1-1-2011

Exploring Traffic Safety and Urban Form in Portland, Oregon

Kristie Werner Gladhill
Portland State University

Exploring Traffic Safety and Urban Form

in Portland, Oregon

by

Kristie Werner Gladhill

A thesis submitted in partial fulfillment of the
requirements for the degree of

Master of Science
in
Civil and Environmental Engineering

Thesis Committee:
Christopher M. Monsere, Chair
Kelly Clifton
James Strathman

Portland State University

ABSTRACT

Street layout and design, once established, are then not easily changed. Urban form affects community development, livability, sustainability, and traffic safety. There has been an assumed relationship between urban form and traffic safety that favors designs with less through streets to improve safety. An empirical study to test this assumed relationship was carried out for crash data for Portland, Oregon.

This thesis presents an empirical methodology for analyzing the relationship between urban form and traffic safety utilizing a uniform grid for the spatial unit. Crashes in the Portland, Oregon city limits from 2005-2007 were analyzed and modeled using negative binomial regression to study the effect of urban form and street layout through factors on exposure, connectivity, transit accessibility, demographic factors, and origins and destinations. These relationships were modeled separately by mode: vehicle crashes, pedestrian and bicycle crashes. Models were also developed separately by crash type and by crash injury severity.

The models found that urban form factors of street connectivity and intersection density were not significant at 95% confidence for vehicle and pedestrian crash rates, nor for different crash severity levels, indicating that high connectivity grid street layout may have comparable safety to loops and

lollipops, in contrast to results in earlier studies. Elasticity for all models was dominated by VMT increases. Business density, population and transit stops were also significant factors in many models, underlining the importance not only of street layout design, but also planning to direct development to influence where businesses, employment, and housing will grow and handle traffic volumes safely.

DEDICATION

I thank my parents for valuing education and instilling that drive in me as well, along with perseverance to carry things through. Most of all, though, I thank my husband Richard Gladhill for his un-flagging support and love in this and so many other aspects of our family's life together.

ACKNOWLEDGEMENTS

I would like to thank my advisor Dr. Chris Monsere and committee members Dr. Kelly Clifton and Dr. James Strathman for their support, guidance, and direction throughout this project. I would also like to thank fellow graduate student Nicole Wheeler for her incredibly generous help with ArcGIS, editing, and reviewing. Thanks also go to Nathan McNeil for sharing geo-coded business data in the Portland area, and April Cutter for help with ArcGIS data.

I could not have done this without the support and help of so many in the Portland State University Department of Civil Engineering, the ITS lab, and OTREC. Having resources for study, ideas, friendships, and work to support me through this time have been a tremendous basis to work from.

Table of Contents

LIST OF TABLES

LIST OF FIGURES

1.0 INTRODUCTION

Streets have many functions, both for traffic movement, and as a framework for the neighborhood and local community (1). Street layout and design are typically defined in the planning and design phases for a particular land area. Once infrastructure is established it is then not easily changed. The design and layout of streets affect not only how the local neighborhood and community develop, but also traffic safety. Even though human factors dominate crash causality, it is important to understand secondary effects (2) and interactions with the built environment to aid planning and design of new development. It is important to study the actual safety relative to urban form to check if long held assumptions are valid, so that further development can appropriately consider safety in balance with connectivity and accessibility, along with street design factors such as width, traffic control, and presence of sidewalks. With this understanding, safety can be improved.

Recent studies in San Antonio, Texas (3) and Calgary, Alberta (4), (5) have started to include street layout in the crash analysis, and have shown lower crash rates within the limited through way neighborhoods.

1.1 Problem Statement

Traffic safety has long been a concern in street layout and design. Designers such as Frederick Law Olmsted rejected the grid layout in the 19th century, and started recommending curvilinear streets as part of an idealized suburban lifestyle. Twentieth century designers recommended limited access with limited through way designs for residential areas to reduce through traffic, sacrificing interconnectivity for perceived safety. Although rural road safety benefits from wide lanes and good visibility; urban streets with these features may encourage higher speeds, and thus less safe conditions in an environment that mixes vehicular, pedestrian, and bicycle traffic.

This assumed that a relationship between urban form and traffic safety would favor designs with less through streets to improve safety. Although this premise has been adopted and even codified into national and local standards (1), few studies have looked empirically at differences in traffic safety for different street layout designs. If safe street layout and design features could be identified, designers would be able to recommend urban form(s) that would provide better traffic safety and build safety into the infrastructure, even as connectivity, mobility, and accessibility are also considered. Infrastructure designed for the best safety practices and connectivity for vehicle, pedestrian, and bicycle traffic would provide a long term framework for safe and well functioning transportation networks that build and sustain communities.

1.2 Research Objectives and Scope

The objective of this research is to empirically quantify the relationship between urban form (defined by exposure, connectivity, transit accessibility, demographics, and origin and destination measures) and traffic safety (defined by reported motor vehicle crashes). This study was undertaken to test whether grid layout, which provides high connectivity and alternate through routes, is any more or less safe than loops and lollipops. The study looked at reported crashes from 2005-2007 within the Portland, Oregon 2007 city limit boundary. A uniform grid was used for spatial analysis to include all crashes without needing to give special consideration to crashes on analysis zone boundaries.

1.3 Organization

In the following chapter, this thesis will explore prior work related to traffic safety and urban form in a literature review. Chapter 3 describes the data sources and methodology used for this study. Chapter 4 presents qualitative analysis. Chapter 5 covers quantitative analysis, model building using negative binomial regression models for crashes, model results, and elasticity. Conclusions and recommendations for further work are made in Chapter 6.

2.0 LITERATURE REVIEW

Safety in design of urban streets has been an issue since Roman times (1). Modern studies have looked at not only the street design specifications for right of way and lane layout; but also how the form of the street layout may affect safety along with land use, demographic, and socioeconomic data. This review focuses on published literature that has dealt with empirically modeling urban form and safety.

Perhaps one of the first studies was conducted by Marks who studied five years of crash records on the Los Angeles County street system, encompassing 86 subdivision tracts over 4,320 acres (6). The study was limited to right angle crashes, which were nearly 84 percent of vehicular crashes within subdivisions, and were therefore felt to be representative of most crashes. Major streets bordering the subdivisions were not included in the study, nor was there any adjustment for traffic volume made in the analysis in this study.

Marks found that most crashes occurred at intersections, and some design features increased crash rates. Four leg intersections had much higher crash frequencies than three leg intersections, which have three conflict points compared to sixteen conflict points in a four legged intersection. Features such as jogs in alignment, skew, or "Y" shape at intersections were associated with increased crash frequencies. Obstructions to visibility such as bridges or railroad tracks also increased crash frequencies. Intersections spaced too closely were found to be a factor in crash frequencies. Limited access tracts had much lower crash rates than grid layout areas.

Based on these results, Marks recommended elements for safe design which included limited access design; avoidance of continuous through streets; collector streets exiting onto only one major street; preferred use of T-type intersections over four legged intersections; and avoiding multi-leg, Y, skewed, or jog intersections. With guidelines such as these, it was stated that subdividing could be done for safety.

Kim and Levine showed the value of using GIS data for crash analysis studying crashes on Oahu (7). In their analysis, they found that approximately 43 percent of crashes occurred at intersections. Using TIGER (Bureau of Census Topographically Integrated Geographic Encoding and Referencing) data for streets, crashes were assigned to the nearest intersection within 363 block groups for the entire island, including urban areas in Honolulu as well as rural and agricultural uses (8). Crashes were found to be concentrated in built-up, urban areas. Freeways themselves were relatively safe, but freeway ramps and crossroads were particularly dangerous. There was variation in the spatial pattern of crashes by time of day, day of week, and different vehicle types. Spatial GIS data for crash analysis was recommended as a tool that could be used to help develop meaningful community safety plans.

Land use activity, pedestrian friendliness, and infrastructure were found to be more effective at reducing road hazard than traffic controls or posted speed limitations in a study by Ossenbruggen et al (9). The study looked at 892 crashes from 87 sites on rural and suburban two lane undivided highways from 1993 to 1997 in Strafford County, New Hampshire. In addition to roadway measurements, qualitative data on land use activity, street life, and vehicle pedestrian interactions were taken at the crash sites. It was found that pedestrian friendly sites were associated with the least hazard, even with high traffic volume. Thus village sites, which were mixed-use areas with sidewalks, were less hazardous than residential or shopping areas without sidewalks. The infrastructure itself and multi-purpose activities seemed to be more effective at warning drivers of the need to proceed cautiously than sites which required traffic control devices to stop or interrupt traffic flow on the main road.

Hadayeghi studied crash frequency and severity in Toronto, Ontario to develop macro-level crash prediction models based on traffic demand, network, economic, and demographic variables (10). Major and minor roadway length in the analysis zone, intersection count, employment, and household population increases also increased crash rates. Crash rates decreased with higher posted speed and higher congestion levels. It was suggested that geometry of neighborhood design may also be important, though it had not been available for this study.

Ladron de Guevara (11) showed that planning-level models could be developed that would be useful in MPO (metropolitan planning organization) forecasting. Crashes from 1998 to 1999 in 859 TAZ (transportation analysis zones) in Tucson, AZ were analyzed using negative binomial regression for different levels of crash severity. Demographic and road characteristic variables were studied. Exposure to risk was felt to be better represented in a planning forecasting model by population rather than VMT (vehicle miles traveled) since population would have better future estimates, and was more likely to be available by TAZ than VMT in many jurisdictions. The fatal crash model found both population and population over 17 to be significant factors, along with intersection density. Injury crash parameters also included population density and intersection density, along with employment and miles of arterial and collector roadways.

Al-Masaeid and Suleiman pointed out that reducing the need for travel reduces exposure and traffic, and thus reduces crash risk. They looked at land use, population, VMT, and street network factors in crashes 2001-2002 in the

Syrian capital of Damascus (12). The land-use factors included percent commercial frontage, green area, industrial, or public buildings. Grid networks had more intersections and higher crash rates than comparable limited access elements. A higher percentage of commercial frontage and public buildings correlated with higher crash rates, suggesting that moving commercial frontage away from major thoroughfares may reduce crash rates. There was multicollinearity found between some of the street network and urban planning variables.

Kim et al in 2006 studied crashes on Oahu further using negative binomial regression analysis to relate crash rates to land use, population and economic activity (13). A grid of uniform sized cells was set up for analysis, rather than TAZ or census block groups. This study showed that grid cell spatial units can be used to statistically model crash rates. Population had a positive relationship to crash frequency, but job count in a particular cell was an even stronger factor to explain crash rates.

Kim et al also looked at using accessibility measures and other demographic and land use attributes to predict crash rates (13). The accessibility measures included road length, bus stops, bus route length, intersections, and dead ends. Uniform 0.1 square mile grid cells were again used for the analysis, with each crash assigned to the nearest intersection. Negative binomial regression was not producing a model fit, so logistic regression was used.

Results showed statistical significance for crash rates and vulnerable populations (elderly and children), disability, and job count. None of the demographic factors were significant, but business and land-use factors such as high density residential and military land uses were statistically significant. Bus stops and bus route length correlated with increased pedestrian crashes. Population size was only associated with bicycle crashes. Multicollinear relationships were found with variables such as population size and vulnerable populations (i.e. elderly and children). The researchers concluded that crash predictions could be useful to identify locations needing safety improvement strategies which could be implemented through enforcement, engineering, and education.

Clifton et al studied pedestrian-vehicle crashes in Baltimore, MD to test the hypothesis that the built environment affects injury severity in such crashes (14). In this area, 25 percent of households did not have access to a vehicle though there were numerous public transit options available. This study used more than 4500 pedestrian-vehicle crashes from 2000-2004 that were geocoded to nearest intersection. Analysis (probit model) found that pedestrian connectivity and transit access were the only significant built environment variables related to safety. The better these were the less likely severe injuries were to occur. Areas with low connectivity and transit access seemed most in need of safety intervention, since injuries there were more severe. In vehicle pedestrian crashes, children were more likely to sustain injury; the elderly were more likely to be fatalities; and male pedestrians were more likely to sustain injury, especially if substance abuse was involved.

Dumbaugh and Rae questioned whether the design concepts which have become engrained into policies have been empirically tested (3). Using census block groups for San Antonio crashes 2004 to 2006, they analyzed 150,626 reported crashes along with street network and demographic variables. In a methodological assumption, they assigned crashes on peripheral streets to each adjacent spatial unit. This assumption results in "double counting" since a crash would be included in more than one spatial unit. They found that total crash frequency increased with VMT, young or older drivers, and rose fifteen percent with each additional arterial mile. Population density was associated with fewer crashes, attributed to less travel demand when located close to services in densely populated areas. Injury crashes also increased significantly with additional arterials and four leg intersections. Higher income was associated with decreased crash rates. Fatal crashes rose with road and street network elements that increased vehicle speed.

In conclusion, Dumbaugh and Rae felt that the relationship between community design and traffic safety was important and complex, requiring community-level design solutions beyond simply preventing residential cut through traffic, but which would give attention to how different land uses and street network configurations influence speed and driver behavior. Traffic re-located away from residential areas reduced neighborhood traffic volume. However, arterials designed with wide lanes and long sight distances for higher speeds should have access management, limiting entry and exit points into the high speed traffic. Commercial zones with many entries and exits into the traffic stream should instead be located on lower speed thoroughfares.

Rifaat and Tay looked at injury crash rates (4) and crash severity modeling (5) for Calgary, Alberta. The street network was classified as being one of four patterns: loops and lollipops, mixed, warped parallel, or gridiron, (see Figure 1, Figure 2, Figure 3, and Figure 4). Other factors studied included roadway characteristics, demographics, land use, and socioeconomic factors. Crashes on boundary roadways were not considered due to boundary problems which would have further complicated the model, and since the peripheral traffic was considered to largely be non-local through traffic. Crashes were converted to EPDO (equivalent property damage only) crashes, and different models tried. Limited access patterns were associated with lower crash rates than the gridiron, with warped parallel consistently having the lowest crash rates with different models.

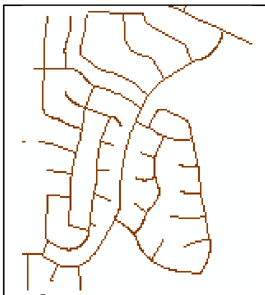


Figure 1 Loops and Lollipops

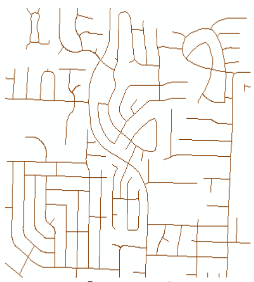


Figure 2 Mixed

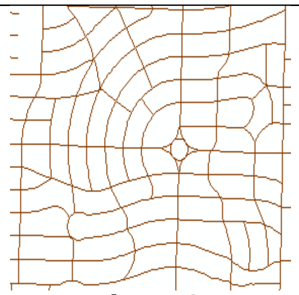


Figure 3 Warped Parallel

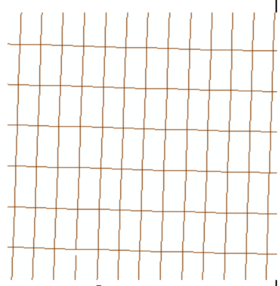


Figure 4 Grid

The papers reviewed are summarized in Table 1. These have explored many urban form factors and their relationship to crashes in many different cities. Exposure, connectivity, accessibility, demographic factors, land use, and origin and destination factors were of interest. Multicollinearity amongst factors has led to difficulties with the modeling, and is thus an issue to be aware of in this type of study.

Choice of spatial unit has also created difficulties. TAZ and block groups were often used as spatial units. These spatial analysis units were not of equal size, requiring them to be normalized for comparison. TAZ and block groups are typically bounded by roadways, which makes treatment of crashes on the peripheral roadways an issue: one study counted them in each adjacent zone, effectively double counting those crashes; another study removed those crashes from the study. A consistent way to handle all crashes is needed. Kim utilized a uniform spatial grid and showed ArcGIS to be a useful crash analysis tool, but crash location was still typically tied to the nearest intersection which may skew analysis of urban form. It would be illuminating to study crashes at their specific geo-coded location, all included, all equally weighted.

This study will look at crash data for Portland, Oregon, considering traffic safety related to exposure, connectivity, transit accessibility, demographic, and origin and destination factors. It will utilize a uniform spatial grid as did Kim et al (13) for spatial unit, to allow inclusion of all crashes equally weighted without double counting. This should provide insight into overall local traffic safety, rather than only within developed neighborhoods.

Table 1 Summary of Studies Reviewed

Resear-chers	Spatial Unit	Crash data	Other factors, parameters	Analysis Method	Findings	Conclusions
Marks 1957 LA County	86 sub-division tracts	5 years crash data, - only right angle crashes, - major bordering streets not included	Did not control for traffic volume, nor consider land use arrangements	Compari-son of crash rates	Most crashes were at intersections: 4 leg intersections higher crash frequency than 3 leg intersections. Irregularities increased crash rates.	Elements for safe design: limited access design avoid continuous through streets collectors exit onto only one major street prefer T over 4 leg intersections avoid irregularities
Kim, Levine, Nitz 1995 Oahu	363 census block groups	1990 Oahu crashes		Spatial mapping	Spatial pattern of crashes varied by time of day, day of week, vehicle type.	Spatial mapping of crashes could be useful for community safety improvement planning.
Kim and Levine 1996 Oahu	Looked at point, segment, and zonal analysis	Crash data geocoded		Compared actual to predicted crash rate, spatially mapped	~43% crashes occurred at intersections.	GIS data useful in looking at spatial relationships of crash data
Ossen-bruggen Pendharkar Ivan 2001 Stafford Count, NH	87 sites on rural and suburban two lane undivided highways	1993-1997, 892 crashes	Road measurements and land use, driver behavior	Ranked site hazard by crash rate Logistic regression	Village (mixed use) pedestrian friendly sites less hazardous than residential and shopping sites without pedestrian amenities such as sidewalks.	Land use activity, pedestrian friendliness, and infrastructure more effective at reducing road hazard than traffic controls or posted speed limitations.
Hadayeghi 2003 Toronto, Ontario, CA	463 traffic zones	1996 crash data, geocoded	Socio-economic, demographic, traffic demand, network data	Negative binomial regression	Crash rate related to roadway length, number of intersections, employment, household population, posted speed limit, and higher congestion levels.	Predictive models developed relating crash rate to various parameters. - Geometry of a neighborhood design may also be significant.

Table 1 Summary of Studies Reviewed, continued

Resear-chers	Spatial Unit	Crash data	Other factors, parameters	Analysis Method	Findings	Conclusions
Ladron de Guevara, Washington, Oh 2004 Tucson, AZ	859 TAZ	1998-1999 crash frequency and severity	Population for exposure risk; demographic, including schools, job density; bus stops; bike routes, road miles	Negative binomial regression	Fatal crash model found population significant. Injury crash significant parameters: population density, intersection density, employment, miles arterials and collector roadways.	Planning level models can be developed and useful for MPO safety forecasting. Population appears to better represent exposure to risk than VMT for this type model.
Al-Masaeid and Suleiman 2004 Damascus, Syria	14 urban zones	2001-2002	VMT, population, land use, street network	Multivariate regression analysis	Grid networks have more intersections and higher crash rates than limited access elements: urban crashes are exponentially proportional to intersection density, total street length.	Reducing the need for travel reduces exposure and traffic, thus reduces crash risk. Commercial frontage away from major thoroughfares may reduce crashes.
Kim, Brunner, Yamashita 2006 Oahu	Uniform grid	Vehicle, bike, and pedestrian crashes	Land use, employment, economic, population, demographic	Negative binomial regression	Fatal and injury crash parameters differed slightly, related to population, age, intersection density, employment, and miles of arterial and collector roads.	Grid cell characteristics can be used to statistically model crash rates
Kim, Pant, Yamashita 2010 Oahu	Uniform grid	2002-2004 crashes assigned to nearest intersection, freeway crashes excluded	Demographic, land use, accessibility measures	Logistic regression	Statistical significance for vulnerable populations, disability, job count, land use. Pedestrian crashes increased with more bus stops, bus route length. Population only associated with bicycle crashes.	Crash predictions useful in developing locations needing safety improvement strategies through enforcement, engineering, or education. Multicollinear relations with variables such as population size, vulnerable population.

Table 1 Summary of Studies Reviewed, continued

Researchers	Spatial Unit	Crash data	Other factors, parameters	Analysis Method	Findings	Conclusions
Clifton, Burnier, and Akar 2009 Baltimore, MD	disaggregate data, 1/4 mile buffer around each crash location	Pedestrian vehicle crashes 2000-2004 geocoded to nearest intersection	Street network, transit access, land use, vehicle type, weather, road condition, sex, population, substance abuse	Probit model	Transit access and pedestrian connectivity were the only built environment variables significant in the analysis. Used Herfindahl-Hirschmann index measure of land use mix.	Areas with low connectivity and transit access need greater safety interventions, injuries there are more severe.
Dumbaugh and Rae 2009 San Antonio, TX	747 census block groups plus buffer to include periphery streets	2004-2006 crashes, on and off roadway; peripheral crashes included and possibly double counted	Parcel-level land use data, demographic data; roadway network data (street miles, number of 3, 4 leg intersections)	Negative binomial regression	Traffic re-located away from residential areas reduces neighborhood traffic volume. Arterials designed with wide lanes and long sight distances for higher speeds should have limited access, with commercial traffic on lower speed thoroughfares.	There is an important relationship between community design and traffic safety. Designing pedestrian scale, livable streets emphasizes access over mobility, and has better traffic safety.
Rifaat and Tay 2009 Calgary, Alberta Canada	227 community areas	2003-2005 two vehicle crashes; - no crashes on boundary roadways	4 street patterns: - gridiron, - warped parallel, - loop & lollipop, - mixed	Logistic regression	Roadway and demographic data provided control relationships to crash rates. - Crashes on boundary roadways not considered due to boundary problem	Compared to gridiron, loops and lollipops design has decreased crash injury risk.
Rifaat and Tay 2010 Calgary, Alberta Canada	227 community areas	2003-2005 two vehicle crashes; crashes on boundary roadways not considered	*4 street patterns - gridiron, - warped parallel, - loop & lollipop, - mixed * Road condition, demographic, socioeconomic, land use;	Negative binomial regression	Crash data converted into EPDO (equivalent property damage only) crashes. * AADT estimated from ITE trip generation models. * Control factors affected crash rates; socioeconomic and demographic factors also statistically significant	Limited access street patterns had lower crash rates than gridiron layouts.

3.0 DATA SOURCES AND PREPARATION

Prior work had shown that important data for studying the relationship between urban form and traffic safety were of course crash data, as well as choice of spatial unit, and how crashes along roadways at the periphery of spatial units were handled. Other important factors fell into categories of exposure (traffic volume), connectivity (street length, number and types of intersections), accessibility to transit, demographics (population), and origin and destination (employment and businesses). Data sources and preparation of these sources are discussed in the following sections.

3.1 Data Sources

3.1.1 Crash Data

This study looked at reported crashes from 2005-2007 within the Portland, Oregon city limit (as defined in 2007). Crash data were available from the City of Portland, geo-coded by crash location. Further crash details were found using OrTSDA, the Oregon Traffic Safety Data Archive, which is a mirror of the statewide Crash Data System (CDS) maintained by ODOT (Oregon Department of Transportation). Crash data from a three-year period were used in order to have a large data set while limiting the likelihood of structural changes over the study period.

Crash reporting varies from state to state. Oregon is a self-reporting state for crashes, where the individual drivers are required to file an Oregon Traffic Accident and Insurance Report within 72 hours if they are involved in a crash that results in injury, death, more than $1,500 damage to their vehicle, or more than $1,500 damage and towing of another vehicle. While police officers do complete and file reports, many non-injury, property-damage-only (PDO) crash reports do not include a police officer's written report. These reporting disparities mean that many PDO crashes are not reported in Oregon.

Freeway crashes were eliminated from the data. A freeway is a very different type of transportation infrastructure, which would be expected to have very different effects on traffic safety than the local streets which were the focus of this study.

3.1.2 Exposure

Exposure data tend to be crucial in crash analysis: the more exposure to use of the transportation system, the greater the probability of crashes. Metro, the regional government in the Portland metropolitan area, provided ArcGIS layers from the regional travel demand model that included 2005 exposure data on volume to capacity ratios (v/c), peak hour volume, VMT (vehicle miles traveled), and average free speed. Although available for only some streets, these data were felt to give relative data for analysis and comparison.

3.1.3 Connectivity

Connectivity data were a chance to represent the street network layout. The street network scale developed by Rifaat and Tay was applied to all 1284 spatial grid cells; some had too little street network to have one of the four network types assigned. Other data expected to shed light on connectivity included road network length for street or arterials, as well as counts of intersections and how many legs to the intersections,

Metro maintains a rich geo-spatial database, the Regional Land Use and Information System (RLIS). The RLIS 2009 dataset at PSU (Portland State University) was used for data on streets and roadway. Road network data included layers with lines showing the location of streets, minor arterials, major arterials, and freeways.

3.1.4 Transit Accessibility

Transit accessibility was expected to also inform this study. High transit usage could mean less vehicles on the road, and thus less exposure. Transit riders are often pedestrians either before or after their transit portion of a trip. Transit ridership data were obtained from 2007 TriMet data in the PORTAL database archive at PSU. Transit stops, routes, and schools were available in the RLIS 2009 dataset.

3.1.5 Demographics, Origins, and Destinations

Population has been shown to be a strong factor in previous crash data studies, and is often used in transportation modeling and forecasting. Employment is a strong factor in trip generation, since most workers need to commute to the workplace. The number of businesses was of interest since businesses attract not only customers and employees to make trips, but the number of business establishments also affects the number of driveways along roadways which increases likelihood of conflict and potential for crashes.

The RLIS database included demographic data on population, housing units, and dwelling units. Metro shape files provided 2005 employment data (number of employees) by TAZ (transportation analysis zones), and modeled block size in raster layers.

Business data were obtained from www.ReferenceUSA.com for grocery, clothing, goods, services (beauty, laundry, mail, bank), fitness, entertainment, food, schools & academies, religious institutions. The businesses were then geo-coded by address so that the number of businesses could be counted

3.2 Preparation

3.2.1 Definition of Spatial Analysis Unit

ArcGIS was used to illuminate and aggregate the data. A uniform spatial grid of ~0.1 square miles (1670 feet long on a side) was set over the Portland metropolitan area to be used as the spatial unit for analysis, each spatial unit inherently of the same area. This allowed all crashes to be included without double counting, since there would be little likelihood that spatial grid boundaries would fall on roadway locations, unlike the use of TAZ or census block group spatial units, which typically are bounded by roadways. The grid was limited to whole cells within the 2007 Portland City Limits for a total of 1284 cells.

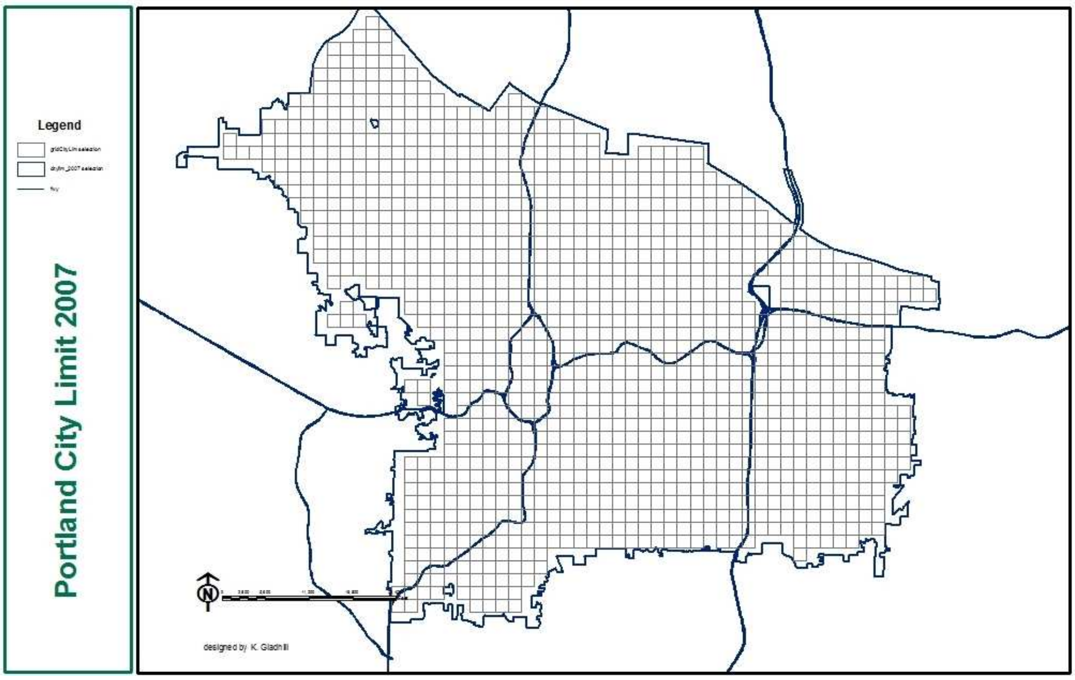


Figure 5 Spatial Data Grid within the Portland City Limits

3.2.2 Data Aggregation

Crash data were allocated into the spatial grid cells using ArcGIS. Exposure, connectivity, transit accessibility, demographic, and origin and destination data were also aggregated into the grid cells: the point data were summed by count, line data summed for total length, and polygon data summed proportional to the percentage of the polygon's area that was within a grid cell.

$$\text{Point data:} \quad \text{cell value} = \sum_{k=0}^{n} x$$

$$\text{Line data:} \quad \text{cell value} = \sum_{k=0}^{n} (line\ length\ within\ grid)$$

$$\text{Polygon data:} \quad \text{cell value} = \sum_{k=0}^{n} \quad pctArea$$

where *pctArea = (% polygon in grid cell) * (factor value for polygon)*

As an example, a spatial grid cell containing 30% of the area of a TAZ spatial unit with a factor value of 120 would get 30% of the TAZ value for a factor, or 36. The proportion of that factor for all other TAZ units represented in the grid cell was similarly calculated, and the total summed to compute a value for that factor for that grid cell. Table 2 lists the data variables and how they were aggregated into the spatial grid cells.

Intersections were counted using ArcGIS to determine points where street lines intersected, and an algorithm then eliminated duplicates and tallied the number of intersection legs.

Table 2 Data aggregated into Uniform Spatial grid Cells for Study

DATA	**Description**	**Aggregation**	**SOURCE**
Crash data	Portland, OR 2005-2007	sum of point data	City of Portland, OrTSDA
CONNECTIVITY			
Street Network	4 = grid 3 = warped parallel 2 = mixed 1 = loops and lollipops 0 = could not determine	values assigned to each spatial grid cell	evaluated by researcher inspecting each spatial grid cell
Intersections	Count of total intersections	intersections per spatial grid cell	RLIS 2009, intersection analysis
Streets	Total street length, arterial length, major arterial length, freeway length	sum of street length (line data) in spatial grid cell	RLIS 2009
TRANSIT ACCESSIBILITY			
Transit stops	Count of transit stops	sum of point data	RLIS 2009
Transit route	Total transit route length	sum line data in spatial grid cell	RLIS 2009
Ons and Offs	Total transit count of riders getting on and off	sum of values in spatial grid cell	PORTAL 2007 data
EXPOSURE			
v/c	Volume to capacity ratios on some roadways	average of values in spatial grid cell	Metro 2005 Transportation Model
VMT	Vehicle miles traveled	sum of values in spatial grid cell	Metro 2005 Transportation Model
Avg free speed	2006 data	average of values in spatial grid cell	Metro 2005 Transportation Model
Schools	2006 data	sum of point data	RLIS 2009
DEMOGRAPHICS, ORIGINS, and DESTINATIONS			
Population	2006 data	apportioned ratio of area in spatial grid cell to factor value	RLIS 2009
Dwelling units	2006 data	apportioned	RLIS 2009
Households	2006 data	apportioned	RLIS 2009
Employment	2006 data	apportioned	Metro
Business	2010	sum of point data	ReferenceUSA.com
Block Size	Raster file	raster converted to points, sum of points in spatial grid cell	Metro

4.0 DESCRIPTIVE ANALYSIS

This section will take an initial look at the data, with analysis of the crash data, and a qualitative look at exposure, connectivity, transit accessibility, demographic, and origin and destination data using chloropleths from ArcGIS.

4.1 Crash Data

The study dataset had a total of 21,492 non-freeway crashes within the city limits of Portland for the years 2005-2007. This number of crashes is considerably lower than more than 150,000 in the San Antonio study (3). This may be largely due to the underreporting issue with Oregon crash data. Looking at the crash data aggregated into the uniform spatial grid cells revealed that spatially crashes were concentrated in areas with heavy traffic: downtown, along high volume arterials, and adjacent to freeways (see Figure 6). Figure 7 shows a histogram of crash count per spatial grid cell for the study dataset; most cells had low crash counts.

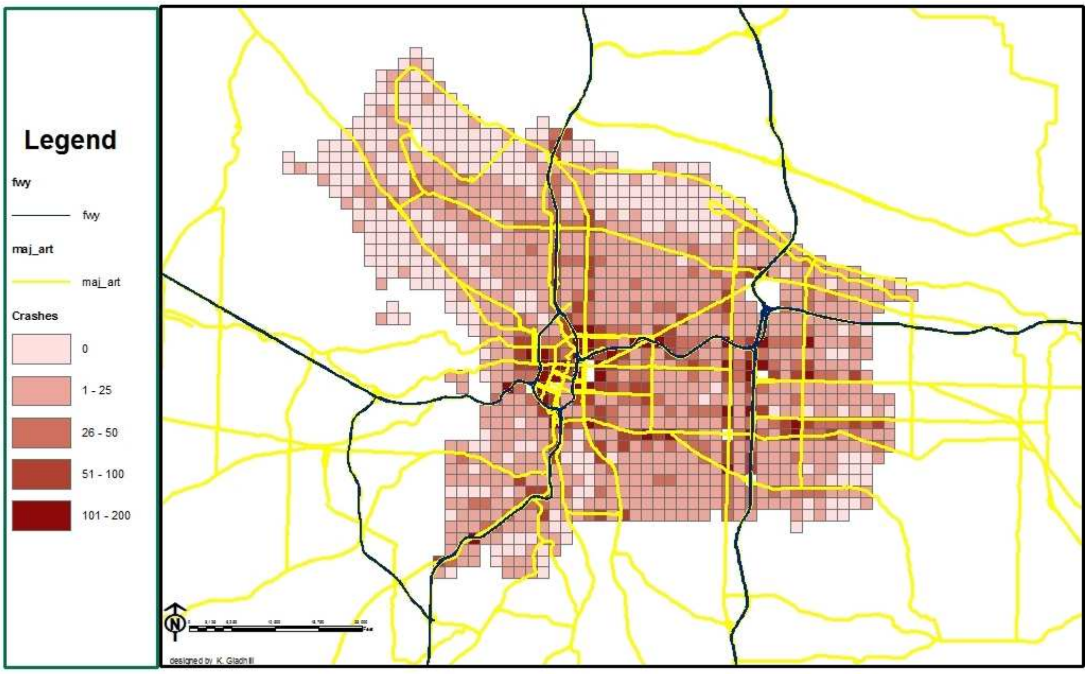


Figure 6 Total Non-freeway Crashes Portland 2005 – 2007

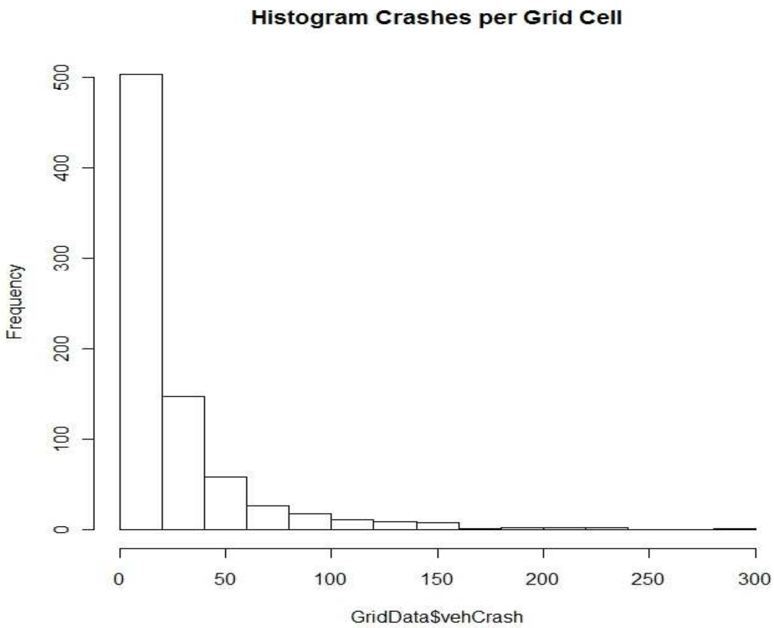


Figure 7 Histogram of 2006-2007 Crashes in Portland City Limits

Analysis of the crash data by crash types, see Figure 8, clearly showed that the four primary crash types were angle; stopped, where both vehicles were going the same direction; straight crashes where both vehicles were going the same direction; and turn crashes where one vehicle was turning in front of an oncoming vehicle going straight in the opposite direction.

Figure 9 illustrates each of these crash types. Angle crashes would have increased likelihood of occurrence the more cross streets to a roadway, causing a vehicle to cross in front of another vehicle, so would be expected to increase the more intersections. Stopped crashes, most likely rear ending, would be increased with traffic control bringing vehicles to a stop at intersections. Straight crashes would likely be caused by going too fast, thus overtaking another vehicle; or due to a sudden deceleration that the following vehicle did not respond to in time. Turn crashes would have increased likelihood with increased cross streets and driveways, where one vehicle would turn across an on-coming vehicle's path.

PDO (property damage only) crashes were most common (see Table 3) at 60.4 percent of the total crashes. Minor injury level C accounted for 21.6 percent of the total crashes, which was more than half of all injury crashes. Fixed object and pedestrian crash types accounted for 49 of the 82 fatalities.

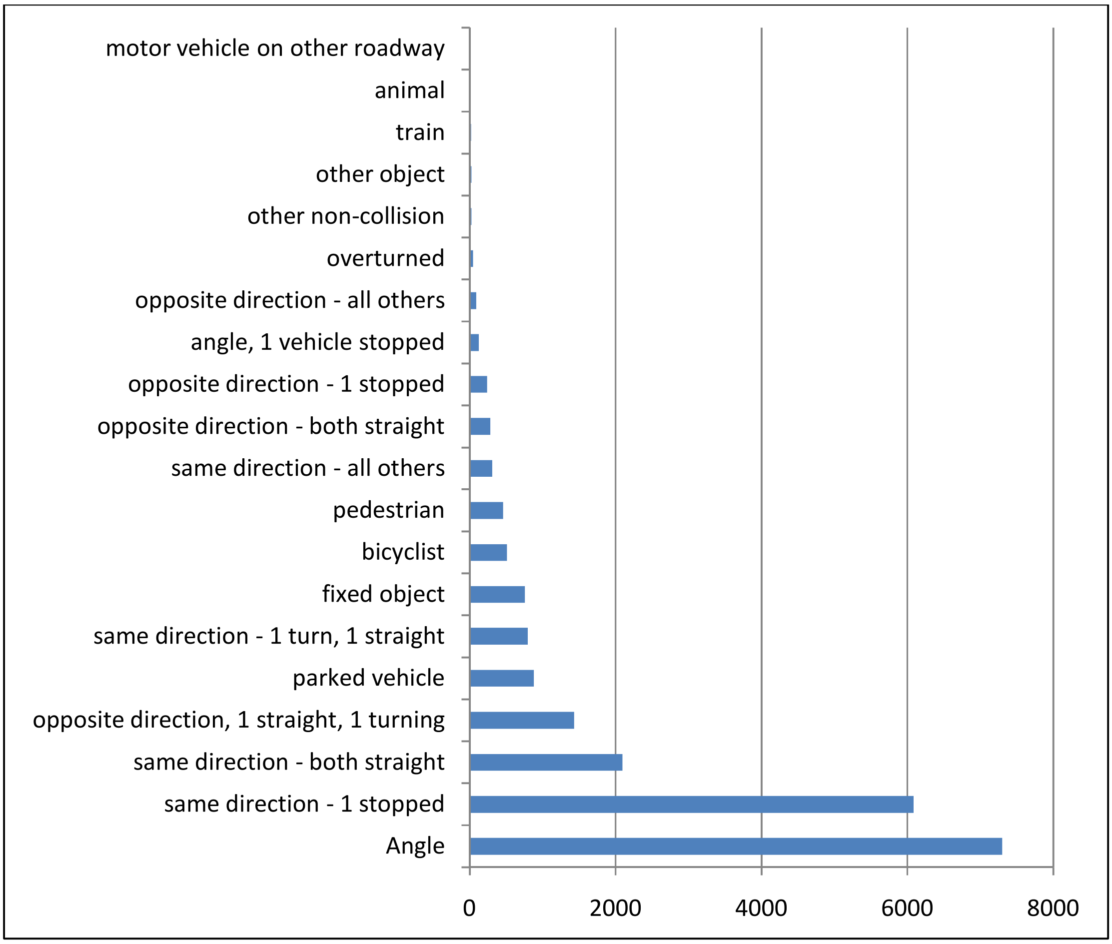


Figure 8 2005 – 2007 Total Crashes by Crash Type

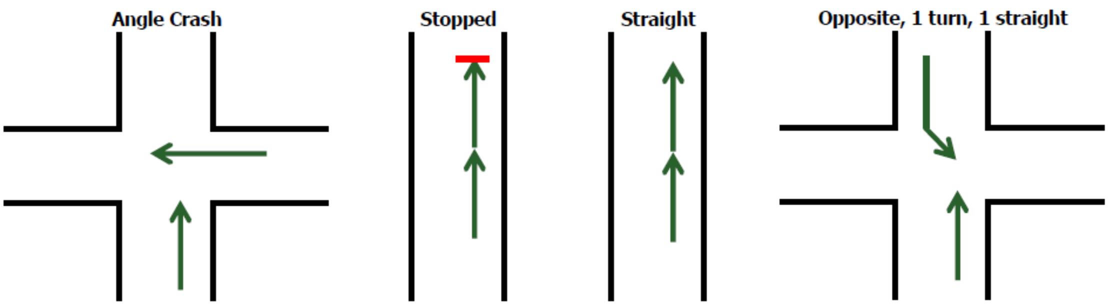


Figure 9 Primary Crash Types

Table 3 2005 - 2007 Portland Non-freeway Crashes in Spatial Grid Cells

	Total	**Fatality**	**Injury**	**INJ A**	**INJ B**	**INJ C**	**PDO**
Angle	7303	9	2769	197	1200	1372	4525
same direction - 1 stopped	6086		2675	75	743	1857	3411
same direction - both straight	2094	4	507	17	148	342	1583
opposite direction, 1 straight, 1 turning	1431	1	617	48	242	327	813
parked vehicle	881	3	177	22	81	74	701
same direction – 1 turning, 1 straight	797	2	174	5	46	123	621
fixed object	757	24	293	54	136	103	440
bicyclist	511	7	477	71	256	150	27
pedestrian	457	25	427	83	204	140	5
same direction – all others	310		50	2	9	39	260
opposite direction, both straight	285	6	127	23	58	46	152
opposite direction, 1 stopped	242		30	2	9	19	212
angle - 1 vehicle stopped	128		25	3	3	19	103
opposite direction all others	92		20		6	14	72
overturned	43	1	31	6	17	8	11
other non-collision	25		12	3	7	2	13
other object	22		5		2	3	17
train	20		7	2		5	13
animal	4		1			1	3
motor vehicle on other roadway	4						4
TOTAL	**21492**	82	8424	613	3167	4644	12986
		0.4%	39.2%	2.9%	14.7%	21.6%	60.4%

INJ A = injury level A, incapacitating
INJ B = injury level B, non- incapacitating
INJ C = injury level C, possible injury
PDO = property damage only

More information on injury severity can be found at:
http://www.oregon.gov/ODOT/TD/TDATA/car/docs/2007CodeManualVersion2.0.pdf

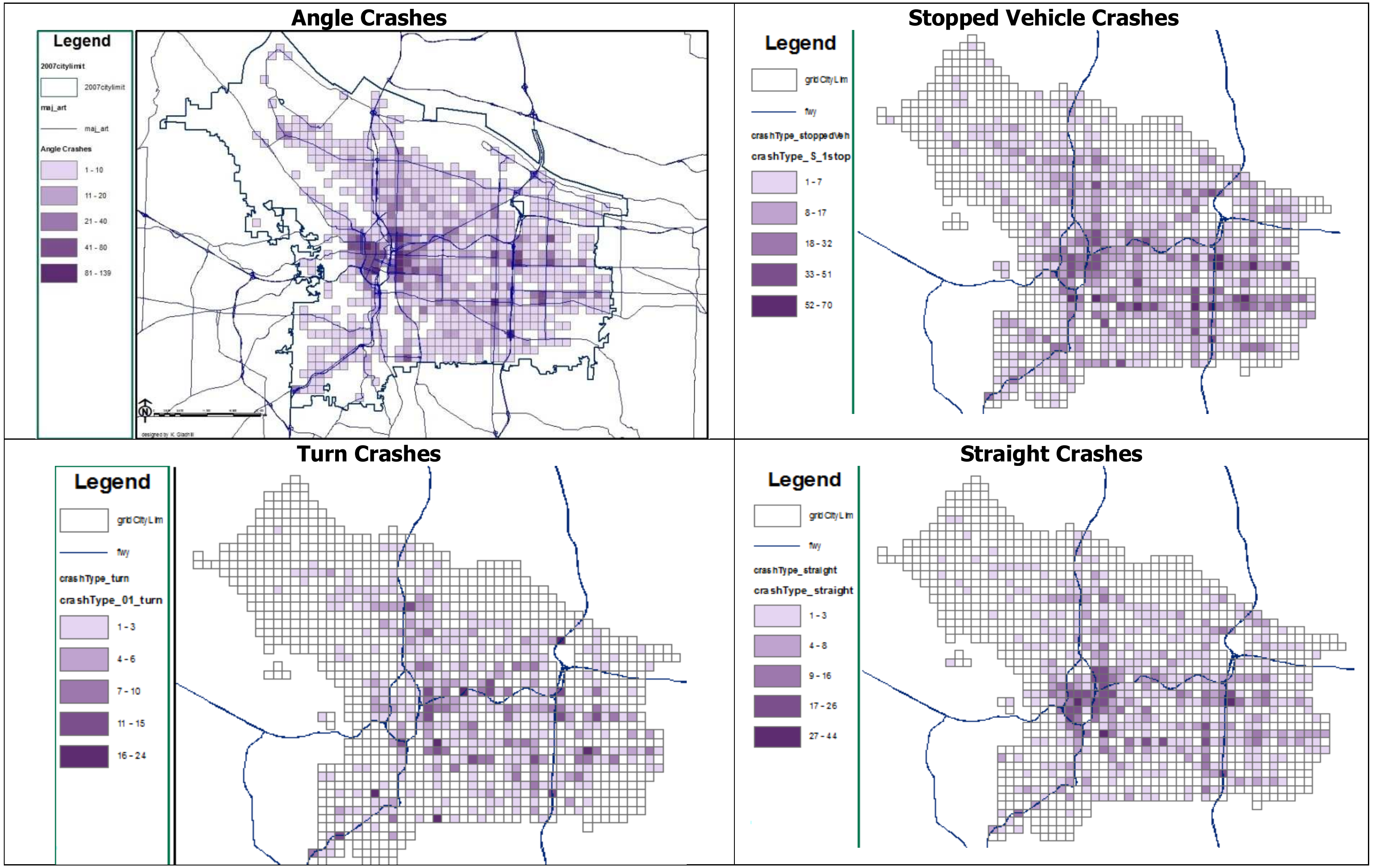


Figure 10 2005 – 2007 Vehicle Crashes by Crash Type, Spatial Maps

Figure 10 shows that the spatial distribution of crashes differed for different crash types. Angle crashes, stopped vehicle crashes, and crashes where both vehicles were going straight showed similar distributions to total vehicle crashes with crashes concentrated along major arterial corridors. Turn crashes were more randomly distributed, although major arterial corridors are still discernible. High frequency in specific spatial grid cells indicates that specific locations may have problems which contribute to likelihood of each particular crash type.

The crash dataset included 503 pedestrian crashes and 523 bicycle crashes. These crashes are summarized by motor vehicle operator error, see Table 4 and Table 5. Failure to yield right of way (ROW) was the leading cause cited in the crash data for both pedestrian and bicycle crashes; presumably the vehicle failing to yield. Whereas greater than 60 percent of all vehicle crashes were PDO, pedestrian and bicycle crashes involved injury more than 93 percent of the time. Pedestrian crash fatality outcomes were more than ten times that of overall crash fatalities, and bicycle crash fatality rates were three times that of overall crashes.

Spatially, pedestrian crashes were concentrated downtown and along major arterial routes in the city, see Figure 11; bicycle crashes were more randomly distributed, see Figure 12.

Table 4 2005 – 2007 Pedestrian Crashes

	Total	**Fatality**	**Injury**	**INJ A**	**INJ B**	**INJ C**	**PDO**
Failed to Yield Right of Way	294	13	277	42	140	95	4
Non-motorist illegally in roadway	111	7	103	28	42	33	1
Too fast for conditions	16	1	15	3	11	1	
Other driving error	13		13	3	7	3	
Inattention	12	2	10	2	3	5	
Disregarded Red-Amber-Green traffic signal	11		11	2	8	1	
Careless	10		10	5	5		
Speed too fast for conditions	8	1	7	6		1	
Not visible	7	1	6	1	4	1	
Other	6	2	4		2	2	
Improper passing	4	2	2		2		
Improper turn	2		2	2			
Reckless	2		2			2	
Followed too closely	4		4	2	2		
Passed stop sign	1		1	1			
Fatigue	1		1		1		
no code applicable	1		1	1			
TOTAL	**503**	29	469	98	227	144	5
		5.8%	93.2%	19.5%	45.1%	28.6%	1.0%

Table 5 2005 – 2007 Bicycle Crashes

	Total	**Fatality**	**Injury**	**INJ A**	**INJ B**	**INJ C**	**PDO**
Failed to Yield Right of Way	343	4	318	38	178	102	21
Disregarded Red-Amber-Green	47	1	43	10	17	16	3
Non-motorist illegally in roadway	28		27	5	13	9	1
Other	20		20	4	10	6	
Other driving error	18	2	15		9	6	1
Passed stop sign	18		18	4	10	4	
Improper turn	13		13	3	9	1	
Followed too closely	11		10	5	3	2	1
Improper lane change	6		6	1	4	1	
Inattention	6		6	3	1	2	
Careless	4		4	2	2		
Improper passing	3		3	1	1	1	
Reckless	3		3			3	
Speed too fast for conditions	2	1	1		1		
Not visible	1		1		1		
TOTAL	**523**	8	488	76	259	153	27
		1.5%	93.3%	14.5%	49.5%	29.3%	5.2%

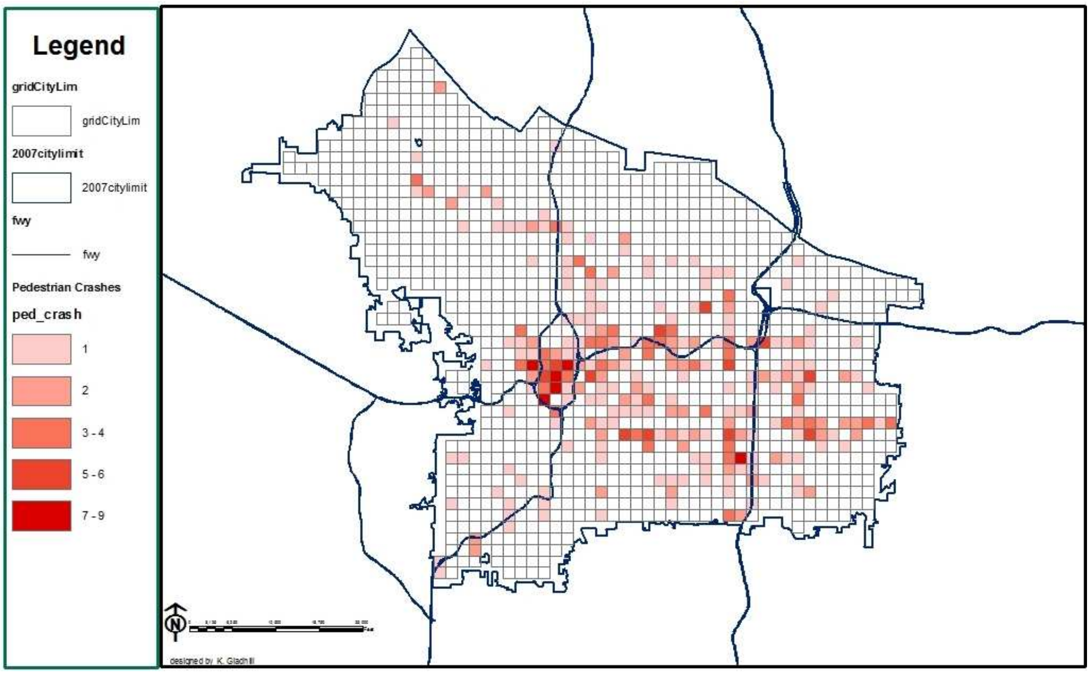


Figure 11 2005 – 2007 Pedestrian Crash Spatial Map

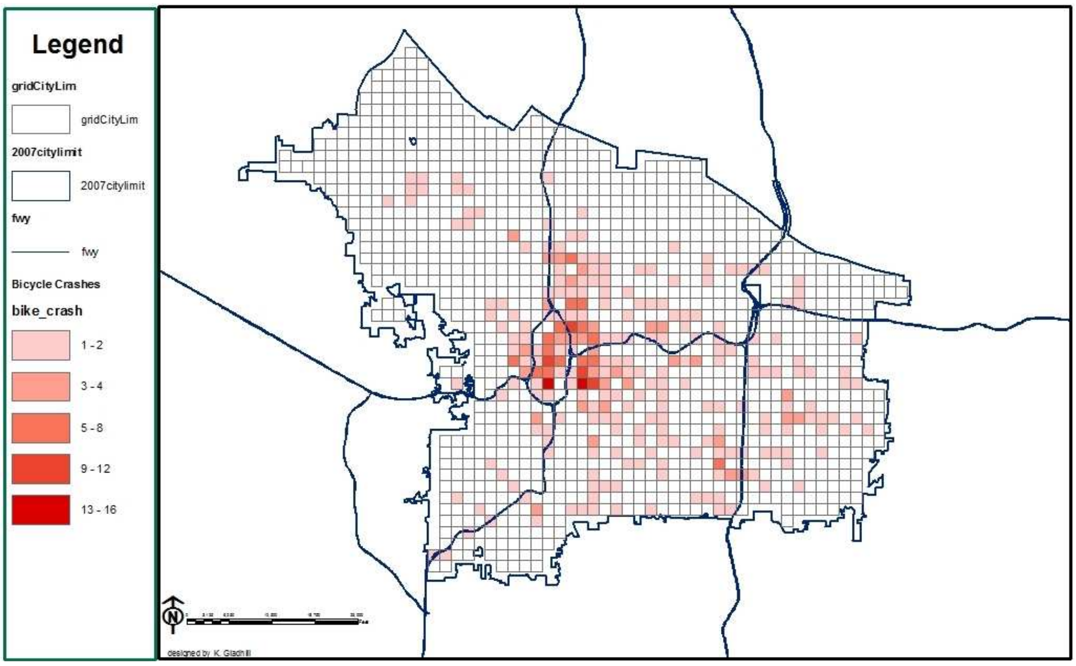


Figure 12 2005 – 2007 Bicycle Crash Spatial Map

Vehicle crashes are shown in Figure 13. PDO crashes (see Figure 14), show essentially the same distribution as total vehicle crashes, which is expected considering that PDO were more than 60% of the vehicle crashes. Figure 15 shows all fatal and injury crashes, still in a similar distribution to total vehicle crashes. Fatal crashes were rarer, as seen in Figure 16. Adding injury level A crashes in Figure 17 starts to again make major corridors discernible.

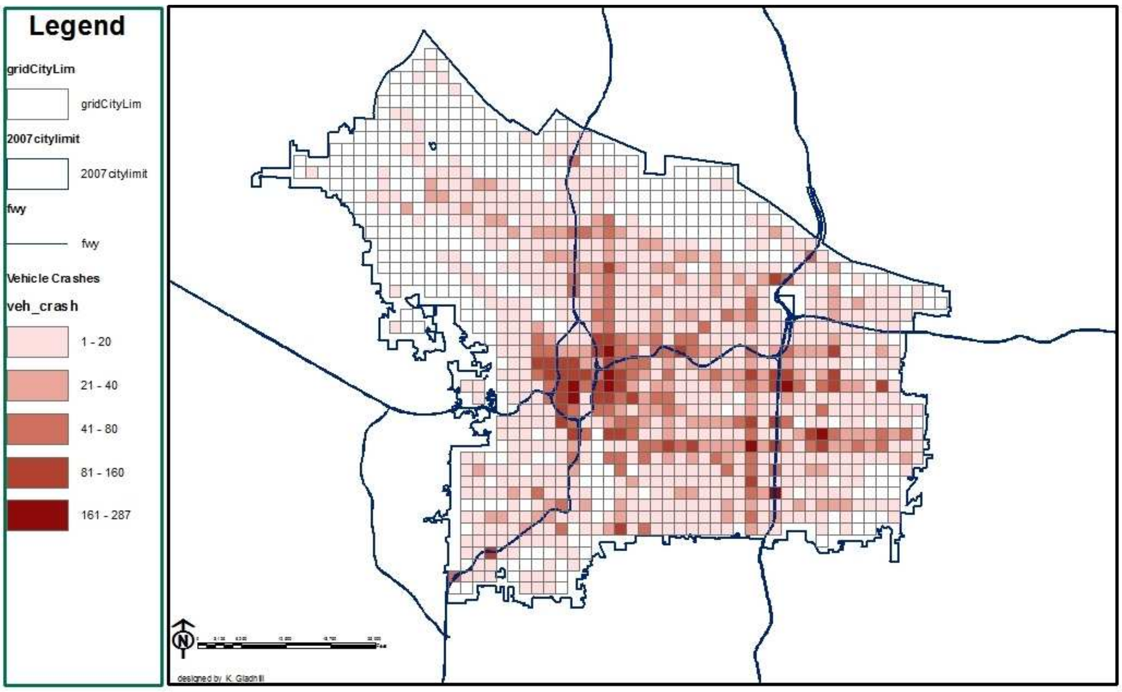


Figure 13 2005 – 2007 Vehicle Crashes

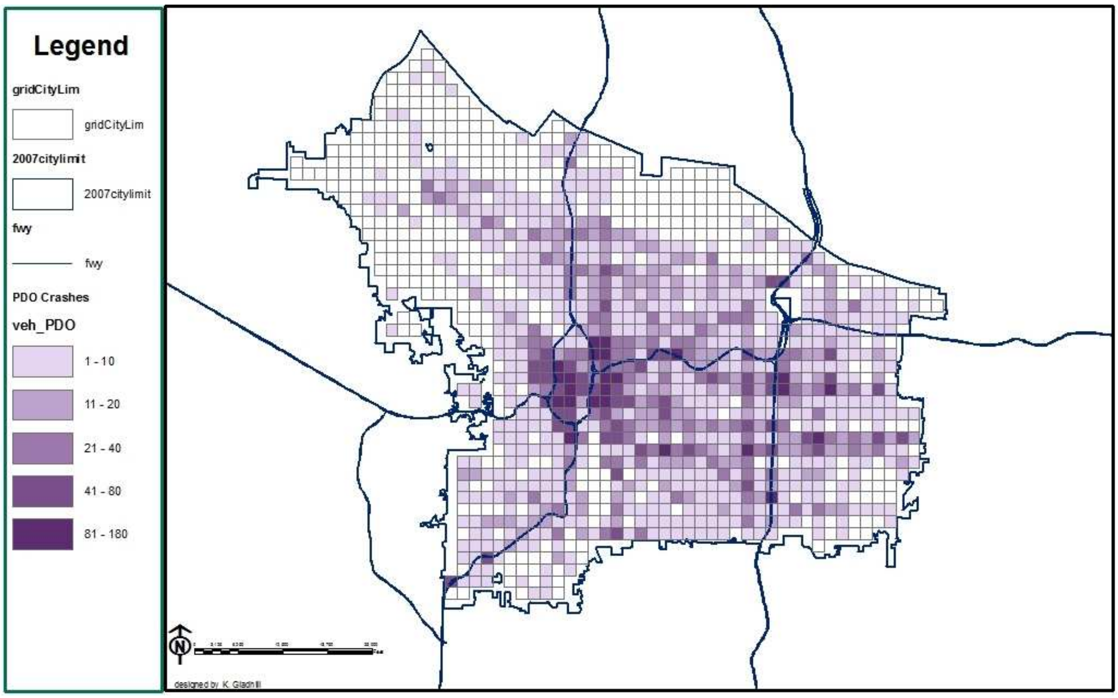


Figure 14   2005 – 2007 Vehicle Crashes, PDO Severity

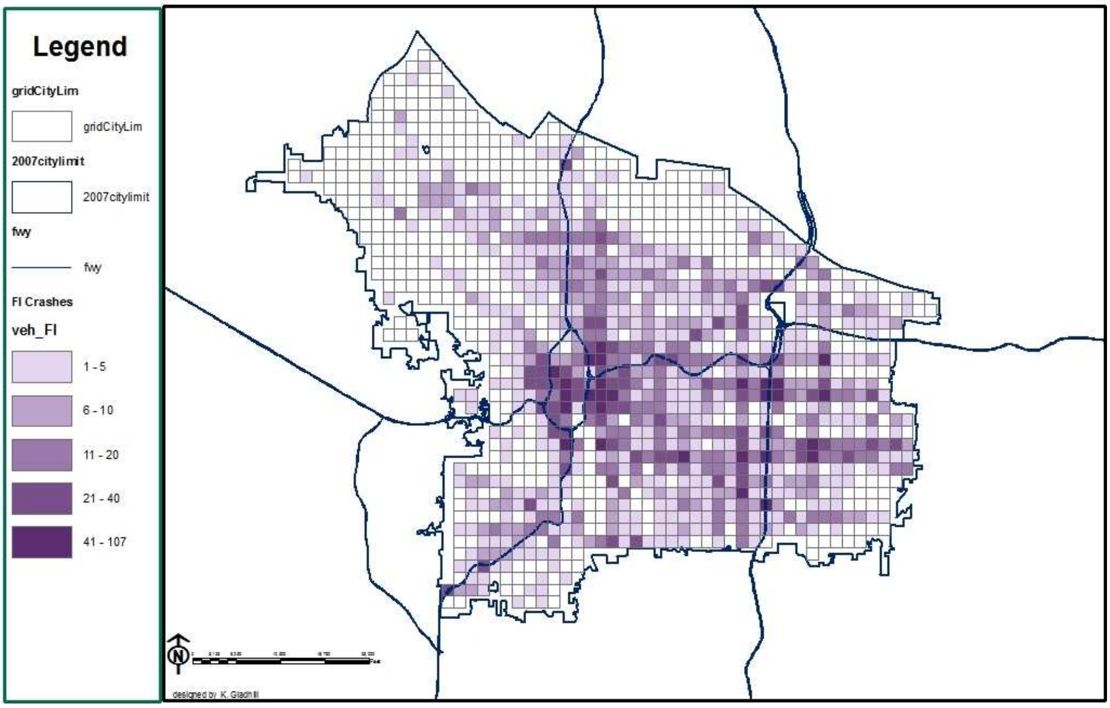


Figure 15   2005 – 2007 Vehicle Crashes, FI Severity

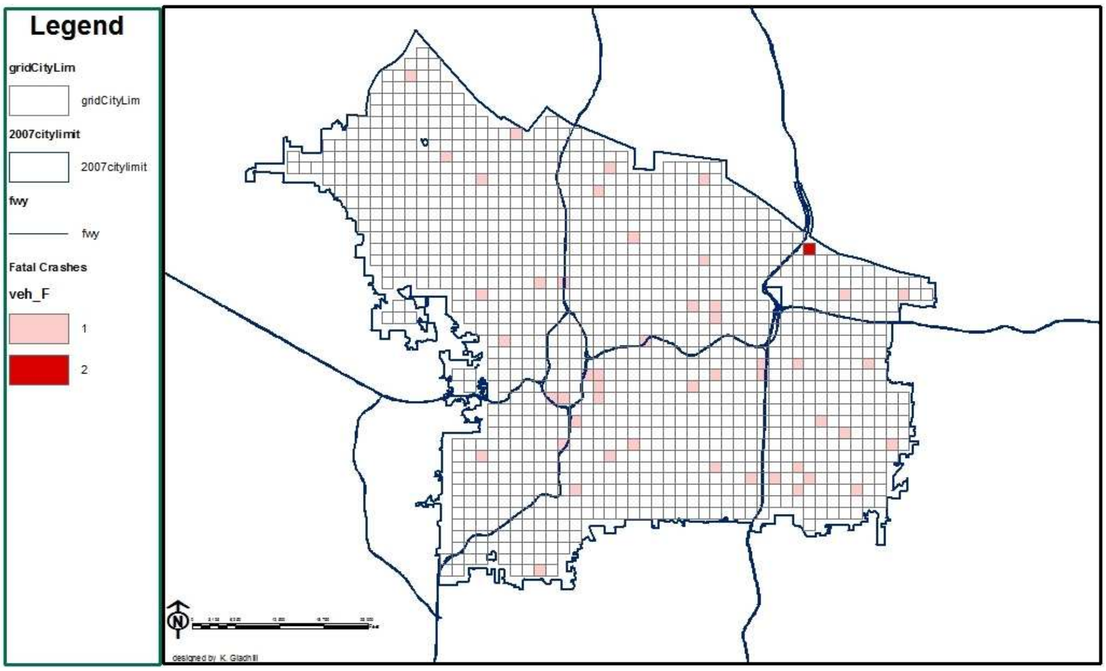


Figure 16 2005 – 2007 Vehicle Crashes, Fatal

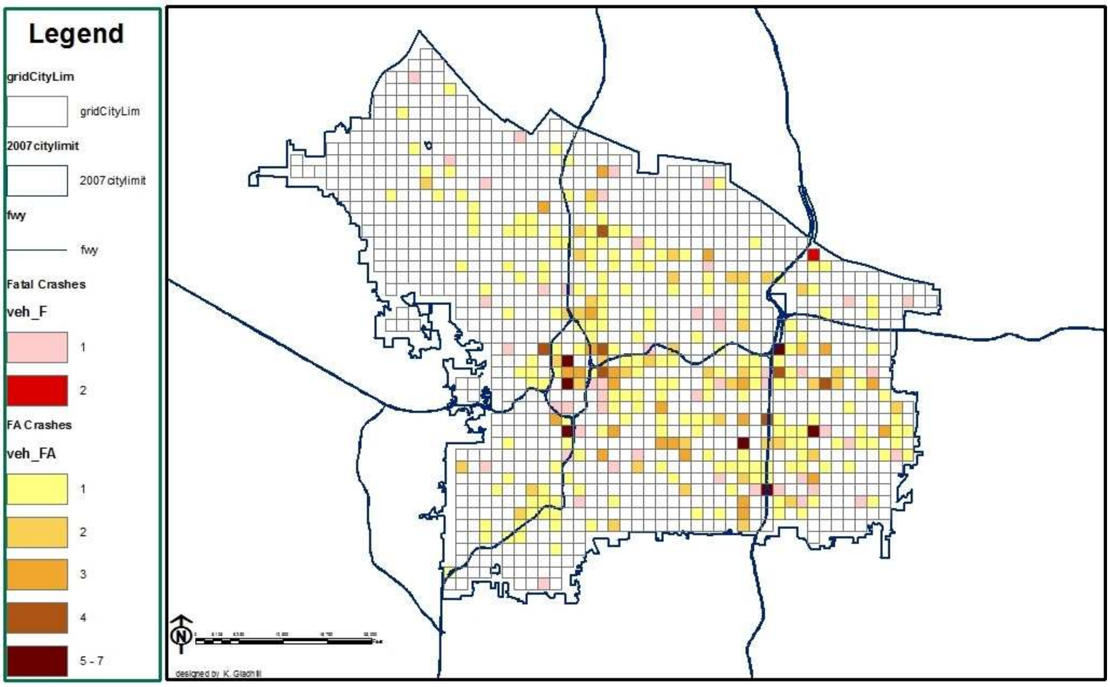


Figure 17 2005 – 2007 Vehicle Crashes, Fatal and Injury A Severity

4.2 Exposure

Higher exposure on the transportation network typically increases crash risk. This study considered exposure factors for vehicles, not for pedestrians nor bicycles. Exposure data were represented by volume to capacity ratios, VMT (vehicle miles traveled), average free speed, and presence of schools. Volume to capacity ratio (v/c) values within a spatial grid cell were averaged for each grid cell, see Figure 18. Downtown did not have the highest v/c ratios. Instead, some particular arterial corridors, such as Powell Boulevard in the southeast and Barbur to the southwest can be seen in the v/c spatial map. Note that some spatial grid cells had no v/c data. These cells were eliminated from the modeling dataset, reducing the cells for modeling consideration from 1284 to 928 total cells

Average free speed is shown in Figure 19, with major arterials having higher speeds and the highest average free speeds in outlying areas. These average free speeds are based on posted speeds, and are inputs to the travel demand model rather than modeled values.

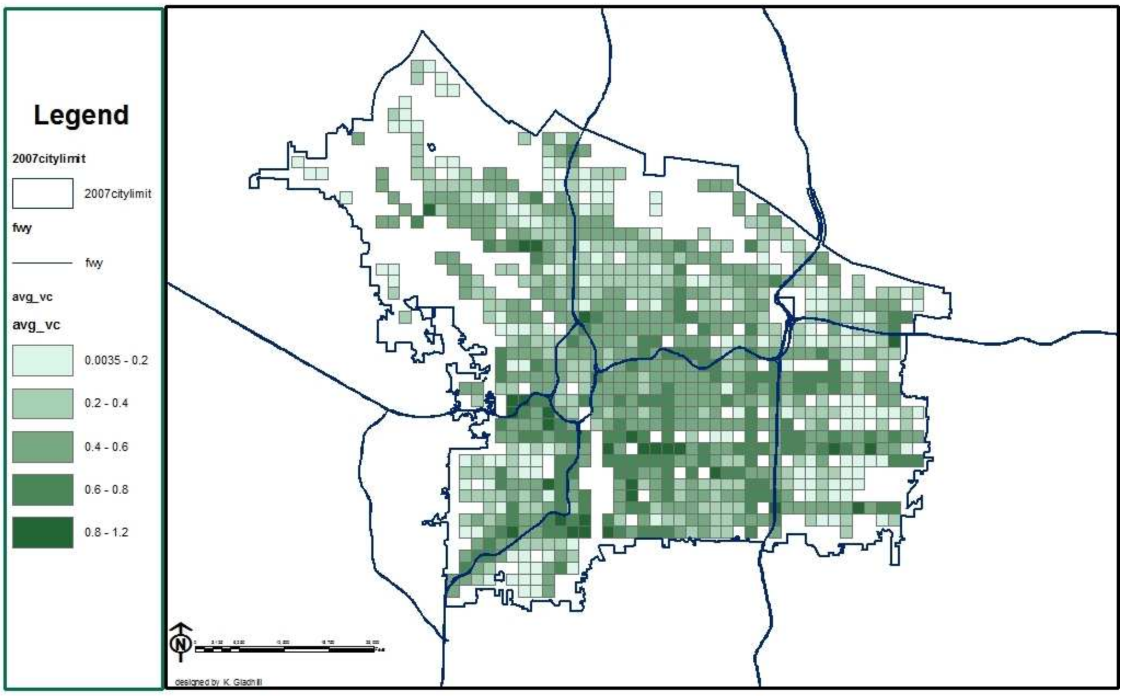


Figure 18 2005 Average of Volume to Capacity Ratios within a Spatial Grid Cell

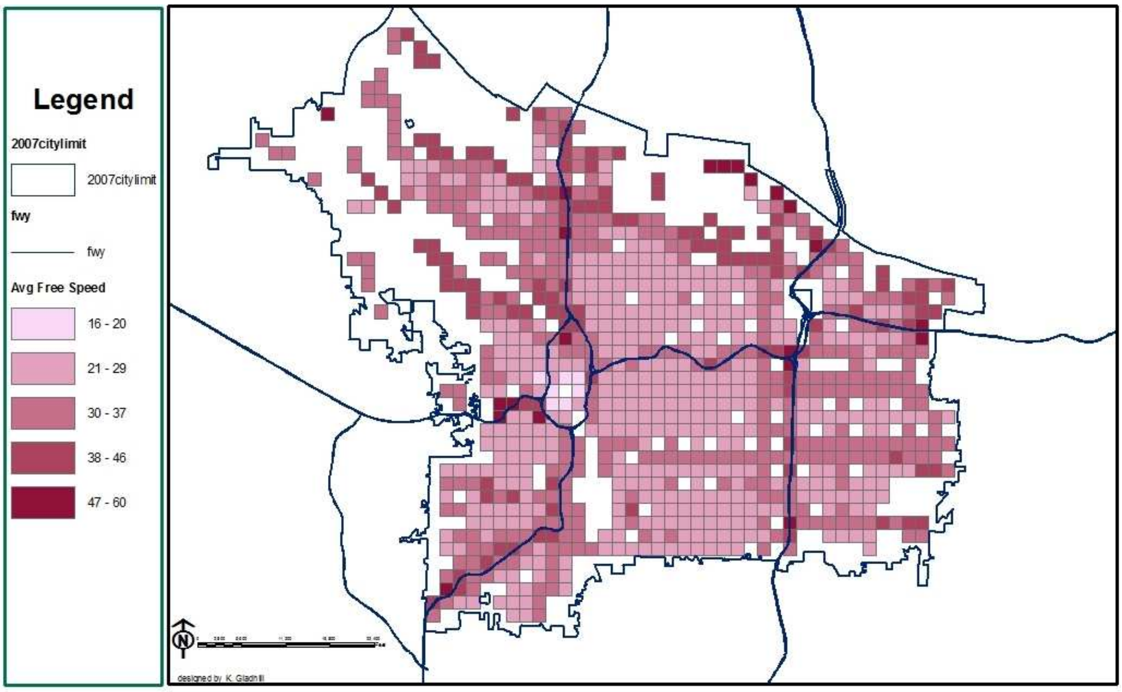


Figure 19 2005 Average Free Speed

4.3 Connectivity

Connectivity and urban form were looked at using several factors One was the four level street network scale developed by Rifaat and Tay (4) of loops and lollipops, mixed, warped parallel, or grid (see Figures 1 through 4), referred to as the "street network" factor in this study, Some spatial grid cells did not have a discernible type, and were assigned a value of zero. These cells were also removed from the modeling dataset, taking the set down to 792 spatial cells, still a considerable sample size. Figure 22 is a histogram of the frequency that the four different street network values were assigned.

The percentage of intersections in each spatial grid cell that were four leg intersections was also calculated. This would be 100% for full grid street layout, decreasing to a lower percentage as the street network becomes less grid like. Zero could be achieved if only 3 leg intersections are designed into a loops and lollipops style of development. The percentage four leg intersections could be calculated, and was thus less subjective than applying the street network scale in Figure 20. Percentage four leg intersections also has continuous rather than discrete numeric values, generally better for modeling.

The percentage four leg intersections correlated well with the Rifaat and Tay street network scales visually, see Figure 20 and Figure 21. Both factors were used in modeling.

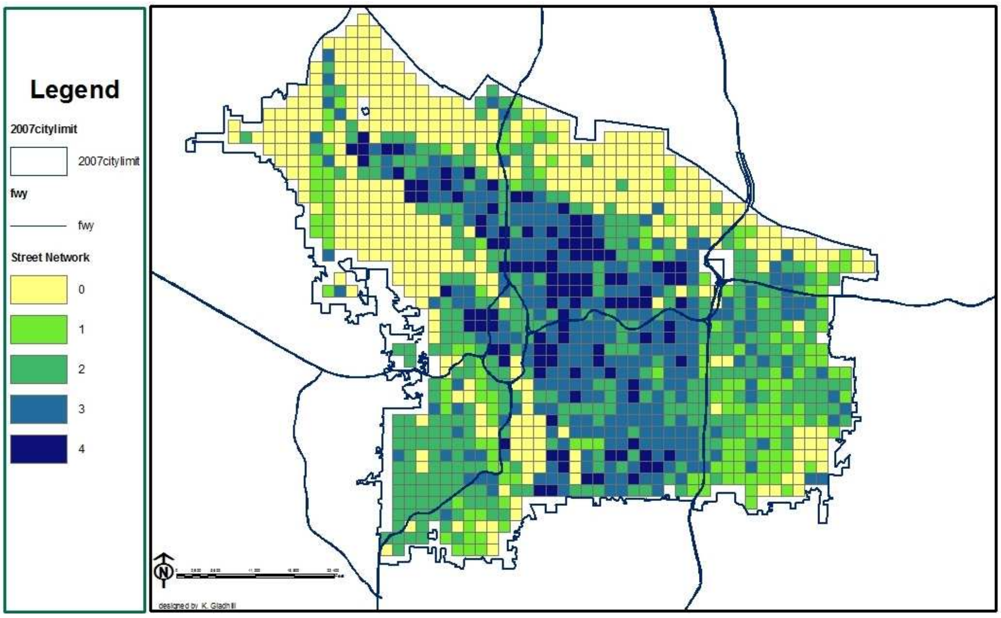


Figure 20 Street Network Value Assignments
1 = Loops and Lollipops 2 = Warped Parallel 3 = Mixed 4 = Grid

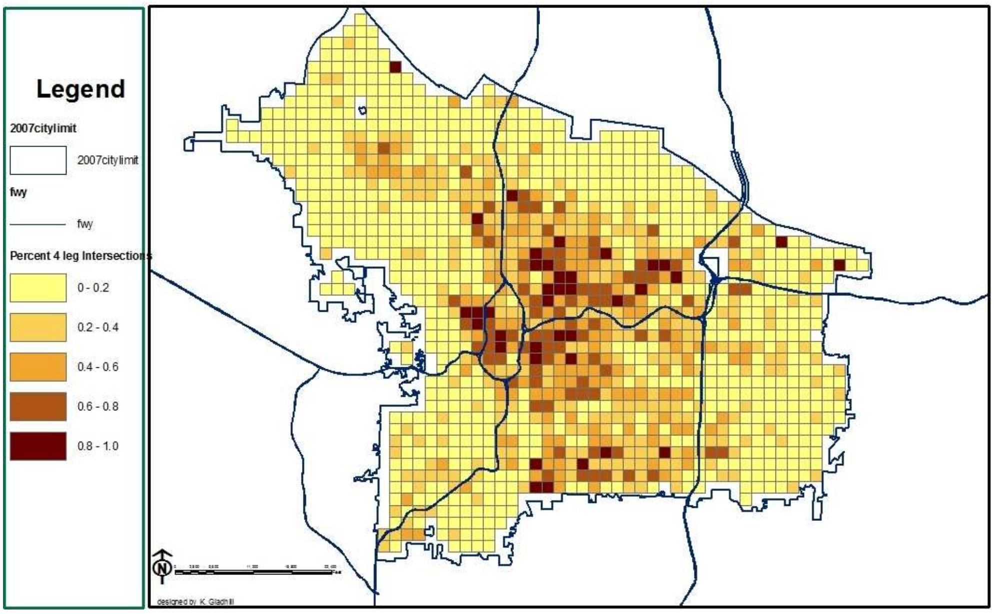


Figure 21 Percentage of Intersections which are Four Leg Intersections

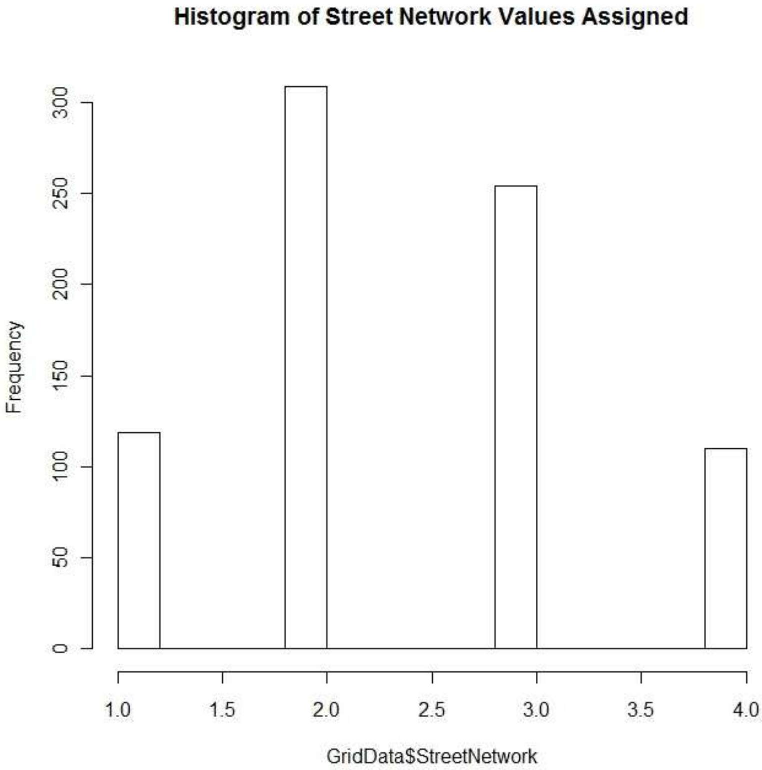


Figure 22 Histogram of Street Network Values
1 = Loops and Lollipops 2 = Warped Parallel 3 = Mixed 4 = Grid

Modeling also included total major arterial length as a factor, since it did not correlate to other street and intersection factors, see Appendix E, Figure E-1.

4.1 Transit Accessibility

Several factors were considered to represent transit accessibility. As can be seen in Figure 23, transit stops are widespread throughout the Portland city limits along major arterial routes and concentrated in the downtown area. Outlying areas have sparse or no transit service. This corresponds to population and housing density, shown in later figures, indicating a well planned transit system which has more service in areas of greater demand.

Total boardings were also considered, as well as route length.

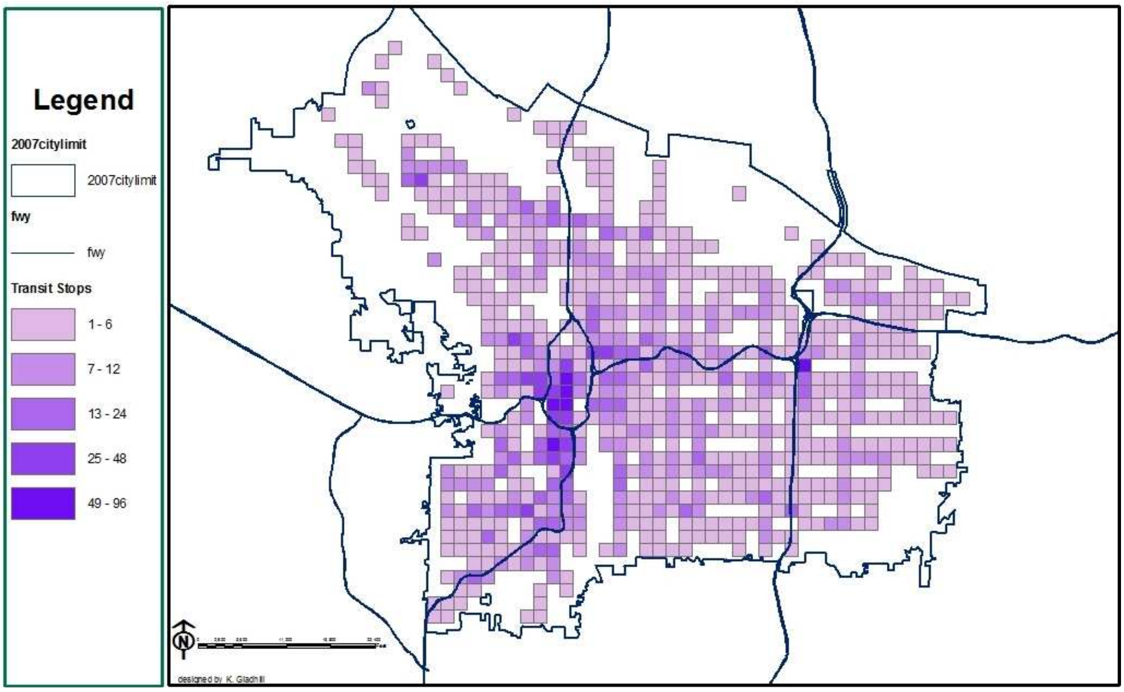


Figure 23 2007 Transit Stops Spatial Map

4.2 Demographics

The highest population was concentrated in the downtown area, (see Figure 24). There was a large area with high population density east of the Willamette River, and sparse population in outlying areas. Dwelling units (Figure 25) and households showed the same distribution.

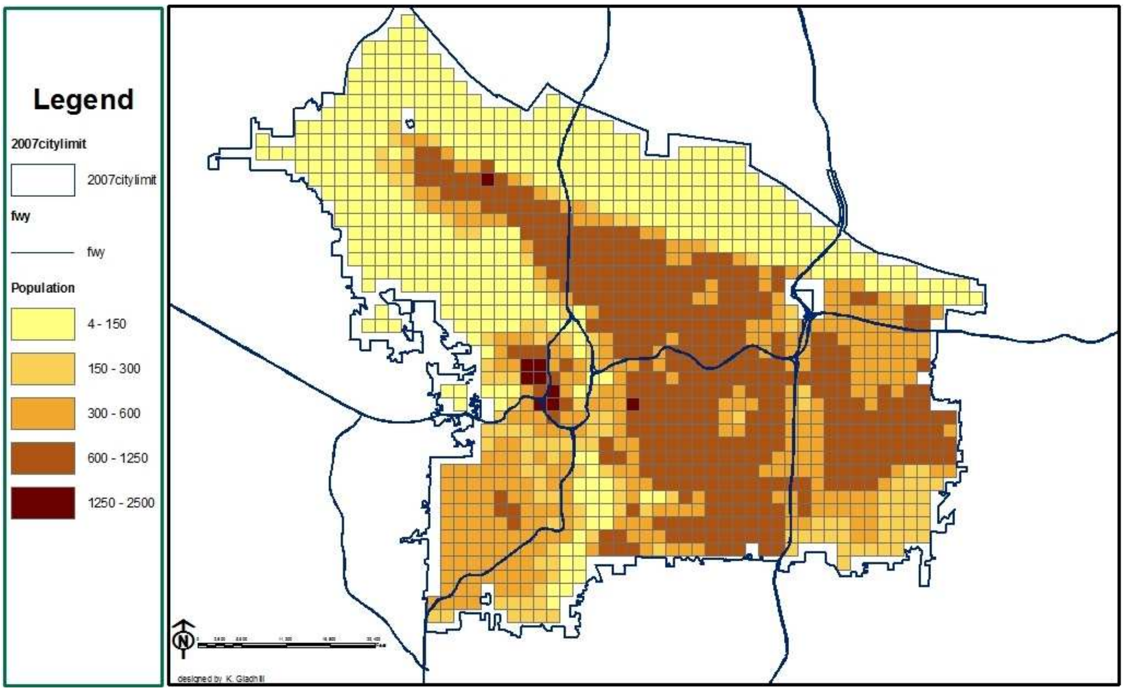


Figure 24 2006 Population

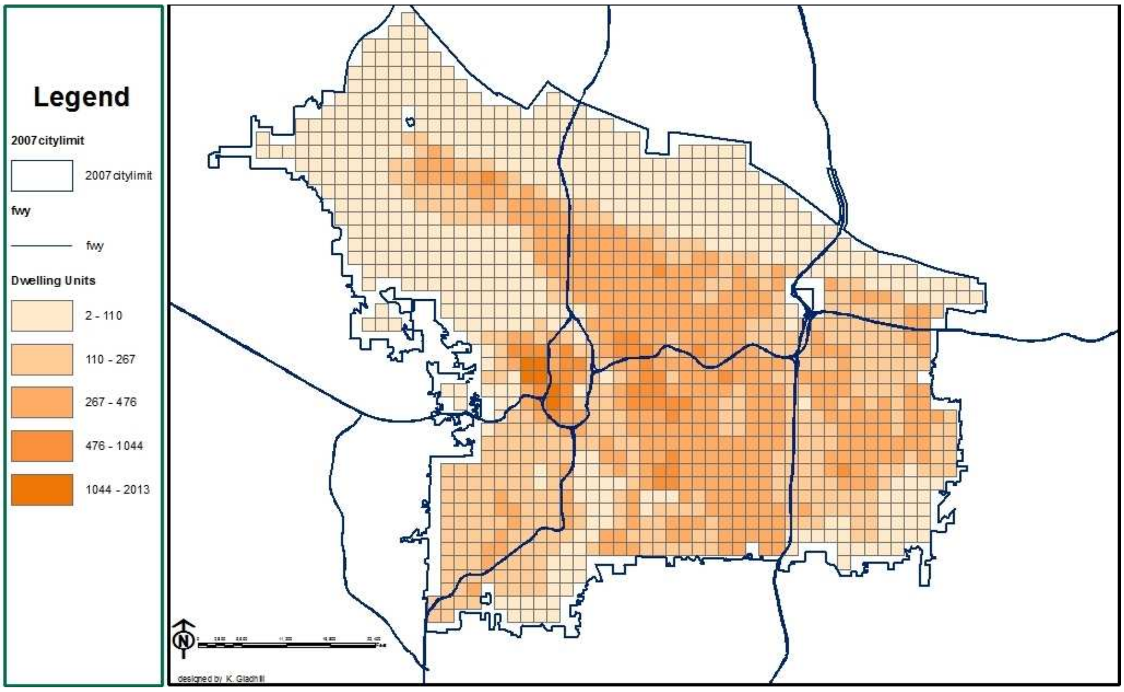


Figure 25 2006 Dwelling Units Spatial Map

4.3 Origins and Destinations

Employment (Figure 26) and the number of business establishments (Figure 27) were included to represent origins and destinations, since there is travel to get to and from work, as well as to patronize businesses of all sorts. Employment, or the number of employees, was highly concentrated around the downtown area, with some satellite areas. The number of business establishments also was concentrated heavily in the downtown area, but had a more diverse spread throughout the city, corresponding to the densely populated areas seen in Figure 24.

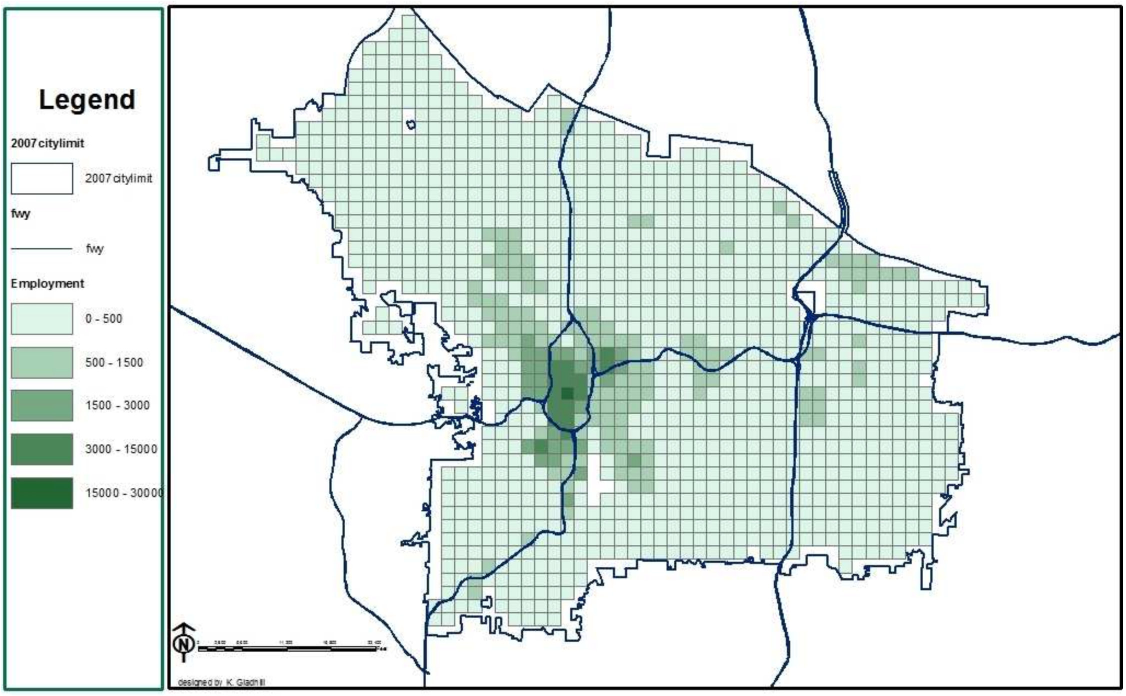


Figure 26 2005 Employment Spatial Map

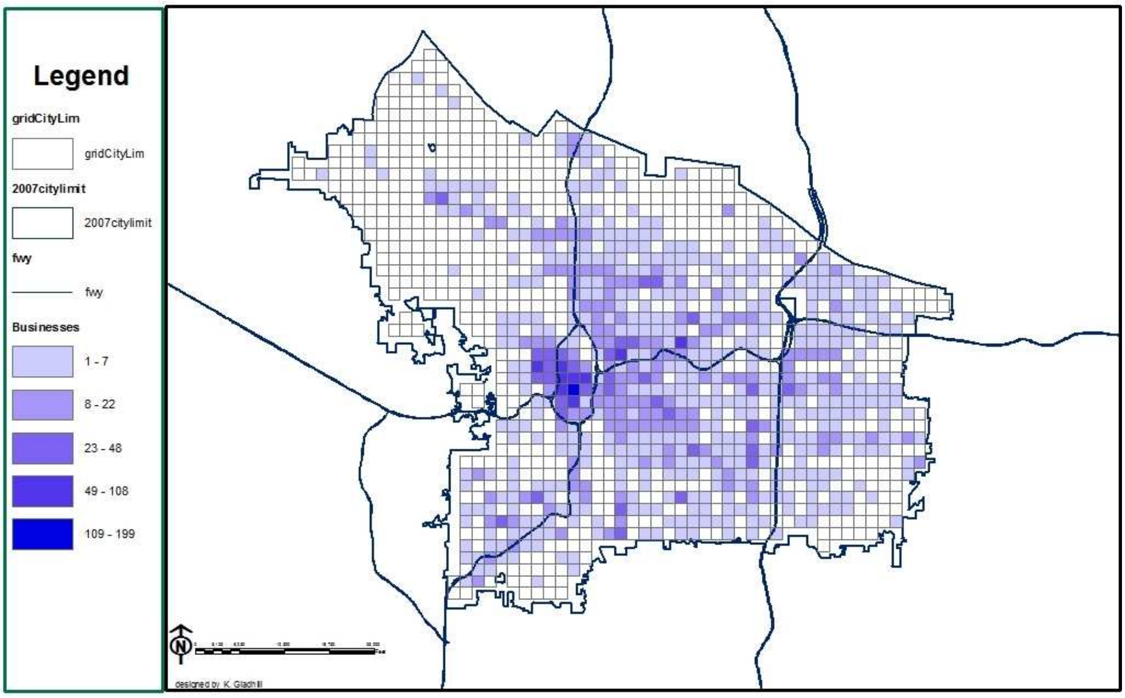


Figure 27 2010 Business Spatial Map

5.0 MODEL DEVELOPMENT AND RESULTS

5.1 Development

5.1.1 Summary of Input Data

Spatial grid cells with street length less than 3,300 feet, which would correspond to twice the length of the side of a cell, were eliminated as a lower bound on street length in a cell for urban form to be evident. Cells with no volume model data (VMT, v/c) were also eliminated. This left a dataset with 928 spatial grid cells for modeling. A further reduction was made to eliminate cells which had street network of zero, indicating that I had been unable to determine which of the four Rifaat and Tay street network categories it corresponded to. With this final reduction, the dataset was 792 spatial grid cells. A summary of the model input data for 792 spatial grid cells can be found in Table 6. Minimum, mean, maximum, standard deviation, and totals are shown.

5.1.2 Selection of Independent Variables

Many of the independent variables were correlated. Using Pearson's correlation and qualitative analysis, these were reviewed and reduced to a smaller set of factors to be used for modeling (see Appendix E). For example, since population was highly correlated with household and dwelling units, population was chosen to use in the modeling, representing all three demographic factors. Bus stops and routes were the majority of the transit stops and routes, and all highly correlated, so there was no need to duplicate these factors by including them separately: only the transit stops factor was left in the model. Transit boardings and alightings, (ons and offs), were highly correlated; the sum of ons and offs was included in the model building. Street length and intersection factors showed cross correlation. Street network street length, major arterial length, percent four leg intersections, total intersections, and block size represented connectivity factors.

Table 6 Modeling Input Data Summary

	Min.	Mean	Max.	Median	std dev	TOTAL
CRASH DATA						
Vehicle Crashes	0	24.82	287	13	34.56	19658
Pedestrian Crashes	0	0.5947	9	0	1.222	471
Bicycle Crashes	0	0.649	16	0	1.527	514
Angle Crash Type	0	8.963	139	4	15.48	7099
Turn Crash Type	0	1.726	20	0	2.92	1367
Straight Crash Type	0	2.518	44	1	4.62	1994
Stop Crash Type	0	7.304	70	3	11.16	5785
Fatal Crashes	0	0.05051	1	0	0.219	40
Fatal, Injury A Crashes	0	0.5896	7	0	1.01	467
Injury Crashes	0	9.068	107	5	12.71	7182
Fatal & Injury Crashes	0	9.119	107	5	12.74	7222
PDO Crashes	0	15.7	180	8	22.48	12436
EXPOSURE						
VMT	2.392	2045.78	14623.3	10328	2680	1.62E+6
v/c	0.0035	0.438	1.215	0.4547	0.201	
Avg Free Speed	15.5	30.23	60	29	5.83	
schools count	0	0.2803	4	0	0.586	222
CONNECTIVITY						
Street Network scale	1	2.448	4	2	0.909	
intersections	0	23.67	81	22	11.34	18750
fourLegPct	0	0.311	1	0.2697	0.229	
nonFwy Street	3649	12682	28697	12481	3789	10E+6
Major Arterial length	0	1535.7	18547.6	212.8	2471	1.22E+6
block Size	9	61.39	325	55	33.09	
ACCESSIBILITY						
Transit Stops	0	5.987	87	4	8.09	4742
On + Off	0	173305	8715715	27212	592000	1.4E+08
DEMOGRAPHIC						
Population	4.1	559.3	2475.7	567.7	303	442939
ORIGINS AND DESTINATIONS						
Employment	0.39	149.25	346.32	157.24	936	319731
Business density	0	5.77	108	2	10.73	4570

5.1.3 Model Building

Since crashes are essentially a failure event, crash data do not follow a normal distribution. Poisson modeling can be considered, but the likelihood of many zero values recommends that a check be made for over dispersion. Poisson modeling with the study data confirmed that the data were over dispersed, so negative binomial regression would be an appropriate model (15), of the form:

$$N_{crash} = \exp(a + b*x_b + c*x_c + \ldots + n*x_n)$$

where N_{crash} = number of crashes

a through n = coefficients

x_i = factor i value

VMT provided a better fit as log(VMT), equivalent to putting VMT before the exponent in the formula, which was consistent with crash models that are have volume outside the exponential term:

$$N_{crash} = VMT * \exp(a + b*x_b + c*x_c + \ldots + n*x_n)$$

where N_{crash} = number of crashes

VMT = vehicle miles traveled

a through n = coefficients

x_i = factor i value

The street network factor was included in the negative binomial models as a factor, having discrete values of:

1 = loops and lollipops

2 = mixed

3 = warped parallel, and

4 = grid street layout

which compared each of the other levels to level 1 = loops and lollipops.

Negative binomial models were developed for vehicle crashes, pedestrian crashes, and bicycle crashes separately. Preliminary analysis had indicated that the top four crash types were angle crashes, turning crashes, straight, and stopped vehicle crashes. Separate models were developed for each of these crash types.

Separate models were also developed for levels of crash severity. Modeling was unsuccessful on fatalities alone due to the small number of non-zero data points. Crash severity was looked at for several groupings: "FA" grouped fatal and injury level A (incapacitating) together, the most severe crash injuries; "FI" grouped fatal and all three injury levels (A = incapacitating, B = non-incapacitating, C = possible injury) together; "I" grouped the three injury levels, and PDO (property damage only crashes).

Negative binomial regression models were developed in a step-wise fashion by adding one factor at a time using the cumulative regression method described by Banfro and Hauer (16) to decide whether the addition of a factor improved the model. This method involved inspection of the regression diagnostic plots (cumulative residual plots) for each model as the new factor was added. The standard deviation for the 2 sigma limits was calculated as follows:

$$\sigma^{*2} = \sigma^2(n)[1 - \sigma^2(n)/\sigma^2(N)]$$

where σ^* = standard deviation,
$\sigma(n)$ = sigma for the current value
$\sigma(N)$ = sigma for all values

The cumulative residual plots were judged as to whether the cumulative residuals were within the $\pm$ 2 sigma limits, the 2 sigma limits were getting narrower, the cumulative residuals were centered around zero, and the final cumulative of residuals was closer to zero, since the cumulative residuals should theoretically sum to zero. If the cumulative regressions looked better with the second factor, the factor was kept in the model. If not, the factor was dropped from the model. Figure 28 shows that adding the block size factor into the vehicle crash model improves the cumulative regression. R code for developing these CURE plots is included in the Appendix.

Other model fit plots were also examined. As seen in Figure 29, the standard deviation residuals Q-Q plot should be approaching a straight line.

Finally, a negative binomial regression was also run with all the modeling factors, followed by a case with the significant factors from the all factors model. The model from the "add a factor" approach was then compared to the "all factors" and "all significant factors" models by comparing cumulative residual charts for non-freeway street length, and the best one selected for the results.

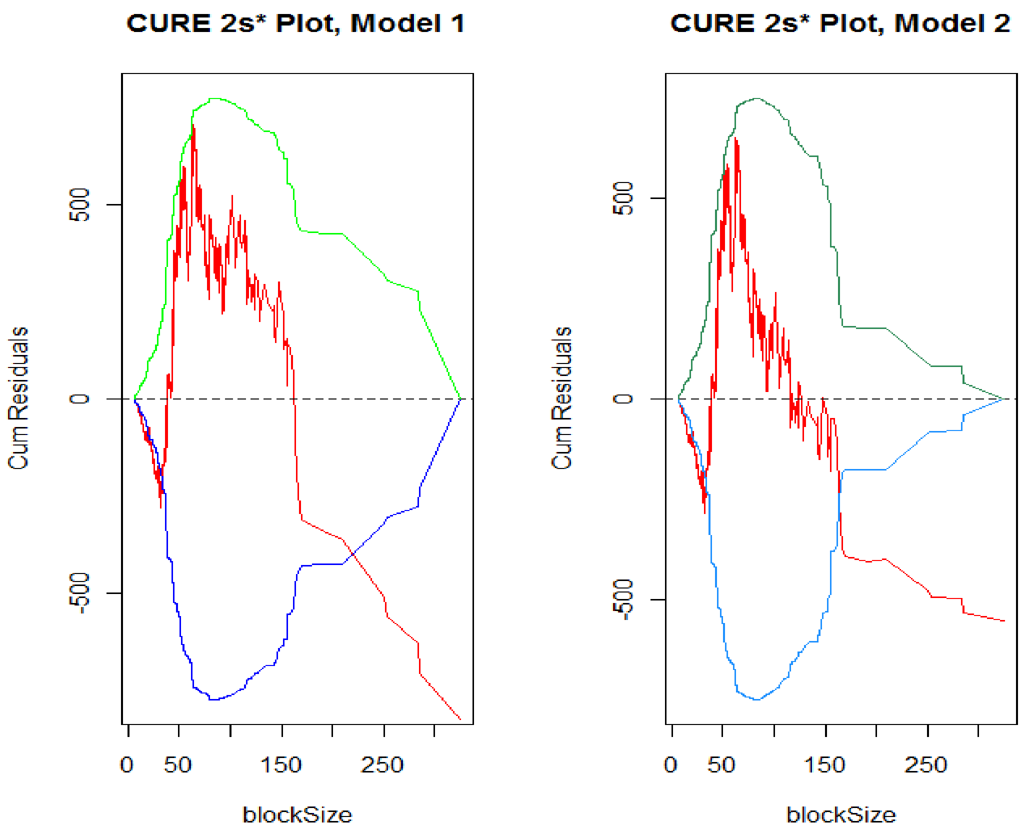


Figure 28 Cumulative Residual CURE 2 sigma Plots for Comparison

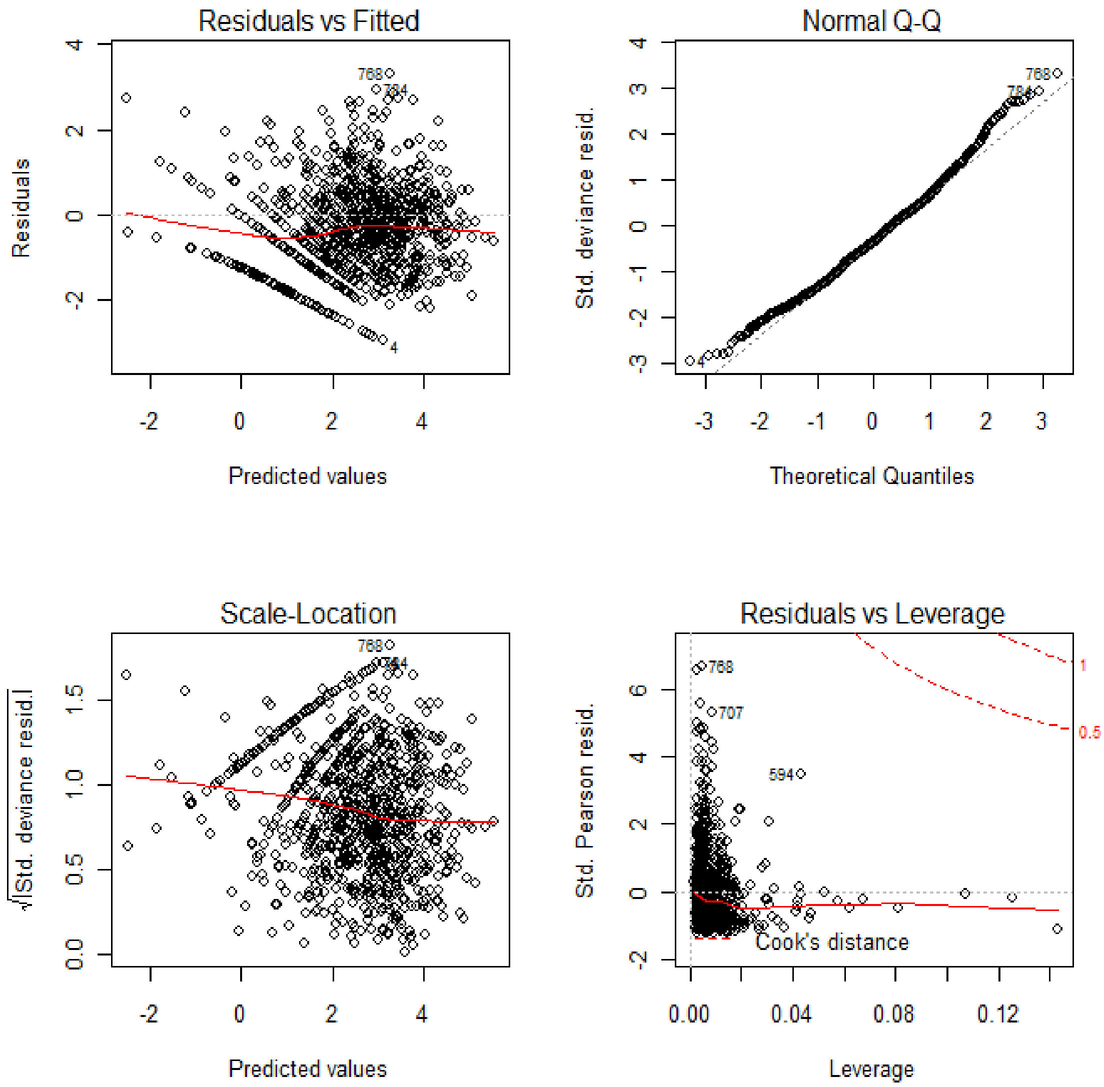


Figure 29 Plots for Negative Binomial Regression Fit

5.2 Model Results

Table 7 shows goodness of fit for the negative binomial modeling results. All models had the same sample size, 792 spatial grid cells. Note that the models have better fit where there are more crash data: vehicle crashes, angle and stopped vehicle crashes; all injury crashes, and PDO crashes. The standard error is low, and estimated R2 values look good for crash modeling

Table 7 Negative Binomial Modeling Goodness of Fit

	AIC	theta	std error	2 x log-likelihood	null deviance	null dev degrees of freedom	residual deviance	res dev degrees of freedom	R^2 estimate
Crash Group									
Vehicle	5854	2.037	0.119	-5920	2298	791	875	776	0.62
Pedestrian	1393	1.43	0.27	-1367	909	791	575	780	0.37
Bicycle	1482	1.03	0.167	-1461	884	791	590	782	0.33
Crash Type									
angle	4376	1.599	0.11	-4346	1986	791	881	778	0.56
turn	2478	0.76	0.07	-2453	1037	791	713	781	0.31
straight	2656	1.248	0.12	-2628	1539	791	745	779	0.52
stop	4081	1.163	0.078	-4055	1711	791	839	778	0.51
Crash Severity									
FA	1533	1.785	0.387	-1517	821	791	663	785	0.19
FI	4541	1706	0.116	-4510	1812	791	889	778	0.51
I	4526	1.715	0.117	-4494	1822	791	885	777	0.51
PDO	5128	2.143	0.136	-5098	2329	791	862	778	0.63

5.2.1 Estimated Coefficients

The model coefficients are presented in Table 8: negative coefficients are in red type; statistical significance is indicated by shading, see key for the table.

5.2.1.1 Exposure

Looking at exposure factors in Table 8, VMT is significant in all the models. This would be expected, since the relationship of increased crash rates with increased traffic volume are well established. The volume to capacity ratio, or v/c, is not significant for all models, however. Models of bicycle crashes, angle, and turn crashes, and the higher severity injuries (FA combines fatalities and incapacitating injury severity) did not show this factor as significant.
V/c could be seen as an indication of congestion, and as such indicates that congestion is less important for crashes of those types. It makes sense that congestion, where more stop and go behavior increases the risk of incidents with one vehicle overtaking another vehicle, would contribute to increased straight and stop crashes.

Average free speed was significant for all three different crash mode models, and for turn and PDO crashes. What is interesting is that the coefficient is negative, implying that a higher speed would decrease the crash risk; typically higher speeds are associated with increased crash rates. This could tie in with congestion: if speeds are lower due to increased congestion, that could explain why lower speeds were less safe. Alternatively, it could be due to the quality of the data source, which is based on posted rather than actual speed limits.

Schools were significant for a few crash models: vehicle crashes, angle crashes, FI and I. This again had a negative coefficient, meaning the more schools in an area, the less crashes predicted. This could indicate that efforts to improve traffic safety around schools are indeed effective, slowing drivers down and making them more alert to the potential for pedestrians in the area. With this in mind, it's interesting to note that schools are not significant for the pedestrian crash model, which is the type of crash most school traffic safety policies are targeting.

5.2.1.2 Connectivity Factors

No connectivity factors were statistically significant for vehicle or pedestrian crashes. Several factors were significant for bicycle crashes, though: street network layout of warped parallel and grid network compared to loops and lollipops; intersection density, and total street length. This indicates that bicycles may have a greater crash risk with more streets, particularly more grid like. This may be due to the need to cross intersections, which is indicated by the intersection coefficient being the highest coefficient in the bicycle model.

The percentage of four leg intersections was significant for angle crashes. This is logical due to the fact that four leg intersections have more opportunities for angle crashes. Major arterial length was significant for stopped vehicle crashes and injury crashes; block size was also significant for stopped vehicle crashes. This makes sense because major arterials are more likely to have traffic signal control, and shorter block lengths would mean more frequent stop lights, thus more opportunity for rear end collisions.

5.2.1.3 Transit accessibility, Demographics, and Origin Destination

Transit accessibility and origin and destination factors were significant in many models. Transit stops were significant in vehicle, crash severity, and crash type models. Transit stops correlate to employment (see Appendix E), but employment was not significant in most models, only the vehicle crash model, where is had a negative coefficient. Pedestrian crashes showed transit ons and offs to be significant; which is logical since transit riders are pedestrians immediately before and after their transit trip: the higher pedestrian volumes would be expected to increase the possibility for pedestrian crashes.

Population was not significant for bicycle crashes nor for FA injury severity crashes, but was for all other crash models. Business density was significant in all the models. Many trips are due to people getting to and from work, as well as frequenting businesses, so finding these factors significant for crash rates indicates another aspect of exposure to the transportation network.

Table 8 Model Coefficients

	Vehicle	Ped	Bicycle	angle	turn	straight	stop	FA	FI	I	PDO
Intercept	-1.909	-4.974	-4.369	-2.810	-5.417	-5.828	-4.558	-4.100	-2.846	-2.677	-2.475
Exposure											
log(VMT)	0.534	0.381	0.527	0.463	0.743	0.771	0.684	0.412	0.515	0.491	0.535
average v/c	0.661	1.375		0.115		0.678	1.653		0.700	0.730	0.680
AvgFreeSpeed	-0.012	-0.015	-0.054	-0.009	-0.026	-0.009	-0.003	0.008	-0.004	-0.006	-0.013
Schools	-0.093			-0.129					-0.113	-0.113	-0.082
Connectivity											
Street Mixed	0.140	0.136		0.256	0.395	0.060	0.065		0.084	0.107	0.163
Warped Parallel	0.116	0.138	0.518	0.246	0.498	0.226	0.034		0.108	0.162	0.167
Street Grid	0.075	0.049	0.659	0.299	0.459	-0.147	0.062		0.075	0.166	0.186
Intersections	0.005	0.005	0.719	0.005		-0.004	0.003		0.001	-0.008	0.001
FourLegPct	0.299			0.847		0.041		0.518	0.107	-0.070	0.146
non-Freeway Street	1.7E-05		8.0E-05	1.9E-05	2.6E-05	2.0E-05	-1.4E-5		1.4E-05	1.3E-05	3.3E-05
Major Arterial length					-5.8E-05						
Block Size										4.2E-03	
Transit Accessibility											
Transit Stops	0.020	-0.002	0.012	0.016	0.016	0.011	0.023	0.014	0.015	0.015	0.015
Ons & Offs	5.6E-08	2E-07									
Demographic											
Population	4.9E-04	1.3E-03		7.3E-04	6.9E-04	3.7E-04	6.6E-04	-8.9E-5	6.7E-04	7.3E-04	4.8E-04
Origin-Dest											
Employment	-1.77E-04										
Business	0.033	0.029	0.018	0.032	0.016	0.035	0.023	0.012	0.026	0.026	0.030
Significance	0.001		0.01		0.05		0.1				

5.2.2 Elasticity

Elasticity gives a measure of relative effect of a factor on the outcome. Elasticity was calculated as follows (17):

$$E_{xj} = \beta_j x_j$$

where E_{xj} is the elasticity for attribute *j*,
β_j is the model coefficient and
x_j is the mean value for attribute *j*.

5.2.2.1 By Vehicle, Pedestrian, and Bicycle Models

Looking at elasticity for vehicle crashes in Figure 30, VMT had the largest affect. For every additional one percent VMT, there would be an additional 1.7% increase in vehicle crashes. Street network, non-freeway street length, population and business density had the next strongest affects on crash rate, 0.3 elasticity for mixed and warped parallel compared to loops and lollipops; a lower elasticity of 0.2 for grid layout. Population and business density elasticity were both under 0.3.

The highest elasticity for pedestrian crashes was 1.3 for VMT, and 0.74 for population. Bicycle crashes have high elasticity for VMT, warped parallel and grid street network, intersections, and non-freeway street length. Average free speed had a high negative elasticity of -0.44., indicating that the model predicts less crashes with increased speed. This could be explained if the lower free speed is due to heavier congestion and thus higher likelihood of bicycle vehicle interaction, or due to bicyclists avoiding high speed roadways if a lower speed alternate route is available.

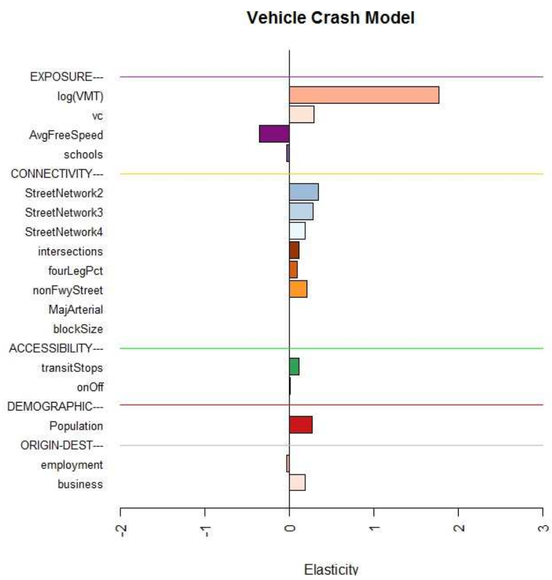
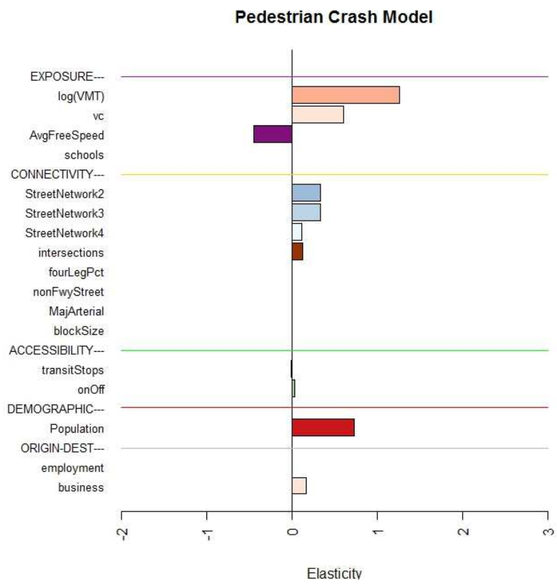


Figure 30a



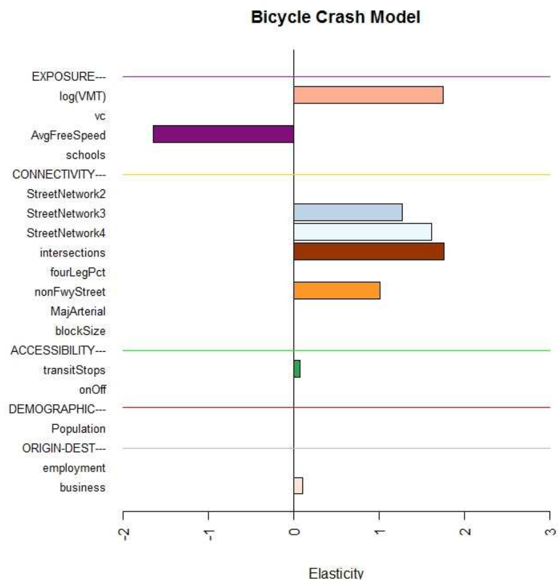


Figure 30b **Figure 30 c**

Figure 30 Elasticity for Vehicle, Pedestrian, and Bicycle Crashes

5.2.2.2 Vehicle Crash Types

Elasticity for crash type models are shown in Figure 31. For all the major crash types, VMT had the highest elasticity. Angle and turn crashes showed positive elasticity for the street layout factors. These crashes are more likely to occur with more intersections; particularly angle crashes at four leg intersections, so this is not surprising. Straight and stopped vehicle crashes had less elasticity for street layout and connectivity factors.

All four crash type models showed a positive elasticity for population and businesses. They also all showed a negative elasticity with average free speed, particularly turn crashes. Lower speeds may encourage drivers pull out of driveways or cross streets when they think they have adequate gap to pullout, but they really don't; whereas higher speed roadways may intimidate drivers to wait for safer gaps. The business count may indicate something about driveways: more businesses typically require more access points, unless design specifically limits access. So the negative elasticity for average free speed and positive elasticity for business density may be related.

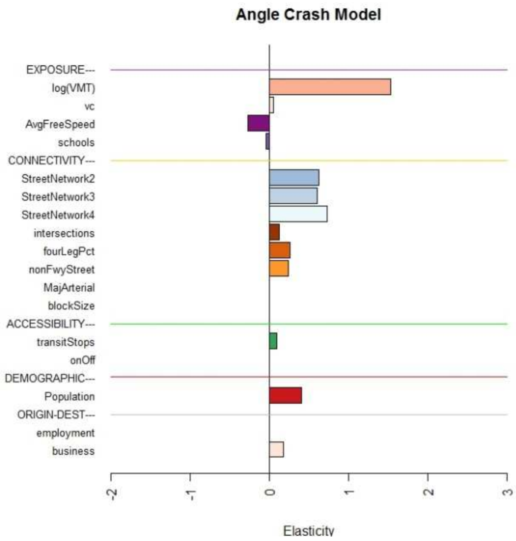


Figure 31a

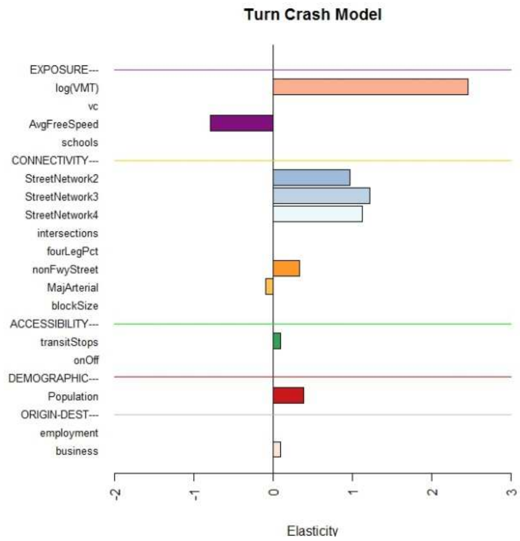


Figure 31b

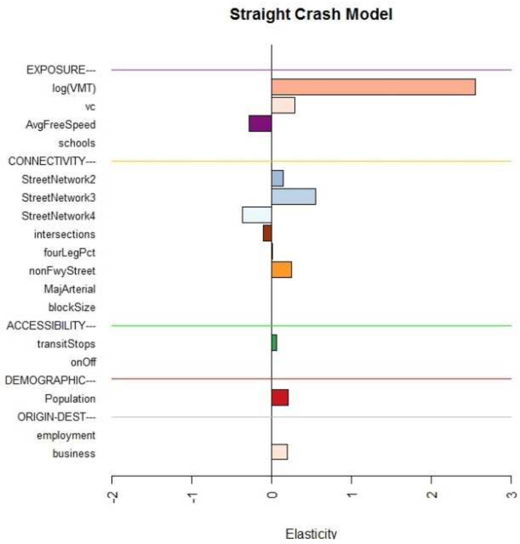


Figure 31c

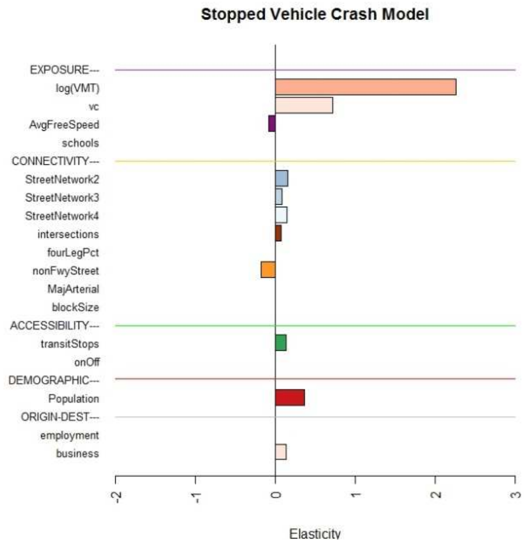


Figure 31d

Figure 31 Elasticity for Vehicle Crash Types

5.2.2.3 Crash Severity

Elasticity for models by crash severity is shown in Figure 32. Again, VMT had the highest elasticity for all crash severities. V/c showed some positive elasticity in the less severe injury categories, but was not in the model for FA, since the factor did not improve that model. FA crashes had few significant factors: VMT, average free speed, percent four leg intersections, transit stops, and business density. This set of significant factors for FA crashes, including the only positive elasticity for average free speed, brings to mind a busy roadway where many things are going on: lots of traffic, moving fast, with transit vehicles and riders in the mix; four legged signalized intersections where drivers may be going through beyond their signal phase; and high business density meaning more access points where vehicles are entering and leaving the transportation network. There are many things going on that could take a driver's attention away from the task of safe driving in this situation, with high speed contributing to less reaction time.

Street network factors had positive elasticity for all three less severe crash models, as do population and businesses. Average free speed was again a negative elasticity.

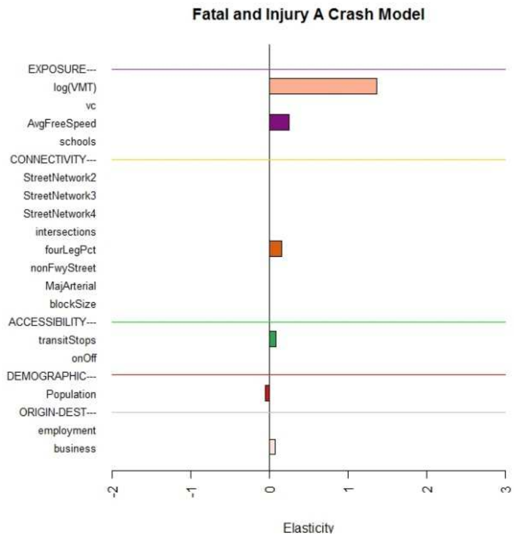


Figure 32a

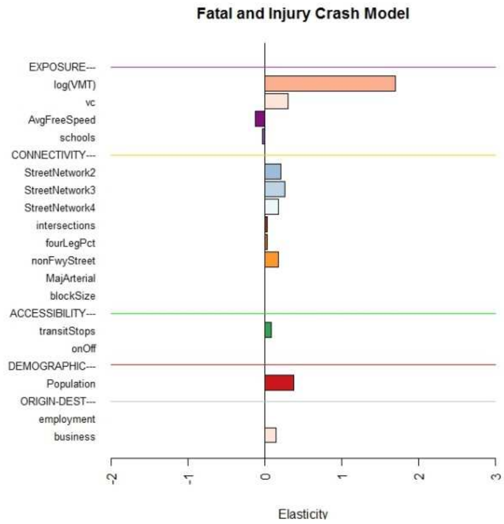


Figure 32b

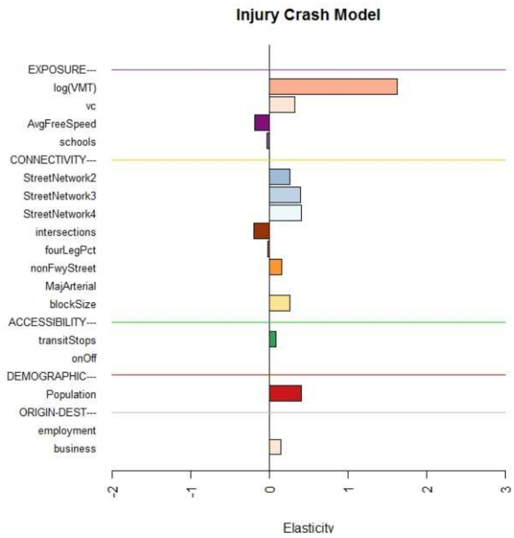


Figure 32c

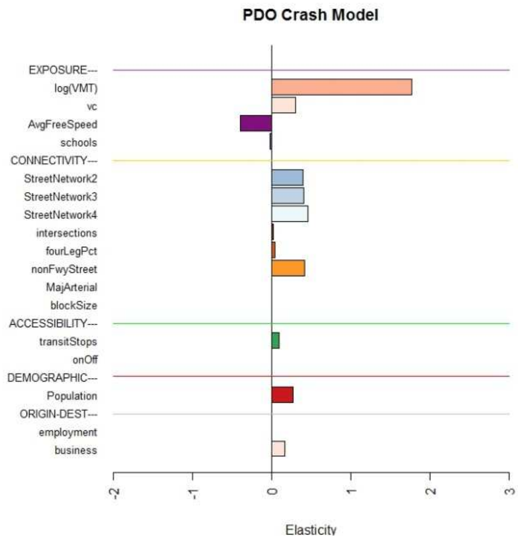


Figure 32d

Figure 32 Elasticity for Crash Severity

5.2.2.4 *Elasticity by Factors*

Looking at elasticity by factors in Figure 33, the elasticity for VMT was high for all models, particularly for turn, straight, and angle crashes. This indicated that increased traffic in an area greatly increased the probability of crashes. Straight crashes had the highest elasticity. This may be due to increased VMT in urban areas usually being due to increased congestion, causing vehicles to pack themselves into tighter space, with less distance between vehicles. Decreased distance between vehicles gives a driver less time to respond to a sudden deceleration of a vehicle in front of them, which could lead to a straight or stopped vehicle crash.

Volume to capacity (v/c) had highest elasticity in the pedestrian and stopped vehicle crash models. Increased v/c could mean more congestion, leaving less space for pedestrians and vehicles to stay separated. The less severe and non-injury crash models also showed sensitivity to v/c.

Average free speed had negative elasticity for every model except FA. Bicycle crashes had the strongest negative elasticity, which seems counter intuitive. This may be due to bicyclists tending to prefer lower speed routes to riding alongside higher speed traffic. In lower speed conditions, bicycle riders may take advantage of their size and squeeze between or past vehicles stopped at traffic signals or stop sings, putting them at greater risk of collision. The most severe injuries, FA crashes, had the only positive elasticity for average free speed, strongly suggesting a link between speed and severe injury.

Elasticity for street network compared to loops and lollipops can be seen in Figure 34. These factors had especially high elasticity for bicycle and turn crashes, as was discussed going through elasticity by model. Bicycle crashes also had high elasticity for intersection density, percent four leg intersections, and total street length (see Figure 35).

Population elasticity was highest for the pedestrian crash model. This makes sense since the more people there are, the more pedestrians . Population elasticity was higher for the less severe crash types.

Similarly business had positive elasticity as well, with less of an effect on FA crashes than other severity levels. The positive elasticity all models may be due to the fact the a higher count of businesses probably means more smaller businesses, which would each have a separate location, and thus more driveways for people to get to and from the businesses. Driveways have been associated with increased crash risk.

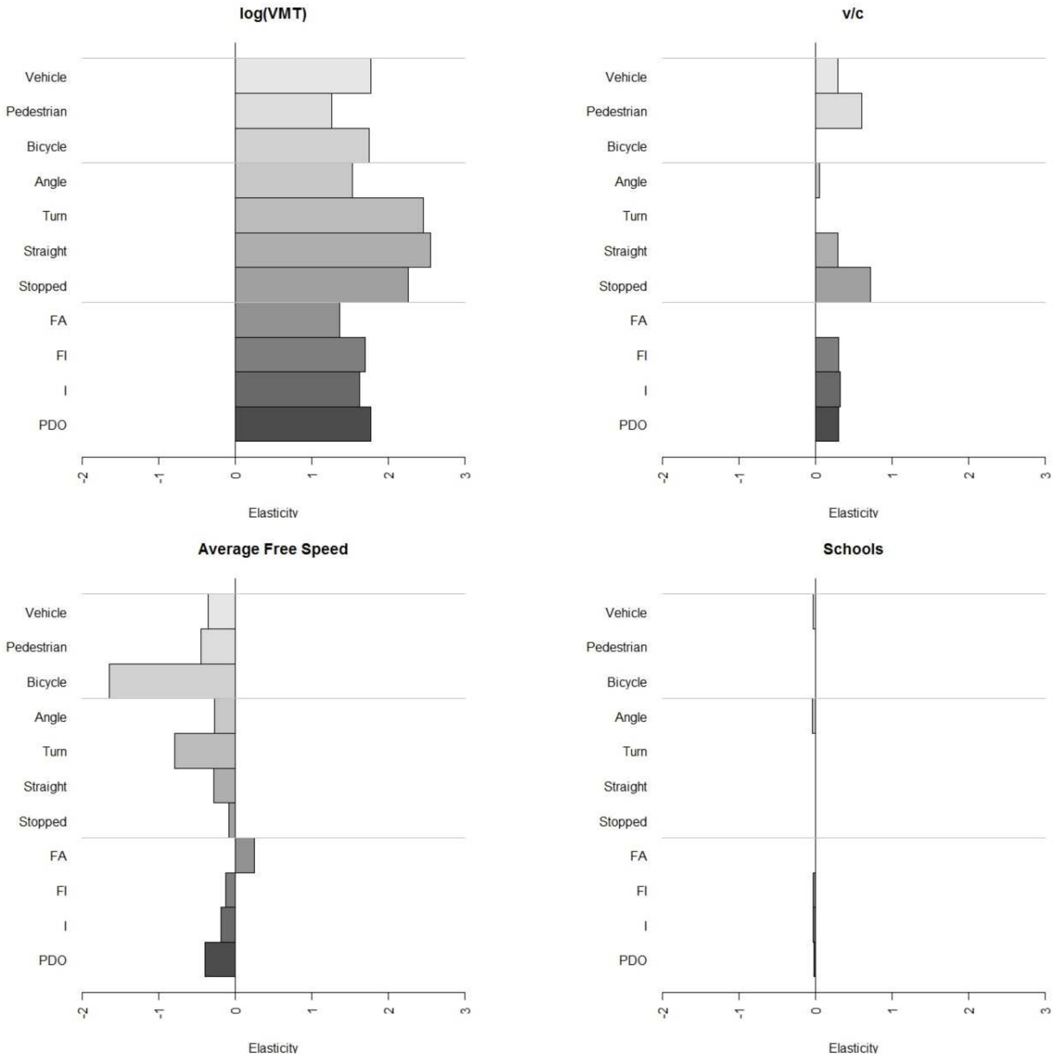


Figure 33 Elasticity for Exposure Factors

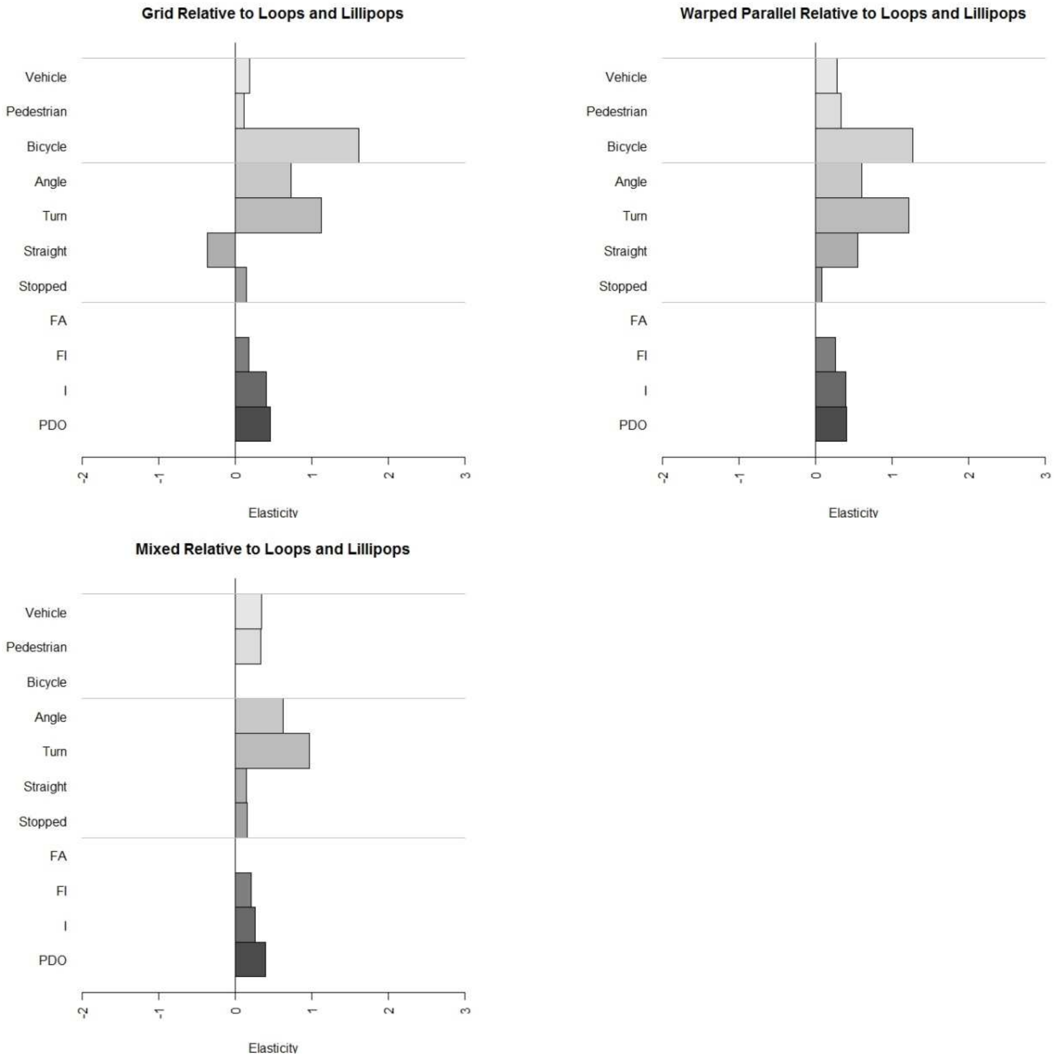


Figure 34 Elasticity for Street Network Factors

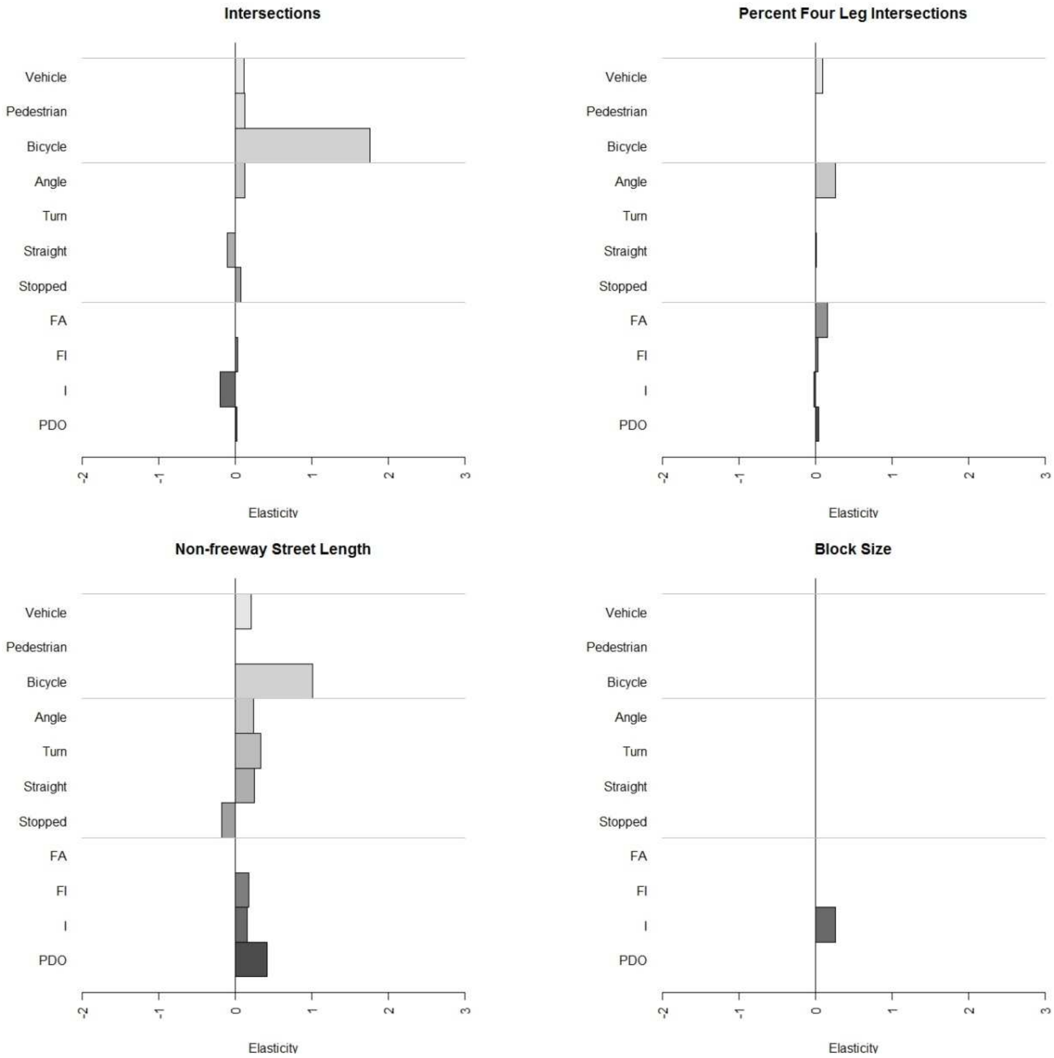


Figure 35 Elasticity for Connectivity Factors

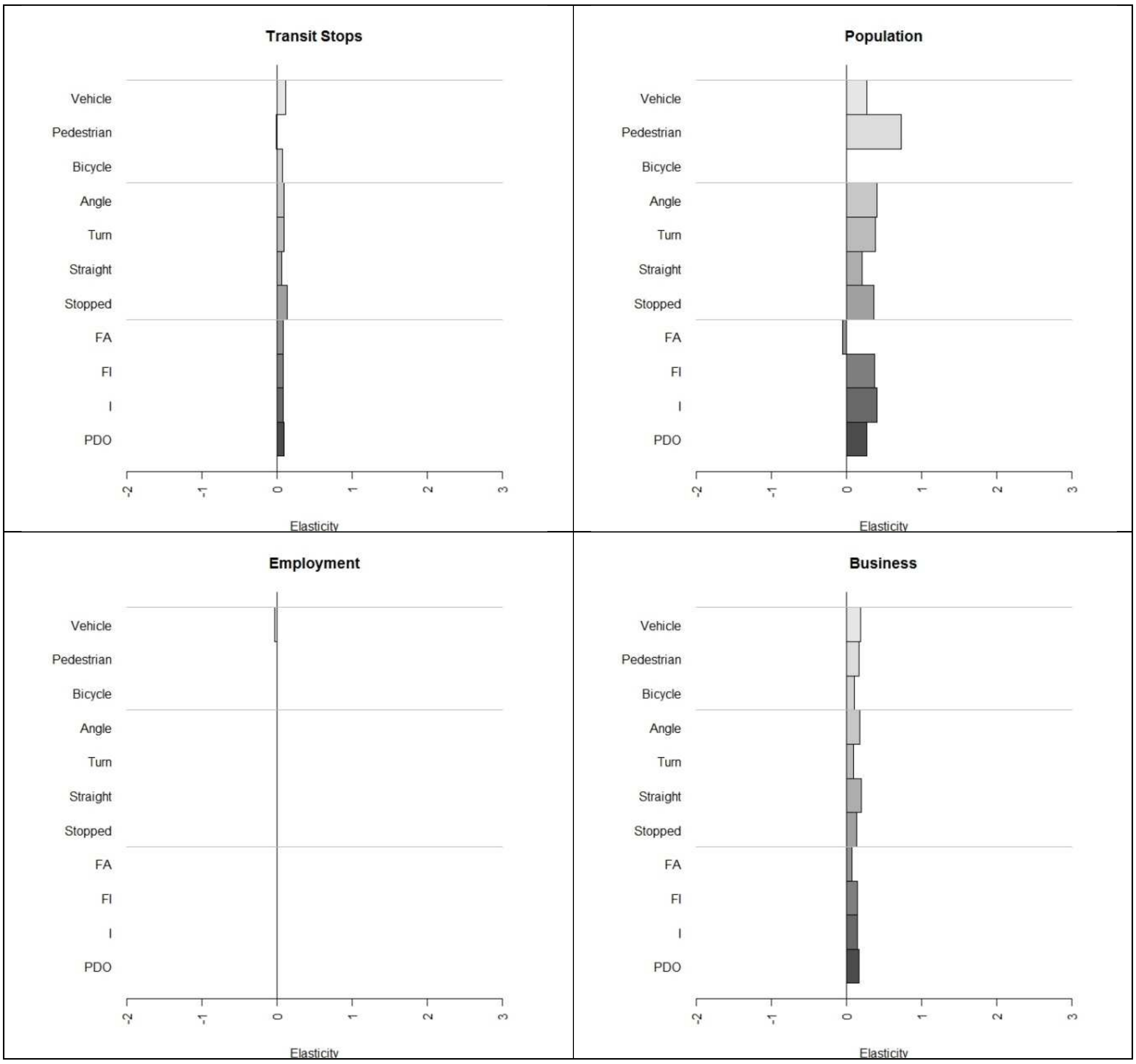


Figure 36 Elasticity for Transit Accessibility, Demographic, Origin and Destination Factors

6.0 CONCLUSIONS

Street layout and design, once established, are then not easily changed. Urban form affects community development, livability, sustainability, and traffic safety. There has been an assumed relationship between urban form and traffic safety that favors designs with less through streets to improve safety. An empirical study to test this assumed relationship was carried out for crash data for Portland, Oregon, considering factors for exposure, connectivity, transit accessibility, demographics, and origin and destination measures.

This study looked at traffic safety and urban form for the city of Portland, Oregon using a uniform spatial grid to provide an impartial way to assign crashes to the analysis spatial units. Data were assigned to each spatial grid cell by summing point data, line data, or apportioning data from a different underlying polygon spatial unit into the grid cell spatial unit.

In qualitative analysis of chloropleths showing the spatial distribution of the crashes in the grid cell aggregation, major arterials and high volume roadways clearly stood out as having more crashes. Comparing the 20,705 non-freeway vehicle crashes to 503 pedestrian and 523 bicycle crashes indicated much higher and more severe injury rates for pedestrian and bicycle crashes than in vehicle crashes: the pedestrian fatality rate is more than ten times that for vehicles, and the bicycle fatality rate is nearly four times that for vehicle crashes. The pedestrian or bicyclist involved in a crash with a vehicle has little protection against the order of magnitude greater mass and momentum of the vehicle.

Qualitative analysis of the exposure factors showed higher volumes along major arterials and in the downtown area, as would be expected. Connectivity factors for street network using the four Rifaat and Tay designations (4) were assigned to each spatial grid cell, and visually correlated to a calculable factor of percentage four leg intersections. Transit stop locations gave a view of transit accessibility. Demographic factors of population, households, and dwelling units were highly correlated; only population was included in modeling. Distribution of employment and businesses throughout the study area could also be seen in the chloropleths for those factors. Employment was highly concentrated in the downtown area and other specific clusters. Business density was also concentrated downtown and on the near east side, but was distributed throughout the city more so than the employment, indicating smaller businesses employing fewer workers.

Negative binomial regression models were built separately for groups of vehicle, pedestrian, and bicycle crashes. Models were also built for the top four crash types, as well as by crash severity. The selected models showed that exposure factors were significant in all models studied: higher traffic volumes and congestion increased crash likelihood. Exposure factors also had the highest elasticity, indicating that crash rates have strong sensitivity to these volume factors. Average free speed had a surprising negative elasticity, particularly for bicycle crashes.

Elasticity for connectivity factors showed bicycle, angle, and turn crashes to be particularly sensitive to a more grid like street network, compared to loops and lollipops. Street network was a factor in less severe crashes, but was not seen as a factor in incapacitating or fatal crashes. FA crashes were the only model to have a positive elasticity for average free speed, indicating that speed can be a strong factor in severe injury crash rates.

Urban form factors of street connectivity and intersection density were not significant at 95% confidence for vehicle and pedestrian crash models, nor for different crash severity levels. This indicates that street layout in terms of grid versus loops and lollipops does not have a statistically significant effect on vehicle crash safety, so connectivity does not have to be sacrificed in the name of safety for vehicles or pedestrians. Other factors, such as VMT, v/c, population, and business density, are far more influential.

Several origin and destination factors were significant in the models. Business density was significant for all crash models. Business density could indicate the number of access points into the transportation network. Driveways increase crash rates, so the significance of the business factor may be due to increased driveways and access points. If so, limited access design could help control and mitigate crash risk along corridors.

Population was also significant in many models. Although logical, this dependence of crashes on business and population densities raises concerns about vibrant, economically vital areas where businesses, pedestrians, bicyclists, and transit thrive alongside vehicle traffic. Thriving neighborhoods are at the heart of successful development. These results should be seen to highlight the important effect planning and directing development for where businesses, employment, and housing will grow potentially has on safety, and stress that design and planning include plans for traffic safety. Portland Metro is working on planning for major corridors which handle large traffic volumes to serve multiple transportation modes as "complete" streets that are safe for all modes, and attract people to spend time and enjoy their streets.

This study makes a contribution to the study of traffic safety and urban form in having found that connectivity factors for street layout are not statistically significant for vehicle and pedestrian crash rates. This substantiates that grid street layout designs, which provide high connectivity and thus alternative routes to allow large traffic volumes to pass through an area, are not sacrificing traffic safety compared to limited access loops and lollipops, despite long held assumptions that limited access designs are better in terms of traffic safety.

The methodology of aggregating crashes and factors into a uniform spatial grid for traffic safety analysis, an approach suggested by Kim (13), provides a way to include all crashes for analysis regardless of whether they were located along arterials that would border TAZ or block group spatial units. Previous researchers have looked at the relationship of traffic safety and urban form using TAZ or block group spatial units. These spatial unit choices were problematic regarding how to handle crashes on the peripheral roadways. The uniform spatial grid methodology gives equal weight to all crashes for an unbiased look at the overall traffic safety, so that safety for neighborhoods streets and arterials can both be considered.

6.1 Further Work

The model could be applied to data for years later than the study, when they are available, to see how well the models predicted Portland City crashes. A larger study area could be considered if geo-coded crash data are available, and using connectivity factors such as percent four leg intersections, total intersection density, and total street length, rather than needing to manually assign the Rifaat and Tay street network scale.

Data on volume for more streets would be expected to improve the model. Arterial congestion data would be particularly interesting to study along with speed data to see if there are interactions or relationships there that affect crash rates, particularly for bicycle crashes. This study had data for vehicle roadways; further studies with bicycle and pedestrian facility and volume data could be illuminating for pedestrian and bicycle crash safety.

Future work could consider looking at spatial proximity effects on the analysis. Spatial correlation likely exists, but we did not develop models to account for this. Even though freeway roadways and crashes were eliminated from the study, a dummy variable indicating whether a spatial grid cell included freeway could be included to look for whether areas adjacent to freeways have the same or higher crash rates than non-freeway adjacent areas.

Specific locations could be looked at in more detail, and a mean time to failure (i.e. crash) analysis approach.

7.0 REFERENCES

1. Southworth M, Ben-Joseph E. Streets and the shaping of towns and cities. New York: McGraw hill; 1997.

2. Kim K, Brunner I, Yamashita E. Influence of Land Use, Population, Employment, and Economic Activity on Accidents. In Transportation Research Record: Journal of the Transportation Research Board 2006;1953:56-64.

3. Dumbaugh E, Rae R. Safe Urban Form: Revisiting the Relationship Between Community Design and Traffic Safety. Journal of the American Planning Association 2009;75(3):309.

4. Rifaat SM, Tay R. Effects of Street Patterns on Injury Risks in Two-Vehicle Crashes. In Transportation Research Record: Journal of the Transportation Research Board 2009;2102:61-67.

5. Rifaat S, Tay R. Effect of Street Pattern on Road Safety: Are Policy Recommendations Sensitive to Different Aggregations of Crashes by Severity? Presented at the 89th Annual Meeting of the Transportation Research Board; 2010

6. Marks H. Subdividing for Traffic Safety. Traffic Quarterly 1957;11(2):308-325.

7. Kim K, Levine N. Using GIS to improve highway safety. Computers, Environment and Urban Systems 1996;20(4-5):289-302.

8. Kim K, Levine N, Nitz L. Development of a prototype traffic safety geographic information system. In Transportation Research Record 1995;(1477):41-47.

9. Ossenbruggen P, Pendharkar J, Ivan J. Roadway safety in rural and small urbanized areas. Accident Analysis and Prevention 2001;33(4):485-498.

10. Hadayeghi A. Macrolevel Accident Prediction Models for Evaluating Safety of Urban Transportation Systems. In Transportation research record 2003;0361-1981(1840):87.

11. Ladrón de Guevara F, Washington S, Oh J. Forecasting Crashes at the Planning Level: Simultaneous Negative Binomial Crash Model Applied in Tucson, Arizona. In Transportation Research Record: Journal of the Transportation Research Board 2004 Jan;1897(-1):191-199.

12. Al-Masaeid HR, Suleiman G. Relationships between urban planning variables and traffic crashes in Damascus. Road and Transport Research 2004;13(4):63-73.

13. Kim K, Pant P, Yamashita E. Accidents and Accessibility: Measuring the Influences of Demographic and Land Use Variables in Honolulu, Hawaii. Presented at the 89th Annual Meeting of the Transportation Research Board; 2010

14. Clifton KJ, Burnier CV, Akar G. Severity of injury resulting from pedestrian-vehicle crashes: What can we learn from examining the built environment? In Transportation Research Part D: Transport and Environment 2009 Aug;14(6):425-436.

15. Lord D, Mannering F. The Statistical Analysis of Crash-Frequency Data: A Review and Assessment of Methodological Alternatives. Transportation Research, Part A 2010;44(5):291-305.

16. Hauer E, Bamfo J. Two Tools for Finding What Function Links the Dependent Variable to the Explanatory Variables. Proceedings of ICTCT 97 Conference 1997 Nov;

17. Strathman JG, Dueker KJ, Zhang J, Williams T. Analysis of Design Attributes and Crashes on the Oregon Highway System. Oregon Department of Transportation, Research Unit; 2001.

APPENDIX A: R code for the CURE plots

```
##################################################
# K. Gladhill, R. Conrad, C. Monsere
# October 30.2010
# CURE 2s* Plots Ref Hauer "Statistical Road Safety Modeling" TRR
##################################################
#function cureSigma
cureSigma <-
function(dependentVar,orderVar,orderVarName,model1,model2,limY)
{
  models <- cbind (orderVar, dependentVar, fitted(model1), fitted (model2))
  models <- as.data.frame(models[order(orderVar),])                  #order the
new data frame
  names(models) <- list("Parameter", "CRASH", "M1","M2")           #add names
to the data fields
  ## +2sigma
  ##sigma 1
  res1N<-sum((models$CRASH-models$M1)^2)
  res1n<-cumsum((models$CRASH-models$M1)^2)
  sigma1<-sqrt(res1n*(1-(res1n/res1N)))
  ##sigma 2
  res2N<-sum((models$CRASH-models$M2)^2)
  res2n<-cumsum((models$CRASH-models$M2)^2)
  sigma2<-sqrt(res2n*(1-(res2n/res2N)))

  ## Cumulative and CURE plots
```

```
  # find the max value for the y-axis before opening plot window
  ymax1 <- 2*max(sigma1, na.rm=TRUE)
  ymax2 <- 2*max(sigma2, na.rm=TRUE)
  if (limY < 1)
    {
    par(mfrow = c(1, 2)) #set graph parameter
        plot( models$Parameter, cumsum(models$CRASH-models$M1), type="l",
col="red",
          main="CURE 2s* Plot, Model 1", ylab="Cum Residuals",
xlab=orderVarName)
          abline(h=0, lty=2)
          lines( models$Parameter,-2*sigma1 , col="blue")
          lines( models$Parameter, 2*sigma1 , col="green")

        plot( models$Parameter, cumsum(models$CRASH-models$M2), type="l",
col="red",
          main="CURE 2s* Plot, Model 2", ylab="Cum Residuals",
xlab=orderVarName)
          abline(h=0, lty=2)
          lines(models$Parameter,-2*sigma2, type="l", col="dodgerblue", pch=16)
          lines(models$Parameter, 2*sigma2, type="l", col="seagreen", pch=17)
      par(mfrow = c(1, 1)) #set graph parameter
    }
  if (limY > 0) {
      par(mfrow = c(1, 2)) #set graph parameter
        plot( models$Parameter, cumsum(models$CRASH-models$M1), type="l",
ylim=c(-ymax1,ymax1), col="red",
```

```
          main="CURE 2s* Plot, Model 1", ylab="Cum Residuals",
xlab=orderVarName)
          abline(h=0, lty=2)
          lines( models$Parameter,-2*sigma1 , col="blue")
          lines( models$Parameter, 2*sigma1 , col="green")

        plot( models$Parameter, cumsum(models$CRASH-models$M2), type="l",
             ylim=c(-ymax2,ymax2), col="red",
          main="CURE 2s* Plot, Model 2", ylab="Cum Residuals",
xlab=orderVarName)
          abline(h=0, lty=2)
          lines(models$Parameter,-2*sigma2, type="l", col="dodgerblue", pch=16)
          lines(models$Parameter, 2*sigma2, type="l", col="seagreen", pch=17)
      par(mfrow = c(1, 1)) #set graph parameter
    }
 }
#end function
```

```
cureSigma3 <-
function(dependentVar,orderVar,orderVarName,model1,model2,model3,limY)
{
  models <- cbind (orderVar, dependentVar, fitted(model1), fitted (model2),fitted
(model3))
  models <- as.data.frame(models[order(orderVar),])                    #order the
new data frame
  names(models) <- list("Parameter", "CRASH", "M1","M2","M3")    #add names
to the data fields
  ## +2sigma
  ##sigma 1
  res1N<-sum((models$CRASH-models$M1)^2)
  res1n<-cumsum((models$CRASH-models$M1)^2)
  sigma1<-sqrt(res1n*(1-(res1n/res1N)))
  ##sigma 2
  res2N<-sum((models$CRASH-models$M2)^2)
  res2n<-cumsum((models$CRASH-models$M2)^2)
  sigma2<-sqrt(res2n*(1-(res2n/res2N)))
  ##sigma 3
  res3N<-sum((models$CRASH-models$M3)^2)
  res3n<-cumsum((models$CRASH-models$M3)^2)
  sigma3<-sqrt(res3n*(1-(res3n/res3N)))

  ## Cumulative and CURE plots
  # find the max value for the y-axis before opening plot window
  ymax1 <- 2*max(sigma1, na.rm=TRUE)
  ymax2 <- 2*max(sigma2, na.rm=TRUE)
```

```
 ymax3 <- 2*max(sigma3, na.rm=TRUE)
if (limY < 1)
   {
   par(mfrow = c(1, 3)) #set graph parameter

       plot( models$Parameter, cumsum(models$CRASH-models$M1), type="l",
col="red",
         main="CURE 2s* Plot, Model 1", ylab="Cum Residuals",
xlab=orderVarName)
         abline(h=0, lty=2)
         lines( models$Parameter,-2*sigma1 , col="blue")
         lines( models$Parameter, 2*sigma1 , col="green")

       plot( models$Parameter, cumsum(models$CRASH-models$M2), type="l",
col="red",
         main="CURE 2s* Plot, Model 2", ylab="Cum Residuals",
xlab=orderVarName)
         abline(h=0, lty=2)
         lines(models$Parameter,-2*sigma2, type="l", col="dodgerblue", pch=16)
         lines(models$Parameter, 2*sigma2, type="l", col="seagreen", pch=17)

       plot( models$Parameter, cumsum(models$CRASH-models$M3), type="l",
col="red",
         main="CURE 2s* Plot, Model 3", ylab="Cum Residuals",
xlab=orderVarName)
         abline(h=0, lty=2)
         lines(models$Parameter,-2*sigma3, type="l", col="dodgerblue", pch=16)
```

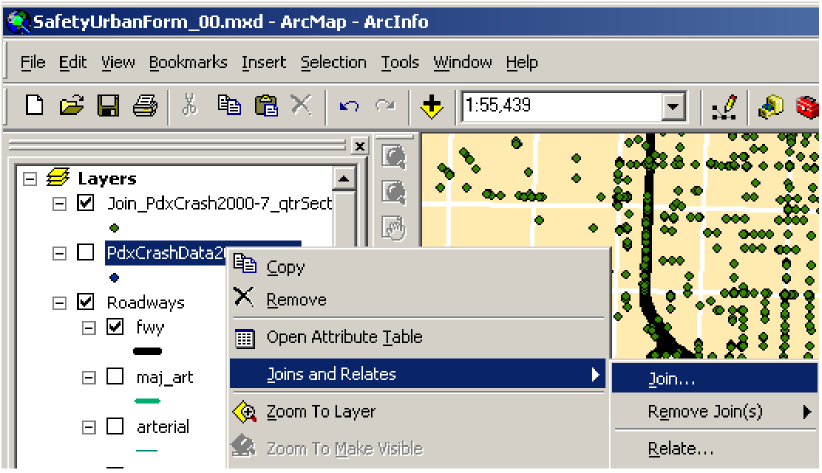
```
        lines(models$Parameter, 2*sigma3, type="l", col="seagreen", pch=17)

      par(mfrow = c(1, 1)) #set graph parameter
    }
  if (limY > 0) {
      par(mfrow = c(1, 3)) #set graph parameter

       plot( models$Parameter, cumsum(models$CRASH-models$M1), type="l",
ylim=c(-ymax1,ymax1), col="red",
        main="CURE 2s* Plot, Model 1", ylab="Cum Residuals",
xlab=orderVarName)
        abline(h=0, lty=2)
        lines( models$Parameter,-2*sigma1 , col="blue")
        lines( models$Parameter, 2*sigma1 , col="green")

       plot( models$Parameter, cumsum(models$CRASH-models$M2), type="l",
          ylim=c(-ymax2,ymax2), col="red",
        main="CURE 2s* Plot, Model 2", ylab="Cum Residuals",
xlab=orderVarName)
        abline(h=0, lty=2)
        lines(models$Parameter,-2*sigma2, type="l", col="dodgerblue", pch=16)
        lines(models$Parameter, 2*sigma2, type="l", col="seagreen", pch=17)

       plot( models$Parameter, cumsum(models$CRASH-models$M3), type="l",
ylim=c(-ymax2,ymax2), col="red",
        main="CURE 2s* Plot, Model 3", ylab="Cum Residuals",
xlab=orderVarName)
```

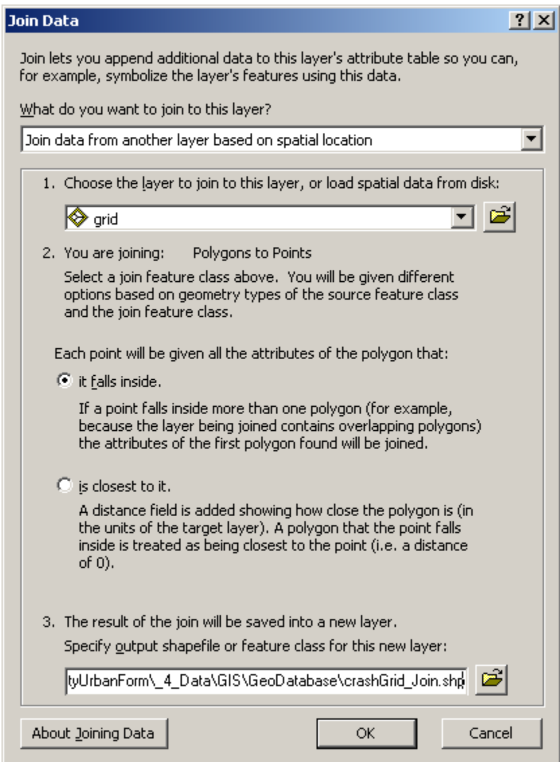
```
        abline(h=0, lty=2)
        lines(models$Parameter,-2*sigma3, type="l", col="dodgerblue", pch=16)
        lines(models$Parameter, 2*sigma3, type="l", col="seagreen", pch=17)

      par(mfrow = c(1, 1)) #set graph parameter
    }
  }
#end function
```

APPENDIX B: Join point data to spatial area in ArcGIS 9.3.1

Join point data to spatial area

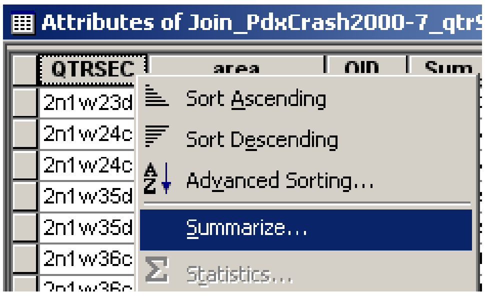
a. Choose layer with point data, start join from there:

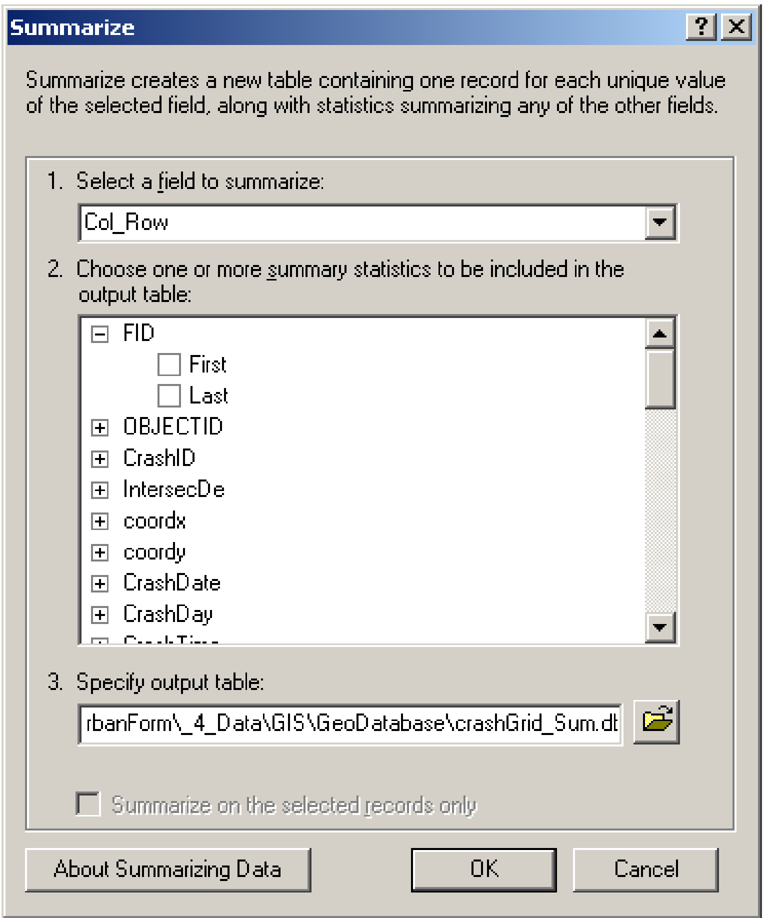


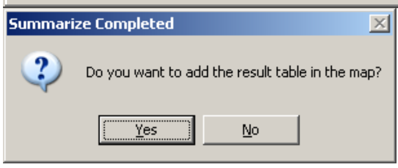


This may take some time.

b. In joined attribute table, choose column to summarize on.
In this case, a count of field "Col_Row" will do, it's the ID for each grid cell.







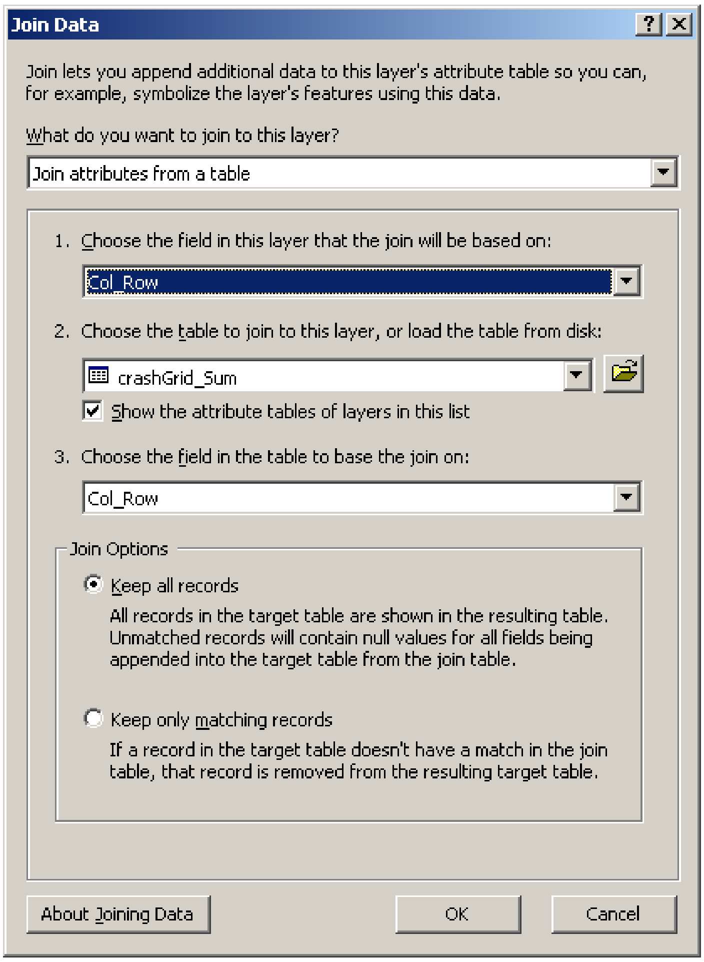
c. Right-click on the table on the "Source" tab to open that table to check data in the summary column:

Attributes of Sum_PdxCrash2000-7_qtrSect

OID	QTRSEC	Count_QTRSEC
0		6735
1	1n1e01a	5
2	1n1e01b	67
3	1n1e01c	3

Record: 1 Show: All Selected

d. Join summary data to the original spatial Layer.
Start with R-click of spatial layer, choose Joins and Relates -> Join



e. Check the attributes table for spatial layer, it now includes columns from the joined table.


Attributes of qtr_sec

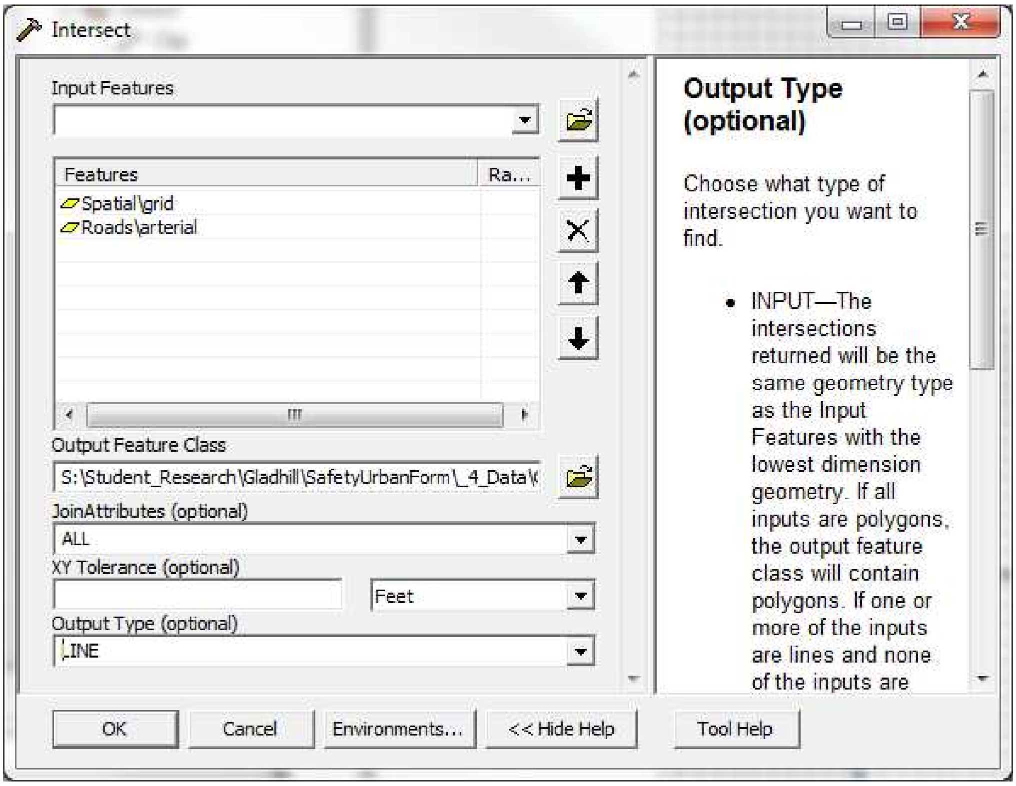
FID	Shape *	QTRNO	QTRSEC	area	OID	QTRSEC	CrashCnt
0	Polygon		4n1w05b	6971845.08693	<Null>	<Null>	<Null>
1	Polygon		4n1w05a	6971558.32219	<Null>	<Null>	<Null>
2	Polygon		4n1w04b	7046081.60644	<Null>	<Null>	<Null>
3	Polygon		4n1w04a	7046597.35093	<Null>	<Null>	<Null>

Record: 1 Show: All Selected Records (571 out of 10

APPENDIX C: Join Line Data to spatial area in ArcGIS 9.3.1

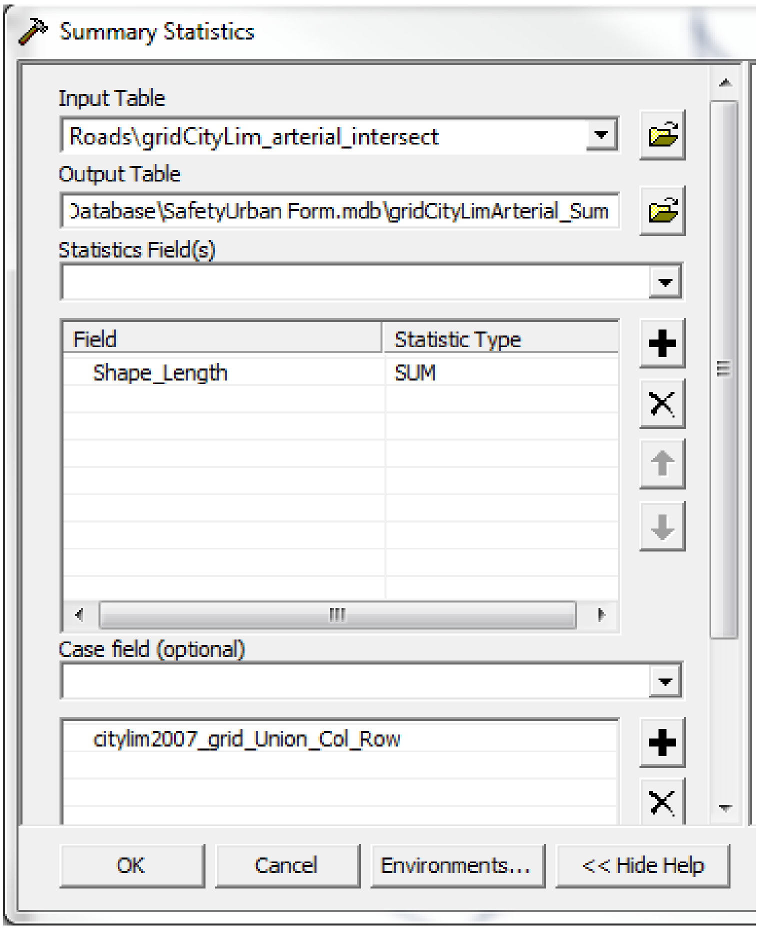
Determine the length of line data in polygons

1. Create a new output layer is containing road segments intersecting grid cells, clipped at each cell.
   Analysis Tools -> Overlay - > Intersect



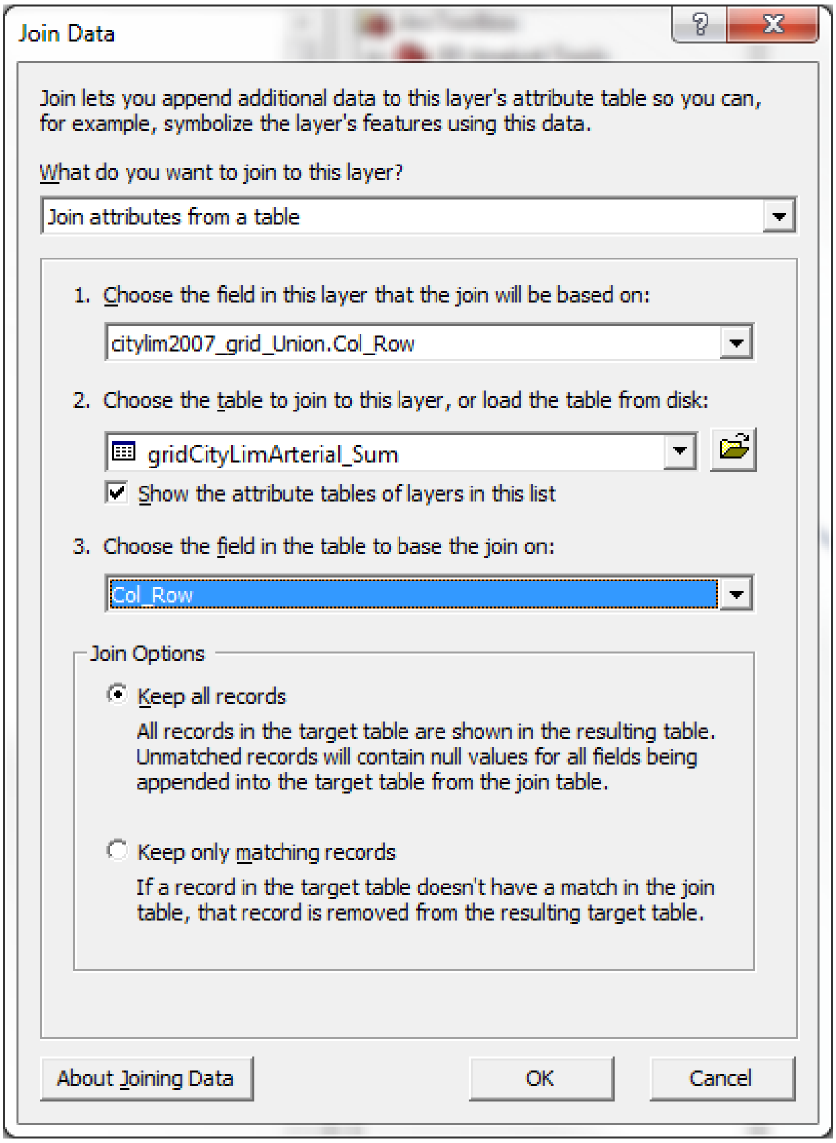
2. Sum the lines that fall within the polygons

   a. ArcToolbox,
      'Analysis Tools -> Statistics ->Summary Statistics



Hit OK, and Close—it will look like nothing happened…that's OK.

b. In the Table of Contents window, Right-click the polygon layer and select : Join and Relates ->Join…



Hit OK—if prompted hit yes

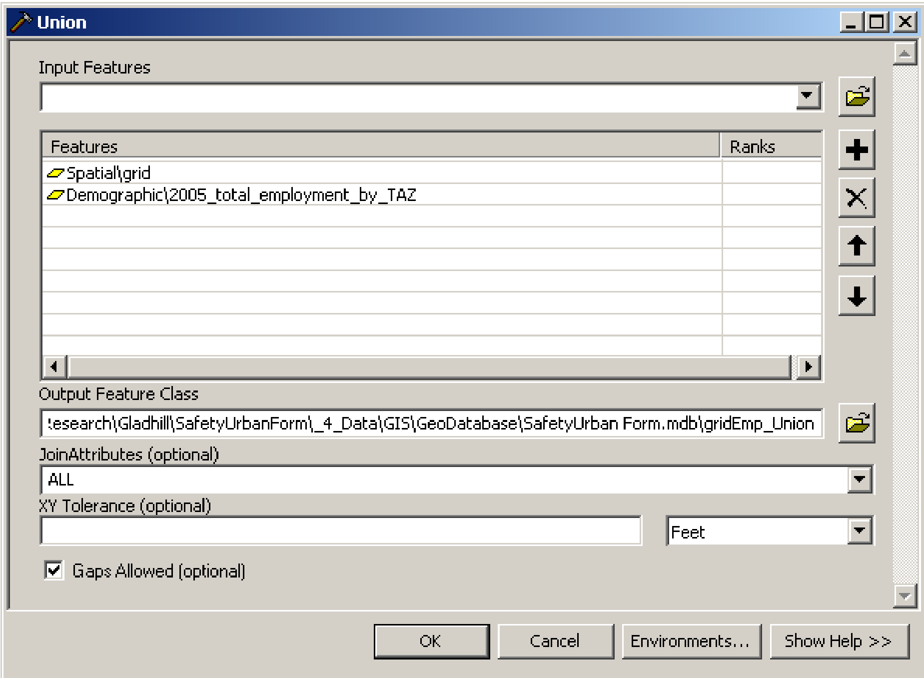
The summary value should now appear as a column in the polygon attribute table..

APPENDIX D: Apportion polygon data to a different polygon spatial area in ArcGIS 9.3.1

Example for employment by TAZ being put into uniform grid layer

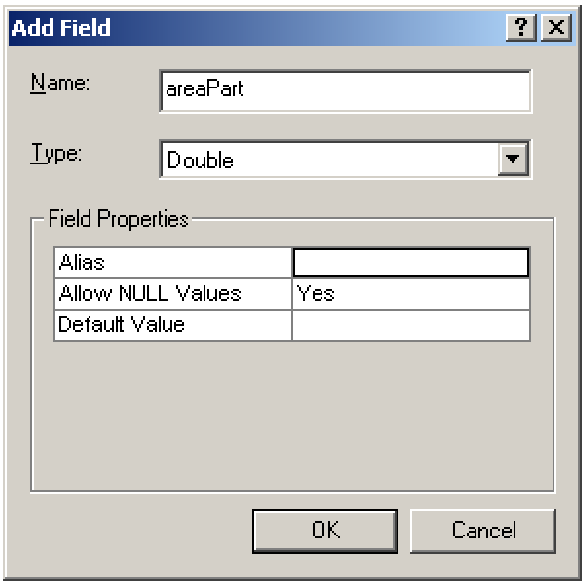
1. Create smaller TAZ/Employment areas contained within each grid cell
   a. Select the "ArcToolbox from the top
      (it's a red icon, looks like a toolbox)
      Analysis Tools -> Overlay - > Union

      For Input, select the "Employment" and the shapefile for the layer with the polygons data are going into.
      Make sure the Output file is named what you want it to be

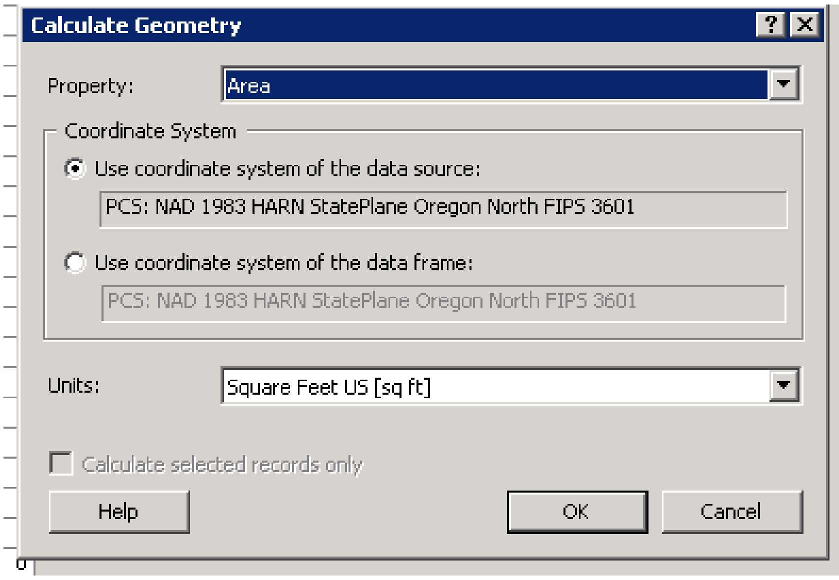


      Hit OK—It will take a bit, but the new shapefile should be automatically added to the map.

2. Clean-up output layer and perform calculations
    a. Right click output layer and select "Open Attribute Table"
    b. Delete all unnecessary columns by selecting from the top of each column).
    c. While still in the Attributes window, select "Options"
       (at the bottom of the dialogue box),
       then "Add Field" (to be the new "area" of the smaller
       TAZ/Employment areas)
    d. Enter information as show below, and hit OK:

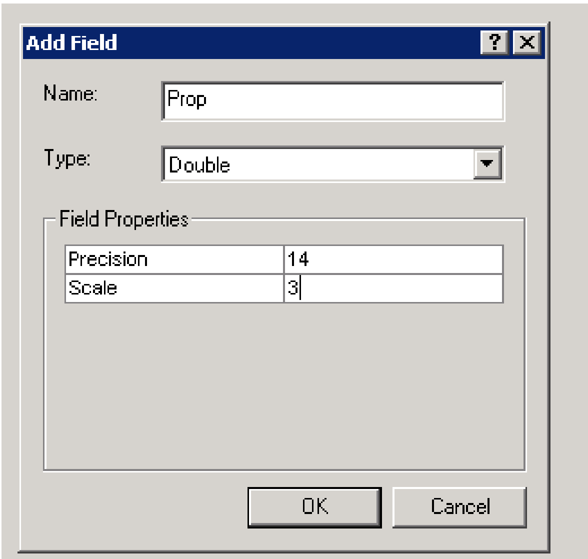


e. To populate the new "AreaPart" Right click the "AreaPart" column, select "Calculate Geometry" hit "yes" and enter information as show below:

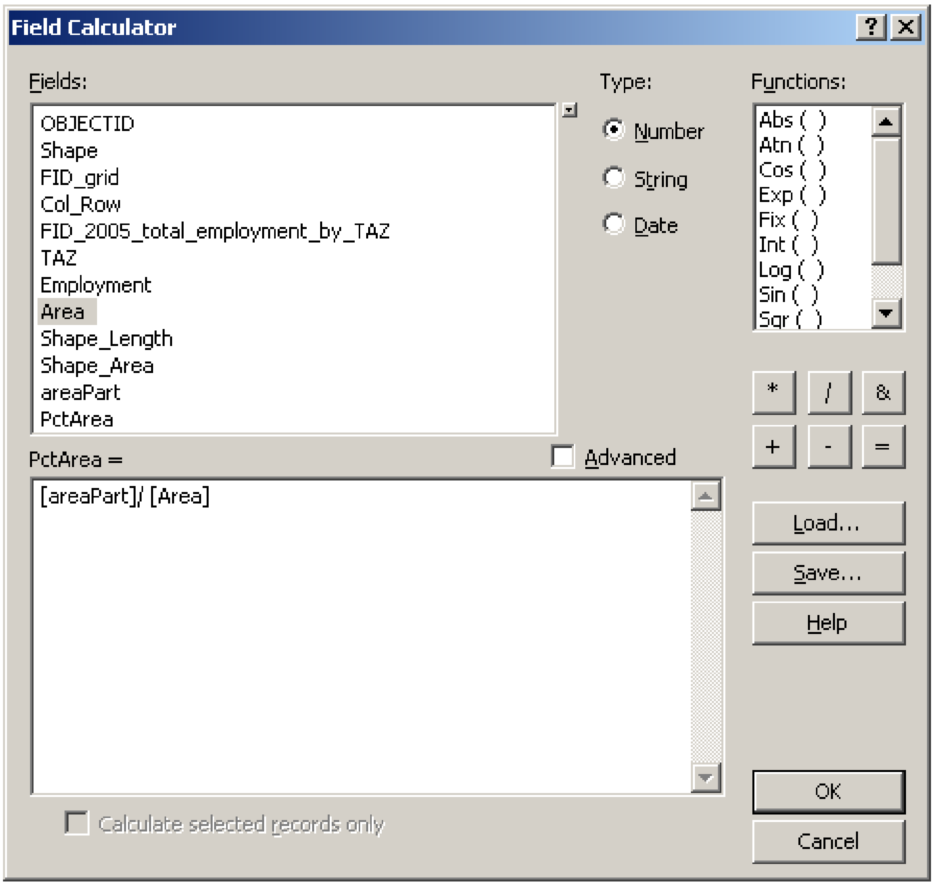


Much of the information above should be there (if not all)—hit OK.

f. Create a new field—this will be to calculate the percentage of the total TAZ/Employment area that is in the grid cell. Field information is as below:



g. Select the new area field, Right-Click and select "Field Calculator" and "Yes". Using the Mouse to select the Values (sometimes if you use keyboard it doesn't work right), select AreaPart / Area and hit OK.

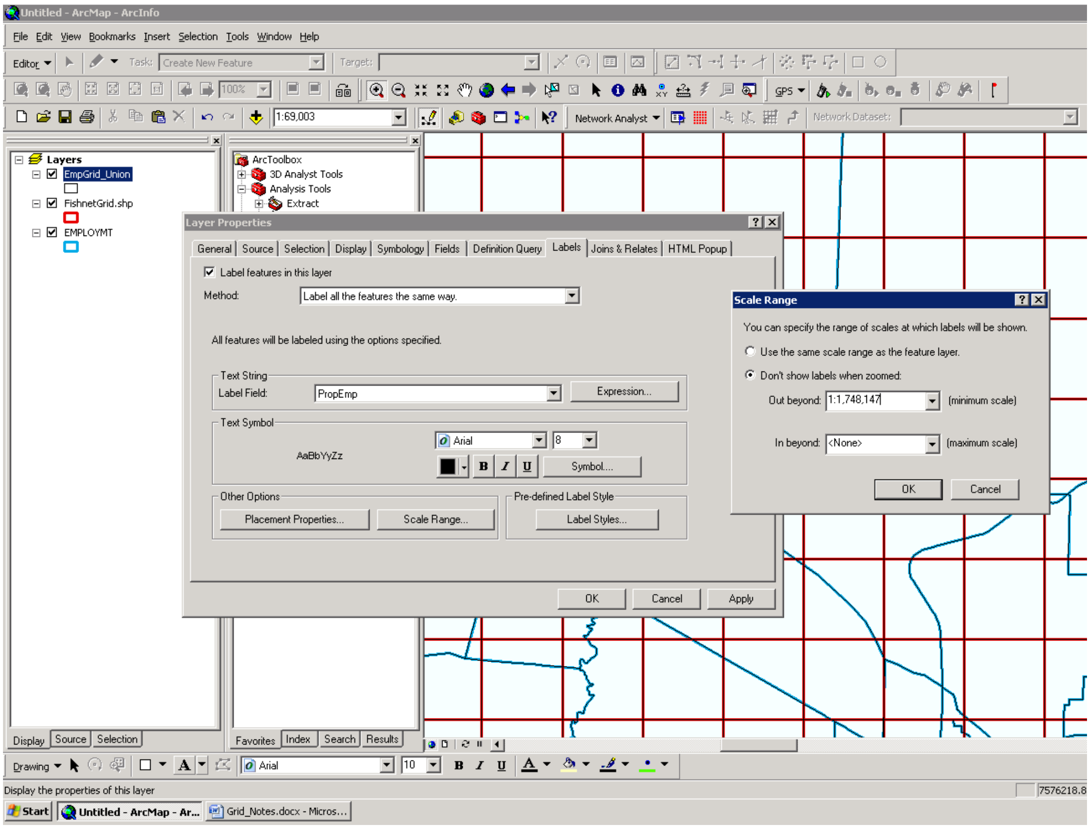


If there is a prompt (from an error) hit "yes"—this "error" is because some cells have zero "Area" because the TAZ/Employment does not cover the entire grid.

h. Create a new field for the Proportion of Employment in the smaller area by selecting "options", "add new field named something like "PortionEmp".

i. Select the "PortionEmp" field, Right-Click and select "Field Calculator" and "Yes".
Select the values: [PctArea] * [Employment]
Hit OK when done, and yes if there is a prompt.

j. Format the EmpGrid_Union and label the "PortionEmp" field by following steps from above and shown below:

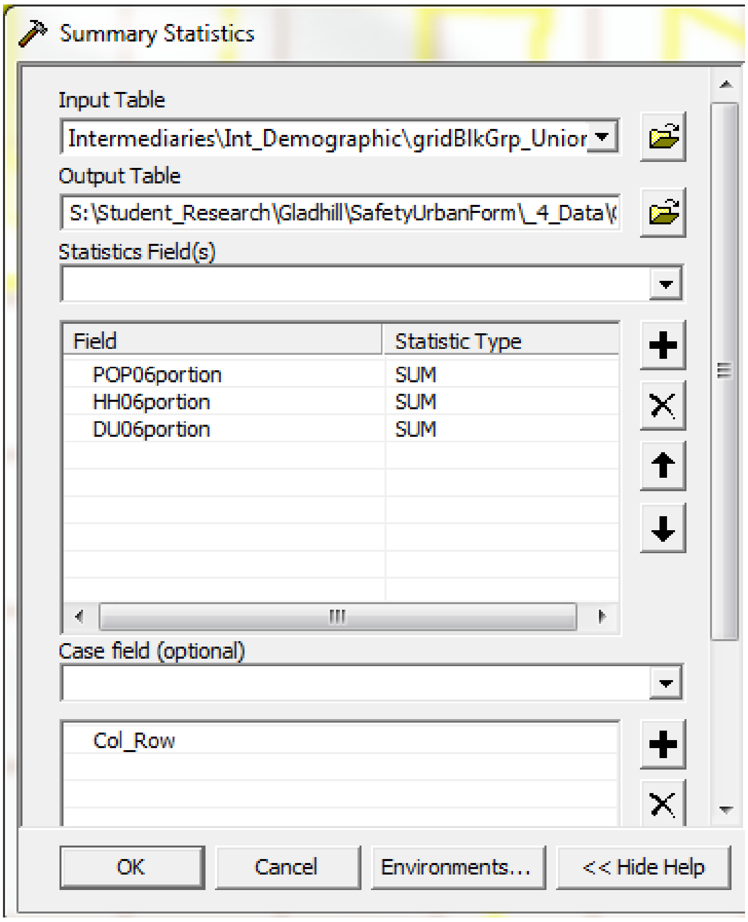


k. Hit Apply and OK

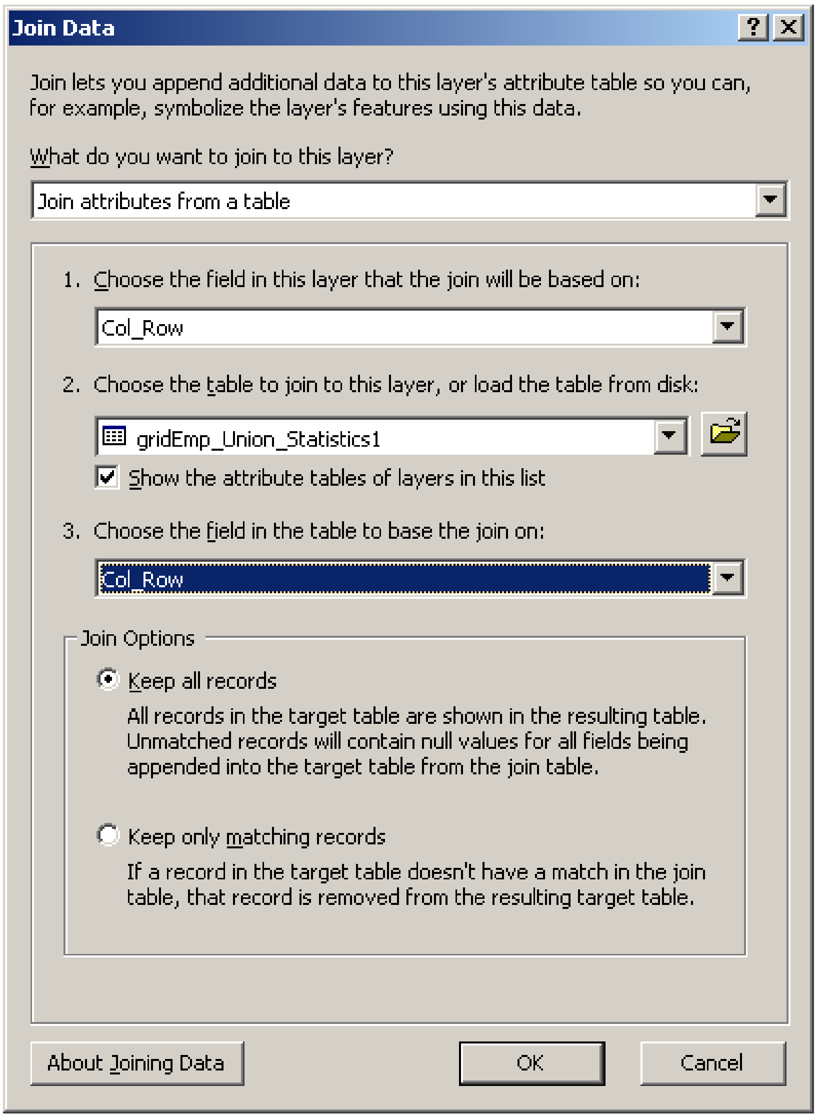
3. Sum the Smaller TAZ/Employment totals that fall within the grid cells
    a. From the ArcToolbox,
       'Analysis Tools -> Statistics ->Summary Statistics
       For the Statistics field, select the 'PortionEmp'—this is the proportion of employment within the smaller TAZ/Employment areas (that fall within a given grid). They will need to be summed up within each grid (so choose 'SUM' for the statistic type in the drop down, as shown above).

       For the case field, you will have to select, the 'Col_Row' field—this is because you want to sum by each grid (denoted by the Col_Row coordinate).

       Hit OK, and Close—it will look like nothing happened…that's OK.



b. In the Table of Contents window, select the 'grid' layer by Right-click, and select:
Join and Relates ->Join...



Hit OK—if prompted hit yes

Note: gridEmp_Union_Statistics1 can be found under Source tab,

APPENDIX E: Pearson's Correlation on Factors – Figure E-1 Correlation for Exposure Factors

	Street Network	4legPct	intersec- tions -	Int3Leg4	Int3Leg	Int4Leg	non-Fwy Street	Arterial	Maj Art	block Size
4legPct	0.619									
intersections	0.666	0.501								
Int3Leg4	0.699	0.604	0.944							
Int3Leg	0.385	-0.057	0.603	0.627						
Int4Leg	0.644	0.825	0.784	0.842	0.112					
nonFwyStreet	0.794	0.660	0.889	0.920	0.541	0.805				
gridStreet	0.754	0.641	0.908	0.924	0.527	0.812	0.970			
gridArterial	0.106	0.190	0.416	0.388	0.143	0.361	0.326			
gridMajArt	0.006	0.130	0.322	0.287	0.034	0.301	0.208	0.889		
blockSize	0.538	0.524	0.905	0.886	0.444	0.802	0.797	0.679	0.620	
transitStops	0.265	0.296	0.367	0.396	0.144	0.407	0.386	0.240	0.187	0.375
busStops	0.266	0.296	0.367	0.396	0.144	0.407	0.386	0.237	0.184	0.374
transitRoute	0.170	0.246	0.348	0.359	0.111	0.372	0.332	0.426	0.388	0.434
busRoute	0.172	0.248	0.347	0.359	0.111	0.372	0.331	0.414	0.374	0.429
ons	0.149	0.198	0.236	0.265	0.054	0.294	0.229	0.211	0.217	0.278
offs	0.151	0.202	0.244	0.273	0.052	0.307	0.233	0.215	0.224	0.288
onOff	0.151	0.201	0.241	0.270	0.053	0.302	0.232	0.214	0.222	0.284
estLoad	0.231	0.282	0.341	0.381	0.113	0.401	0.347	0.267	0.252	0.387
VMT	0.093	0.156	0.363	0.325	0.118	0.299	0.238	0.841	0.844	0.600
VCdist	0.239	0.293	0.478	0.468	0.199	0.430	0.423	0.824	0.729	0.667
avgVC	0.111	0.128	0.172	0.169	0.095	0.146	0.177	0.318	0.261	0.232
AvgFreeSpeed	-0.443	-0.349	-0.339	-0.388	-0.256	-0.338	-0.446	0.213	0.311	-0.189
schools	0.204	0.193	0.236	0.260	0.137	0.239	0.244	0.033	0.006	0.207
Pop06	0.629	0.519	0.569	0.618	0.431	0.516	0.669	0.066	-0.041	0.445
HH06	0.563	0.535	0.560	0.623	0.340	0.577	0.641	0.144	0.052	0.488
DU06	0.555	0.533	0.562	0.624	0.338	0.580	0.638	0.153	0.064	0.495
employment	0.117	0.251	0.244	0.286	-0.032	0.381	0.223	0.217	0.248	0.315

business	0.316	0.398	0.341	0.413	0.031	0.507	0.386	0.186	0.169	0.355

Figure E-2, Pearson's Correlation for Transit accessibility, Exposure, Demographic, and Origin Destination Factors

	transit Stops	busStops	transit Route	busRoute	ons	offs	onOff	estLoad	VMT	VCdist	avgVC	Avg Free Speed	schools	Pop06	HH06	DU06	employment
bus Stops	1.00																
transit Route	0.83	0.83															
bus Route	0.83	0.83	1.00														
ons	0.72	0.72	0.64	0.64													
offs	0.71	0.71	0.65	0.64	0.98												
onOff	0.72	0.72	0.64	0.64	1.00	1.00											
est Load	0.79	0.79	0.69	0.69	0.82	0.78	0.80										
VMT	0.21	0.21	0.43	0.42	0.22	0.22	0.22	0.26									
VC * dist	0.35	0.35	0.49	0.49	0.26	0.25	0.26	0.37	0.85								
avgVC	0.09	0.09	0.17	0.17	0.03	0.02	0.02	0.13	0.38	0.54							
Avg Free Speed	-0.28	-0.28	-0.13	-0.14	-0.17	-0.18	-0.18	-0.19	0.28	0.09	0.29						
schools	0.30	0.30	0.20	0.20	0.15	0.15	0.15	0.18	0.03	0.09	0.03	-0.22					
Pop06	0.26	0.27	0.16	0.16	0.14	0.15	0.14	0.21	0.05	0.22	0.19	-0.45	0.31				
HH06	0.40	0.40	0.31	0.31	0.25	0.27	0.26	0.30	0.13	0.30	0.20	-0.43	0.32	0.92			
DU06	0.40	0.40	0.32	0.32	0.26	0.27	0.27	0.31	0.14	0.31	0.20	-0.42	0.32	0.92	1.00		
employ-ment	0.67	0.67	0.70	0.70	0.69	0.72	0.71	0.57	0.24	0.25	0.00	-0.18	0.16	0.07	0.27	0.29	
business	0.62	0.61	0.55	0.55	0.58	0.58	0.59	0.52	0.18	0.28	0.07	-0.26	0.24	0.31	0.45	0.45	0.74