.* Washington D.C. Island Press.

Southworth, M., & Owens, P. M. (1993). The evolving metropolis: Studies of community, neighborhood, and street form at the urban edge. *Journal of the American Planning Association, 59*(3), 271-287.

Stanislawski, D. (1946). The origin and spread of the grid-pattern town. *Geographical Review, 36*(1), 105-120.

Steven M. R. (1995). Impact of the 65 mph speed limit on accidents, deaths, and injuries in illinois. *Accident Analysis & Prevention, 27*(2), 207-214. doi:10.1016/0001-4575(94)00058-T

U.S. Bureau of Census. (2010). *"Census 2000".* Assessed January, 2011. http://www.census.gov/main/www/cen2000.html.

Voelker, D. H. et al. (2011). *Cliffs Notes statistics quick review* Cliffs Notes.

Wagenaar, A. C. et al. (1990). Effects of the 65 mph speed limit on injury morbidity and mortality * 1. *Accident Analysis & Prevention, 22*(6), 571-585.

Wolfe, C. (1987). Streets regulating neighborhood form: A selective history. *Public Streets for Public use.* New York: Van Nostrand Reinhold.

BIOGRAPHICAL SKETCH

Dixue Li is currently a student at the University of Florida, with a major in urban and regional planning. Her specialization is geographical information system (GIS) and transportation planning. She got her undergraduate degree in environmental and urban planning management from Zhejiang Gongshang University, Hangzhou, China.

During the two years of studies in the master's degree program, she got ICGIS (Interdisciplinary Concentration in Geographic Information Systems) Certificate. In summer 2010, she presented her work in International Association for China Planning Conference. She also worked at City of Gainesville, Public Works Department during the spring semester as an intern.