Final Report

Contract BDK78 977-13

Two Level Approach to Safety Planning Incorporating the Highway Safety Manual (HSM) Network Screening

Mohamed A. Abdel-Aty, Ph.D., P.E.
Pei-Fen Kuo, Ph.D.
Ximiao Jiang, Ph.D.
Jaeyoung Lee
Samer Al Amili

University of Central Florida
Department of Civil, Environmental & Construction Engineering
Orlando, FL 32816-2450

April 2014

DISCLAIMER

"The opinions, findings, and conclusions expressed in this publication are those of the authors and not necessarily those of the State of Florida Department of Transportation."

UNITS CONVERSION

APPROXIMATE CONVERSIONS TO SI UNITS

SYMBOL	WHEN YOU KNOW	MULTIPLY BY	TO FIND	SYMBOL
		LENGTH		
in	inches	25.4	millimeters	mm
ft	feet	0.305	meters	m
yd	yards	0.914	meters	m
mi	miles	1.61	kilometers	km
SYMBOL	**WHEN YOU KNOW**	**MULTIPLY BY**	**TO FIND**	**SYMBOL**
		AREA		
in^2	squareinches	645.2	square millimeters	mm^2
ft^2	squarefeet	0.093	square meters	m^2
yd^2	square yard	0.836	square meters	m^2
ac	acres	0.405	hectares	ha
mi^2	square miles	2.59	square kilometers	km^2
SYMBOL	**WHEN YOU KNOW**	**MULTIPLY BY**	**TO FIND**	**SYMBOL**
		VOLUME		
fl oz	fluid ounces	29.57	milliliters	mL
gal	gallons	3.785	liters	L
ft^3	cubic feet	0.028	cubic meters	m^3
yd^3	cubic yards	0.765	cubic meters	m^3

NOTE: volumes greater than 1000 L shall be shown in m^3

SYMBOL	WHEN YOU KNOW	MULTIPLY BY	TO FIND	SYMBOL
MASS				
oz	ounces	28.35	grams	g
lb	pounds	0.454	kilograms	kg
T	short tons (2000 lb)	0.907	megagrams (or "metric ton")	Mg (or "t")
SYMBOL	**WHEN YOU KNOW**	**MULTIPLY BY**	**TO FIND**	**SYMBOL**
TEMPERATURE (exact degrees)				
^oF	Fahrenheit	5 (F-32)/9 or (F-32)/1.8	Celsius	^oC
SYMBOL	**WHEN YOU KNOW**	**MULTIPLY BY**	**TO FIND**	**SYMBOL**
ILLUMINATION				
fc	foot-candles	10.76	lux	lx
fl	foot-Lamberts	3.426	candela/m^2	cd/m^2
SYMBOL	**WHEN YOU KNOW**	**MULTIPLY BY**	**TO FIND**	**SYMBOL**
FORCE and PRESSURE or STRESS				
lbf	poundforce	4.45	newtons	N
lbf/in^2	poundforce per square inch	6.89	kilopascals	kPa
SYMBOL	**WHEN YOU KNOW**	**MULTIPLY BY**	**TO FIND**	**SYMBOL**
LENGTH				
mm	millimeters	0.039	inches	in

m	meters	3.28	feet	ft
m	meters	1.09	yards	yd
km	kilometers	0.621	miles	mi
SYMBOL	**WHEN YOU KNOW**	**MULTIPLY BY**	**TO FIND**	**SYMBOL**
AREA				
mm^2	square millimeters	0.0016	square inches	in^2
m^2	square meters	10.764	square feet	ft^2
m^2	square meters	1.195	square yards	yd^2
ha	hectares	2.47	acres	ac
km^2	square kilometers	0.386	square miles	mi^2
SYMBOL	**WHEN YOU KNOW**	**MULTIPLY BY**	**TO FIND**	**SYMBOL**
VOLUME				
mL	milliliters	0.034	fluid ounces	fl oz
L	liters	0.264	gallons	gal
m^3	cubic meters	35.314	cubic feet	ft^3
m^3	cubic meters	1.307	cubic yards	yd^3
SYMBOL	**WHEN YOU KNOW**	**MULTIPLY BY**	**TO FIND**	**SYMBOL**
MASS				
g	grams	0.035	ounces	oz
kg	kilograms	2.202	pounds	lb
Mg (or "t")	megagrams (or "metric ton")	1.103	short tons (2000 lb)	T

SYMBOL	WHEN YOU KNOW	MULTIPLY BY	TO FIND	SYMBOL
TEMPERATURE (exact degrees)				
^{o}C	Celsius	1.8C+32	Fahrenheit	^{o}F
SYMBOL	**WHEN YOU KNOW**	**MULTIPLY BY**	**TO FIND**	**SYMBOL**
ILLUMINATION				
lx	lux	0.0929	foot-candles	fc
cd/m^2	candela/m^2	0.2919	foot-Lamberts	fl
SYMBOL	**WHEN YOU KNOW**	**MULTIPLY BY**	**TO FIND**	**SYMBOL**
FORCE and PRESSURE or STRESS				
N	newtons	0.225	poundforce	lbf
kPa	kilopascals	0.145	poundforce per square inch	lbf/in^2

1. Report No. BDK78 977-13	2. Government Accession No.	3. Recipient's Catalog No.
4. Title and Subtitle **Two Level Approach to Safety Planning Incorporating the Highway Safety Manual (HSM) Network Screening**		5. Report Date April 18, 2014 6. Performing Organization Code
7. Author(s) **Mohamed A. Abdel-Aty, PhD, PE; Pei-Fen Kuo, PhD; Ximiao Jiang, PhD; Jaeyoung Lee; Samer Al Amili**		8. Performing Organization Report No.
9. Performing Organization Name and Address **Center for Advanced Transportation and Systems Simulation, University of Central Florida, P.O. Box 162450, Orlando, FL 32816-2450**		10. Work Unit No. (TRAIS) 11. Contract or Grant No.
12. Sponsoring Agency Name and Address **Florida Department of Transportation**		13. Type of Report and Period Covered **Final Report** 14. Sponsoring Agency Code
15. Supplementary Note		

16. Abstract

Compared to microscopic safety studies, macroscopic-focused research is more efficient at integrating zonal-level features into crash prediction models and identifying hot zones. However, macroscopic screening has accuracy limitations. Thus, this study developed a new integrated screening approach to overcome the above-mentioned shortcomings of current screening techniques and to achieve a balance between efforts towards accuracy and efficiency.

For conducting macro level safety analyses, the research team faced several challenges. First, using current Traffic Analysis Zones (TAZs) as basic geographic units caused a high percentage of boundary crashes. The research team used regionalization to develop a new study unit: Traffic Safety Analysis Zones (TSAZs) systems. Approximately 10% of boundary crashes have been integrated in new zones after the regionalization but more than 60% of crashes still occur on the boundary of TSAZs. Hence, a nested structure was proposed to estimate safety performance models separately for boundary and interior crashes. This nested structure allows different contributing factors for different crash types, so this model can provide more accurate and predictable results than a single model. In addition, a Bayesian Poisson Lognormal Spatial Error Model (BPLSEM) was adopted for the SPF analysis. The BPLSEM contains a spatial error term that control for the spatial autocorrelation of crash data. As for the micro level analysis, the research team developed SPFs based on the major function classes of roads in our study area. The research team still used the Full Bayesian Poisson Lognormal models to predict crash frequency but tried four different variable combinations to identify the best model.

After identifying hot spot areas at the macro- and microscopic levels, the research team integrated these macroscopic and microscopic screening results. However, this integration task was challenging because we needed to (1) combine various SPFs from different scales, areas, and roadway types; (2) determine an appropriate weight for each group; and (3) choose a measurement for our final results.

In order to solve the above mentioned problems, this study then developed a new criterion to identify whether a zone has safety issues at the macro- and/or microscopic levels. All TSAZs were classified into twelve categories that include two scale groups (macro or micro) and four safety levels (hot, normal, cold, or no data). Then, the research team defined weights for different scales and roadway types. At the macroscopic level, TSAZs were ranked by their zonal PSIs (Potential for Safety Improvements); at the microscopic level, the calculation of average PSI was more complicated because each TSAZ had several intersections and segments. Both the intersection and segment PSI ranks were averaged. The PSI is used in the HSM for network screening but it is the first time that is used for zonal screening. TSAZs with top 10% PSIs were categorized as "Hot" zones. Finally, the percentile ranks of the PSIs were used in the integration (instead of the original PSIs) because the units of PSI intersections and PSI segments were different. In summary, this study presents an integrated screening method that can be used to overcome the shortcomings of macro- and micro-level approaches. In particular, our results provide a comprehensive perspective on appropriate safety treatments by balancing the accuracy and efficiency of screening. Also, it is recommended that different strategies for each hot zone classification be developed because each category has distinctive traffic safety risks at each of the different levels.

17. Key Word	18. Distribution Statement
Network screening, Hotspots, Regionalization, Macro-level, Safety performance functions	

19. Security Classif. (of this report)	20. Security Classif. (of this page)	21. No. of Pages	22. Price
	Unclassified	166	

EXECUTIVE SUMMARY

Many studies have analyzed at the microscopic level the sites with high traffic safety risk (e.g., segments, intersections, etc.), including the HSM Part B (AASHTO, 2010). Recently, several studies have begun to focus on zonal-based network screening at the macroscopic level. Compared to microscopic safety studies, macroscopic-focused research is more efficient at integrating zonal-level features into crash prediction models and identifying hot zones. However, macroscopic screening has accuracy limitations because it cannot identify and separate hot spots from other sites within a single zone. Thus, this study developed a new integrated screening approach to overcome the above-mentioned shortcomings of current screening techniques and to achieve a balance between efforts toward accuracy and efficiency.

For conducting macro level safety analyses, the research team faced several challenges. First, using current Traffic Analysis Zones (TAZs) as basic geographic units caused a high percentage of boundary crashes because TAZs were delineated for transportation planning but not for traffic crash analysis. In order to solve this problem, the research team used regionalization to develop a new study unit: Traffic Safety Analysis Zones (TSAZs) systems. In other words, this regionalization can alleviate limitations of the TAZ system by aggregating TAZs into a sufficiently large and homogenous zonal system. The research team used the Brown-Forsythe test to select the optimal scale since it minimizes boundary crashes and zones without rare types of crashes (e.g. fatal). Approximately 10% of boundary crashes have been integrated in new zones after regionalization but more than 60% of crashes still occur on the boundary of TSAZs. Hence, a nested structure was proposed to estimate safety performance models separately for boundary and interior crashes. This nested structure allows different contributing factors for different crash types, so this model structure can provide more accurate

and predictable results than a single model. The six types of crashes in each model are varied based on their locations (boundary or interior) and roadways (FACR, other state roads or non-state roads). They are FSB (FACR State road Boundary crashes), FSI (FACR State road Interior crashes), OSB (Other State road Boundary crashes), OSI (Other State road Interior crashes), NSB (Non-state road Boundary crashes) and NSI (Non-state road Interior crashes). In addition, a Bayesian Poisson Lognormal Spatial Error Model (BPLSEM) was adopted for the SPF analysis in this nested structure. The BPLSEM contains a disturbance term for handling the over-dispersion problem, and its spatial error term can control for the spatial autocorrelation of crash data. In addition, the PSI (Potential for Safety Improvements), the difference between the expected crash count and the predicted crash count, was used as our measurement to define hot-zones. The PSI is the approach used in the HSM for microscopic network screening.

As for the micro level analysis, the research team developed SPFs based on the major function classes of roads in our study area (Osceola, Seminole and Orange counties). For these segments, there are rural 2 lanes undivided, rural 2 or 4 lanes divided, urban 2 lanes divided, urban 4 lanes divided, urban 2 or 4 lanes undivided, six or more lanes interrupted (i.e., partial access control) roads, one way roads, and 3 lane with Two Way Left Turn Lane (TWLTL). For the intersection, there are 4 Leg Intersections and 3 Leg Intersections. Overall, these road classes are consistent with the HSM road classification. Moreover, this study includes some new roadway types that are not presented in the HSM, such as six or more lanes interrupted roads. Because there is no existing SPF or reference group data available, a Full Bayesian model was used to estimate the PSI value for different roadway types in the study area. A Poisson log-normal model with random effect was employed. For the segment, the independent variables were Average Annual Daily Traffic (AADT) and segment length. For the intersection, the model fitting procedure was similar as with the segments. The research team still used the Full Bayesian Poisson Lognormal models to predict crash frequency but tried four different variable combinations to identify the best model.

After identifying hot spot areas at both the macro- and microscopic levels, the research team integrated these macroscopic and microscopic screening results. However, this integration task was challenging because we needed to (1) combine various SPFs from different scales, areas, and roadway types; (2) determine an appropriate weight for each group; and (3) choose a measurement for our final results.

In order to solve the above mentioned problems, this study then developed a new criterion to identify whether a zone has safety issues at the macro- and/or microscopic levels. All TSAZs were classified into twelve categories that include two scale groups (macro or micro) and four safety levels (hot, normal, cold, or no data). These categories are: HH, HN, HC, HO, NH, NN, NC, NO, CH, CN, CC, and CO. The first character of the classification represents the macroscopic safety risk, and the second character illustrates the microscopic safety risk.

Then, the research team defined weights for different scales and roadway types. At the macroscopic level, TSAZs were ranked by their zonal PSIs; at the microscopic level, the calculation of average PSI was more complicated because each TSAZ had several intersections and segments. The PSIs of the intersections in each TSAZ were averaged by the number of intersections, and the zones were ranked by their averaged intersection PSI. Simultaneously, the PSIs of segments in each zone were averaged by the total length of the segments in the zone, and zones were ranked by their averaged segment PSI. After that, both the intersection and segment PSI ranks were averaged; the TSAZs were ranked by the final averaged intersection and segment

PSIs. As was the case at the macroscopic level, TSAZs with top 10% micro-level PSIs were categorized as “Hot” zones at the microscopic level.

Finally, the percentile ranks of the PSIs were used in the integration (instead of the original PSIs) because the units of PSI intersections and PSI segments were different. The research team analyzed hot TSAZs for both total crashes and fatal-and-injury crashes in order to be consistent with the HSM. Moreover, by doing so the results also allowed an examination of whether there are any differences with regards to hot zone locations among various crash severity levels. The total crash hot zone screening results display the overall crash distributions within the study area, whereas the fatal-and-injury crash hot zone screening results represent the more severe crash distributions.

In summary, this study presents an integrated screening method that can be used to overcome the shortcomings of macro- and micro-level approaches. In particular, our results provide a comprehensive perspective on appropriate safety treatments by balancing the accuracy and efficiency of screening. Also, it is recommended that different strategies for each hot zone classification be developed because each category has distinctive traffic safety risks at each of the different levels.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

LIST OF ACRONYMS

AADT	Average Annual Daily Traffic
ADT	Average Daily Traffic
AIC	Akaike Information Criterion
BG	Block Groups
BPLSEM	Bayesian Poisson Lognormal Spatial Error Model
CAR	Crash Analysis Reporting
CB	Census Blocks
CC	Macro Cold / Micro Cold
CH	Macro Cold / Micro Hot
CN	Macro Cold / Micro Normal
CO	Macro Cold / Micro No Data
CRP	Continuous Risk Profile
CT	Census Tracts
DB	Database
DIC	Deviance Information Criterion
DOT	Department of Transportation
DUI	Driving Under the Influence
EB	Empirical Bayes
FAC	Full Access Control
FACR	Full Access Control Road
FB	Full Bayesian
FDOT	Florida Department of Transportation
FI	Fatal-and-Injury
FIU	Florida International University
FSB	FACR State road Boundary crashes
FSI	FACR State road Interior crashes
GIS	Geographic Information Systems
HC	Macro Hot / Micro Cold
HH	Macro Hot / Micro Hot

HN	Macro Hot / Micro Normal
HO	Macro Hot / Micro No Data
HSM	Highway Safety Manual
LOSS	Level of Safety Service
MAD	Mean Absolute Deviation
MAUP	Modifiable Areal Unit Problem
MPO	MetroPlan Orlando
MSPE	Mean Square Prediction Error
NB	Negative Binomial
NC	Macro Normal / Micro Cold
NH	Macro Normal / Micro Hot
NN	Macro Normal / Micro Normal
NO	Macro Normal / Micro No Data
NSB	Non-State road Boundary crashes
NSI	Non-State road Interior crashes
O-D	Origin-Destination
OSB	Other State Boundary crashes
OSI	Other State Interior crashes
OSO	Osceola, Seminole and Orange Counties
PAVECOND	Pavement Condition
PDO	Property Damage Only
PSI	Potential for Safety Improvement
RCI	Roadway Crash Inventory
SA	SafetyAnalyst
SECTADT	Section Average Annul Daily Traffic
SES	Socio-Economic Status
SPF	Safety Performance Function
TAD	Traffic Analysis District
TAZ	Traffic Analysis Zone
TSAZ	Traffic Safety Analysis Zone

TSP	Transportation Safety Planning
TWLTL	Two Way Left Turn Lane
VMT	Vehicle-Miles Traveled

1. INTRODUCTION

Previous studies have been conducted to identify high risk traffic zones; an examination of these investigations has shown that they could have benefitted from a macro-level screening of the study area. All Highway Safety Manual (HSM) volumes (including Part B) are location based, which is a more microscopic level of screening. In Part B of the HSM, network screening is key; however, such screening is based on locations/sites. Transportation Safety Planning (TSP) is macroscopic and deals with loosely defined "areas." These areas might include a few intersections, segments, etc. We focus on two levels of analysis: microscopic and macroscopic. They are both relevant to this type of investigation, even though the HSM does not address this point. The results of TSP could affect transportation safety at the network level, and could also be useful in policy decision making. For example, in the HSM Part B, locations are ranked according to expected crash frequency or rate, and then diagnosed, countermeasures are selected, etc. Using safety planning, we can first rank the TAZ's or other areas with high safety risk (potentially using the same approaches as expressed in the HSM). Then, within these areas we can screen the network. This would be a two-level approach using both TSP and the HSM. The benefit is that we would be able to improve whole areas facing multiple problems, which is especially pertinent because prioritization is part of the HSM Part B's listed objectives. This two-level approach will improve screening by simultaneously examining the problem at both the macroscopic and individual site levels. It is able to provide area-wide treatment in some zones, and a combination of integrated treatments in others. Therefore, the objectives of this research are as follows:

- To integrate a macro-level component with network screening methodologies in order to identify and rank those zones and sites where improvements have the potential to reduce the number of crashes.
- To provide a broader picture of safety scenarios within the context of both zone and micro-level network components (e.g., intersections, segments, etc.) for safety analysts.
- To present two-level screening results as an added resource for planners because the process would allow for the filtering of zones with a high safety risk, and thus will provide a window for proactive safety management in both long and short-range transportation plans.
- Practical implications for integrating safety and planning.
- Improvement to the HSM Part B.
- A framework to find the appropriate mix of micro- and macro-level screening within the context of the HSM Part B procedures.

This report is divided into ten chapters. A review of existing network screening methods is provided in Chapter 2. Data collection and preparation are presented in Chapter 3. Chapter 4 lists the main challenges and corresponding solutions for macro-level screening. Developing SPFS and defining hotspots at the macroscopic level are discussed in Chapters 5 and 6, respectively. Chapters 7 and 8 are dedicated to a micro-level analysis. Chapter 9 provides the integration results. Finally, conclusions and recommendations are provided in Chapter 10.

2. LITERATURE REVIEW

This literature review focuses on the two major aspects of interest in this project: macro- and micro-level safety analysis.

2.1. Micro-level Traffic Safety Analysis

Network screening is a process for reviewing a transportation network to identify and rank sites with respect to safety risk, then provide a ranking from most likely to least likely to realize a reduction in crash frequency with the implementation of a countermeasure.

There is a growing body of literature that has focused on the development of traffic safety network screening. The majority of these studies are specifically at the microscopic level, which deals with the safety screening of road segments or intersections, as well as other types of spots. Several methods have been developed in the last a few decades. The principal network screening methods include:

- Table C method
- Level of Safety Service (LOSS) method
- Empirical Bayes (EB) method
- Continuous Risk Profile (CRP) for highway segments
- Screening based on high proportions
- Detection of safety deterioration over time

The Table C method (Ragland et al., 2007) identifies sites that have experienced a considerably greater number of crashes per unit of ADT than the average. For highway segments, roadway units are screened by sliding a window of 0.2 miles in increments of 0.02 miles. Alternatively, for intersections the influence area is 250 feet from the intersection; all crashes within the influence area are considered to be intersection crashes. The criteria for an area to be considered a hotspot are as follows: 1) the observed crash frequency is more than the average for the rate group with 99.5% confidence level in either a three, six, or twelve month period; and 2) four or more crashes in the given time period.

Level of Service of Safety was proposed by Kononov et al. (2003). The LOSS method is similar to the Table C method in that the observed crash frequency is compared to an expected crash frequency, and then the level of deviation is measured. The Table C method considers whether the deviation is large enough for a statistically significant indication that more crashes occurred than would be expected for the average site. In the LOSS method, the deviation from the expected for an average site is shown by creating four categories of service levels. The expected LOSS for similar sites is determined according to safety performance functions (SPFs) using traffic volume, number of lanes, lane width, and so forth. Due to this use of SPFs, the LOSS method is superior to the Table C method; it removes the use of constant crash rates.

The EB method began with its application to traffic safety by Abbess et al. (1981). The EB method is actually a suite of screening methods based on the EB strategy of estimating the long-term expected crash frequency for a location. This method was a preferred method in the Highway Safety Manual. The EB estimate of expected crash frequency for a location is a weighted combination of the predictions obtained from an SPF and the observed crash frequency for the given location. The weights are calculated based on an EB that makes use of the over-dispersion parameter (which is an outcome of the SPF development, using a negative binomial model).

The Continuous Risk Profile (CRP) method was introduced by Chung et al. (2007). CRP deals only with observed crashes. The key concept behind the CRP method is that a continuous profile plot of risk along a roadway can be helpful to identifying sites of high risk.

Screening based on high proportions was suggested by Heydecker et al. (1991). This method identifies and ranks the sites that have a certain proportion of a specific crash type relative to a total number of crashes that is higher than some average or threshold proportion value for similar road types.

Last but not least, Hauer (1996) developed a methodology that detects safety deterioration over time. In this method, two tests are conducted. The first detects any potentially gradually increasing trend in mean crash frequency. The second detects any potential for a sudden increase in mean crash rate, and can be ranked if necessary.

The Highway Safety Manual (HSM, 2010) summarized many of the previously-mentioned methods and presented a clear process for developing a micro-level network screening program, as shown in Figure 2-1.

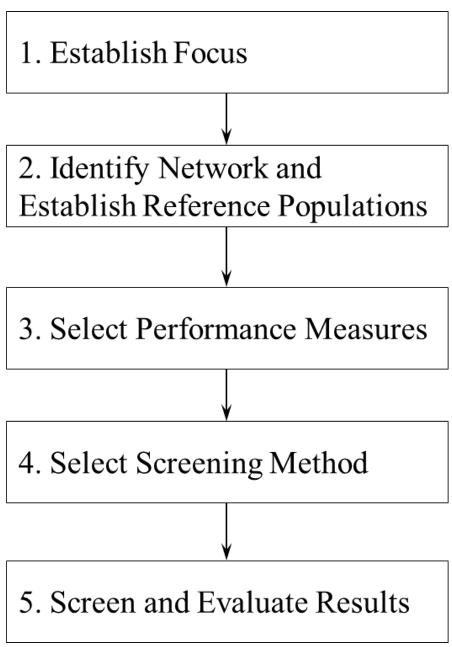


Figure 2-1 Network screening steps

As can be seen from Fig. 1, there are five steps in our micro-level network screening process. First of all, establishing focus identifies the purpose or the intended outcome of the network screening analysis. The second step, identifying network and establishing reference populations, specifies the types of sites and facilities being screened (e.g., segments, intersections, etc.) and groups together similar sites or facilities. In the next step, fourteen performance measures are provided to measure the expected crash frequency or other equivalent values obtained from the site. Following this step, several methods are provided for screening. There are three principle screening methods: the ranking method, the sliding window method and the peak searching method. The final step in the network screening process is conducting the screening analysis and evaluating the results.

2.2. Macro Level Traffic Safety Analysis

Several researchers have conducted studies analyzing traffic safety at the macroscopic level. Table 2-1 summarizes the majority of previous macro-based work. As shown in Table 2-1, researchers primarily have focused on total, severe, fatal and bicycle/pedestrian related crashes in various areal units such as TAZs, census tracts, block groups, ZIP areas, and so forth.

Recently, Siddiqui et al. (2011) found that the pedestrian and bicycle model with spatial correlation (hierarchical Bayesian) performs better than alternative methods. Abdel-Aty et al. (2013) compared TAZs, block groups, and census tracts models and concluded that MAUP (the Modifiable Areal Unit Problem) presents in the macroscopic crash modeling and the new zone system exclusively developed for the safety analysis is required.

A novel approach was proposed by Siddiqui et al. (2011) to account for the spatial influence of neighboring zones on pedestrian and bicycle crashes which occur specifically on or near zonal boundaries. It was found that crash models (which separately account for boundary and interior crashes) had better goodness-of-fit measures compared to models with no specific consideration for crashes located at or near zone boundaries. Siddiqui (2012) described how motor vehicle crashes were classified as 'on-system' and 'off-system' crashes; two sub-models were fitted in order to calibrate the safety performance function for these crashes. In conclusion, it was evident by comparing this on- and off-system sub-model-framework to other candidate models that it provided superior goodness-of-fit for both total and severe crashes.

Table 2-1 Summary of macroscopic studies

Year	Author(s)	Target	Areal Units
1995	Levin, Kim, & Nitz	Total crash	Block Group
1998	Blatt & Furman	Fatal crash (all, young, male/female, drinker & child fatality)	ZIP
2000	LaScala, Gerber, & Gruenewald	Pedestrian crash	Census Tract
2000	Stamatiadis, & Puccini	Fatal crash	ZIP
2002	Ng, Hung, & Wong	Fatal, Pedestrian crash	TAZ
2003	Hadayeghi, Shalaby, & Persaud	Total, Severe crash	TAZ
2003	Amoros & Laumon	Injury crash	County
2003	Noland	Fatal, Injury crash	State
2003	Clark	Fatal crash	ZIP
2004	De Guevara, Washington, & Oh	Fatal, Injury, PDO crash	TAZ
2004	Noland & Quddus	Fatal, Injury crash	Census Ward
2004	Noland & Quddus	Bicycle, Pedestrian crash	Standard Statistical Regions
2004	Noland & Oh	Fatal, Injury crash	County
2004	MacNab	Injury cash	Local Heath Areas
2006	Hadayeghi, Shalaby, Persaud, & Cheung	Total, Severe crash (during morning peak)	TAZ
2006	Aguero-Valverde & Jovanis	Injury, Fatal crash	County
2006	Kim, Brunner & Yamashita	Total, Vehicle-vehicle, Pedestrian and Bicycle crash	Grid-based Structure
2006	Romano, Tippetts, Blackman, & Voas	Fatal crash	ZIP
2007	Loukaitou-Sideris, Liggett, & Sung	Pedestrian fatal crash	Census Tract
2008	Quddus	Fatal and Injury crash	Census Ward
2009	Wier, Weintraub, Humphreys, Seto, & Bhatia	Pedestrian crash	Census Tract
2010	Hadayeghi, Shalaby, & Persaud	Total and Severe crash	TAZ
2010	Naderan & Shashi	Total, Severe and PDO crash	TAZ
2010	Cottrill & Thakuriah	Pedestrian crash	Census Tract
2010	Huang, Abdel-Aty, & Darwiche	Total and Severe crash	County
2011	Abdel-Aty, Siddiqui, & Huang	Total, Severe, Peak-hour and Pedestrian/Bicycle crash	TAZ
2011	Ukkusuri, Hasan, & Aziz	Pedestrian crash	Census Tract
2011	Siddiqui & Abdel-Aty	Pedestrian crash	TAZ
2012	Siddiqui, Abdel-Aty, & Choi	Bicycle and Pedestrian crash	TAZ
2012	Siddiqui	Total, Severe, Pedestrian and Bicycle crash	TAZ
2013	Abdel-Aty, Lee, Siddiqui, & Choi	Total, Severe and Pedestrian crash	TAZ, BG, CT

2.3. Comparison between Micro- and Macro-level Analyses

Previous studies have contributed to the understanding and application of traffic safety network screening. However, there is much room for improvement at both the micro- and macroscopic levels of network screening. For example:

At the microscopic level:

- Some performance measures require the employment of Safety Performance Functions (SPFs) in the HSM. These SPFs might be built with incomplete crash data. Fortunately, the research group has collected complete crash data (both long form and short form) in Florida. The validation of SPFs in the HSM with complete data is necessary.
- Since many network screening methods have been developed, the selection of screening methods is critical because some methods may not be capable of identifying the actual hotspots. Discovering and/or developing the most appropriate methods can help make FDOT more efficient in their traffic safety planning.
- The application of methods in the HSM is time consuming. Making the network screening methods readily applicable for FDOT engineers would be desirable.

At the macroscopic level:

- Even though much research has been done at the macroscopic level with regards to traffic safety, studies of macro-level network screening methods are rare. There is a need for the development of a complete methodology and set of guidelines for macro-level zonal screening.
- Previous studies have compared the performance of crash prediction methods with different sets of demographic and other spatial factors. However, none developed methods for selecting variables for macro-level models that could optimize predictions.

- Boundary and autocorrelation issues in spatial data analysis are very important in macro-level studies. Though many scholars have proposed various solutions to these issues, there is still much room for improvement.

In addition, although many safety studies have been conducted either at the macroscopic level or in microscopic network screening, no studies have tried to integrate macro- and micro-level screening. Therefore, the integration of these two levels of network screening would be desirable. The major objective of this project is to develop a method of simultaneously investigating safety issues at the macroscopic and individual site levels. The results of this project will help FDOT engineers and planners make decisions regarding the distribution of funding in traffic safety planning.

3. DATA COLLECTION AND PRELIMARY ANALYSIS

For the macroscopic analysis, all of the data were prepared based on TAZs. Since crash data contains coordinate data, each crash can be displayed on a GIS map. The crashes were counted based on TAZs. The US Census Bureau does not provide demographic data based on TAZs. Instead, it provides data as a census block (CB). Since TAZs consists of one or more CBs or combinations of CBs, the CB-based data can be aggregated into TAZ-based data. On the other hand, demographic and socioeconomic data from MetroPlan Orlando were provided based on TAZs. Roadway data from the FDOT Roadway Crash Inventory (RCI) were also prepared based on TAZs.

3.1. TAZ GIS Map

The TAZ GIS maps were obtained from MetroPlan Orlando. First of all, three TAZ maps were merged using GIS because TAZ GIS maps of three counties (Orange, Seminole, and Osceola) were provided separately from the data provided by MetroPlan Orlando (Figure 3-1).

Table 3-1 describes the basic information from the TAZs as the number of zones and their respective areas. The average areas of Orange and Seminole Counties are relatively small (1.414 mi^2 and 1.519 mi^2, respectively).because most of the zones in Orange and Seminole Counties are urban or suburban, while many zones in Osceola County are rural. Osceola County also has a larger average area (8.107 mi^2). After mergence, the total number of TAZs was found to be 1,116 in OSO.

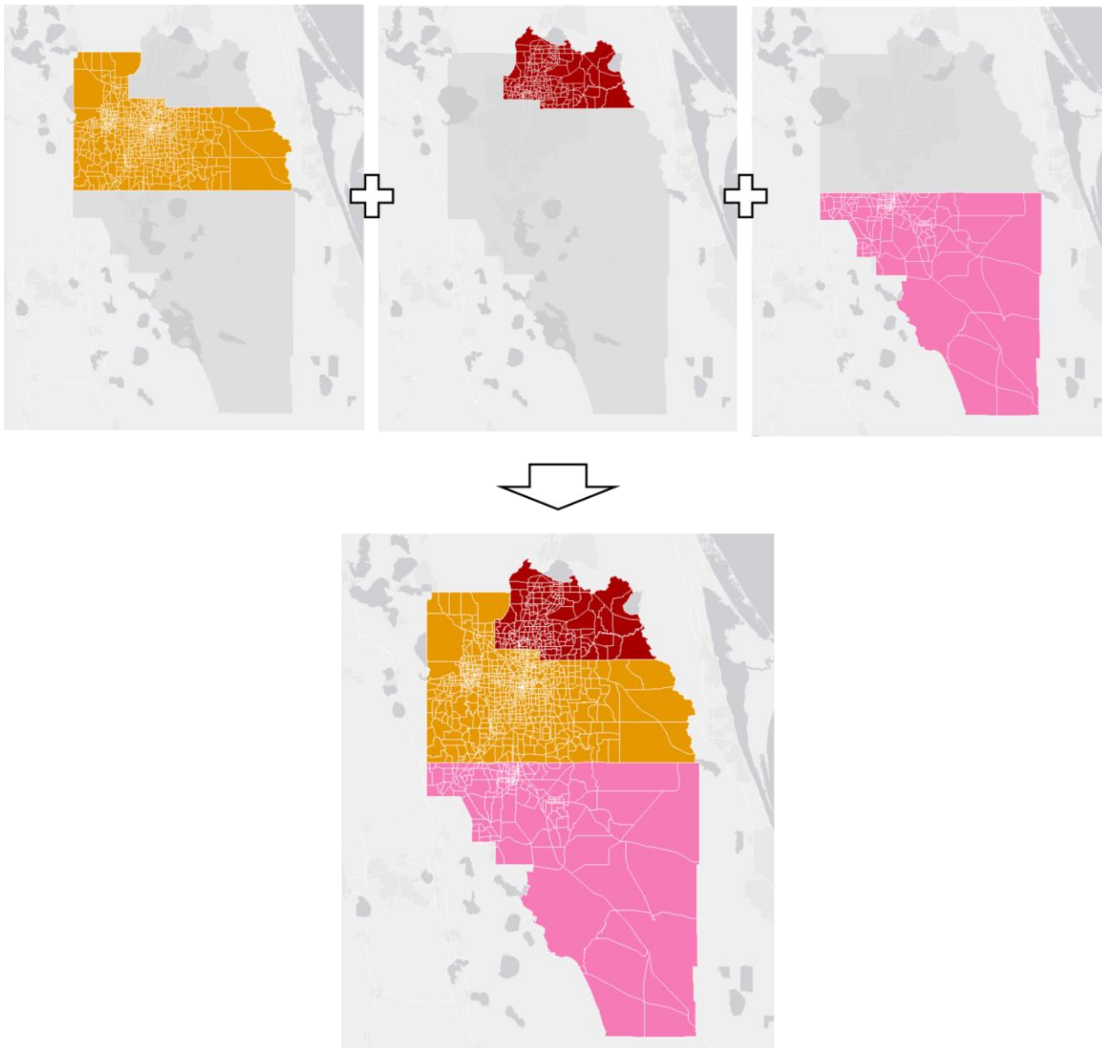


Figure 3-1 Merging of the TAZ GIS maps of Orange, Seminole, and Osceola Counties

Table 3-1 Descriptive statistics of TAZs in OSO (2008-2009)

County	Number of zones	Area by each TAZ (mi^2)			
		Average	Stdev	Min	Max
Orange	711	1.414	4.553	0.021	71.662
Seminole	220	1.519	2.932	0.057	24.634
Osceola	185	8.107	23.770	0.061	170.664
Merged zones	1,116	2.544	10.710	0.021	170.664

3.2. Crash Data

Figure 3-2 summarizes the overall process of crash data collection. Two forms of crash reports are used in the State of Florida: short form and long form crash reports. In Florida, a long form is used when the following criteria are met:

- Death or personal injury
- Leaving the scene involving damage to attended vehicles or property (F.S. 316.061(1))
- Driving while under the influence (F.S. 316.193)
- Officer's discretion

The short form is used to report other types of PDO traffic crashes.

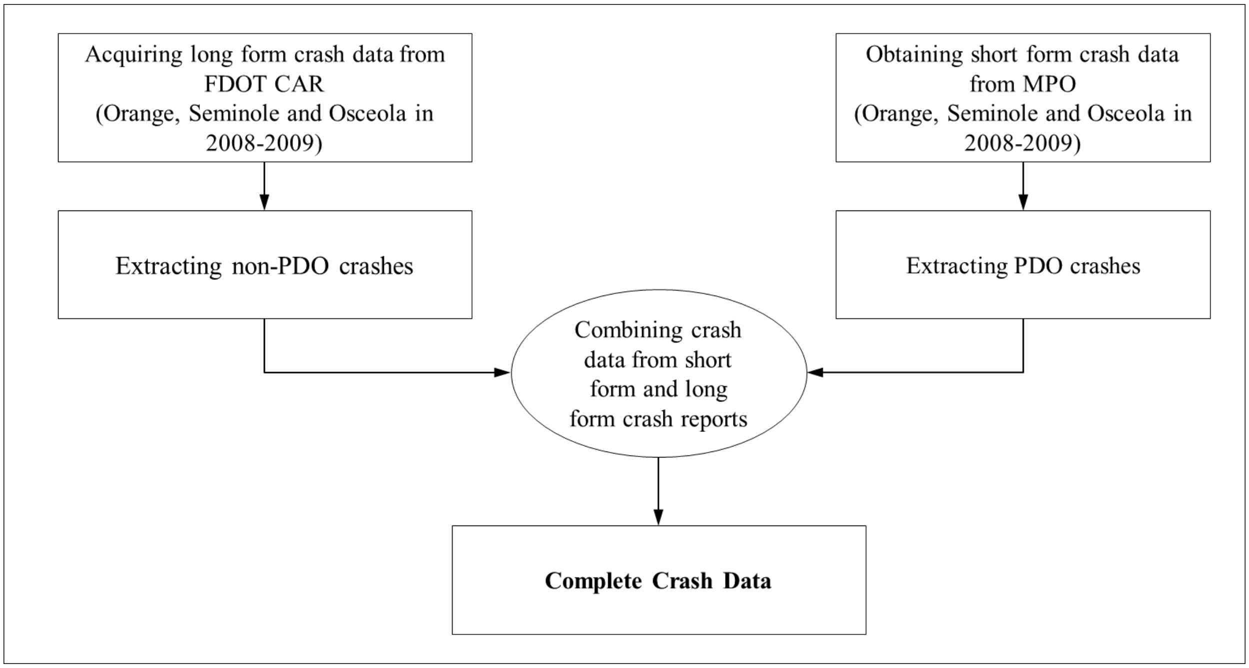


Figure 3-2 Crash data collection process

Crashes reported on the long form involve either a higher injury severity level or criminal activities such as hit-and-runs or DUIs. Only long form crashes are coded and archived in CAR (Crash Analysis Reporting) DB. Therefore, previous researchers could only get access to long form crashes for their crash analyses. Fortunately, this research group was able to collect short form crash data for 2008-2009 in OSO from MetroPlan Orlando and other sources. Therefore, the research team was able to use more complete crash data for this research project.

Table 3-2 shows the descriptive statistics based on TAZs of traffic crashes in OSO. On average, 80.503 total crashes and 58.341 PDO (Property Damage Only) crashes occurred in one TAZ over the two year period used for this study. Severe crashes occurred much less than PDO crashes, with 2.272 severe crashes happening in a TAZ, on average.

Table 3-2 Descriptive statistics of traffic crashes based on TAZs in OSO (2008-2009)

Crash types	Mean	Stdev	Min	Max
Total crash	80.503	90.207	0	745
PDO crash	58.341	68.903	0	586
Severe crash	2.272	2.778	0	21
Fatal crash	0.376	0.779	0	6
Bike crash	0.591	1.058	0	8
Pedestrian crash	0.811	1.351	0	12
Rear-end crash	20.838	35.112	0	437
Angle crash	5.871	7.985	0	66
Head-on crash	1.837	3.041	0	27
Sideswipe crash	3.272	4.922	0	39
Off-road crash	4.076	5.762	0	62
Rollover crash	0.652	1.608	0	20
Left-turn crash	4.487	7.855	0	73
Right-turn crash	1.315	2.729	0	26

3.3. Roadway/Traffic Data

Roadway data were collected from two different sources: FDOT RCI and MetroPlan Orlando. Traffic data such as total AADT and truck AADT were acquired only from FDOT RCI. FDOT RCI provides various types of roadway data including functional classification, number of lanes, speed limits, median types, pavement conditions, etc. Unfortunately, because FDOT focuses mainly on interstate highways and state roads, many of the collector and local roadway data could not be provided by RCI. The roadway data from MetroPlan Orlando incorporated all types of roadways, including collectors and local roadways. However, MetroPlan Orlando roadway data does not have any traffic-related factors except for speed limits; thus, only posted speed limit data from MetroPlan Orlando were used. As shown in Figure 3-3, it is evident that

MetroPlan Orlando has considerably more complete roadway information, especially for local roads.

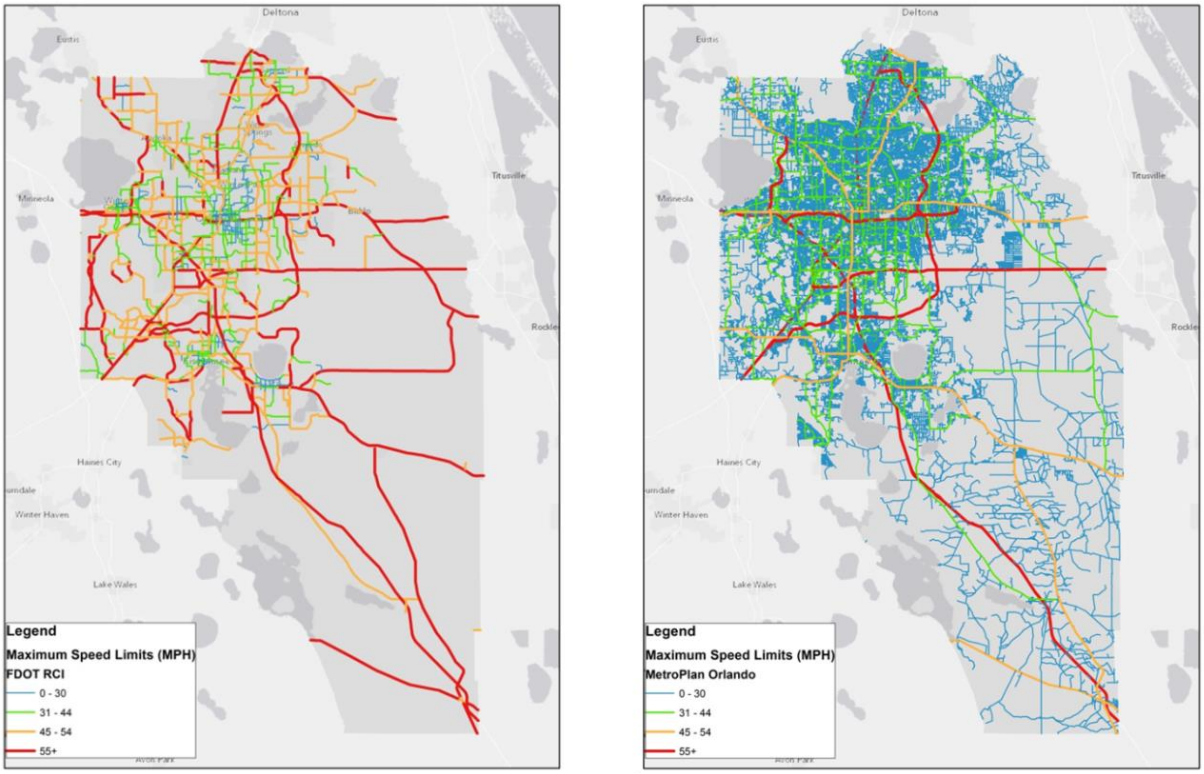


Figure 3-3 Comparison of roadway sections between FDOT (left) and MetroPlan Orlando (right)

Other data such as functional classification, speed limits, median types, pavement conditions, overall AADT, and truck AADT were collected from FDOT RCI and were used in this analysis. Table 3-3 summarizes the roadway variables in OSO. Figures 3-4, 3-5, 3-6, 3-7, and 3-8 present functional classifications, pavement conditions, posted speed limits, overall AADT, and truck AADT, respectively.

Table 3-3 Descriptive statistics of roadway/traffic variables based on TAZs in OSO (2012)

Variables	Mean	StDev	Min	Max	Source
No of traffic signals	0.737	1.268	0.000	9.000	FDOT RCI
No of intersections	13.909	11.958	0.000	78.000	
Roadway length of the bridge	0.066	0.206	0.000	3.078	
Roadway length of interstate/expressways	0.190	0.498	0.000	4.309	
Roadway length of principle arterials (except for interstate/expressways)	0.451	1.703	0.000	42.315	
Roadway length of minor arterials	0.424	0.814	0.000	14.689	
Roadway length of collectors	0.793	1.217	0.000	13.104	
Roadway length of local roads	0.064	0.263	0.000	4.349	
Roadway length with the speed limit less than 25MPH	9.139	12.546	0.000	195.969	MetroPlan Orlando
Roadway length with the speed limit 45MPH or over	1.649	4.057	0.000	72.168	
Roadway length without the median	2.101	2.619	0.000	45.825	
Roadway length with the traversable median	0.577	0.937	0.000	13.074	
Roadway length with the non-traversable median	0.617	1.088	0.000	23.121	
Roadway length with the pavement condition: poor	0.150	0.448	0.000	5.126	
Roadway length with the pavement condition: fair	0.269	0.669	0.000	7.008	
Roadway length with the pavement condition: good	2.532	3.733	0.000	70.780	
Vehicle-miles traveled (VMT)	93948.387	100954.250	0.000	839660.390	
Truck vehicle-miles traveled (Truck VMT)	2857.102	5036.736	0.000	86779.091	

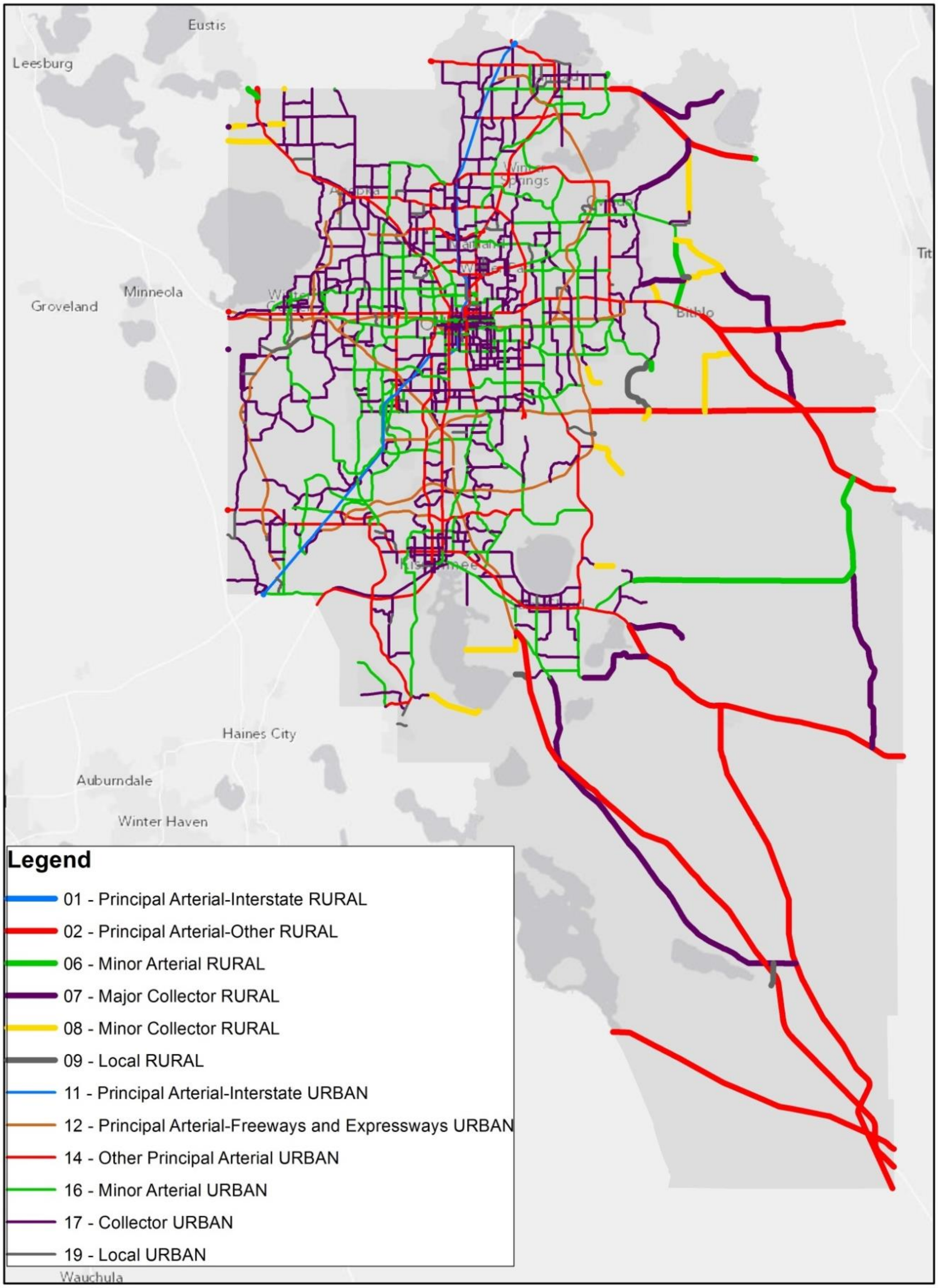


Figure 3-4 Roadway functional classifications in OSO (2012)

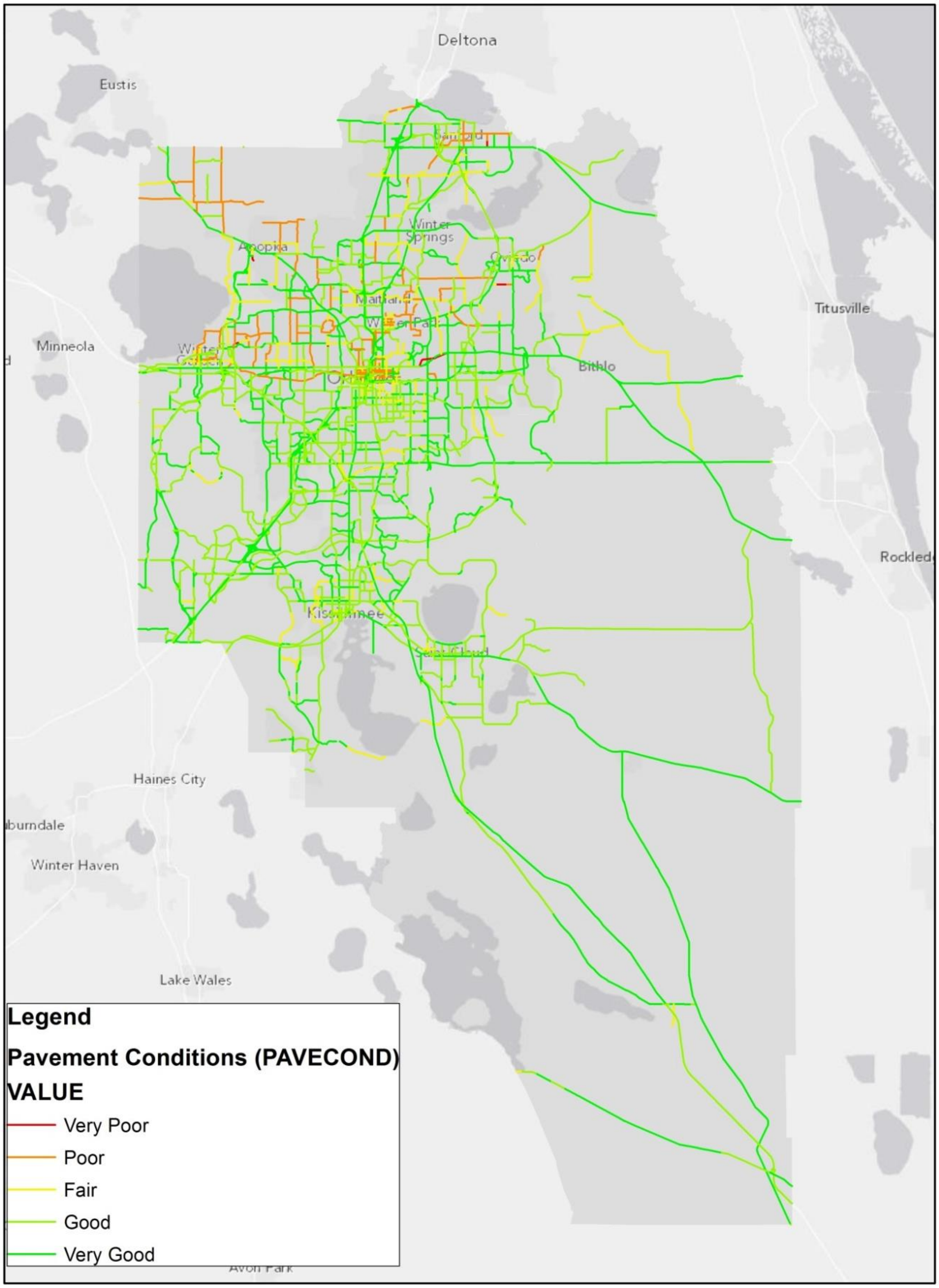


Figure 3-5 Roadways by pavement condition in OSO (2012)

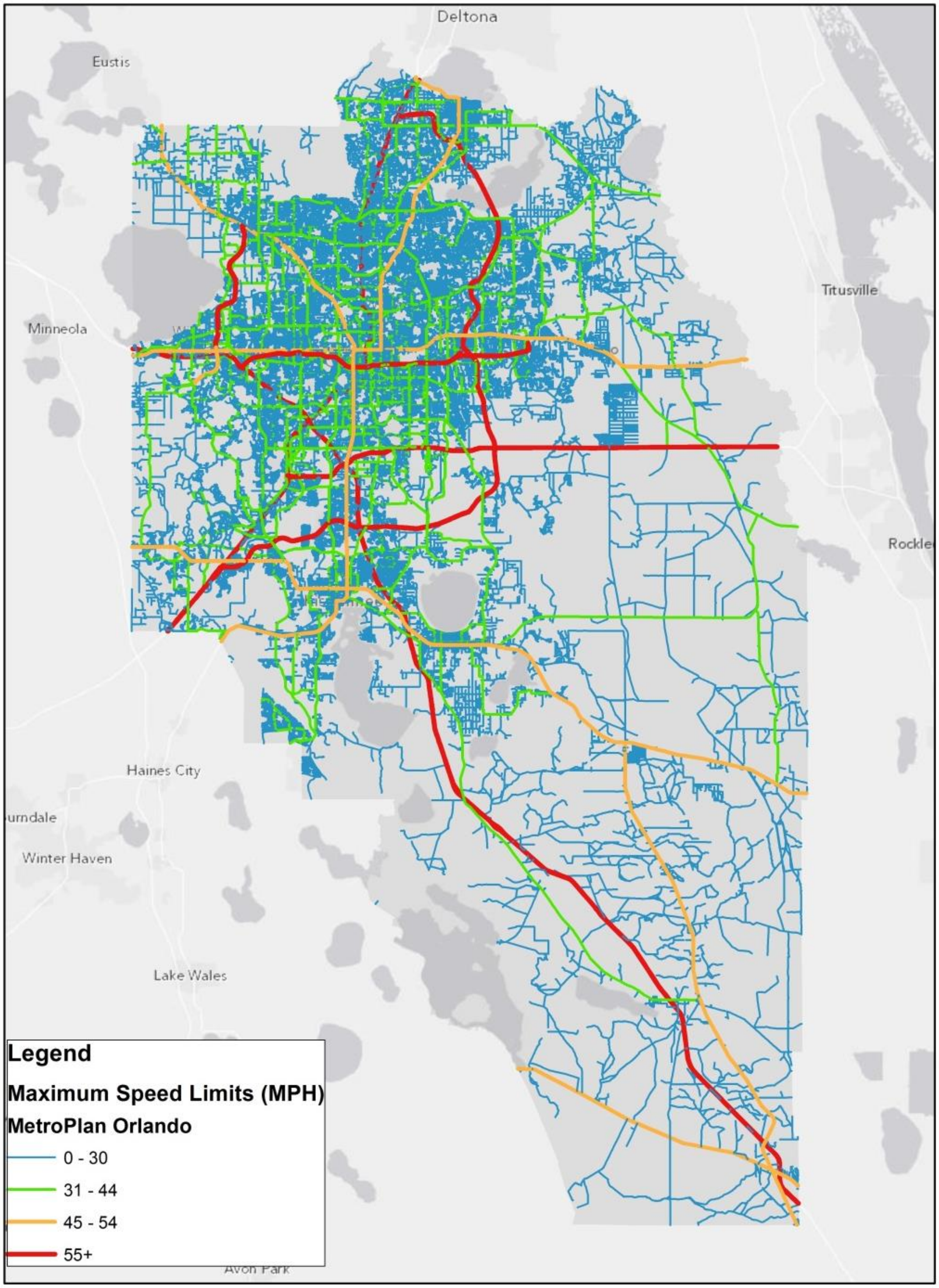


Figure 3-6 Roadways by posted speed limits in OSO (2012)

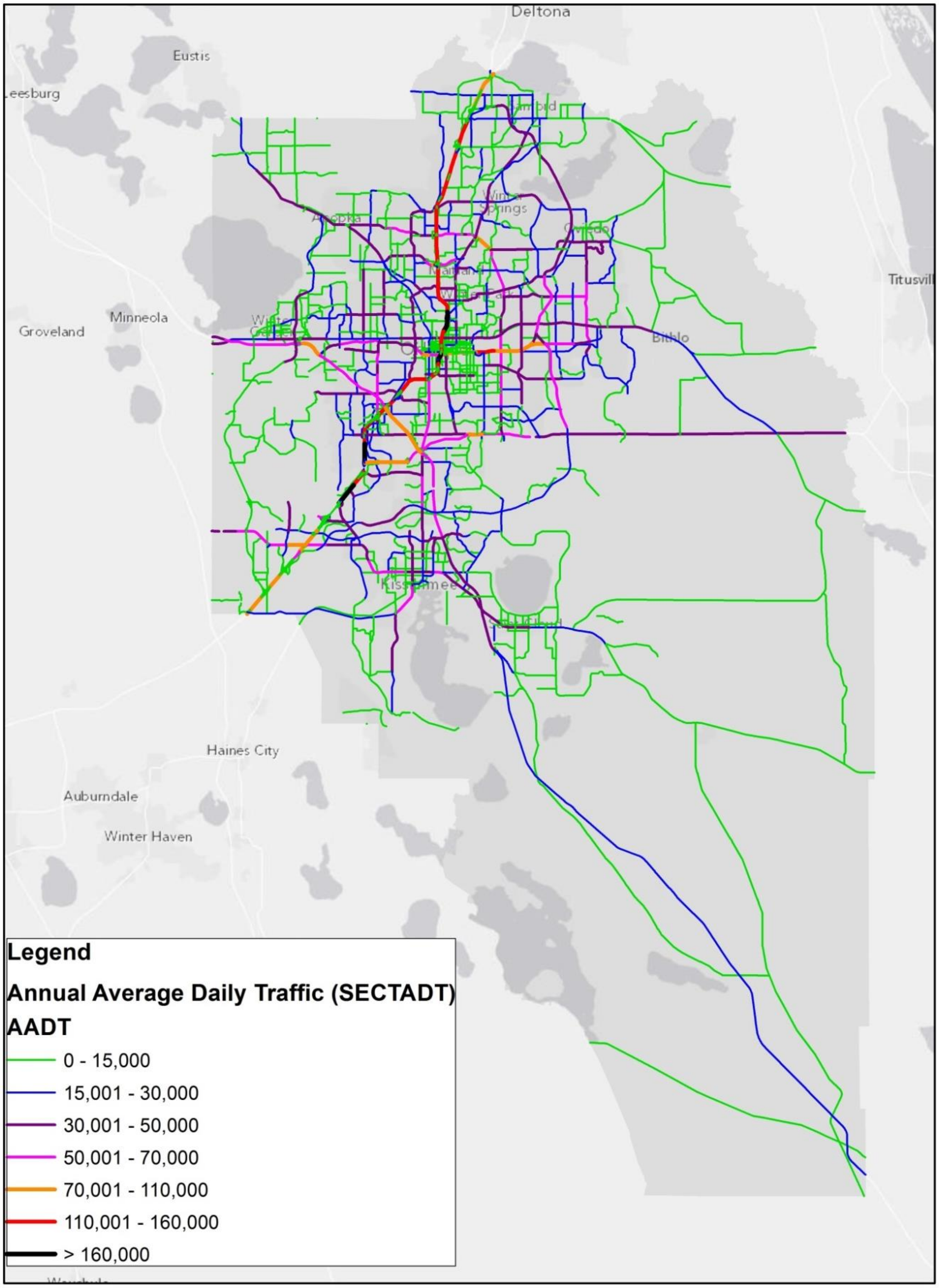


Figure 3-7 Roadways by overall AADT in OSO (2012)

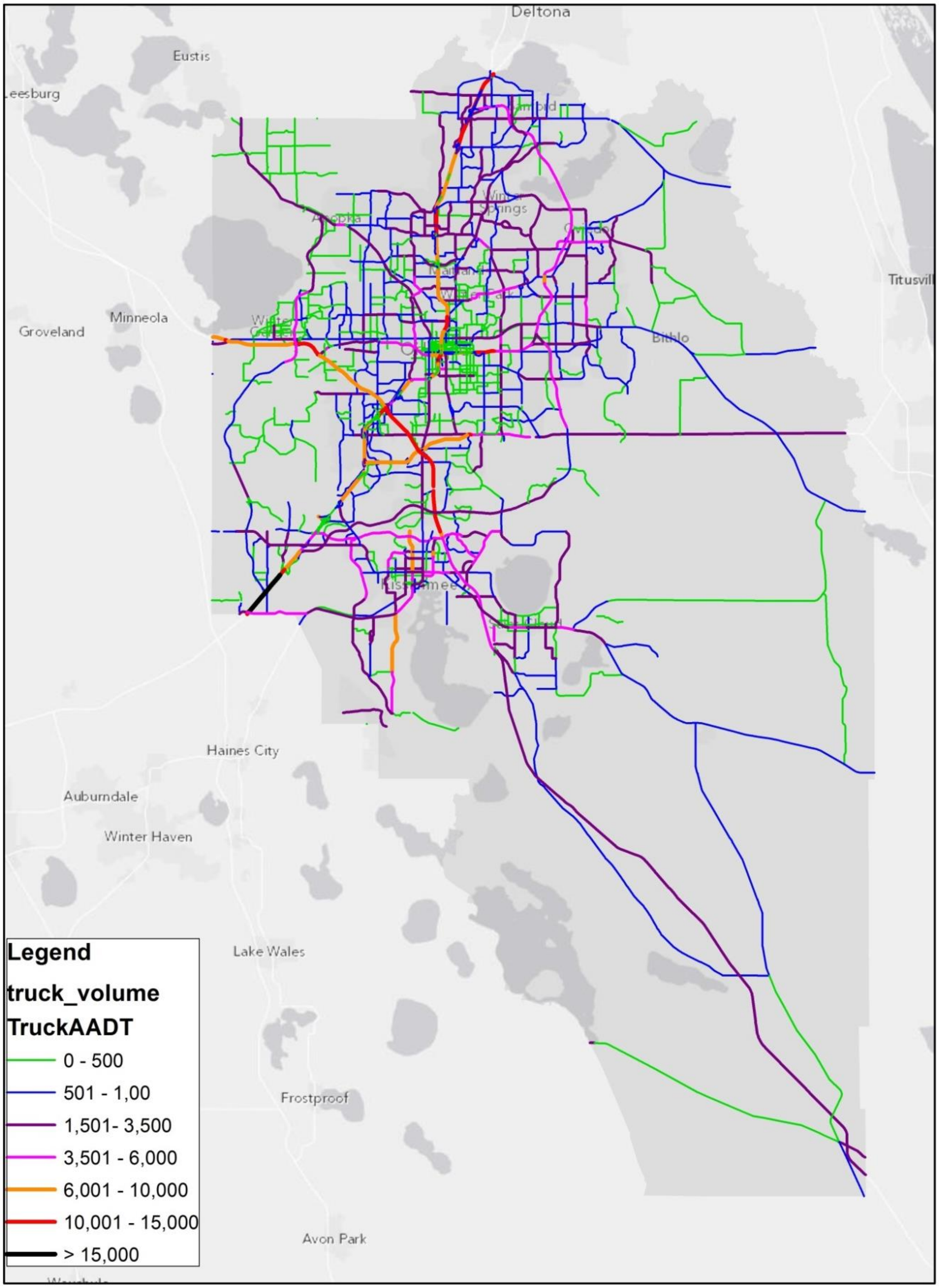


Figure 3-8 Roadways by truck AADT in OSO (2012)

3.4. Demographics and Socioeconomic Data

Both demographic and socioeconomic data based on TAZs were acquired from MetroPlan Orlando, and CB-based data were collected from the US Census Bureau. The US Census does not offer data based on TAZs. Instead, it provides the data on a census block (CB) basis. TAZs consist of one or more CBs, or combinations of CBs (Figure 3-9). Thus, CB-based data can be aggregated into TAZ based data. This process is summarized in Figure 3-10.

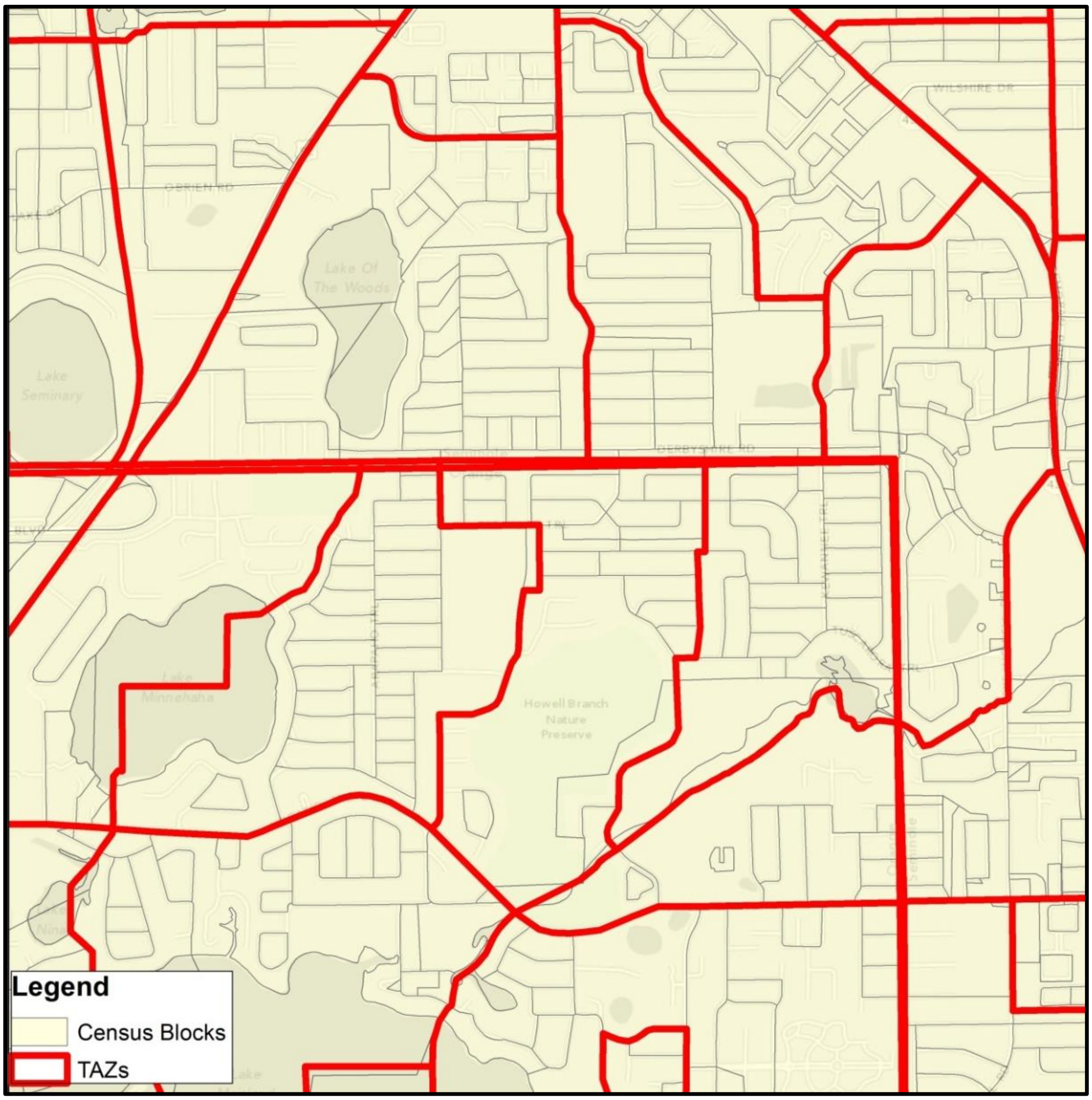


Figure 3-9 Census blocks within TAZs

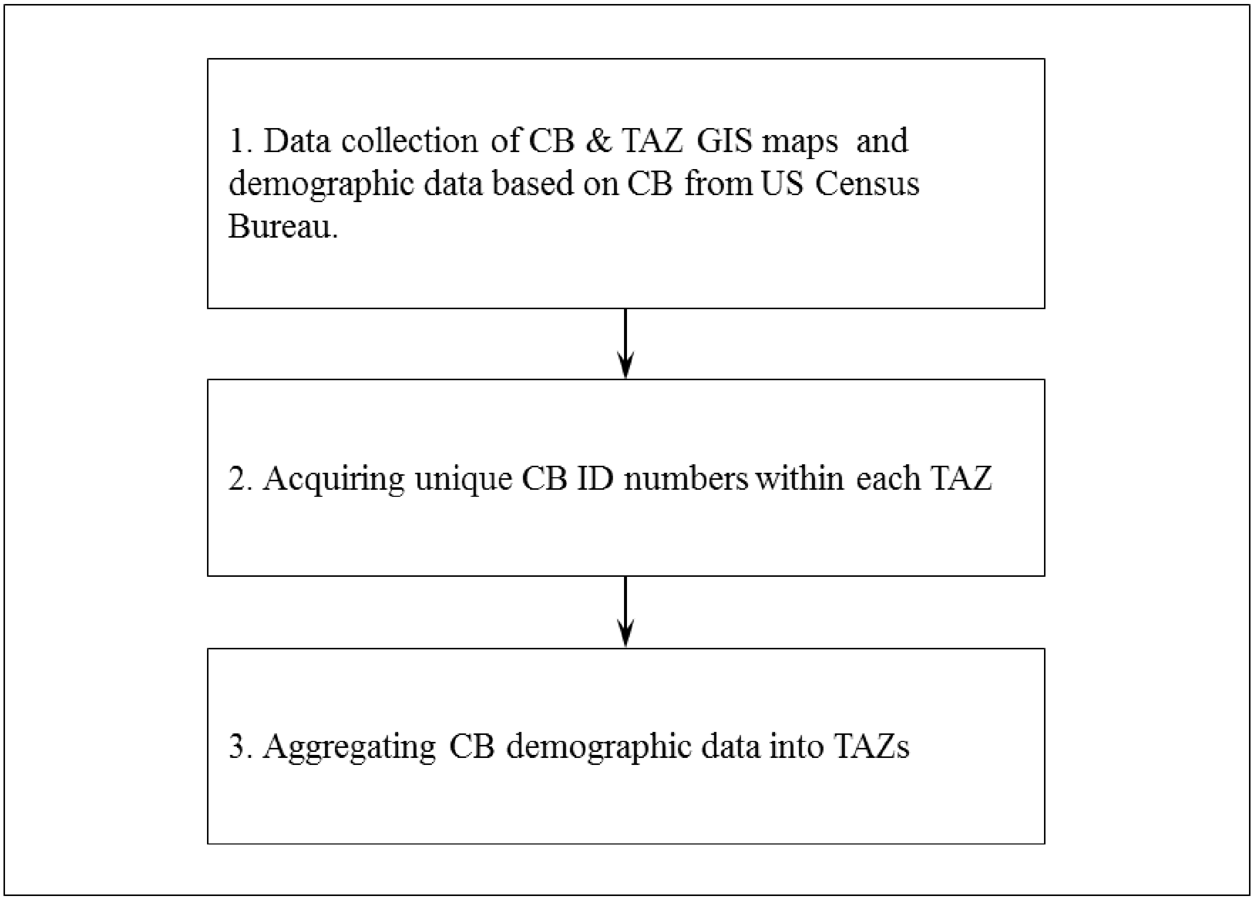


Figure 3-10 Process of conversion of CB data to TAZ-based data

After the process of data conversion, demographic data such as total population, proportion of African Americans, proportion by age group, and proportion by gender were obtained based on TAZs. These data are summarized in Table 3-4.

Table 3-4 Descriptive statistics of demographic data based on TAZs in OSO (2010)

Variable	Average	StDev	Min	Max
Total population	852.592	1193.685	0	14401
Population of African American	153.895	352.764	0	3613
Population of Hispanic People	234.435	515.084	0	9815
Population of age group (5-14 years old)	166.199	257.519	0	3380
Population of age group (15-19 years old)	66.528	141.423	0	3160
Population of age group (20-24 years old)	72.282	152.885	0	2255
Population of age group (25-64 years old)	458.322	637.082	0	7666
Population of age group (over 64 years old)	89.262	133.005	0	1196
Male population	418.392	588.079	0	6892
Female population	434.201	612.307	0	7509

MetroPlan Orlando offered vehicle ownership data by household. They also provided hotel, motel, and timeshare rooms, employment, and school enrollment. The basic descriptive statistics of the socioeconomic data from MetroPlan Orlando are presented in Table 3-5.

Table 3-5 Descriptive statistics of demographic and socioeconomic data based on TAZs in OSO (2010)

Variable	Average	StDev	Min	Max
Households without vehicle	44.581	67.582	0	576
Households with one vehicle	270.487	337.795	0	2808
Households with two or more vehicles	434.875	516.957	0	3396
Rooms of hotel, motel and time shares	296.201	1355.119	0	14341
Total employment	875.103	1717.683	0	27088
School enrollment	1042.441	1821.272	0	27732

4. MAIN CHALLENGES FOR MACRO-LEVEL ANALYSES

In this chapter, several challenge for conducting macro-level analyses are identified. First, the limitations of current TAZs as basic geographic units for macro-level analyses are examined. Second, regionalization is used to develop our new zone system: Traffic Safety Analysis Zone (TSAZ). Finally, we address spatial autocorrelation and boundary crash issues.

4.1. Limitation of TAZs

TAZs have been widely used for basic geographical units in many studies (Ng et al., 2002; Hadayeghi et al., 2003; De Guevara et al., 2004; Hadayeghi, Shalaby, Persaud & Cheung, 2006; Hadayeghi et al., 2010; Naderan & Shashi, 2010; Siddiqui & Abdel-Aty, 2011; Abdel-Aty et al., 2011; Siddiqui & Abdel-Aty, 2011; Siddiqui et al., 2012; Siddiqui, 2012). Usually TAZ systems are preferred by traffic researchers over other areal units such as census tracts, block groups, or ZIP areas because the TAZ is the only unit that is also a transportation-related geographical unit. Nevertheless, TAZs have historically been delineated for transportation planning fields to develop long term transportation plans, and not for traffic crash analysis.

TAZs are usually defined based on several criteria (Baass, 1981) in an effort to achieve homogenous socioeconomic characteristics for each zone’s population. These criteria include:

1. Minimizing the number of intrazonal trips;
2. Recognizing physical, political, jurisdictional and historical boundaries;
3. Generating only connected zones and avoiding zones that are completely contained within another zone;

4. Devising a zonal system in which the number of households, population, area, or trips generated and attracted are nearly equal in each zone; and
5. Basing zonal boundaries on census zones.

Possible limitations of TAZs for crash analysis arise from numbers 2 and 3 above. First, TAZs were designed to discover O-D (origin-destination) pairs of trips generated by each zone. In other words, planners want to minimize the number of trips inside a particular zone because they cannot track intra-zonal trips (criteria number 2, above); however, macroscopic traffic crash analysts want to trace the crashes that occur inside each zone so that they can relate the zonal characteristics with the level of traffic risk. Therefore, TAZs may be too small for use in analyzing traffic crashes at the macroscopic level. Moreover, the small size of the zone results in many zero crash frequencies for specific types of crashes such as fatal, bicycle, pedestrian, and so forth.

Secondly, TAZs usually are divided based on physical boundaries such as major arterials (criteria number 3, above). As a result, it is difficult to count crashes that occur on TAZ boundaries, which are major arterials in many cases. It is also hard to collect roadway data on boundaries. As depicted in Figure 4-1, state roads (indicated by red lines) divide the study regions into several TAZs. In this case, crashes on state roads do not belong to any of the adjacent TAZs because they are on the boundary between these TAZs. Fortunately, much of this boundary issue was solved after combining several of the zones with homogeneous crash patterns, because state roads and TAZ boundary crashes were then inside a single zone (Figure 4-2). However, for the remaining crashes that occur on the boundaries of TSAZs, other solutions are still required.

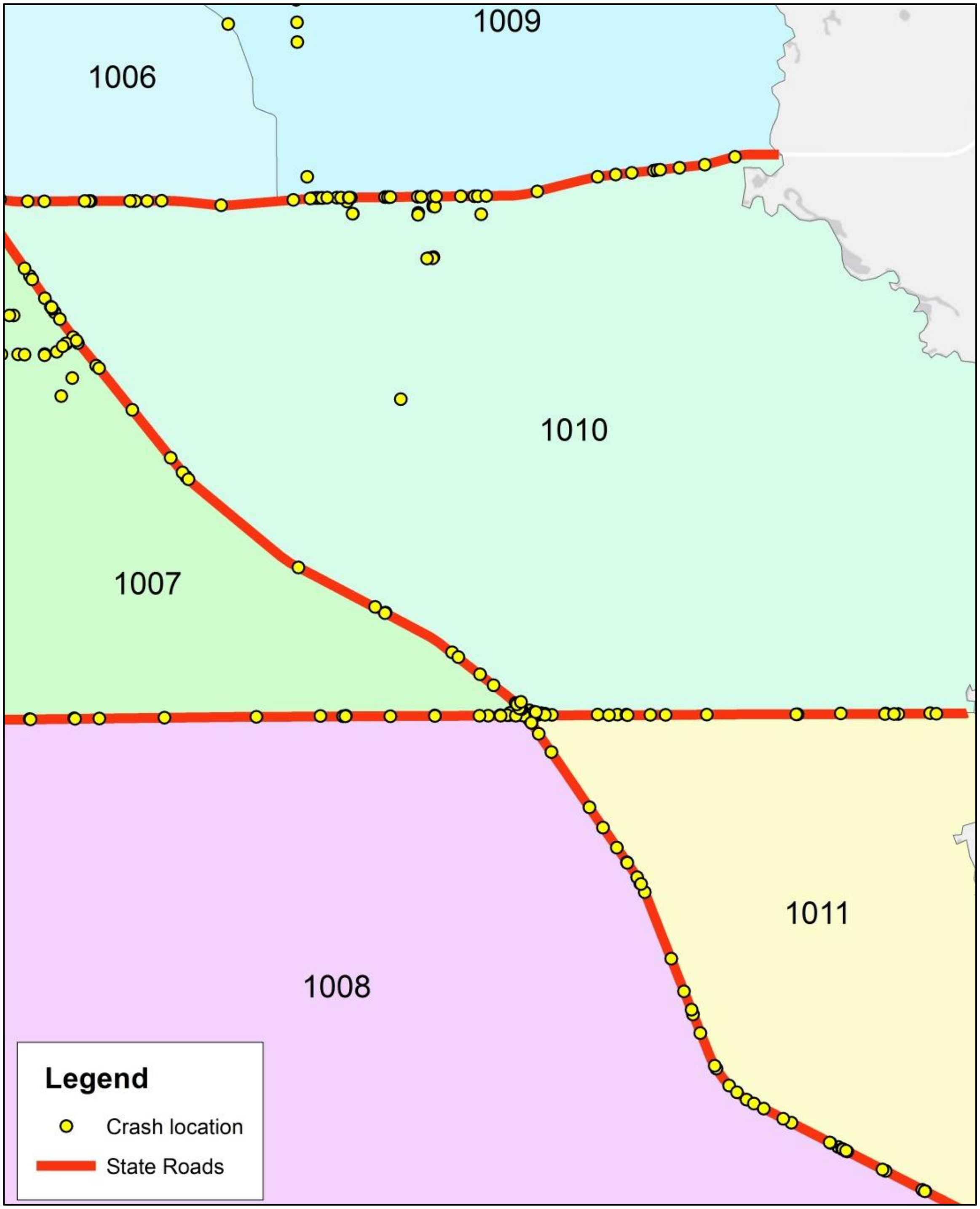


Figure 4-1 Boundary issues in TAZs

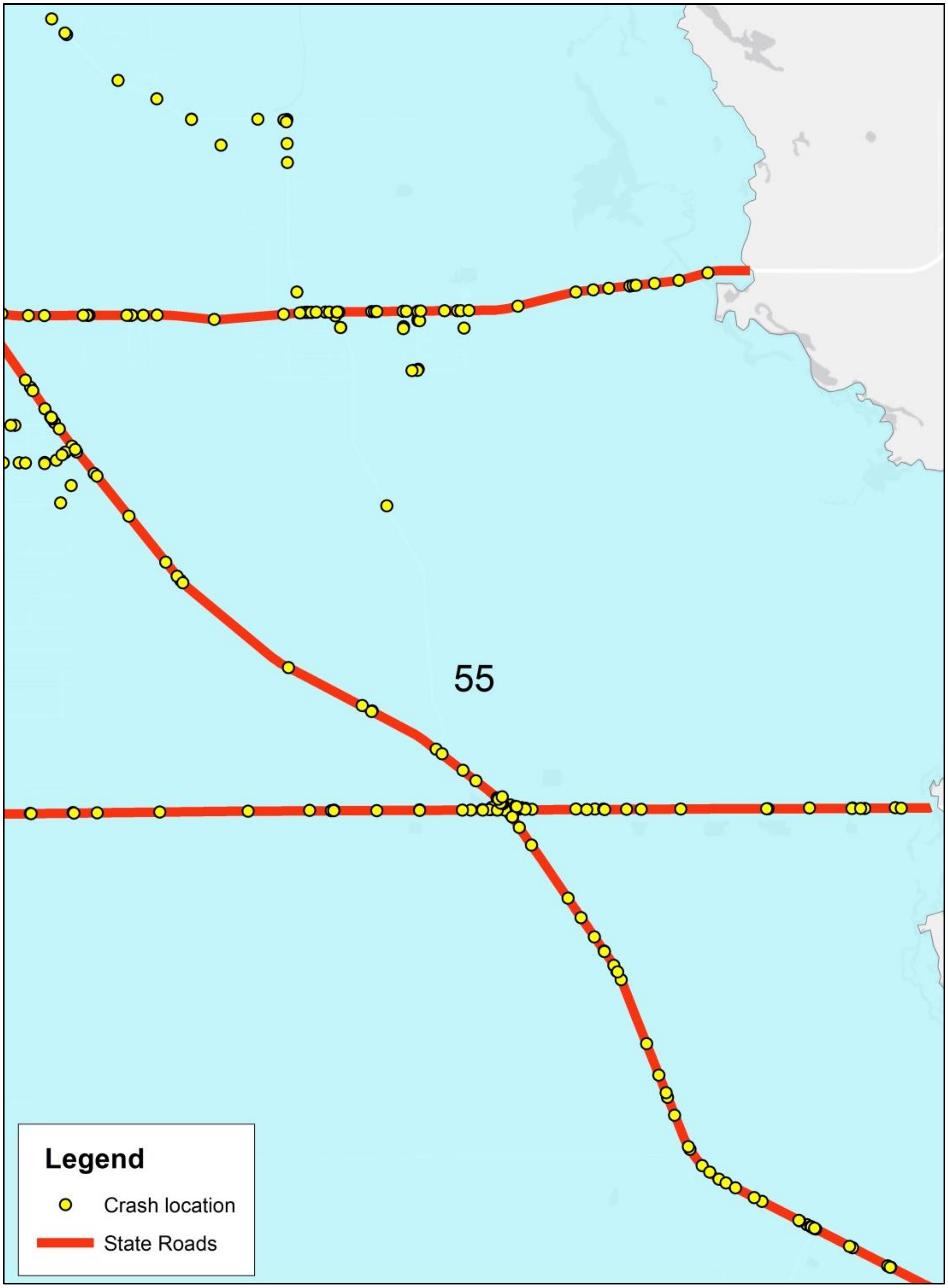


Figure 4-2 Boundary crashes/roadways after regionalization

In conclusion, TAZs may not be the best choice for traffic crash analysis units. Therefore, alternatives to TAZs for crash analysis purposes are required.

4.2. Regionalization

Regionalization refers to the process of combining a number of areal units into a smaller number of areas, while simultaneously optimizing an objective function (Guo & Wang, 2011). The research team used the regionalization program REDCAP, developed by Guo at the University of South Carolina.

The process of regionalization is shown in Figure 4-3. First, numbers of crashes were counted based on TAZs, and the total lengths of the roadways in the TAZs were calculated using GIS. Then, crash rates by miles were calculated for each TAZ, and the calculated crash rates were used as the objective function of the regionalization.

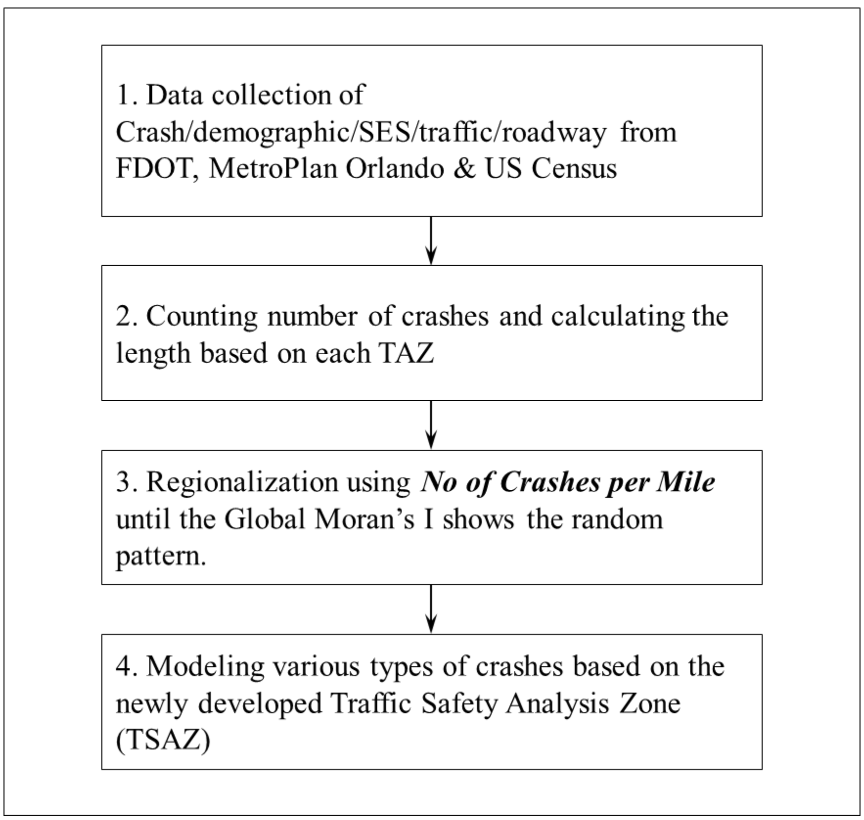


Figure 4-3 Regionalization and modeling process

One of the objectives of regionalization in this research project was to combine TAZs as much as possible until all adjacent zones with similar crash patterns were aggregated. In order to achieve this objective, the research team used spatial autocorrelation statistics and Moran's *I* index.

As discussed in the previous section, regionalization can alleviate the issues with and limitations of current zone systems by aggregating TAZs into sufficiently large and homogenous zones, based on major zonal characteristics.

In order to develop the zone system for macroscopic traffic safety analysis, the research team used the regionalization technique and determined the optimal zone scale using Brown-Forsythe tests.

4.2.1. Background of Regionalization

MAUP (the Modifiable Areal Unit Problem) is present when artificial boundaries are imposed on continuous geographical surfaces and the aggregation of geographic data causes a variation in the statistical results (Openshaw, 1984).

Assuming that areal units within a particular study are specified differently, it is possible that very different patterns and relationships could be shown up (O'Sullivan and Unwin, 2002).

MAUP was first investigated by Gehlke and Biehl (1934). These researchers found that the correlation coefficient increases as the unit area increases. According to Openshaw (1984), MAUP is composed of two effects: scale effects and zoning effect. Scale effects result from the different levels of spatial aggregation. For example, traffic crash patterns are differently described in lower aggregation spatial units such as TAZs and higher aggregation units such as counties or states. Zoning effects are described according to the different zoning configurations and at the same level of spatial aggregation.

4.2.2. Regionalization Process

As discussed in the previous section, the major limitations of the Traffic Analysis Zone (TAZ) system for macroscopic safety analyses are: 1) small zone sizes, and 2) boundary problems. Regionalization can alleviate some of these limitations for macroscopic safety analyses by aggregating TAZs into a sufficiently large and homogenous zonal system. Regionalization is a process of aggregating large numbers of units into smaller numbers of regions while optimizing an objective function (Guo & Wang, 2011). The objective function for this research was that the sum of the squared differences could be expressed as follows:

$$\text{Minimize SSD} = \sum_{r=1}^{k} \sum_{i=1}^{n_r} (x_i - \bar{x})^2 \quad (4\text{-}1)$$

where k is the number of regions, n_r is the number of data objects in region r, x_i is a variable value at observation i and $\bar{x}$ is the regional mean for the variable. The constraint of the objective function is that only adjacent regions can be aggregated.

The study area (Orange, Seminole, and Osceola counties) is comprised of 1,039 urban and 77 rural TAZs. Nevertheless, the gross urban area is only 1,001.394 mi^2, while the total rural area is 1,837.952 mi^2. Hence, TAZs in urban areas are much smaller than those in rural areas (Table 4-1). As such, the research team decided to focus on urban zones for regionalization; the rural zones were already relatively large. Also, the rural zones adjacent to urban zones could be aggregated during the regionalization process.

Table 4-1 Numbers and areas of urban/rural zones

Location	Zones	Total mi^2	Average mi^2	Stdev	Min	Max
Urban	1039	1001.394	0.964	1.377	0.021	12.403
Rural	77	1837.952	23.870	34.113	0.889	170.664

In the regionalization process, the optimal number of zones was determined by using a Brown-Forsythe test. The number of crashes per mile was used as the target variable to regionalize the TAZs. The research team also tried other measures, such as crashes per VMT and crashes per square mile. However, crash rate normalized by VMT was problematic because traffic volume data were only available for state roads and, thus, there were many zones without traffic data. Crashes per square mile were also tested, but the regionalization was most significantly affected by zone size rather than number of crashes. In contrast, crashes per mile did not suffer from such issues and the regionalization result also seemed more reasonable. Therefore, the research team decided that crashes per mile was the best target variable for the regionalization.

4.2.3. Brown-Forsythe Test

The F_{BF} test was used to evaluate whether the variance of variables of interest, such as crash rate, were equal when the scale of the zones changed. The underlying assumption of the test was that there was a greater variance in the crash rates among the smaller zones and a lower variance among the larger zones. A high variance value would mean that the crash risks were local, whereas a low variance would mean that they captured more global characteristics. The optimal zone scale would ensure that the variance of crash rate was somewhere in between. Root et al. (2011) and Root (2012) used this method in medical studies for disease analysis.

F_{BF} statistics are calculated according to the following formula:

$$F_{BF} = \frac{\frac{\sum_{i=1}^{t}(\bar{D}_i - \bar{D})^2}{(t-1)}}{\frac{\sum_{i=1}^{t}\sum_{j=1}^{n_1}(\bar{D}_{ij} - \bar{D}_1)^2}{(N-t)}} \tag{4-2}$$

where n_i is the number of samples in the *i*th zone system, N is the total number of samples for all zone systems, and t is the number of neighborhood groups. Given that y_{ij} is the crash rate of the jth sample from the *i*th zone system and $\bar{y}_i$ is the median of crash rate from the *i*th zone system, then $D_{ij}=\left|y_{ij} - \bar{y}_i\right|$ is the absolute deviation of the *j*th observation from the *i*th zone system median, $\bar{D}_i$ is the mean of D_{ij} for zone system *i*, and $\bar{D}$ is the mean of all D_{ij}. The test assumed that the variances of the different zones were equal under the null hypothesis. An F test with (t - 1, N - t) degrees of freedom was used to test for statistical significance.

There were two steps to the F_{BF} test. First, the variance between each zone system from N200 to N1000 and the largest zone system (N100) was compared for a total of nine separate calculations of F_{BF}, as shown in the F_{BF1} column of Table 4-2. Second, the variance between each neighborhood group from N900 to N100 and the smallest zone system (N1000) was compared (F_{BF2}). A significant value of F_{BF1} would imply that the zone system did not reflect the global pattern of crash data; in essence, each zone was so small that it only captured local crash patterns. On the contrary, a significant value of F_{BF2} would indicate that the zone data were not local; they were so large that local level crash patterns were undetectable. The zone systems between lower and upper limits would identify a spatial scale at which local-level variations would still be detectable but also would capture larger zonal-level crash characteristics.

4.2.4. Optimal zone scale for TSAZ

The F_{BF} test results for homogeneity of variance for total crash rate (total crashes per mile) under various zone scales are listed in Table 4-2. The F_{BF1} test statistics show that zone systems smaller than N700 (i.e., N800, N900 and N1000) have significantly different variances from that of N100. Thus, zone systems smaller than N700 are too small to capture global crash patterns. On the other hand, F_{BF2} test statistics indicate that zone systems larger than N500 (i.e., N400, N300, N200 and N100) are so large that they cannot capture local crash characteristics. Given the results, systems with 500-700 zones should be considered optimal for total crashes. As a result, the research team selected a scale of 500 zones as the new zone system for overall crashes because it minimized boundary crashes and zones that did not have rare types of crashes. The new zone system for overall crashes was named Traffic Safety Analysis Zone (TSAZ) for this study.

Table 4-2 Brown-Forsythe test for determining TSAZ scale

Zones	Crashes per miles		Brown-Forsythe test			
	Mean	Var	F_{BF1}	CV	F_{BF2}	CV
N1000	6.98	63.02	5.32	2.407	-	-
N900	6.59	55.09	4.38	2.511	0.54	6.635
N800	6.27	44.94	3.53	2.639	1.31	4.605
N700	**5.99**	**40.05**	**2.92**	**2.802**	**1.77**	**3.782**
N600	**5.65**	**35.18**	**2.02**	**3.017**	**2.6**	**3.319**
N500	**5.32**	**29.99**	**1.45**	**3.319**	**3.61**	**3.017**
N400	4.71	24.99	1.31	3.782	4.2	2.802
N300	3.91	18.76	0.84	4.605	4.76	2.639
N200	3.18	12.53	0.4	6.635	5.23	2.511
N100	2.67	9.06	-	-	5.32	2.407

4.2.5. Comparison of TAZ and TSAZ

As a result of the regionalization, the original 1,039 TAZs in urban areas and 77 TAZs in rural areas were aggregated into 428 and 72 TSAZs for overall crashes in urban and rural areas, respectively. The descriptive statistics for the TAZ and the new TSAZ systems is presented in Table 4-3.

Table 4-3 Areas of TAZ and TSAZ

Zone system		No of zones	Average area (mi^2)	Stdev	Min	Max
TAZ	**Total**	1116	2.544	10.710	0.021	170.664
	Urban	1039	0.964	1.377	0.021	12.403
	Rural	77	23.870	33.890	0.889	170.664
TSAZ	**Total**	500	5.678	15.493	1.051	170.664
	Urban	428	2.337	1.624	1.051	12.403
	Rural	72	25.541	34.502	1.265	170.664

As a result of the regionalization, zones with zero crashes were reduced from 1.5% to 0.8%. Also, zones without severe crashes were significantly reduced from 30.6% to 14.2% (see Table 4-4).

Table 4-4 Zones without crashes in TAZ and TSAZ

Zone system	Total crashes		Severe crashes	
	Freq	**%**	**Freq**	**%**
TAZ	17	1.5%	341	30.6%
TSAZ	4	0.8%	71	14.2%

Table 4-5 compares boundary crashes in TAZs and TSAZs. Before the regionalization, 76.5% of crashes occurred on the boundaries of TAZs; this was reduced to 61.9% after the regionalization.

Table 4-5 Boundary crashes in TAZ and TSAZ

Zone system	Total crashes		
	Boundary	**Total**	**Boundary %**
TAZ	68,451	89,527	76.5%
TSAZ	55,411	89,527	61.9%

Figure 4-4 shows TAZs in the overall study area before the regionalization. As seen in the downtown map (inside the black-lined square), the TAZs in highly urbanized areas are very small whereas the TSAZs in the downtown are much larger (Figure 4-5).

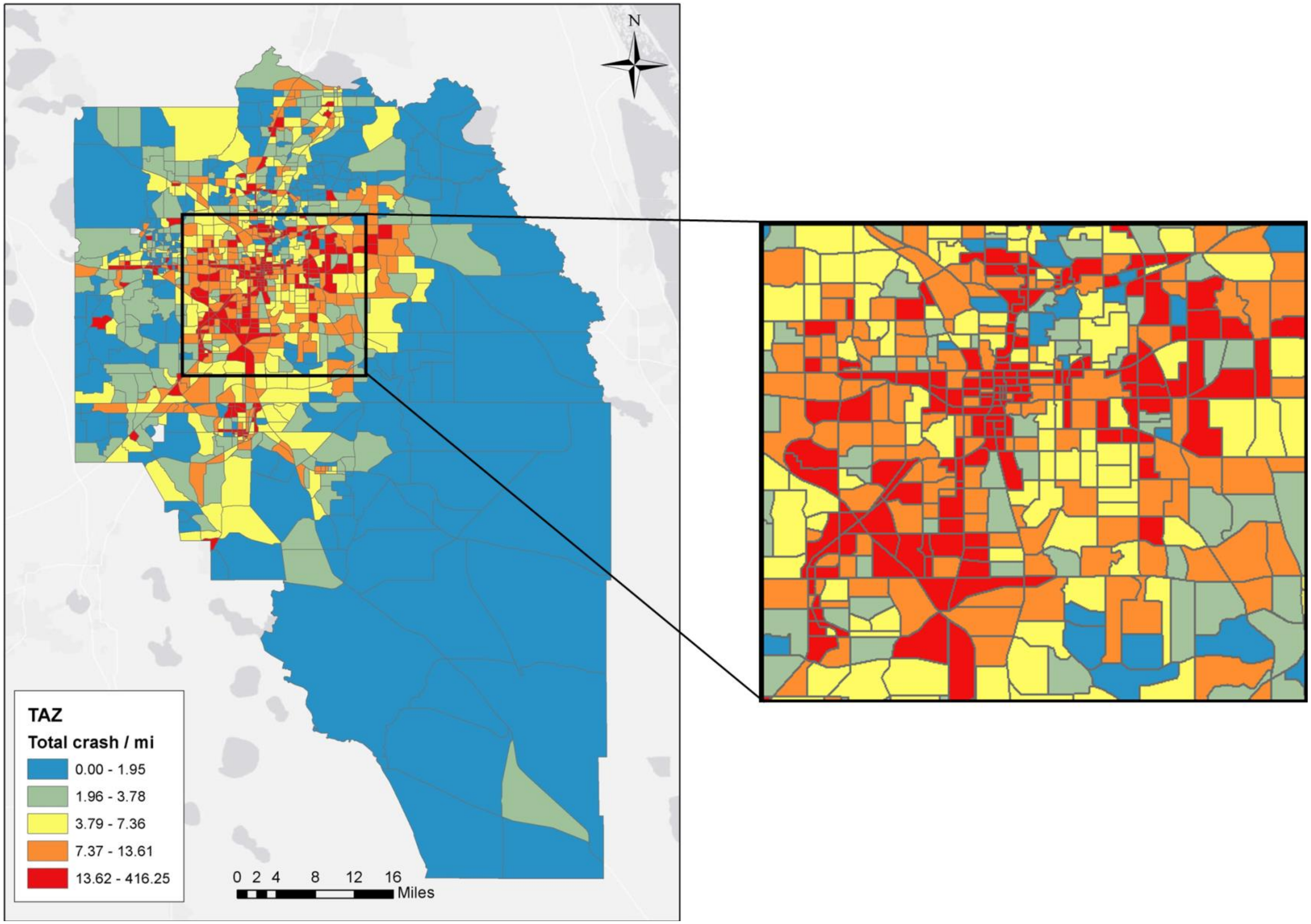


Figure 4-4 Total crashes per mile based on TAZs in the overall study area (left) and TAZs in downtown Orlando (right)

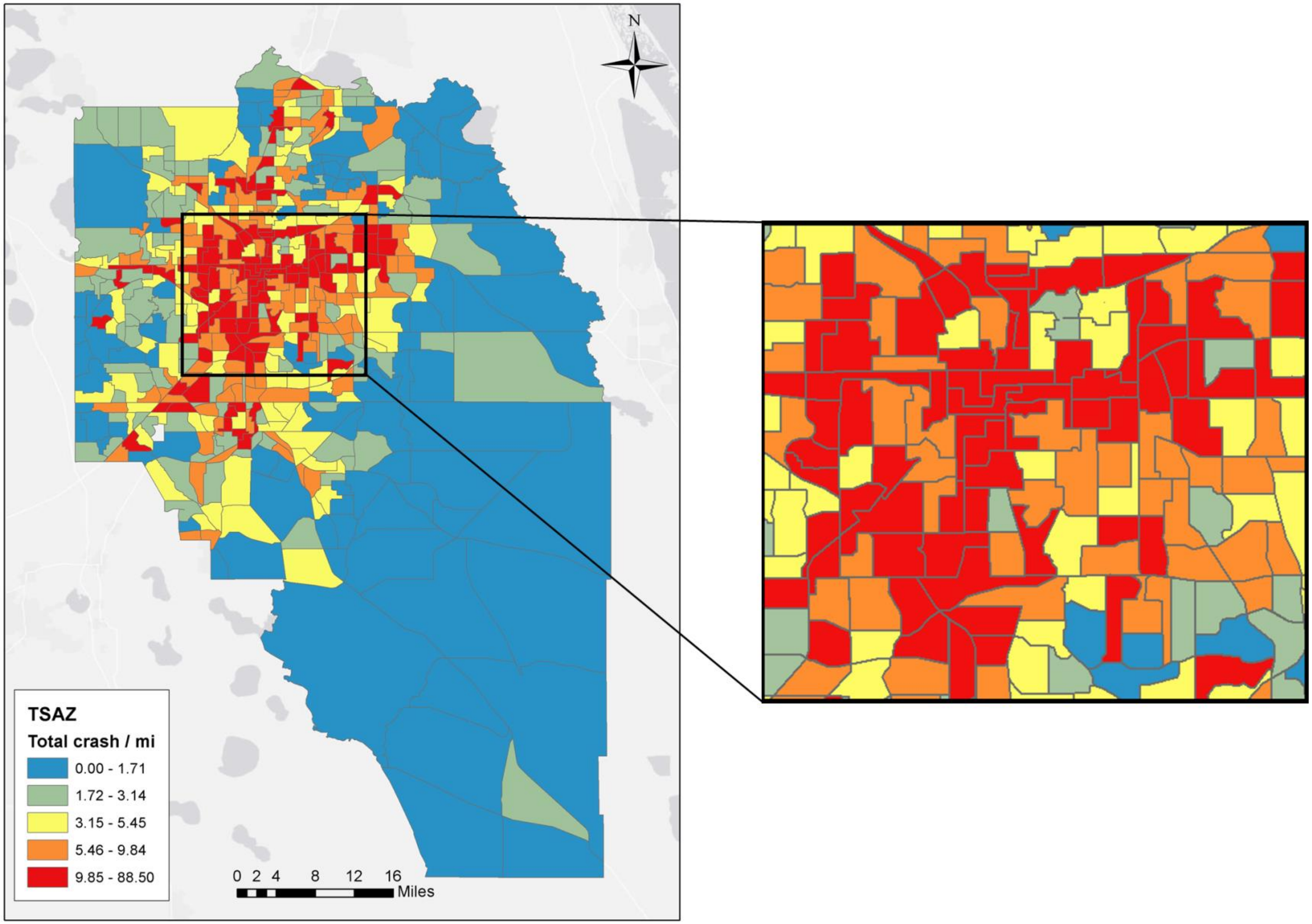


Figure 4-5 Total crashes per mile based on TSAZs in the overall study area (left) and TSAZs in downtown Orlando (right)

4.3. Boundary crashes

Approximately 10% of boundary crashes were reduced after the regionalization, but more than 60% of the crashes still occurred on the boundaries of TSAZs (Table 4-5). The problem is that most crashes occur on roadways that are boundary between TAZs, which makes it difficult to determine the zone characteristics related to those crashes. Luckily, the research team had extensive experience in dealing with this boundary issue. It was revealed that using previous research studies to estimate the safety models separately for boundaries and interior crashes was a more reasonable approach, and made it easier to develop models with a better goodness-of-fit (Siddiqui & Abdel-Aty, 2012).

For total crashes, the research team divided crashes into boundaries and interior crashes. Then, boundary crashes were further classified by crash roadway type. These categories included either 1) on-state highway system (on-system road); or 2) off-state highway system (off-system road). Moreover, on-system road crashes were separated into: 1) full access control (FAC) on-system road (i.e., interstate highway and expressway) crashes; or 2) other on-system road crashes. As a result, four separate safety models are developed in the next chapter, based on these classifications (Figure 4-6).

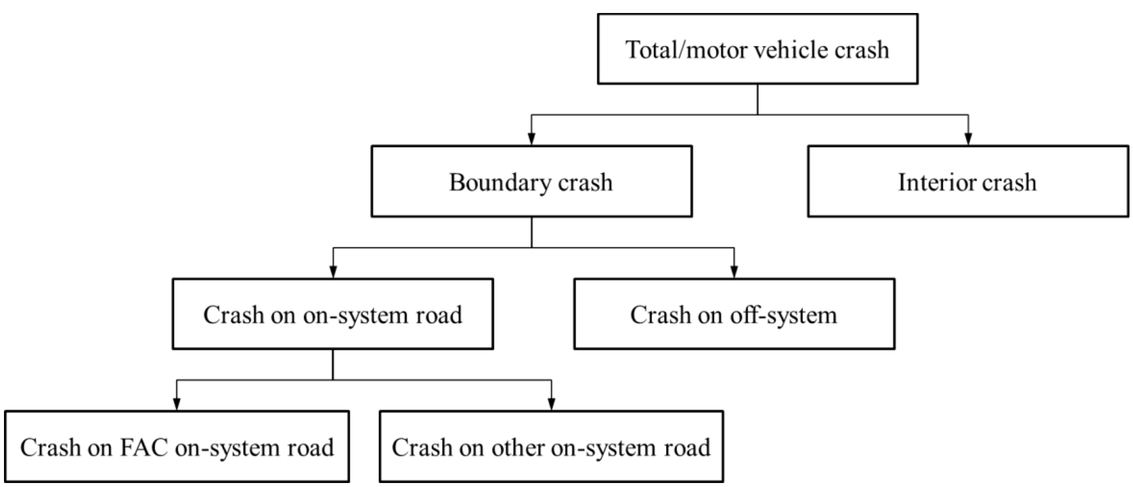


Figure 4-6 Nested structure of total/motor vehicle crash models

Table 4-6 summarizes the descriptive statistics of the total crashes category by classification; Figure 4-7 shows the proportions of crashes by classification. Interior crashes that occurred completely inside a TSAZ comprise 31.4% of the data, and boundary crashes that occurred within 100 ft of a TSAZ boundary make up 66.3%. Among boundary crashes, FAC on-system road crashes were the smallest category (7.9%), whereas other on-system road crashes comprised the largest category (38.1%). Off-system crashes made up 22.6% of the total crashes.

Table 4-6 Descriptive statistics of total crashes by classification

<table>
<tr><th colspan="3">Crash classification</th><th>Total</th><th>Mean</th><th>Stdev</th><th>Min</th><th>Max</th></tr>
<tr><td rowspan="3">Boundary</td><td rowspan="2">On-system</td><td>FAC</td><td>7054</td><td>14.11</td><td>29.85</td><td>0</td><td>195</td></tr>
<tr><td>Other</td><td>34129</td><td>56.23</td><td>112.27</td><td>0</td><td>963</td></tr>
<tr><td colspan="2">Off-system</td><td>20241</td><td>40.48</td><td>60.67</td><td>0</td><td>505</td></tr>
<tr><td colspan="3">Interior</td><td>28116</td><td>68.26</td><td>104.08</td><td>0</td><td>840</td></tr>
</table>

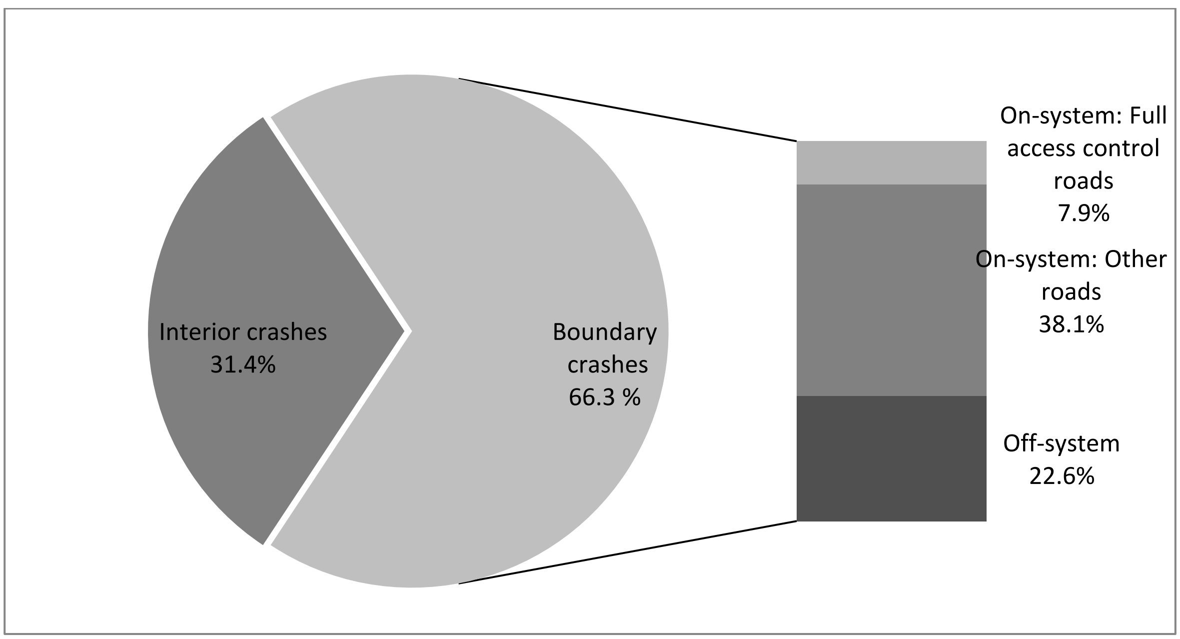


Figure 4-7 Proportions of total crashes by classification

4.4. Spatial autocorrelation

Spatial autocorrelation is the term for the tendency of data from locations but near one another in space to be more similar than data from locations remote to each other (O'Sullivan & Unwin, 2002). Most statistical models assume that the values of observations in each sample are independent or randomly distributed. A positive spatial autocorrelation between zones, however, may violate this assumption if the samples were collected from nearby areas (Lai et al., 2008). In section 4.4.1, the spatial autocorrelation effects in the residuals were explored using Moran's *I*. After the existence of spatial autocorrelations in the residuals was identified, the research team included the term in the SPF to account for the spatial autocorrelations, and then compared the corrected SPF using several spatial error terms from different conceptualizations to determine the method with the best performance (Section 4.4.2).

4.4.1. Detection of spatial autocorrelations in the residual

TSAZ (Traffic Safety Analysis Zone) is the zone system used for macro-level safety analysis. The research team developed TSAZ through the regionalization process, as was explained in the previous chapter. The Poisson log-normal model in the Bayesian framework that was adopted for estimating SPF in the current research is specified as follows:

$$log(\lambda_i) = X_i\beta + \theta_i \tag{4-3}$$

$$\theta_i = Normal\ (0, \tau_\theta) \tag{4-4}$$

where λ_i is the expected crash count in the *i*th TSAZ,

X_i is a row vector of explanatory variables showing characteristics of the *i*th TSAZ,

β are the coefficient estimates of model covariates,

θ_i is the unstructured over-dispersion or unobserved heterogeneity component in the *i*th TSAZ, and

τ_θ is the precision parameter, which is the inverse of the variance and given by a prior gamma distribution.

In order to identify the existence of spatial autocorrelations in the residuals (θ_i), Moran's I was used. Moran's I is one of the measures of spatial autocorrelation developed by Moran (1950), and is calculated by the following formula:

$$I = \frac{n}{\sum_{i=1}^{n}(y_i - \bar{y})^2} = \frac{\sum_{i=1}^{n}\sum_{j=1}^{n} w_{ij}(y_i - \bar{y})(y_j - \bar{y})}{\sum_{i=1}^{n}\sum_{j=1}^{n} w_{ij}} \quad (4\text{-}5)$$

where n is the number of areal units indexed by i and j,

y is the value of interest,

$\bar{y}$ is the mean of y, and

w_{ij} is an element of the matrix of spatial weights.

The research team examined the spatial weights calculated using three different conceptualizations:

1) Inverse distance: w_{ij} is the inverse distance between zones i and j;

2) Inverse distance squared: w_{ij} is the inverse distance squared between zones i and j; and

3) First order polygon contiguity: w_{ij} = 1 if zones i and j are adjacent based on the 1^{st} order contiguity, otherwise w_{ij} = 0.

A positive value of Moran's I index stands for a positive spatial autocorrelation, whereas a negative value indicates a negative autocorrelation. The value ranges from -1 to +1, where -1 means that regions are perfectly dispersed and +1 indicates that the regions are perfectly correlated. On the contrary, if the index is close to zero, this indicates a random pattern. Moran's I index can be converted to a Z-score for the statistical test, in which values greater than 1.96 or smaller than -1.96 show that there is a statistically significant spatial autocorrelation in the region.

Table 4-7 presents the Moran's I calculated from the residuals of SPF and the corresponding z-values' p-value for each conceptualization method. All of the values of Moran's I showed positive spatial autocorrelations, and they were all statistically significant. The first order rook polygon contiguity had the largest Moran's I, whereas the inverse distance had the smallest.

Since statistically significant spatial autocorrelations in residuals both in total and severe crash SPFs were detected, we must account for spatial autocorrelations in the estimations of SPF. One possible solution to account for spatial autocorrelation is to include a spatial random effect component in the model formulation (Eq. 4-6), which is discussed in the next section.

Table 4-7 Moran's *I* of residuals by spatial autocorrelation conceptualization

Model	Conceptualization	Moran's *I*	*z*	*p*
Total crashes	Inverse distance	0.075	11.376	<0.001
	Inverse distance squared	0.126	8.045	<0.001
	First order rook polygon contiguity	0.178	6.681	<0.001
Severe crashes	Inverse distance	0.033	5.089	<0.001
	Inverse distance squared	0.060	3.909	<0.001
	First order rook polygon contiguity	0.134	5.032	<0.001

4.4.2. Comparison of SPFs with different spatial effect conceptualizations

As mentioned in the previous section, the spatial error term (φ_i) was included in the SPF to account for the spatial autocorrelation, using the following equation:

$$log(\lambda_i) = X_i\beta + \theta_i + \varphi_i \quad (4\text{-}6)$$

The spatial distribution was implemented by specifying an intrinsic Gaussian Conditional Autoregressive (CAR) prior with a $Normal\ (0, \tau_{\varphi})$ distribution. The mean of φ_i is defined by:

$$\bar{\varphi}_i = {\sum_{i \neq j} \varphi_j \times w_{ij}}\Big/{\sum_{i \neq j} w_{ij}} \tag{4-7}$$

where values for w_{ij} are defined in Table 4-8 by the different spatial autocorrelation conceptualizations.

Table 4-8 Definition of w_{ij} by different spatial autocorrelation conceptualizations

Conceptualization	w_{ij}
No spatial error term	$\varphi_i = 0$
First order rook polygon contiguity	$w_{ij} = 1$, if zone *i* and *j* are adjacent; $w_{ij} = 0$, otherwise
Inverse distance	$w_{ij} = 1/d_{ij}$
Inverse distance squared	$w_{ij} = 1/d_{ij}^2$

Each model's Deviance Information Criterion (DIC) was computed for comparison. The following equation was used to calculate DIC (Spiegelhalter et al., 2002):

$$DIC = 2 \times \bar{D} - \widehat{D} \tag{4-8}$$

where $\bar{D}$ is the posterior mean of deviance D,

$\widehat{D}$ = $2 \times (p(y|\theta))$, and

$\bar{\theta}$ is the posterior mean of θ.

Models with a smaller DIC are preferred to models with a larger DIC. (Spiegelhalter et al., 2003). Table 4-9 summarizes the DIC from the total and severe crash models with different spatial autocorrelation conceptualizations. It was found that only the spatial error term conceptualized by the first order rook polygon slightly improved the performances of both the total and severe crash models. Thus, the final models will have a spatial error component based on first order rook polygon contiguity, as will be described in the following chapters.

Table 4-9 Comparison of DICs by different spatial autocorrelation conceptualizations

Conceptualization	DIC	
	Total crash model	Severe crash model
No spatial error term	4122.53	2270.81
First order rook polygon contiguity	4122.32	2247.51
Inverse distance	4121.07	2270.79
Inverse distance squared	4121.66	2272.65

5. DEVELOPMENT OF SPFS FOR MACRO-LEVEL ANALYSES

This section describes the overall procedure used to perform the macro-level analyses. A series of SPFs for both total and FI crashes were developed. In order to solve the boundary crash missing problem, a complex structure was used to estimate separately both boundary and interior crashes on various types of roads (see Section 5.1). In Section 5.2, the research team suggested a method of accounting for boundary crashes.

5.1. Nested modeling structure

This study adopted a nested structure which allows different contributing factors for different crash types (such as boundary or interior crashes, and crashes located on different roadway types). The research team expected to achieve more accurate and predictable results from this nested structure than what could be obtained from a single model.

The nested structure includes six sub-models which are named based on their location (i.e., near the zone boundary or in the interior) and roadway type (i.e., freeway-and-expressway, other state roads, and non-state roads). The nested structure is presented in Figure 5-1. Both total crashes and fatal-and-injury crashes were modeled using this same nested structure. In addition, a Bayesian Poisson Lognormal Spatial Error Model (BPLSEM) was adopted for the SPF analysis in this nested structure. This model has a disturbance term for handling the over-dispersion problem, and its spatial error term controls for the spatial autocorrelation of crash data. See Appendix A for more details on this model formulation. The research team assumed that factors contributing to crashes on full access control (FACR) such as interstate highways and expressways are different than those that contribute to crashes on other state roads. Thus, the research team constructed a nested structure with six models, as shown in Figure 5-1.

The six types of crashes in each model vary based on their location (boundary or interior) and roadways (FACR, other state roads, or non-state roads). They are FSB (FACR State road Boundary crashes), FSI (FACR State road Interior crashes), OSB (Other State road Boundary crashes), OSI (Other State road Interior crashes), NSB (Non-state road Boundary crashes) and NSI (Non-state road Interior crashes). Figure 5-2 contains examples of these six crash types.

Meanwhile, some zones have zero probability of having specific types of crashes. For instance, zone #1 in Figure 5-2 has no FACR or other state roads. The expected numbers of FSB, FSI, OSB and OSI in zone #1 should all be zero, regardless of zonal characteristics. It is meaningless to include zones without FACR or state roads in the estimations for FSB, FSI, OSB and OSI models. Therefore, the research team excluded zones without specific types of roads when estimating models for crashes occurring on those types of roads.

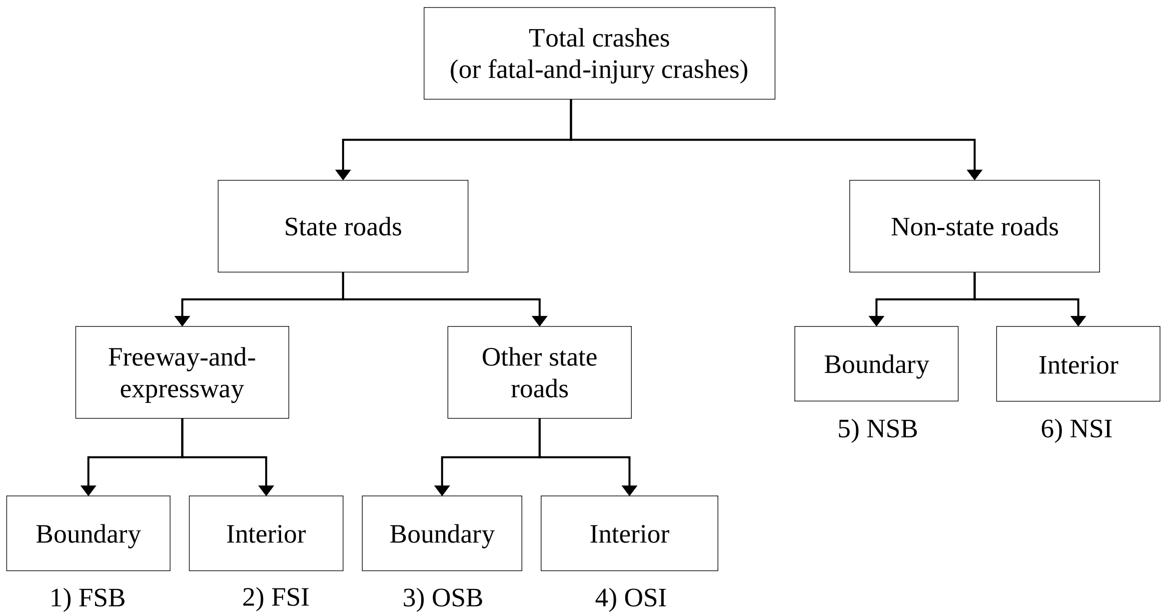


Figure 5-1 Nested structure for macroscopic crash modeling (with six sub-models)

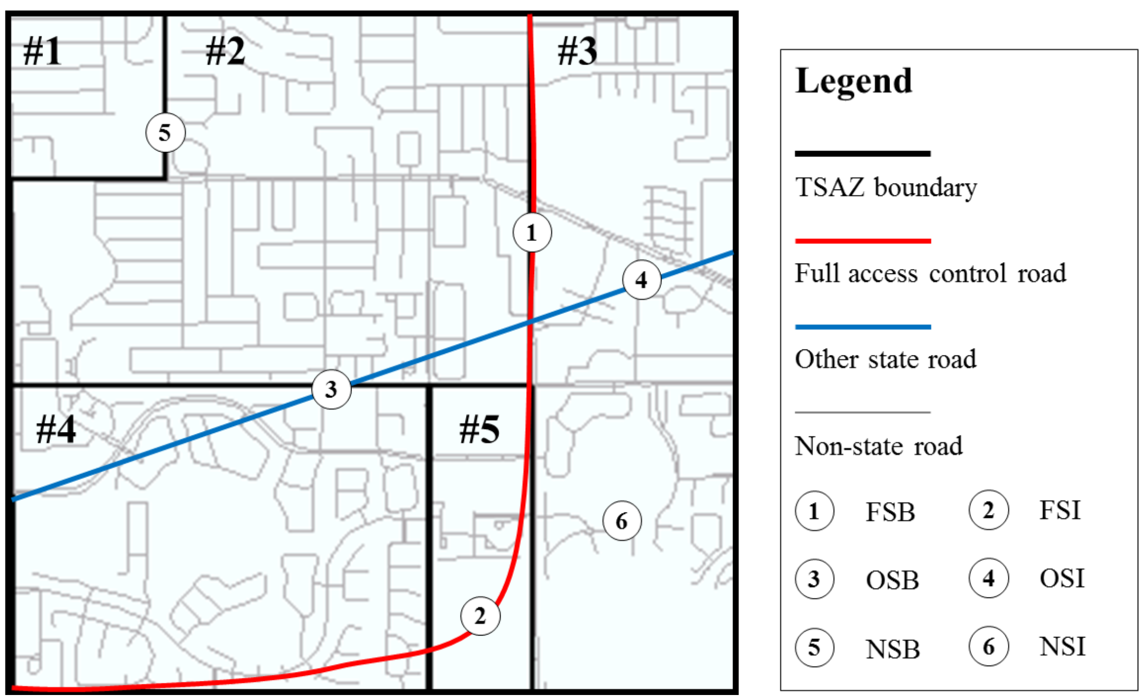


Figure 5-2 Examples of crashes by locations used in the nested structure

5.2. Accounting for Boundary Crashes

It was assumed that interior crashes were influenced only by the characteristics of the zone in which the crashes were located. Thus, the models for the interior crashes were developed using individual zonal characteristics. In contrast, crashes occurring near or on zone boundaries, known as boundary crashes, were hypothesized to be influenced not only by the zonal characteristics of where the crashes occurred, but also by the characteristics of the adjacent zones. Therefore, the models for boundary crashes were estimated using 'transformed' variables possessing information for both the crash zone and any adjacent zones.

Let any TSAZ i share its boundary with adjacent zones j = 1, 2, … k, as shown in Figure 5-3. An original variable x will be transformed to x_{ABC} using the following Eq. 5-1.

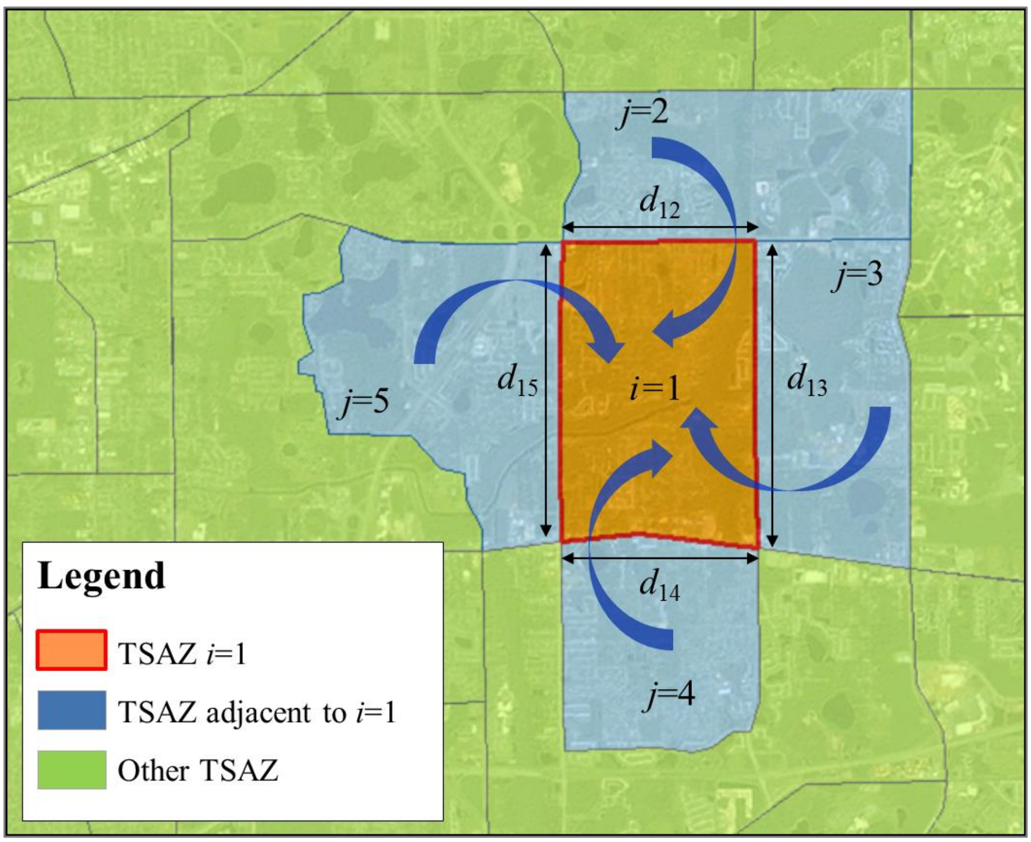


Figure 5-3 Illustration of adjacent zones for crash zone *i*

$$x_{ABCi} = wx_i + (1 - \mathrm{w})\left[{(d_{i1}x_1 + d_{i2}x_2 + \cdots + \ d_{ik}x_k)}\Big/{(d_{i1} + d_{i2} + \cdots + \ d_{ik})} \right] \quad (5\text{-}1)$$

where,

x_{ABCi} = transformed variable x for i^{th} zone,

x_i = variable x for i^{th} zone,

x_j = variable x for the zones adjacent to i^{th} zone (j=1, 2, … k),

d_{ij} = length of the shared boundary between zones i and j,

$d_{i1} + d_{i2} + \ldots \ldots \ldots + d_{ik}$ = perimeter of zone i, and

w= weight.

The first term in Eq. 5-1 represents the characteristics of the i^{th} zone; the second term denotes the weight-averaged features of the adjacent zones. The weighted average is based on the length of the shared boundary between the adjacent zones.

The weight component (w) reflects the actual influence on boundary crashes from zone i and its adjacent zones. For instance, if the boundary crash was uniformly affected by the features of the crash zone (i^{th} zone) and the adjacent zones, the weight would be 0.5. Meanwhile, if the boundary crash was solely influenced by the crash zone, the weight would be 1.0.

With the intention of finding the appropriate weights, the research team developed a wide array of Negative Binomial (NB) models with nine weights (from 1.0 to 0.1 by 0.1) for each sub-model. The models with the lowest Akaike Information Crietrion (AIC) values were selected. The AIC was developed by Akaike (1974), and is calculated as follows:

$$AIC = 2k - 2\ln(L) \quad (5\text{-}2)$$

where,

k is the number of parameters in the model, and

L is the maximum likelihood for the model.

The AIC is an index which compares the relative qualities among various models; it is widely used for model selection. The AIC copes with the tradeoff between goodness-of-fit and the complexity of the model. Among the candidate models, the model with the minimum AIC was chosen. Tables 5-1 and 5-2 present the AICs of candidate models for total and fatal-and-injury crashes, respectively.

The optimal weights for the FSB, OSB, and NSB are 0.7, 0.9, and 0.7, respectively, for the total crash model, and 0.7, 0.8, and 0.8, respectively, for the fatal-and-injury model. These optimal weights were used to estimate the final SPF. However, future studies should use 0.8 as an optimal weight for boundary crashes and 1.0 for interior crashes because no significant difference was observed between the models with weights equal to 0.7, 0.8, or 0.9.

Table 5-1 AIC table of candidate total crash models

Weights	1) FSB	2) FSI	3) OSB	4) OSI	5) NSB	6) NSI
1.0	1835.41	**1083.89**	3353.83	**1514.34**	4114.86	**4134.00**
0.9	1833.67	1084.92	**3351.47**	1516.03	4102.09	4137.71
0.8	1833.30	1087.79	3352.41	1520.01	4099.41	4150.41
0.7	**1833.23**	1093.23	3355.75	1526.46	**4099.20**	4170.31
0.6	1833.75	1102.28	3362.91	1535.29	4101.13	4199.80
0.5	1835.28	1115.46	3374.43	1545.71	4105.22	4236.48
0.4	1838.38	1132.19	3389.40	1556.25	4111.41	4273.68
0.3	1843.35	1150.42	3405.48	1565.51	4119.36	4306.42
0.2	1849.75	1167.25	3420.14	1572.85	4128.59	4334.83
0.1	1856.54	1180.64	3432.07	1578.46	4138.74	4361.82

Table 5-2 AIC table of candidate fatal-and-injury crash models

Weights	1) FSB	2) FSI	3) OSB	4) OSI	5) NSB	6) NSI
1.0	1373.05	**827.93**	2466.38	**1020.80**	2951.42	**2861.46**
0.9	1373.39	833.08	2463.98	1024.11	2947.82	2867.83
0.8	1372.71	837.80	**2462.60**	1027.26	**2947.23**	2882.72
0.7	**1372.33**	844.98	2462.81	1032.17	2947.50	2905.83
0.6	1372.59	855.25	2466.17	1038.96	2948.88	2937.01
0.5	1374.10	868.68	2473.68	1047.12	2951.56	2971.64
0.4	1377.64	884.32	2485.10	1055.58	2955.51	3002.94
0.3	1383.48	900.12	2498.58	1063.18	2960.42	3027.24
0.2	1390.60	913.69	2511.61	1069.26	2965.87	3045.49
0.1	1397.36	923.51	2522.59	1073.77	2971.52	3060.60

In summary, boundary crash types were greatly affected by the crash zone (70%-90%) and rarely influenced by adjacent zones (10%-30%). Moreover, it was proven that interior crashes were affected only by the characteristics of the zone wherein the crash occurred because the optimal weight for all interior crash models is 1. However, these optimal weights are more applicable to crash modeling with TSAZ-based data. The optimal weights may be different if a model is developed based on different scale zones (i.e., census tracts, traffic analysis districts, block groups, traffic analysis zones, etc.).

6. MODELING RESULTS AND IDENTIFICATION OF HOT ZONES FOR MACRO-LEVEL ANALYSES

A new model, the Nested Bayesian Poisson Lognormal Spatial Error model, was proposed to account for boundary crash effects (referred to here as NBPLSEM) based on the Traffic Safety Analysis Zones (TSAZs). Section 6.1 explains certain details about the modeling results and Section 6.2 shows the process of hot zone identification.

6.1. Modeling Results

The research team developed the crash prediction model by using NBPLSEM, with the optimal weights suggested in the previous chapter. The modeling results for total and fatal-and-injury crashes are presented in Tables 6-1 and 6-2, respectively. As seen in these tables, each sub-model had different sample sizes. This is because some zones had zero probability of having specific types of crashes. For instance, one zone had no state roads. In this case, the expected numbers for the FSB, FSI, OSB, and OSI in this zone should all be zero, regardless of zonal characteristics. It is not reasonable to include this type of zone in crash prediction models. Therefore, the research team excluded zones without specific types of roads when estimating models for crashes occurring on those types of roads.

In addition, the total crash model and the fatal-and-injury model both show that each sub-model has its own variable set. All models seem to have reasonable and explainable coefficients. For example, with the FSB model the exposure variable (vehicle-miles-traveled) was positively associated with the crash count. Also, the coefficient of the proportion of young people (15-24 years old), the natural logarithm of employment and school enrollment, and the proportion of roads with a 20 mph or lower max speed were all positive. The first two variables are self-explanatory. The third variable, the proportion of roads with a low speed limit, refers to the proportion of residential roads. It was determined that if a zone has a higher proportion of residential roads, there will be more local drivers entering the freeway and expressway from residential roads. Thus, the zone may have more crashes on freeways and expressways. Also, the results show that the spatial effects were significant. They reveal the existence of spatial autocorrelations among the explanatory variables with associated total and fatal-and-injury crashes.

Interestingly, Ψ (the apportionment of the variability in the error component due to spatial autocorrelation) is always larger in the boundary models than in the interior models of the same roadway type. For example, the Ψ values of the FSB and FSI in the total crash model are 0.505 and 0.228, respectively. From this, the research team concluded that the unobserved heterogeneity in the error component from the spatial effects in the FSB was 50.5%, whereas it was only 22.8% in the FSI. In other words, boundary crashes are more significantly influenced by spatial autocorrelation because they are close to other adjacent zones.

Table 6-1 Nested Poisson Lognormal Spatial Error Model Accounting for Boundary Crashes: total crashes

Variable	1) FSB N=213		2) FSI N=155		3) OSB N=325		4) OSI N=174		5) NSB N=439		6) NSI N=465	
	mean	s.d.	mean	s.d.	mean	s.d.	mean	s.d.	mean	s.d.	mean	s.d.
Intercept	-0.482#	0.791	-1.690**	0.537	-0.602#	0.559	-2.538**	0.747	0.806**	0.485	-0.804**	0.243
Ln of population density			0.096**	0.043	0.122**	0.050					0.100**	0.020
Proportion African Americans											0.911**	0.243
Proportion of Hispanics					1.292**	0.552						
Proportion of young people (15-24 yr)	3.850**	1.645										
Proportion of old people (65yr+)												
Proportion households without car					3.098**	1.439						
Ln of hotel, motel and timeshare rooms	-0.057	0.040			0.092**	0.034	0.514**	0.091			0.049**	0.014
Ln of employment and school enrollment	0.240**	0.108	0.245**	0.066	0.227**	0.079			0.259**	0.064	0.326**	0.033
Proportion of roads with 20 mph or lower max speed	5.943**	1.673										
Proportion of roads with 55 mph or higher max speed												
Roads with poor pavement conditions											0.210**	0.047
Ln of VMT at FSB	0.097**	0.020										
Ln of VMT at FSI			0.173**	0.018								
Ln of VMT at OSB					0.137**	0.016						
Ln of VMT at OSI							0.145**	0.025				
Ln of VMT at NSB									0.051**	0.017		
Ln of VMT at NSI											0.094**	0.010
s.d. of θ_i	0.850	0.108	0.886	0.081	1.213	0.070	1.335	0.091	0.639	0.117	0.527	0.060
s.d. of φ_i	0.890	0.222	0.271	0.128	0.402	0.236	0.180	0.146	1.567	0.222	0.873	0.150
Ψ	0.505	0.087	0.228	0.090	0.236	0.109	0.113	0.076	0.707	0.065	0.620	0.065
DIC	1342.46		846.222		2349.98		1085.5		3007.29		3214.22	

not significant at 20%, ** significant at 5%, * significant at 10%, and all other variables are significant at 20%

Table 6-2 Nested Poisson Lognormal Spatial Error Model Accounting for Boundary Crashes: fatal-and-injury crashes

Variable	1) FSB		2) FSI		3) OSB		4) OSI		5) NSB		6) NSI	
	N=213		N=155		N=325		N=174		N=439		N=465	
	mean	**s.d.**	**mean**	**s.d.**	**mean**	**s.d.**	**mean**	**s.d.**	**mean**	**s.d.**	**mean**	**s.d.**
Intercept	-3.153**	1.043	-4.240**	0.800	-2.525**	1.046	-10.260**	2.632	-0.137#	0.451	-1.796**	0.286
Ln of population density					0.133**	0.067					0.063**	0.022
Proportion African Americans									1.279**	0.425	0.925**	0.261
Proportion of Hispanics					2.146**	0.788						
Proportion of young people (15-24 yr)	4.813**	2.129							-1.558*	0.908		
Proportion of old people (65yr+)												
Proportion households without car					5.968**	2.002						
Ln of hotel, motel and timeshare rooms	-0.091*	0.050	-0.057	0.037	0.113**	0.039	0.132**	0.059			0.042**	0.015
Ln of employment and school enrollment	0.244**	0.121	0.506**	0.104			0.234	0.153	0.226**	0.061	0.282**	0.037
Proportion of roads with 20 mph or lower max speed	3.968*	2.098							-2.690**	1.264		
Proportion of roads with 55 mph or higher max speed					-2.881**	1.165	-5.308**	2.031				
Roads with poor pavement conditions									0.109	0.078	0.193**	0.049
Ln of VMT at FSB	0.166**	0.026										
Ln of VMT at FSI			0.203**	0.022								
Ln of VMT at OSB					0.179**	0.022						
Ln of VMT at OSI							0.271**	0.038				
Ln of VMT at NSB									0.041**	0.015		
Ln of VMT at NSI											0.112**	0.011
s.d. of θ_i	1.173	0.105	0.903	0.108	1.364	0.135	1.558	0.145	0.401	0.183	0.480	0.087
s.d. of φ_i	0.614	0.188	0.303	0.169	0.705	0.451	0.236	0.208	1.789	0.225	0.897	0.180
Ψ	0.339	0.079	0.243	0.112	0.315	0.156	0.124	0.089	0.816	0.087	0.646	0.084
DIC	1037.090		672.768		1788.360		716.974		2337.760		2450.26	

not significant at 20%, ** significant at 5%, * significant at 10%, and all other variables are significant at 20%

6.2. Identification of Hot Zones

The PSI (Potential for Safety Improvement), or excess crash frequency, is a measure of how many crashes can be reduced for a particular site and is suggested in the HSM. The PSI for each zone is the difference between the expected crash count and the predicted crash count (see Figure 6-1).

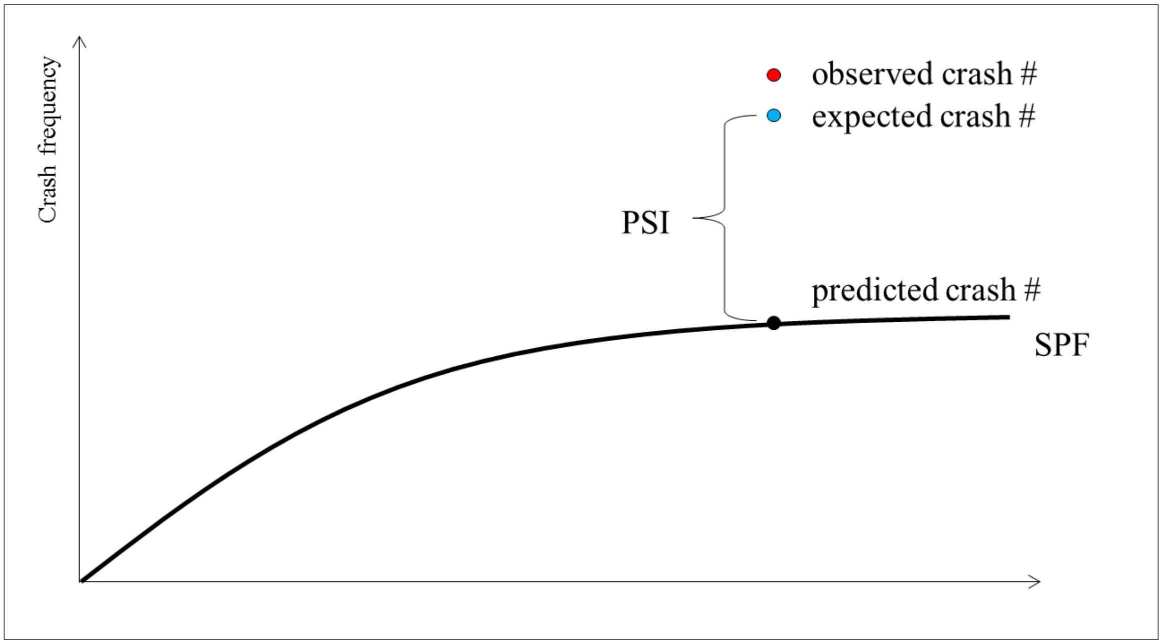


Figure 6-1 Schematic showing definition of PSI

The predicted crash counts were estimated using six sub-models in the nested structure, as was shown previously, and the PSIs were calculated by following the equation proposed by Aguero-Valverde and Jovanis (2009). As suggested in Eq. 6-1, the PSI is the gap between the expected crashes and the predicted crash count. Eqs. 6-2 and 6-3 were derived from Eq. 6-1, for convenience of calculation.

$$PSI = N_{exp} - N_{prd} \quad (6\text{-}1)$$

$$= \exp(\beta_0 + \beta X_i + \theta_i + \varphi_i) - \exp(\beta_0 + \beta X_i) \quad (6\text{-}2)$$

$$= \exp(\beta_0 + \beta X_i)(\exp(\theta_i + \varphi_i) - 1) \quad (6\text{-}3)$$

The PSIs were calculated and the TSAZs were ranked separately for urban and rural areas. Tables 6-3 and 6-4 present the TSAZs with the top 10% PSIs in rural and urban areas, correspondingly. The full tables with all TSAZ figures are included in Appendix C.

Figures 6-2 and 6-3 show the spatial distributions of hot zones with top 10% PSIs in Orange, Seminole, and Osceola counties for both total crashes and fatal-and-injury crashes, respectively. For total crashes, Figure 6-2 indicates that most of the hot zones in rural areas were close to the fringes of urban areas, or they contained major arterials (i.e., SR50) or full access control roads (i.e., SR528, SR91, etc.). With regards to urban areas, hot zones mostly were located from downtown Orlando to eastern Orlando along SR50 (E. Colonial Drive). For fatal-and-injury crashes, hot zones had patterns very close to those of the total crash hot zones (any slight differences are indicated in Figure 6-3). These hot zones were closer to high speed roads (freeways or expressways) in both urban and rural areas. For example, some fatal-and-injury crash hot zones contained I-4 (urban) and SR91 (rural).

Table 6-3 Ranking TSAZs with the top 10% PSIs (rural areas)

Rank	Rank percentile	Total crashes		Fatal-and-injury crashes	
		TSAZ ID	PSI	TSAZ ID	PSI
1	1.4%	367	215.548	367	79.229
2	2.8%	337	152.669	337	70.096
3	4.2%	347	145.548	347	51.083
4	5.6%	406	130.475	281	48.928
5	6.9%	281	118.346	406	45.225
6	8.3%	49	103.374	464	31.660
7	9.7%	361	70.069	49	31.319
8	11.1%	247	61.156	394	26.761

Table 6-4 Ranking TSAZs with the top 10% PSIs (urban areas)

Rank	Rank percentile	Total crash		Fatal-and-injury crash	
		TSAZ ID	PSI	TSAZ ID	PSI
1	0.2%	56	1127.880	202	334.644
2	0.5%	15	971.440	8	272.738
3	0.7%	202	791.730	196	255.596
4	0.9%	8	651.180	2	234.250
5	1.2%	9	648.000	56	233.255
6	1.4%	196	625.459	15	204.740
7	1.6%	192	620.349	89	188.698
8	1.9%	89	595.207	207	179.557
9	2.1%	69	549.530	5	178.469
10	2.3%	104	510.150	43	175.275
11	2.6%	382	498.175	69	171.608
12	2.8%	130	492.320	3	156.202
13	3.0%	224	470.914	12	151.874
14	3.3%	0	433.720	192	150.154
15	3.5%	92	429.485	67	138.363
16	3.7%	67	428.796	62	137.494
17	4.0%	62	413.550	130	134.979
18	4.2%	6	411.870	18	133.330
19	4.4%	43	402.370	104	131.018
20	4.7%	66	385.870	9	129.910
21	4.9%	146	384.350	0	125.090
22	5.1%	178	381.803	66	124.134
23	5.4%	18	376.160	58	118.026
24	5.6%	42	361.726	101	111.759
25	5.8%	212	354.540	65	111.366
26	6.1%	195	350.338	93	110.636
27	6.3%	29	345.127	212	110.133
28	6.5%	35	330.897	16	109.178
29	6.8%	180	327.315	180	104.254
30	7.0%	19	318.380	86	96.362
31	7.2%	207	318.163	57	96.124
32	7.5%	60	302.947	6	94.408
33	7.7%	14	293.278	224	94.395
34	7.9%	2	287.020	38	91.203
35	8.2%	28	280.342	105	87.838
36	8.4%	57	268.780	382	87.799
37	8.6%	3	257.644	195	86.882
38	8.9%	250	253.911	250	82.205
39	9.1%	98	252.000	19	80.906
40	9.3%	5	250.610	233	79.481
41	9.6%	38	248.007	345	79.408
42	9.8%	22	247.428	42	78.342
43	10.0%	93	235.027	333	76.341

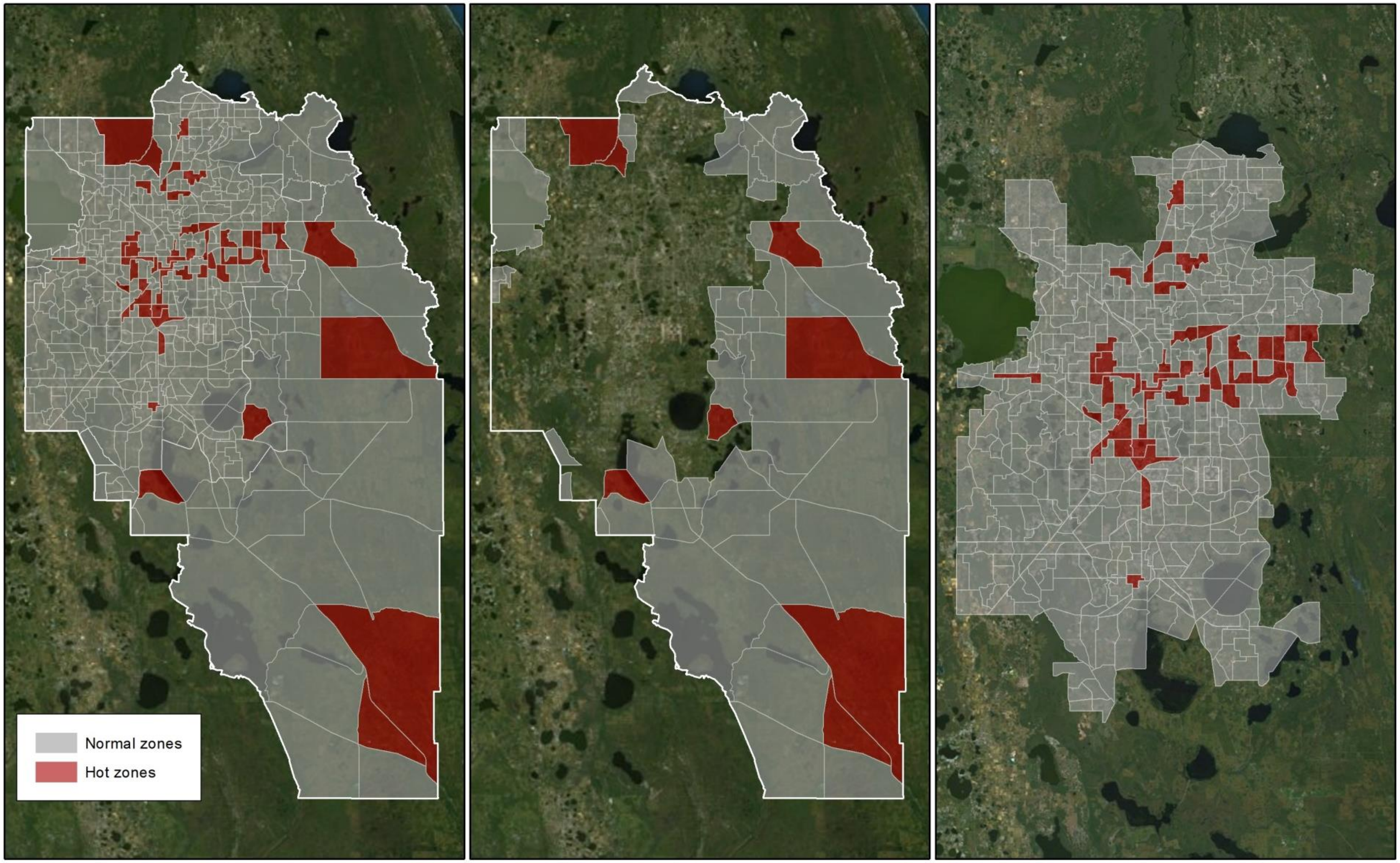


Figure 6-2 Top 10% hot zones for total crashes in both urban and rural areas, rural areas, and urban areas (left to right, respectively)

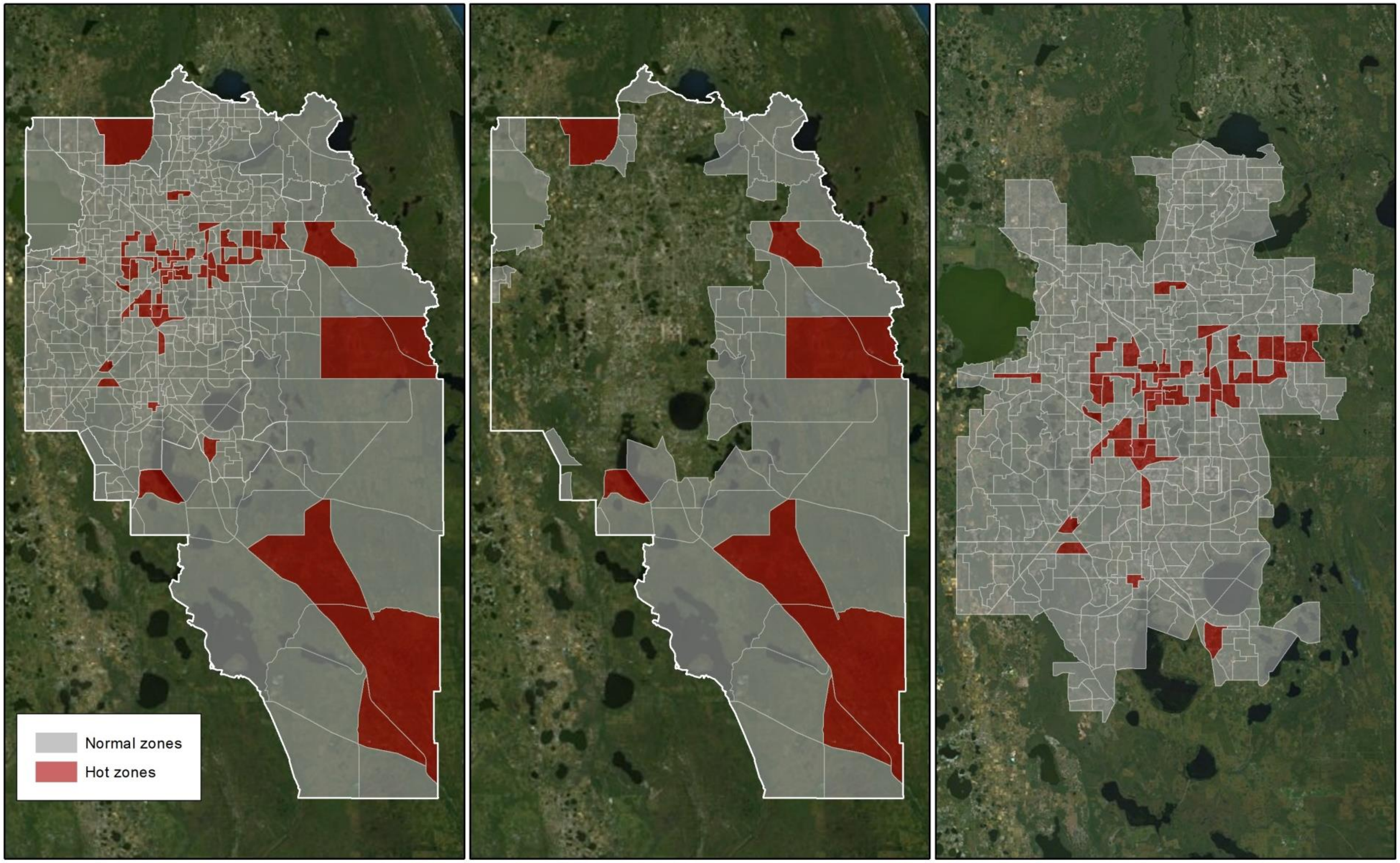


Figure 6-3 Top 10% hot zones for fatal-and-injury crashes in both urban and rural areas, rural areas, and urban areas (left to right, respectively)

7. DEVELOPMENT OF SPFs FOR MICRO-LEVEL ANALYSES

This section describes the overall procedure used to perform the micro-level analyses. A description of the data preparation procedure was provided in Section 7.1. A series of SPFs for segments and intersection were developed separately (see Sections 7.2 and 7.3).

7.1. Data preparation

In order to retain consistency between the macro- and micro-level models, the research team used the same crash data set as described above. GIS techniques were used to collect crash and road characteristics data. This method can be considered both accurate and efficient because it automatically generates crash mapping results, and no extra effort or time was required to examine the merged results. The inventory file of the intersections was based on the Roadway Characteristics Inventory (RCI) dataset. For each signalized intersection, a buffer area with a 250 foot radius was created, and any crashes located within this buffer were counted for that intersection. For the segment analysis, we excluded all intersections and intersection-related areas from the current road network. Table 7-1 shows the statistics of the different roadway types. The AADT data was also collected from the RCI dataset for the years 2008 and 2009. Usually, there were only two roadway IDs per intersection. In some cases, three or more different roadway IDs intersected in the center of a single intersection. In this case, we had three or more roadway IDs instead of two. When that occurred, the average AADT was used for the two roads on the same line/direction. For example, if there were three different roadway IDs intersecting at the same point, two of the three roads would represent the major roads and the third would represent the minor road. The average AADT would then be calculated for the two roadway IDs representing the major road with different AADTs.

Road inventory data and traffic data for each type of road in Florida were collected from the Roadway Characteristics Inventory (RCI) system.

The research team developed SPFs for the major function classes of roads in Osceola, Seminole, and Orange counties. For the segments, there were rural 2 lane undivided, rural 2 or 4 lane divided, urban 2 lane divided, urban 4 lane divided, urban 2 or 4 lane undivided, six or more lane interrupted (partial access control) roads, one way roads, and 3 lane with Two Way Left Turn Lane (TWLTL). For the intersections, there were 4 Legs Intersections and 3 Legs Intersections.

Overall, these road classes are consistent with the HSM road classifications, while the SPFs in the FIU report (Preparing Florida for Development of SafetyAnalyst for All Roads) were fitted for the road classes specified in SafetyAnalyst (SA). Moreover, this study includes certain new roadway types not presented in the HSM, such as 6 or more lane interrupted roads. It should be noted that two categories with similar characteristics were combined, and one dummy variable was added to the model when the sample size was too small. For example, rural 2 lane divided roads were combined with rural 4 lane divided roads. Also, urban 2 lane undivided roads were combined with urban 4 lane undivided roads.

Table 7-1 Collected data for different types of roads

Road segments				
Classification		min	max	total
Rural 2-Lane Undivided Roads (59 segments)	Segment (mi)	0.2154	21.545	188.75588
	Total Crashes	1	19	389
	FI Crashes	0	13	208
Rural 2-Lane Divided Roads (25 segments)	Segment (mi)	0.1009	0.4978	5.221026
	Total Crashes	0	5	26
	FI Crashes	0	5	12
Rural 4-Lane Divided Road (47 segments)	Segment (mi)	0.1035	26.947	149.00231
	Total Crashes	0	175	1125
	FI Crashes	0	70	503
Urban 2-Lane Undivided (1066 segments)	Segment (mi)	0.1001	3.5831	482.79878
	Total Crashes	0	126	7278
	FI Crashes	0	52	3131
Urban 2-Lane Divided (456 segments)	Segment (mi)	0.1005	2.2882	134.44159
	Total Crashes	0	97	3051
	FI Crashes	0	40	1251
Urban 4-Lane Undivided Roads (69 segments)	Segment (mi)	0.1002	1.3211	21.469023
	Total Crashes	0	92	761
	FI Crashes	0	23	344
Urban 4-Lane Divided Roads (778 segments)	Segment (mi)	0.1001	8.6769	654.40990
	Total Crashes	0	348	21762
	FI Crashes	0	144	8958
Six Lanes Interrupted Roads (296 segments)	Segment (mi)	0.1000	3.8197	169.05941
	Total Crashes	0	317	10054
	FI Crashes	0	116	4098
Six Lanes Uninterrupted Roads (102 segments)	Segment (mi)	0.1070	2.6523	58.139970
	Total Crashes	0	279	3699
	FI Crashes	0	100	1507
Eight Lanes Interrupted Roads (22 segments)	Segment (mi)	0.1180	0.7483	7.264682
	Total Crashes	0	112	591
	FI Crashes	0	33	200
Eight Lanes Uninterrupted Roads (76 segments)	Segment (mi)	0.1089	5.7876	43.303028
	Total Crashes	0	228	3385
	FI Crashes	0	90	1524
Urban 3-Lane with TWLTL Arterial Streets (223 segments)	Segment (mi)	0.1000	3.1038	68.799244
	Total Crashes	0	68	1253
	FI Crashes	0	30	531
Urban 5-Lane with TWLTL Arterial Streets (101 segments)	Segment (mi)	0.1005	1.4906	40.530855
	Total Crashes	0	101	1973
	FI Crashes	0	42	924

Urban 7-Lane with TWLTL Arterial Streets (5 segments)	Segment (mi)	0.1158	0.3920	1.092304
	Total Crashes	3	23	61
	FI Crashes	2	10	24
One-Way Roads (53 segments)	Segment (mi)	0.1263	1.1063	17.832387
	Total Crashes	0	58	591
	FI Crashes	0	28	268
Intersection				
Classification		min	max	total
4 Legs Intersection (140 segments)	Total Crashes	2	250	8139
	FI Crashes	0	53	1926
3 Legs Intersection (110 segments)	Total Crashes	1	21	3903
	FI Crashes	1	36	887

(*: FI indicates fatal and injury crashes)

7.2. Model structure for SPFs at the Microscopic Level

Because there was no existing SPF or reference group data available, a Full Bayesian model was used to estimate the PSI values for the different roadway types in the study area. A Poisson log-normal model with random effect was employed for this project. This regression model can be derived from the Poisson model by assuming that the same roadway types share one error term over various years. The detailed framework of the regression model is described in greater detail in Appendix A.

8. MODELING RESULTS AND IDENTIFICATION OF HOT ZONES FOR MICRO-LEVEL ANALYSES

8.1. SPFs for Road Segments

As mentioned in the previous section, the SPFs were fitted using Negative Binomial model by a Full Bayesian approach. Table 8-1 (below) contains the summary statistics. For different road types, the coefficients of intercept, the natural log of the major and minor AADT are different. All coefficients of intercept are negative, while the coefficients of the natural log of the major and minor AADT are positive. When comparing the SPFs for the total crashes and FI crashes, we found that their coefficients were close to each other except in Rural 2-Lane Undivided Roads. For Rural 2-Lane Undivided Roads, the segment length affected the expected total crashes more than the AADT, while it was contrary for the FI crashes.

Table 8-1 SPFs for total crashes

Classification	intercept		Ln (AADT)	
	mean	s.d.	mean	s.d.
Rural 2-Lane Undivided Roads	-0.441	2.868	0.593	0.312
Rural 2/4 Lane Divided Roads	-6.702	1.785	0.700	0.192
Urban 2-Lane Divided Roads	-6.318	0.743	0.757	0.080
Urban 4-Lane Divided Roads	-4.460	0.524	0.654	0.052
Urban 2/4 Lanes Undivided Roads	-2.319	0.381	0.335	0.043
6 or more Lanes Interrupted Roads	-6.638	1.253	0.826	0.119
6 or more Lanes Uninterrupted Roads	-2.007	1.049	0.417	0.093
3 or more Lanes with TWLTL Roads	-3.347	0.844	0.430	0.093
One Way Roads	-4.387	2.868	0.593	0.312

Table 8-2 SPFs for FI crashes

Classification	intercept		Ln (AADT)	
	mean	s.d.	mean	s.d.
Rural 2-Lane Undivided Roads	-1.116	1.265	0.226	0.150
Rural 2/4 Lane Divided Roads	-6.685	1.888	0.599	0.209
Urban 2-Lane Divided Roads	-6.591	0.797	0.707	0.085
Urban 4-Lane Divided Roads	-4.941	0.516	0.618	0.051
Urban 2/4 Lanes Undivided Roads	-3.147	0.432	0.332	0.048
6 or more Lanes Interrupted Roads	-7.327	1.107	0.820	0.105
6 or more Lanes Uninterrupted Roads	-4.061	1.056	0.524	0.093
3 or more Lanes with TWLTL Roads	-3.705	1.048	0.381	0.115
One Way Roads	-7.410	2.621	0.842	0.282

Based on the above models, PSIs can be calculated as the difference between the predicted crash frequency and the expected crash frequency for each road site. Tables 8-3 to 8-20 present the top ten hotspots for total crashes and FI crashes for each roadway type. In Table 8-2, for rural 2-lane roads, the PSI of the first hotspot (ID:77060000) is 10.09, which means that there is the potential to decrease crash frequency by 10.09 at this segment in two years. The rankings of sites based on the PSIs of total and FI crashes were identified separately. In general, the PSI values of the hotspots of the total crashes are higher than the hotspots of the fatal-and-injury crashes for the same roadway type. For example, the maximum PSI for Urban 4-Lanes divided total crashes is 246.3 while the PSI for fatal-and-injury crashes for the same road segment is 71.67.

Table 8-3 Screening output for rural 2–lane undivided segments (total crashes)

Rankk	RD_ID	Length	AADT	COUNT	Pred	Exp	PSI
1	77060000	5.139170	6500	19	6.93	17.02	10.09
2	77000214	3.253740	2300	17	4.744	14.53	9.786
3	75160500	5.807980	1200	16	4.982	13.77	8.788
4	92070000	7.048990	5800	18	7.794	16.36	8.566
5	77060000	4.572620	6500	17	6.635	15.19	8.555
6	77040000	3.908980	6100	16	6.226	14.22	7.994
7	75100000	0.778769	12000	15	5.815	13.2	7.385
8	75100000	1.268880	12000	15	6.026	13.29	7.264
9	75000381	4.545580	3000	14	5.542	12.31	6.768
10	77000214	3.838810	2300	13	4.961	11.24	6.279

(pred: predicted crash frequency. exp: expected crash frequency adjusted by FB method)

Table 8-4 Screening output for rural 2–lane undivided segments (FI crashes)

Rank	RD_ID	Length	AADT	COUNT	Pred	Exp	PSI
1	77000214	3.253740	2300	13	2.386	10.49	8.104
2	77060000	5.139170	6500	12	3.45	10.17	6.72
3	92070000	7.048990	5800	12	3.85	10.31	6.46
4	77040000	3.908980	6100	11	3.111	9.183	6.072
5	77060000	4.572620	6500	10	3.309	8.446	5.137
6	92060000	4.127490	2200	9	2.505	7.251	4.746
7	75100000	0.778769	12000	9	2.96	7.446	4.486
8	75100000	1.268880	12000	9	3.06	7.499	4.439
9	77000214	3.838810	2300	8	2.48	6.442	3.962
10	75000379	1.717150	1200	7	1.877	5.298	3.421

Table 8-5 Screening output for rural 2/4–lane divided segments (total crashes)

Rank	RD_ID	Length	AADT	COUNT	Pred	Exp	PSI
1	75002000	10.833000	38759	138	65.87	137.3	71.43
2	92470000	13.177600	25000	116	71.87	115.6	43.73
3	75005000	4.736780	38759	52	23.65	51.29	27.64
4	75002000	0.463171	34000	38	10.69	36.87	26.18
5	75140000	4.624240	14100	36	11.21	35.01	23.8
6	92030000	6.318820	5900	32	8.22	30.84	22.62
7	92470000	2.612120	26000	35	12.45	34.14	21.69
8	75060000	1.562230	9900	28	5.29	26.58	21.29
9	92030000	5.142790	8114	26	8.34	25.07	16.73
10	75140000	7.595530	15500	36	19.7	35.49	15.79

Table 8-6 Screening output for rural 2/4–lane divided segments (FI crashes)

Rank	RD_ID	Length	AADT	COUNT	Pred	Exp	PSI
1	75002000	10.833000	38759	70	27.39	69.02	41.63
2	92470000	13.177600	25000	53	30.86	52.45	21.59
3	92030000	6.318820	5900	19	4.23	17.63	13.4
4	75140000	4.624240	14100	19	5.32	17.84	12.52
5	75005000	4.736780	38759	21	10.13	20.33	10.2
6	75060000	1.562230	9900	14	2.643	12.53	9.887
7	75140000	7.595530	15500	19	9.124	18.36	9.236
8	75002000	0.463171	34000	14	4.733	13.02	8.287
9	75060000	3.940610	9900	12	3.86	11.02	7.16
10	75140000	2.666190	14100	12	3.881	11.02	7.139

Table 8-7 Screening output for urban 2–lane divided segments (total crashes)

Rank	RD_ID	Length	AADT	COUNT	Pred	Exp	PSI
1	92000103	0.381118	7800	92	3.22	89.02	85.8
2	75000145	0.238122	13500	84	3.76	81.25	77.49
3	75000321	0.330305	4800	61	2.04	57.96	55.92
4	75510501	0.217018	28500	63	6.38	60.96	54.58
5	75000199	0.556225	10000	51	5.38	48.99	43.61
6	75000304	0.382950	23500	48	7.48	46.36	38.88
7	75620000	1.515710	21500	97	57.75	96.54	38.79
8	92010000	0.384625	24000	44	7.62	42.45	34.83
9	75080000	0.116237	18400	40	3.8	37.93	34.13
10	75506503	0.170450	8100	36	2.25	33.55	31.3

Table 8-8 Screening output for urban 2–lane divided segments (FI crashes)

Rank	RD_ID	Length	AADT	COUNT	Pred	Exp	PSI
1	92000103	0.381118	7800	40	1.42	36.87	35.45
2	75000145	0.238122	13500	33	1.67	30.21	28.54
3	75510501	0.217018	28500	19	2.74	17.25	14.51
4	75000321	0.330305	4800	18	0.92	15.31	14.39
5	75080000	0.116237	18400	18	1.71	15.86	14.15
6	75620000	1.515710	21500	32	18.02	31.44	13.42
7	75000192	0.104974	15500	17	1.49	14.81	13.32
8	75000199	0.556225	10000	17	2.22	15.18	12.96
9	75000304	0.382950	23500	16	3.11	14.54	11.43
10	75620000	1.631590	21500	33	21.79	32.57	10.78

Table 8-9 Screening output for urban 4–lane divided segments (total crashes)

Rank	RD_ID	Length	AADT	COUNT	Pred	Exp	PSI
1	75000156	2.557010	16100	305	38.2	302.5	264.3
2	75060000	3.433220	37500	348	122.2	346.8	224.6
3	75000139	1.625300	35500	222	33.7	219.9	186.2
4	75060000	2.065010	37500	234	47.3	232.1	184.8
5	75200000	1.578070	44000	210	37.6	208	170.4
6	75230500	2.378660	32000	224	53	222.3	169.3
7	75000139	1.846740	35500	204	39.3	202.1	162.8
8	75000156	1.210130	16100	169	15.1	166.2	151.1
9	75000178	1.353070	14300	169	15.4	166.3	150.9
10	75620000	2.159570	31000	186	44.6	184.4	139.8

Table 8-10 Screening output for urban 4–lane divided segments (FI crashes)

Rank	RD_ID	Length	AADT	COUNT	Pred	Exp	PSI
1	75060000	3.433220	37500	144	53.22	142.7	89.48
2	75040000	0.987496	31000	95	8.53	91.7	83.17
3	75000156	2.557010	16100	91	17.05	88.72	71.67
4	75060000	2.065010	37500	90	20.38	87.97	67.59
5	75010000	1.666590	26500	75	12.44	72.55	60.11
6	75200000	1.578070	44000	75	16.01	72.92	56.91
7	75690500	3.028690	38000	95	40.37	93.85	53.48
8	75010000	1.596470	26500	67	11.84	64.67	52.83
9	75230500	2.378660	32000	74	23.01	72.44	49.43
10	75000156	1.210130	16100	59	6.65	56.01	49.36

Table 8-11 Screening output for urban 2/4–lane undivided segments (total crashes)

Rank	RD_ID	Length	AADT	COUNT	Pred	Exp	PSI
1	75000059	0.622665	12000	87	5.36	84.38	79.02
2	75000104	1.918320	2500	97	18.7	95.4	76.7
3	75000104	1.924420	2500	96	18.87	94.39	75.52
4	75520000	1.163680	12900	88	11.52	86.03	74.51
5	75000155	0.850311	21500	92	16.12	90.3	74.18
6	75000087	2.519410	6900	126	60.12	125.3	65.18
7	92000103	1.973910	7800	91	29.57	89.89	60.32
8	75040000	0.941903	22000	71	10.18	69.1	58.92
9	92000103	0.370056	7800	65	3.29	62.16	58.87
10	75000015	0.118409	7400	59	2.29	55.83	53.54

Table 8-12 Screening output for urban 2/4–lane undivided segments (FI crashes)

Rank	RD_ID	Length	AADT	COUNT	Pred	Exp	PSI
1	75000059	0.622665	12000	47	2.27	44.13	41.86
2	92000103	1.973910	7800	52	12.43	50.64	38.21
3	75520000	1.163680	12900	45	4.86	42.89	38.03
4	75000104	1.918320	2500	42	7.89	40.41	32.52
5	75000104	1.924420	2500	41	7.96	39.48	31.52
6	92000103	0.370056	7800	34	1.4	31	29.6
7	75000003	0.656769	2100	28	1.33	25.16	23.83
8	75000087	2.519410	6900	49	25.2	48.33	23.13
9	75025501	1.589490	7700	28	7.32	26.75	19.43
10	75080101	1.016250	14000	25	4.09	23.33	19.24

Table 8-13 Screening output for 6 or more lane interrupted roads (total crashes)

Rank	RD_ID	Length	AADT	COUNT	Pred	Exp	PSI
1	75010000	1.425420	61000	271	80.3	269.6	189.3
2	75037000	1.897670	59000	317	148.2	316.1	167.9
3	75002000	0.154878	49000	145	12.1	142.2	130.1
4	75010000	2.071730	57000	305	182.5	304.5	122
5	75002000	0.775936	37000	140	22	137.9	115.9
6	75518000	0.842210	53000	135	32.5	133.4	100.9
7	75010000	0.404445	53500	116	18.14	113.9	95.76
8	75050000	0.993066	28500	119	23.86	117.2	93.34
9	75000139	0.587495	35500	105	16.55	102.9	86.35
10	75000034	1.439420	39500	144	57.08	143	85.92

Table 8-14 Screening output for 6 or more lane interrupted roads (FI crashes)

Rank	RD_ID	Length	AADT	COUNT	Pred	Exp	PSI
1	75037000	1.897670	59000	116	57.22	115	57.78
2	75050000	0.993066	28500	63	10.18	60.42	50.24
3	75010000	1.425420	61000	80	32.59	78.74	46.15
4	75270000	1.486750	51500	70	30.61	68.87	38.26
5	75000139	0.587495	35500	48	7.36	45.45	38.09
6	75050000	1.012960	28500	50	10.44	47.85	37.41
7	75010000	2.071730	57000	106	69.18	105.4	36.22
8	75002000	0.154878	49000	41	5.6	38.28	32.68
9	75037000	0.696127	59000	45	12.79	43.28	30.49
10	75060000	1.548350	49500	62	31.99	61.1	29.11

Table 8-15 Screening output for 6 or more lane uninterrupted roads (total crashes)

Rank	RD_ID	Length	AADT	COUNT	Pred	Exp	PSI
1	75280000	0.243255	180500	279	25.1	275.4	250.3
2	75008000	2.652340	46000	243	81.5	241.4	159.9
3	75280000	0.919223	140000	121	45.15	119.5	74.35
4	77160000	1.838250	135500	146	70.58	144.9	74.32
5	75280000	0.633673	131500	110	35.83	108.4	72.57
6	75280000	0.338652	124000	96	22.99	93.94	70.95
7	75280000	1.283900	7500	86	14.41	83.42	69.01
8	75008000	0.638744	99500	95	26.01	93.13	67.12
9	77160000	0.872188	135500	102	35.1	100.5	65.4
10	75008000	0.579055	99500	91	24.91	89.11	64.2

Table 8-16 Screening output for 6 or more lane uninterrupted roads (FI crashes)

Rank	RD_ID	Length	AADT	COUNT	Pred	Exp	PSI
1	75008000	2.652340	46000	100	29.98	98.21	68.23
2	75280000	0.243255	180500	81	11.6	78.23	66.63
3	75280000	0.633673	131500	60	16.18	58.15	41.97
4	75280000	0.670044	140000	56	17.15	54.29	37.14
5	75008000	1.037580	112500	54	19.69	52.56	32.87
6	75280000	0.189412	139000	46	12.3	44.17	31.87
7	75280000	0.919223	140000	51	20.35	49.69	29.34
8	77160000	1.838250	135500	59	29.92	58.03	28.11
9	75280000	0.655409	131500	46	16.42	44.53	28.11
10	75280000	0.207153	182000	43	14.35	41.45	27.1

Table 8-17 Screening output for 3 or more lane TWLTL (total crashes)

Rank	RD_ID	Length	AADT	COUNT	Pred	Exp	PSI
1	75025500	0.955905	11500	68	9.54	65.87	56.33
2	75025500	1.114230	11500	85	29.02	83.77	54.75
3	75590000	0.859996	31500	85	29.26	83.78	54.52
4	75000126	0.467922	14500	60	4.7	57.23	52.53
5	75025500	0.356414	11500	46	8.27	44.09	35.82
6	75000091	0.524702	7500	42	3.89	39.45	35.56
7	92010000	0.744987	28500	59	23.16	57.99	34.83
8	75000091	0.768833	7500	41	5.82	38.9	33.08
9	77010000	0.654359	23500	51	18.34	49.91	31.57
10	75030000	0.377940	27500	43	12.44	41.65	29.21

Table 8-18 Screening output for 3 or more lane TWLTL (FI crashes)

Rank	RD_ID	Length	AADT	COUNT	Pred	Exp	PSI
1	75590000	0.859996	31500	42	13.03	40.71	27.68
2	75025500	1.114230	11500	41	13.31	39.73	26.42
3	75000126	0.467922	14500	30	1.97	26.99	25.02
4	92010000	0.744987	28500	35	10.47	33.68	23.21
5	75025500	0.356414	11500	23	4.04	21.09	17.05
6	75012500	1.182130	8200	30	13.09	29.07	15.98
7	75000091	0.768833	7500	19	2.46	16.81	14.35
8	75025500	0.955905	11500	20	3.9	18.23	14.33
9	75000192	0.584452	15500	22	6.46	20.67	14.21
10	75000091	0.524702	7500	18	1.68	15.46	13.78

Table 8-19 Screening output for one–way roads (total crashes)

Rank	RD_ID	Length	AADT	COUNT	Pred	Exp	PSI
1	75000408	0.779566	7800	58	14.63	57.08	42.45
2	75000369	0.205179	7000	42	3.79	40.52	36.73
3	75040102	0.614347	8400	48	10.33	47.06	36.73
4	75080000	0.275668	7500	37	4.56	35.69	31.13
5	75000184	0.621036	4800	35	7.53	34.04	26.51
6	75000098	1.106300	1500	39	13.34	38.2	24.86
7	75080000	0.525832	14000	30	11.69	29.41	17.72
8	75080000	0.555808	14000	30	12.52	29.43	16.91
9	75080000	0.471444	7500	24	6.97	23.23	16.26
10	75000262	0.339616	9000	21	5.82	20.21	14.39

Table 8-20 Screening output for one-way roads (FI crashes)

Rank	RD_ID	Length	AADT	COUNT	Pred	Exp	PSI
1	75080000	0.275668	7500	28	2.16	26.02	23.86
2	75000408	0.779566	7800	28	7.68	26.97	19.29
3	75000098	1.106300	1500	19	4.82	17.82	13
4	75000184	0.621036	4800	17	3.42	15.82	12.4
5	75080000	0.471444	7500	16	3.43	14.86	11.43
6	75040102	0.614347	8400	14	5.365	13.29	7.925
7	75080000	0.525832	14000	15	6.748	14.4	7.652
8	75000369	0.205179	7000	9	1.741	7.865	6.124
9	75000184	0.353668	4800	9	1.798	7.889	6.091
10	75040000	0.396734	13500	10	4.769	9.486	4.717

8.2. SPFs for Intersections

For intersections, the model fitting procedure was similar to that used for segments. The research team used Full Bayesian Poisson Lognormal models to predict crash frequency, but tried four different variable combinations to identify the best model. These four combinations included:

(1) Model 1: two variables - the major AADT and the minor AADT,

(2) Model 2: two variables - the major AADT and the interaction between the major AADT and minor AADT;

(3) Model 3: two variables - the minor AADT and the interaction between the major AADT and minor AADT, and

(4) Model 4: one variable - the sum of the major AADT and minor AADT.

The above models were developed for the total crashes and the fatal-and-injury crashes.

8.2.1. Four-Leg Intersection

In Tables 8-21 and 8-22, the results show that there were no significant differences among these four models; to be consistent with the HSM (2010), model 1 was adopted for both cases.

The crash prediction model for total crashes in 4-leg signalized intersections is:

$$Pred_{it} = Exp\left[-12.08 + 0.9245 * \ln\left(AADT_{major}\right) + 0.5489 * \ln(AADT_{minor})\right] \quad (8\text{-}1)$$

For fatal-and-injury crashes, the selected model for 3-leg signalized intersections is:

$$Pred_{it} = Exp\left[-12.01 + 0.8527 * \ln\left(AADT_{major}\right) + 0.4812 * \ln(AADT_{minor})\right] \quad (8\text{-}2)$$

Table 8-21 Four different variable combination models: total crashes

Variable	MODEL 1		MODEL 2		MODEL 3		MODEL 4	
	$AADT_{major}$ - $AADT_{minor}$		$AADT_{major}$ – ($AADT_{major}$×$AADT_{minor}$)		$AADT_{minor}$ – ($AADT_{major}$×$AADT_{minor}$)		Total Entering Vehicles	
	mean	s.d.	mean	s.d.	mean	s.d.	mean	s.d.
Intercept	-12.08**	1.098	-12.37**	0.9492	-11.74**	0.9009	-14.04**	1.222
Ln $AADT_{major}$	0.9245**	0.106	0.4088**	0.1132				
Ln $AADT_{minor}$	0.5489**	0.0566			-0.333**	0.1207		
Ln ($AADT_{major}$ *$AADT_{minor}$)			0.545**	0.0563	0.888**	0.0875		
Ln (Total Volume Entering)							1.549**	0.1058
MAD	18.6368		18.59177		18.70905		19.83413	
MSPE	731.7612		731.9605		733.8136		795.0376	
DIC	1045.64		1046.33		1046.72		1048.4	

** significant at 5%.

Table 8-22Four different variable combination models: FI crashes

Variable	MODEL 1		MODEL 2		MODEL 3		MODEL 4	
	$AADT_{major}$ - $AADT_{minor}$		$AADT_{major}$ – ($AADT_{major}$×$AADT_{minor}$)		$AADT_{minor}$ – ($AADT_{major}$×$AADT_{minor}$)		Total Entering Vehicles	
	mean	s.d.	mean	s.d.	mean	s.d.	mean	s.d.
Intercept	-12.01**	1.324	-12.23**	1.361	-12.1**	1.245	-13.81**	1.32
Ln $AADT_{major}$	0.8527**	0.1252	0.4231**	0.1702				
Ln $AADT_{minor}$	0.4812**	0.0666			-0.376**	0.1465		
Ln ($AADT_{major}$ *$AADT_{minor}$)			0.4647**	0.0631	0.8591**	0.1137		
Ln (Total Volume Entering)							1.407**	0.1138
MAD	5.425141		5.430646		5.425902		5.614377	
MSPE	53.81742		53.68212		53.86145		55.6275	
DIC	819.266		819.3		819.801		820.99	

** significant at 5%.

Based on the above model, PSI can be calculated as the difference between the predicted crash frequency and the expected crash frequency for each road site. Tables 8-23 and 8-24 present the top ten hotspots for total crashes and FI crashes for each roadway type.

Table 8-23 Top 10% hotspots for urban 4-leg signalized intersections: total crash models

Rank	Roadway ID 1	Roadway ID 2	Roadway ID 3	Roadway ID 4	PSI
1	75270000	75270000	75000103	75000139	118.9
2	75230500	75230500	75270000	75270000	75.83
3	75000208	75000208	75270000	75270000	66.4
4	75035001	75035001	75039000	75000279	63.44
5	75250000	75250000	75590000	75000161	57.87
6	75000155	75000155	75270000	75270000	55.32
7	92090000	92090000	92000055	92000054	49.96
8	75010000	75010000	75000319	75000017	43.3
9	75060000	75060000	75620000	75620000	39.91
10	75010000	75010000	75000156	75600000	38.45
11	75200000	75200000	75510500	75510501	38.26
12	77510000	77510000	77120000	77120000	34.89
13	75050000	75050000	75000087	75000087	32.18
14	75020000	75020000	75190000	75190000	30.81

Table 8-24 Top 10% hotspots for urban 4-leg signalized intersections: FI crash models

Rank	Roadway ID 1	Roadway ID 2	Roadway ID 3	Roadway ID 4	PSI
1	75270000	75270000	75000103	75000139	22.01
2	75000155	75000155	75270000	75270000	19.45
3	75000208	75000208	75270000	75270000	15.8
4	75190000	75190000	75250000	75250000	13.16
5	75250000	75250000	75000086	75000087	13.05
6	75020000	75020000	75190000	75190000	11.49
7	75060000	75060000	75000142	75000142	11.09
8	75080000	75080000	75003000	75003000	11.01
9	75050000	75050000	75190000	75190001	10.66
10	92090000	92090000	92605000	92000076	10.4
11	75250000	75250000	75590000	75000161	9.794
12	75060000	75060000	75040000	75040000	9.617
13	92090000	92090000	92000055	92000054	9.586
14	75230500	75230500	75190001	75190001	9.074

8.2.2. Three-Leg Intersections

In Table 8-25, the results show that there were no significant differences among the four models; to be consistent with the HSM (2010), model 1 was adopted for both cases. The crash prediction model for total crashes in a 3-leg signalized intersection is:

$$Pred_{it} = Exp\left[-8.492 + 0.876 * \ln\left(AADT_{major}\right) + 0.206 * \ln\left(AADT_{minor}\right)\right] \quad (8\text{-}3)$$

As for fatal-and-injury crashes (Table 8-26), only models 3 and 4 had significant coefficients at a 95% confidence level. For calculation convenience, in the following sections the research team selected model 4 to predict the PSI.

For fatal-and-injury crashes, the selected model for 3-leg signalized intersections is:

$$Pred_{it} = Exp\left[-8.523 + 0.908 * (\ln\left(AADT_{major}\right) + \ln\left(AADT_{minor}\right)\right] \quad (8\text{-}4)$$

Table 8-25 Four different variable combination models: total crashes

Variable	MODEL 1		MODEL 2		MODEL 3		MODEL 4	
	$AADT_{major}$ - $AADT_{minor}$		$AADT_{major}$ – ($AADT_{major}$×$AADT_{minor}$)		$AADT_{minor}$ – ($AADT_{major}$×$AADT_{minor}$)		Total Entering Vehicles	
	mean	s.d.	mean	s.d.	mean	s.d.	mean	s.d.
Intercept	-8.492**	1.401	-8.286**	1.489	-8.507**	1.309	-9.427**	1.77
Ln $AADT_{major}$	0.876**	0.126	0.583**	0.178				
Ln $AADT_{minor}$	0.206**	0.080			-0.687**	0.146		
Ln ($AADT_{major}$ *$AADT_{minor}$)			0.243**	0.088	0.885**	0.111		
Ln (Total Volume Entering)							1.113**	0.154
MAD	14.800		14.694		14.815		14.609	
MSPE	467.965		463.582		467.795		453.553	
DIC	780.194		780.289		780.237		779.919	

** significant at 5%.

Table 8-26 Four different variable combination models: fatal-and-injury crash models

Variable	1) MODEL		2) MODEL		3) MODEL		4) MODEL	
	$AADT_{major}$ - $AADT_{minor}$		$AADT_{major}$ – ($AADT_{major}$*$AADT_{minor}$)		$AADT_{minor}$ – ($AADT_{major}$*$AADT_{minor}$)		Total Entering Vehicles	
	mean	s.d.	mean	s.d.	mean	s.d.	mean	s.d.
Intercept	-8.33**	1.802	-8.48**	1.675	-8.419**	1.79	-8.523**	1.774
Ln $AADT_{major}$	0.891**	0.157	0.879**	0.178				
Ln $AADT_{minor}$	0.026#	0.081			-0.875**	0.194		
Ln ($AADT_{major}$ *$AADT_{minor}$)			0.026	0.079	0.900**	0.159		
Ln (Total Volume Entering)							0.908**	0.154
MAD	28.67648579		28.6788		28.68016364		28.66561771	
MSPE	1380.108356		1378.210243		1378.579235		1370.715375	
DIC	583.995		579.267		579.414		584.116	

not significant at 20%, ** significant at 5%.

Based on the above model, PSI can be calculated as the difference between the predicted crash frequency and the expected crash frequency for each road site. Tables 8-27 to 8-28 present the top ten hotspots for total crashes and FI crashes for each roadway type.

Table 8-27 Top 10% hotspots for urban 3-leg signalized intersections: total crash models

Rank	Roadway ID 1	Roadway ID 2	Roadway ID 3	PSI
1	75270000	75270000	75000106	74.24
2	75270000	75270000	75505500	64.66
3	75037000	75037000	75000001	60.64
4	75060000	75060000	75160501	57.96
5	77000064	77080000	77080000	52.8
6	75037000	75000288	77170000	44.15
7	75270000	75050000	75050000	41.02
8	75260500	75060000	75060000	36.1
9	75060000	75060000	75000179	34.98
10	92000079	92030000	92030000	33.48
11	75000229	75060000	75060000	33.2

Hotpots for both 4-leg and 3-leg intersections can be seen in Figure 8-1, which shows that several hotspots found in 4-leg and 3-leg signalized intersections were located along the same roads, such as SR 50 and SR 435.

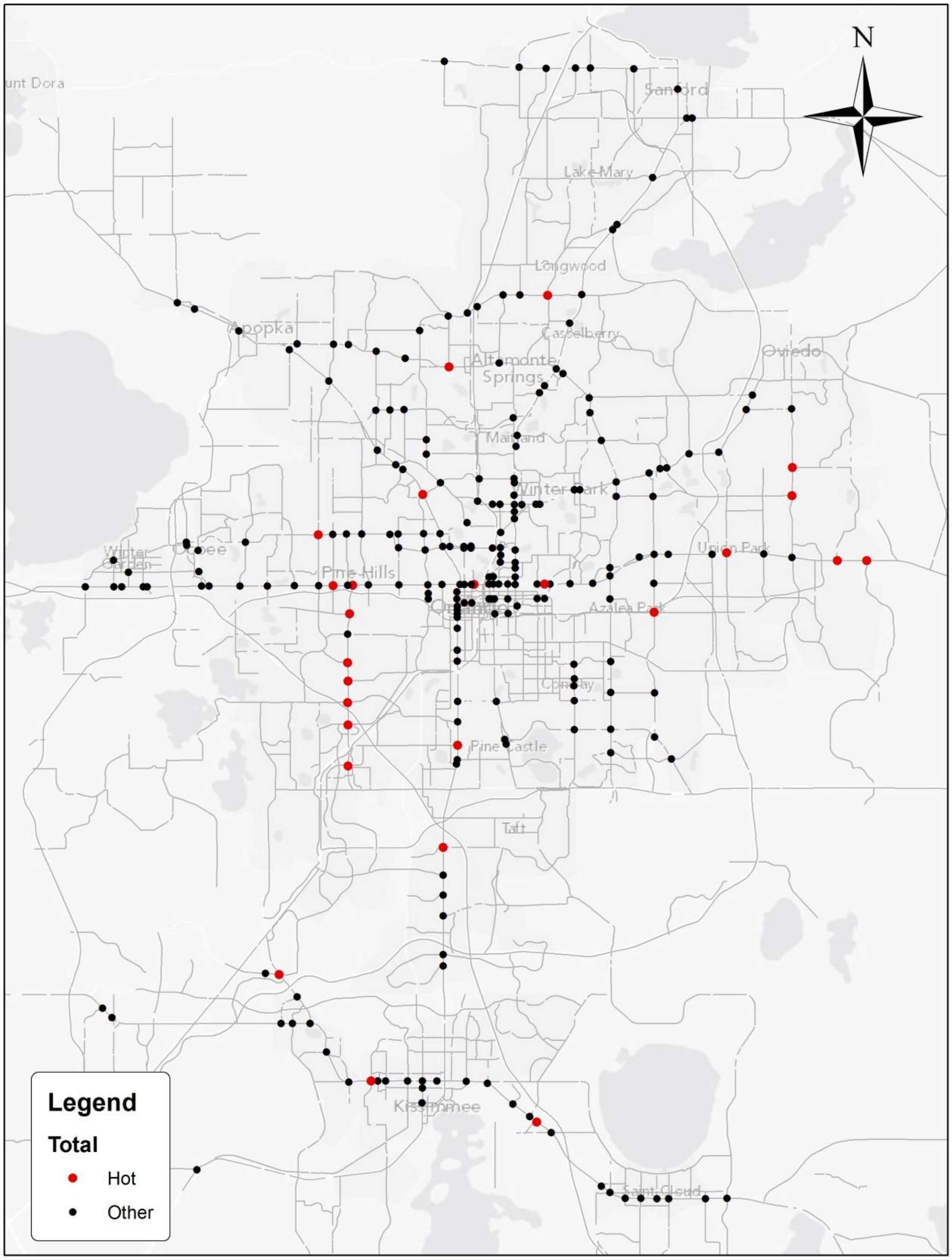


Figure 8-1 Top 10% hotspots for 4-leg and 3-leg signalized intersections for total crashes (red circle: hot 4-leg and 3-leg signalized intersections, black circle: normal 4-leg and 3-leg signalized intersections)

The same procedure mentioned above was adopted to define the hotspots for fatal-and-injury crashes. Table 8-28 lists the top 10% hotspots for urban 3-leg signalized intersections. As was mentioned above, several hotspots were located along the same roads. In addition, the locations of the hotspots for total crashes and fatal-and-injury crashes for both 4-leg and 3-leg intersections were similar (see Figure 8-2).

Table 8-28 Top 10% hotspots for urban 3-leg signalized intersections: fatal-and-injury crash models

Rank	Roadway ID 1	Roadway ID 2	Roadway ID 3	PSI
1	75270000	75270000	75505500	20.99
2	75270000	75270000	75000106	18.37
3	75060000	75060000	75000179	13.68
4	75060000	75060000	75160501	9.777
5	92000079	92030000	92030000	8.634
6	75039000	75039000	75000280	8.625
7	75270000	75050000	75050000	8.11
8	75010000	75010000	75000139	6.977
9	92030000	92030000	92000038	6.834
10	75000229	75060000	75060000	6.489
11	75010000	75010000	75000016	6.409

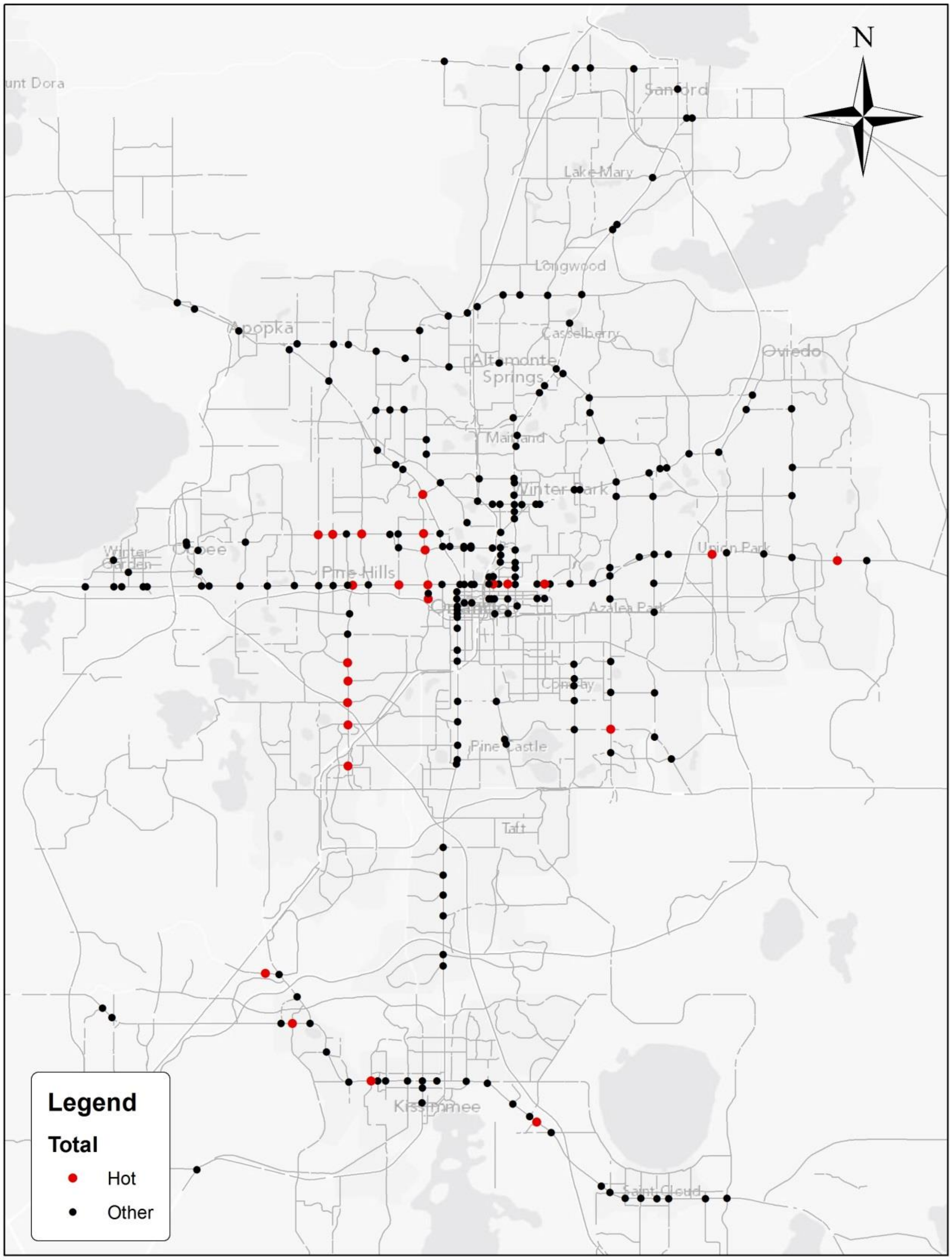


Figure 8-2 Top 10% hotspots for 3-leg signalized intersections for fatal-and- injury crashes (red circle: hot 4-leg and 3-leg signalized intersections, black circle: normal 4-leg and 3-leg signalized intersections)

9. INTEGRATION OF MACRO- AND MICRO-LEVEL SCREENING

In the previous chapters, the research team identified hot spots/areas at both the macro- and microscopic levels. The next step is to integrate these macroscopic and microscopic screening results, and then provide a comprehensive, strategic, and effective traffic safety improvement plan. The integration strategy of combining these two-level screening results is described in Section 9.1, below. In Section 9.2, the integration results for both total and fatal-and-injury crashes are provided as GIS maps and tables. Lastly, a summary that can be drawn from this chapter are provided in Section 9.3.

9.1. Integration Process

This section describes the overall procedure used to integrate the screening results from both the macro- and microscopic levels. Section 9.1.1 describes a brief integration strategy; Section 9.1.2 explains the overall procedure of the integration work.

9.1.1. Integration Strategy

Numerous studies have been done to analyze at the microscopic level certain locations and sites with high traffic safety risk, including the HSM Part B (Hauer, 1996; Heydecker et al., 1991; Kononov et al., 2003; Chung et al., 2007; Ragland et al., 2007; AASHTO, 2010). Recently, several studies have begun to focus on zonal-based network screening at the macroscopic level (Abdel-Aty et al., 2013; Pirdavani et al., 2013). Compared to microscopic safety studies, macroscopic-focused research is more efficient at integrating zonal-level features into crash prediction models and identifying hot zones. However, macroscopic screening has accuracy limitations because it cannot identify and separate hotspots from other sites within a single zone. Thus, a new integrated screening approach is needed to overcome the above-mentioned shortcomings of current screening techniques, and to achieve a balance between efforts towards accuracy and efficiency. In accomplishing such a goal, we can obtain a comprehensive perspective from two levels of screening, and therefore develop more appropriate traffic safety treatments.

However, this integration task is challenging because we need to (1) combine various SPFs from different scales, areas, and roadway types; (2) determine an appropriate weight for each group; and (3) chose a measurement for our final results. Moreover, the integration requires considerable GIS work to determine and visualize the spatial relationships between the segments/intersections and the TSAZs.

In order to identify whether a zone has safety issues at the macro- and/or microscopic levels, all TSAZs are classified into twelve categories which include two scale groups (macro or micro) and four safety levels (hot, normal, cold, or no data). These categories are: HH, HN, HC, HO, NH, NN, NC, NO, CH, CN, CC, and CO (see Table 9-1). The first character of the classification represents the macroscopic safety risk, and the second character illustrates the microscopic safety risk. Thus, HH zones have both macro- and micro-level safety problems; HN zones are risky at the macroscopic level, but their micro-level risk is moderate. Alternately, HC zones have safety problems only at the macroscopic level. NH zones face moderate crash risk at the zonal level, but their microscopic crash risk is quite high. NN zones are intermediate for traffic safety both at the macro- and microscopic levels. Likewise, NC zones have a moderate risk at the macroscopic level but their safety risk at the microscopic level is low. CH zones have high safety risk only at the microscopic level, such as at intersections and segments, while CN zones have a low crash risk at the macroscopic level but an intermediate crash risk at the microscopic level. CC zones are safe at both the macro- and microscopic levels.

HO, NO, and CO zones are dangerous, moderate, and safe, respectively, at the macroscopic level, but they do not have segment or intersection data at all.

Table 9-1 Hot Zone Classification

		Micro Level			
		Hot	Normal	Cold	No Data
Macro Level	Hot	HH	HN	HC	HO
	Normal	NH	NN	NC	NO
	Cold	CH	CN	CC	CO

9.1.2. Integration Procedure

The following integration was conducted based on the macroscopic and microscopic screening results illustrated in the previous chapters (see Figure 9-1). The integration procedure is summarized in Figure 9-2. At the macroscopic level, TSAZs were ranked by their zonal PSIs; TSAZs with top 10% macro-level PSIs were classified as “Hot” zones. TSAZs with bottom 10% zonal PSIs were classified as “Cold” zones, and other TSAZs which were neither “Hot” nor “Cold” were categorized as “Normal.” These percentiles could be changed as needed. At the microscopic level, the calculation of average PSI was more complicated because each TSAZ had several intersections and segments. The PSIs of the intersections in each TSAZ were averaged by the number of intersections, and the zones were ranked by their averaged intersection PSI.

Simultaneously, the PSIs of segments in each zone were averaged by the total length of the segments in the zone, and zones were ranked by their averaged segment PSI. After that, both the intersection and segment PSI ranks were averaged; the TSAZs were ranked by the final averaged intersection and segment PSIs. As was the case at the macroscopic level, TSAZs with top 10% micro-level PSIs were categorized as “Hot” zones at the microscopic level. Finally, TSAZs were classified into twelve categories based on macro- and micro-level screening results. It should be noted that we used the total length of the segments to normalize the segment PSIs because the lengths of the segments could vary. Also, the percentile ranks of the PSIs were used in the integration (instead of the original PSIs) because the units of PSI intersections and PSI segments were different.

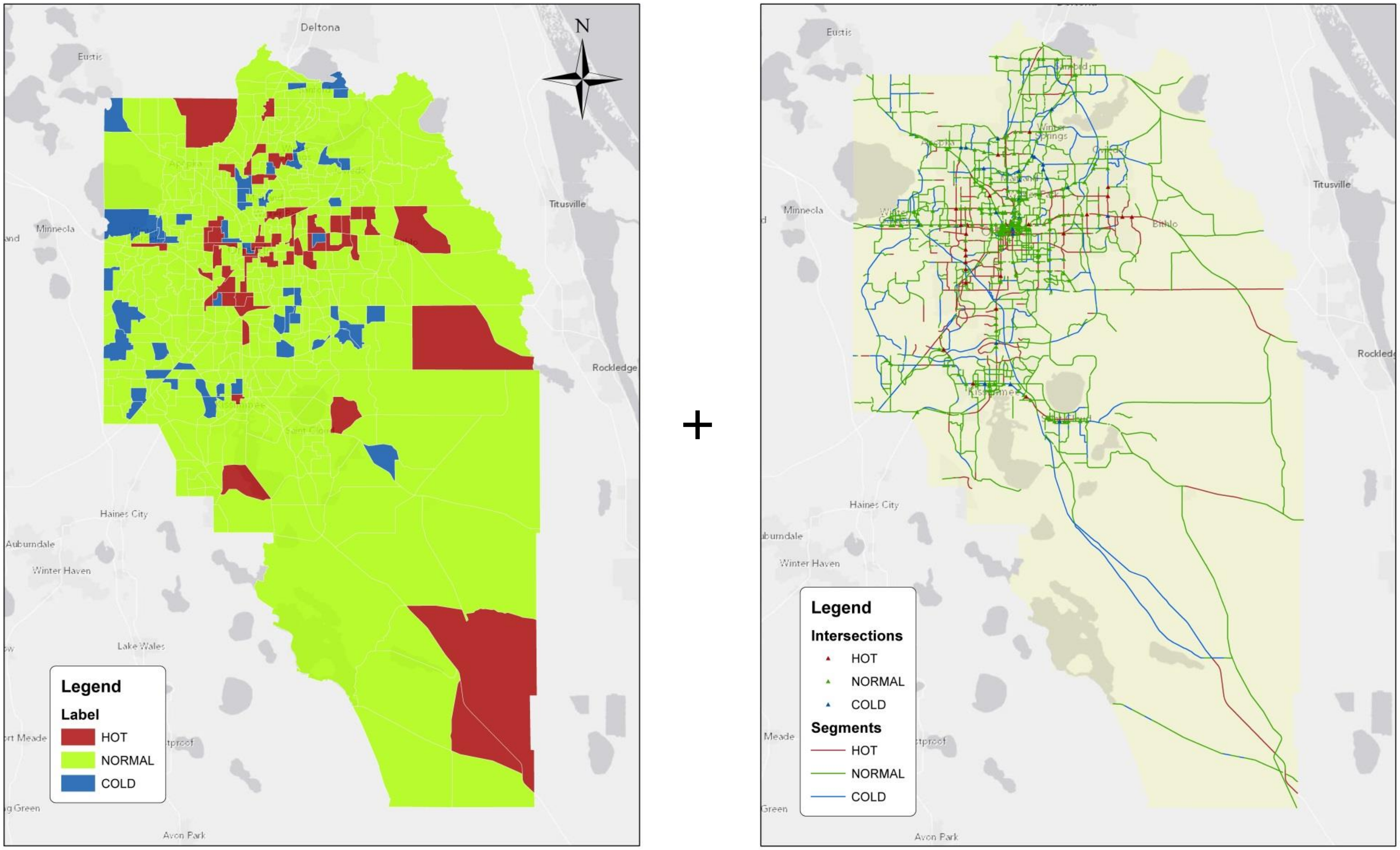


Figure 9-1 Results of macroscopic hot zone screening (left) and microscopic hotspot screening (right)

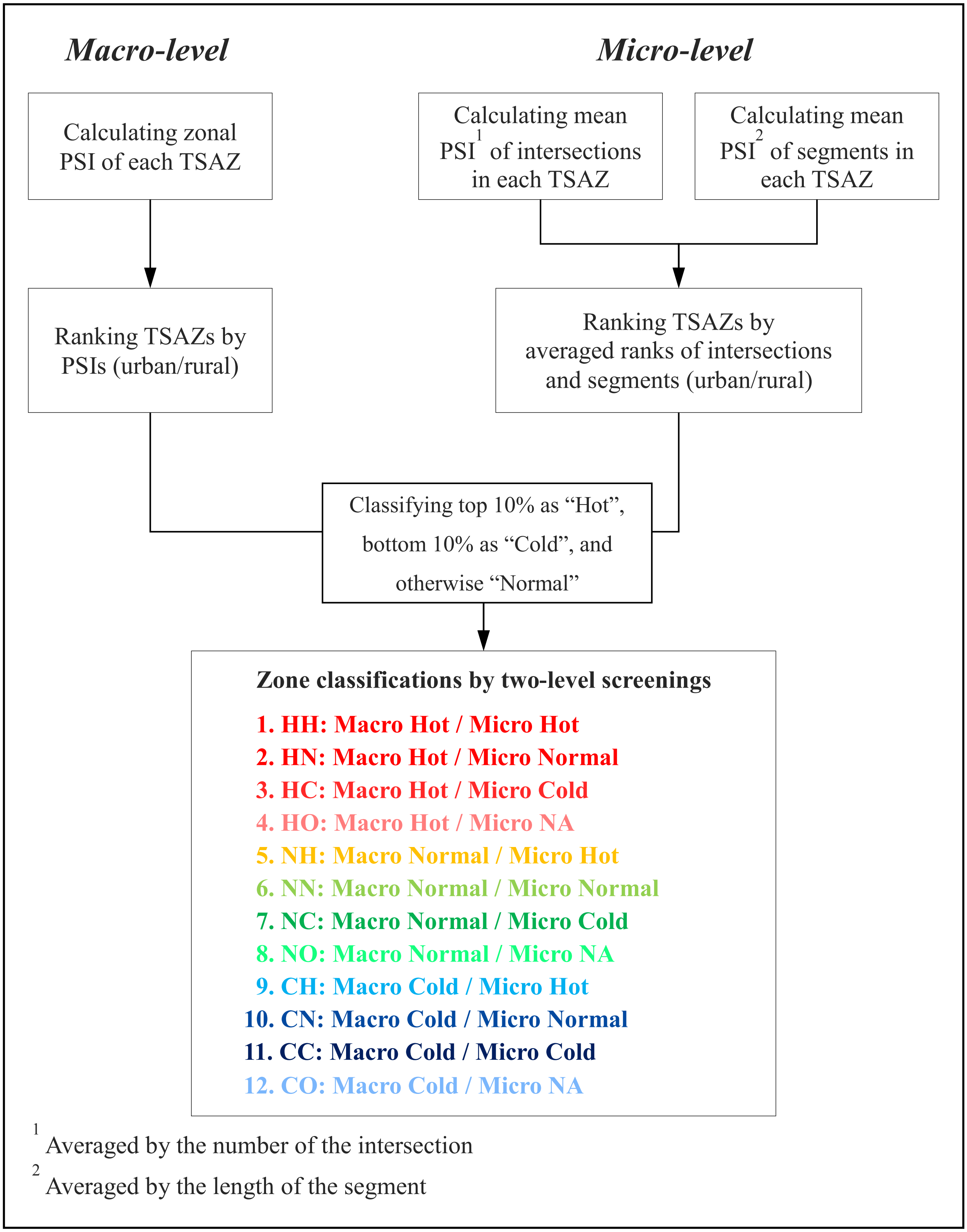


Figure 9-2 Integration process

9.2. Integration Results

Section 9.2 provides the integration results of the macro- and micro-level tests. The research team analyzed hot TSAZs for both total crashes and fatal-and-injury crashes in order to be consistent with the HSM. Moreover, by doing so the results also allowed for an examination of whether there were any differences with regards to hot zone locations among various crash severity levels. The total crash hot zone screening results display the overall crash distributions within the study area, whereas the fatal-and-injury crash hot zone screening results represent the more severe crash distributions. Sections 9.2.1 and 9.2.2 explain the detailed hot zone classifications for total crashes and fatal-and-injury crashes, respectively. Finally, the two screening results are compared in Section 9.2.3.

9.2.1. Total Crashes

Table 9-2 shows the number of zones by hot zone classification for total crashes. Overall, 26 HH zones were identified, which is top priority for safety treatments because this type of zone has a higher crash risk at both the macroscopic and microscopic levels. Moreover, there are 20 HN zones and 21 NH zones, the next highest priority for treatment. HN zones have serious safety problems at the macroscopic level and an intermediate level of risk at the microscopic level, whereas NH zones have a high traffic crash risk at the microscopic level and an intermediate risk at the macroscopic level. It is also necessary to pay attention to HC and CH zones. Both HC and CH zones have contradicting hot zone identifications at different levels. There are three HC zones, each of which is exceedingly risky only at the macroscopic level but safe at the microscopic level. Overall, two CH zones were identified and both are very dangerous at the microscopic level but safe at the macroscopic level. Eight zones were identified as CC, which means they are safe both at the macroscopic and microscopic levels. There is no significant difference in the hot zone identification of urban and rural areas, except for with the NO zones. NO/CO zones, which have no micro-level components, appear in higher percentages in rural areas (25%) than in urban areas (7%). This is because the density levels of major roadway networks in rural areas are much lower than those in urban areas, even though such zones in rural areas are much larger.

Table 9-2 Number of zones by hot zone classification (total crashes)

Classification	Urban		Rural		Sum	
	Zones	%	Zones	%	Zones	%
HH	22	5.1%	4	5.6%	26	5.2%
HN	18	4.2%	2	2.8%	20	4.0%
HC	2	0.5%	1	1.4%	3	0.6%
HO	0	0.0%	0	0.0%	0	0.0%
NH	19	4.4%	2	2.8%	21	4.2%
NN	261	61.0%	32	44.4%	293	58.6%
NC	34	7.9%	6	8.3%	40	8.0%
NO	29	6.8%	17	23.6%	46	9.2%
CH	1	0.2%	1	1.4%	2	0.4%
CN	34	7.9%	5	6.9%	39	7.8%
CC	7	1.6%	1	1.4%	8	1.6%
CO	1	0.2%	1	1.4%	2	0.4%
Sum	428	100.0%	72	100.0%	500	100.0%

Figure 9-3 presents the spatial distribution of TSAZs by hot zone classification for total crashes in urban areas. It was observed that many HH/HN/NH zones are located along State Road 50 (Colonial Drive), State Road 435 (Kirkman Road), State Road 408 (East-West Expressway), US Route 17/92/441 (Orange Blossom Trail), and Interstate 4.

There are two large clusters containing multiple HH zones. The first HH cluster is located in the center of the map, adjacent to Interstate 4, State Road 435 (Kirkman Road), and US 17/92/441 (Orange Blossom Trail). The second cluster is in the East Orlando area. This research shows that several principle arterial roadways (such as State Road 408 and State Road 50) cross the second HH cluster.

On the other hand, it seems that CC zones do not form clusters. Some CC zones are located in the downtown area, whereas other zones are located in suburban areas.

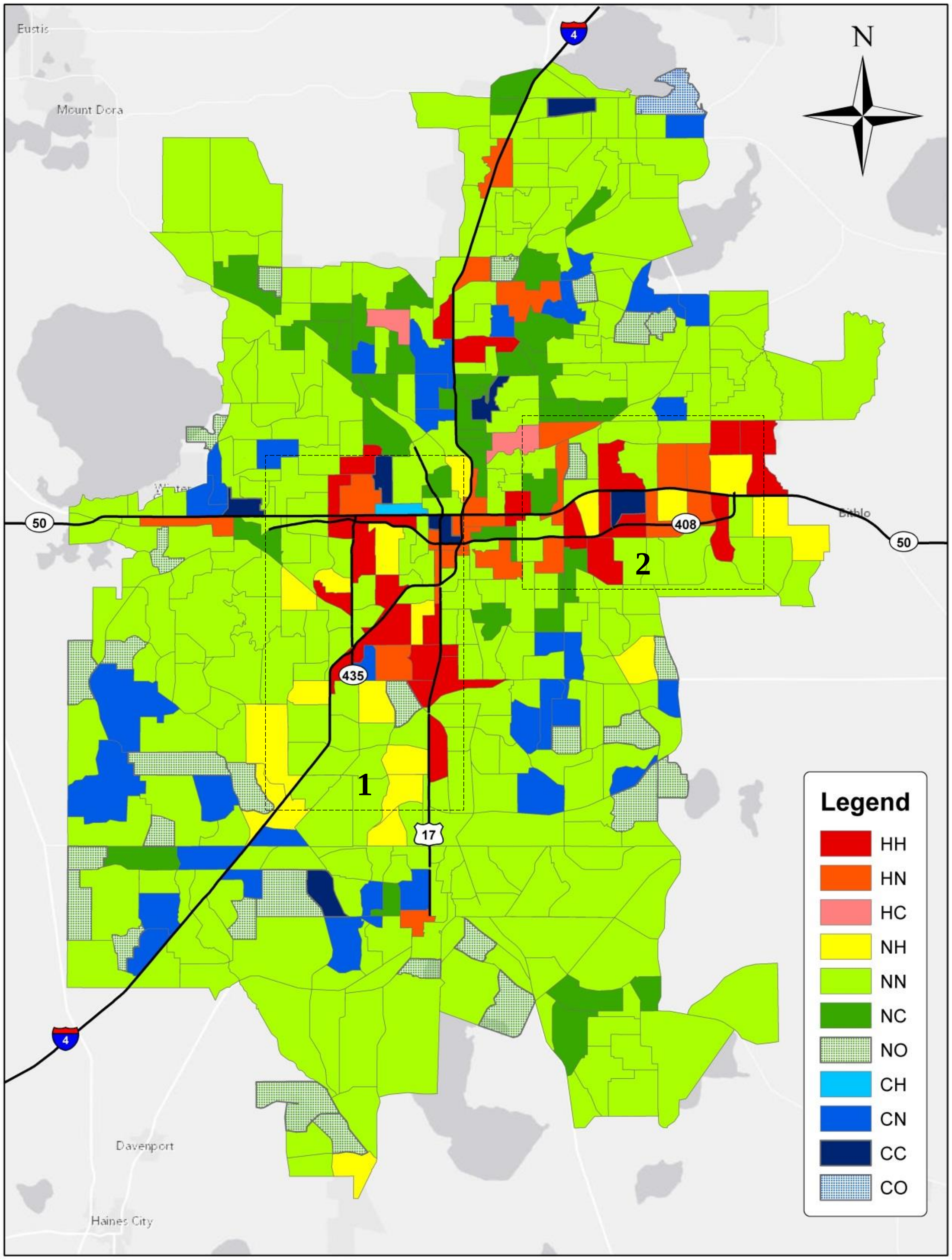


Figure 9-3 Distribution of zones by hot zone classification in urban areas (total crashes)

Figure 9-4 displays the spatial distribution of TSAZs by hot zone classification for total crashes in rural areas. It was found that HH/HN/NH zones could be found near principle arterial roadways including State Road 528 (Beachline Expressway), State Road 520, and State Road 91 (Florida's Turnpike), or adjacent to urban areas.

However, compared to urban areas, HH zones for total crashes in rural areas form no clusters and all are spatially isolated. Two zones are located in the east (near State Road 520). The first HH zone in the northwest has a mixed land use of residential and commercial, and a collector road crossing the zone (County Road 435). The other HH zone is in the southwest corner of the area and its land use is a mixture of residential and commercial. County Road 531 provides the boundary for this zone, and functions as a collector. Only one zone in this rural area is classified as a CC zone for total crashes. This zone is mainly in an agricultural area with some residential buildings.

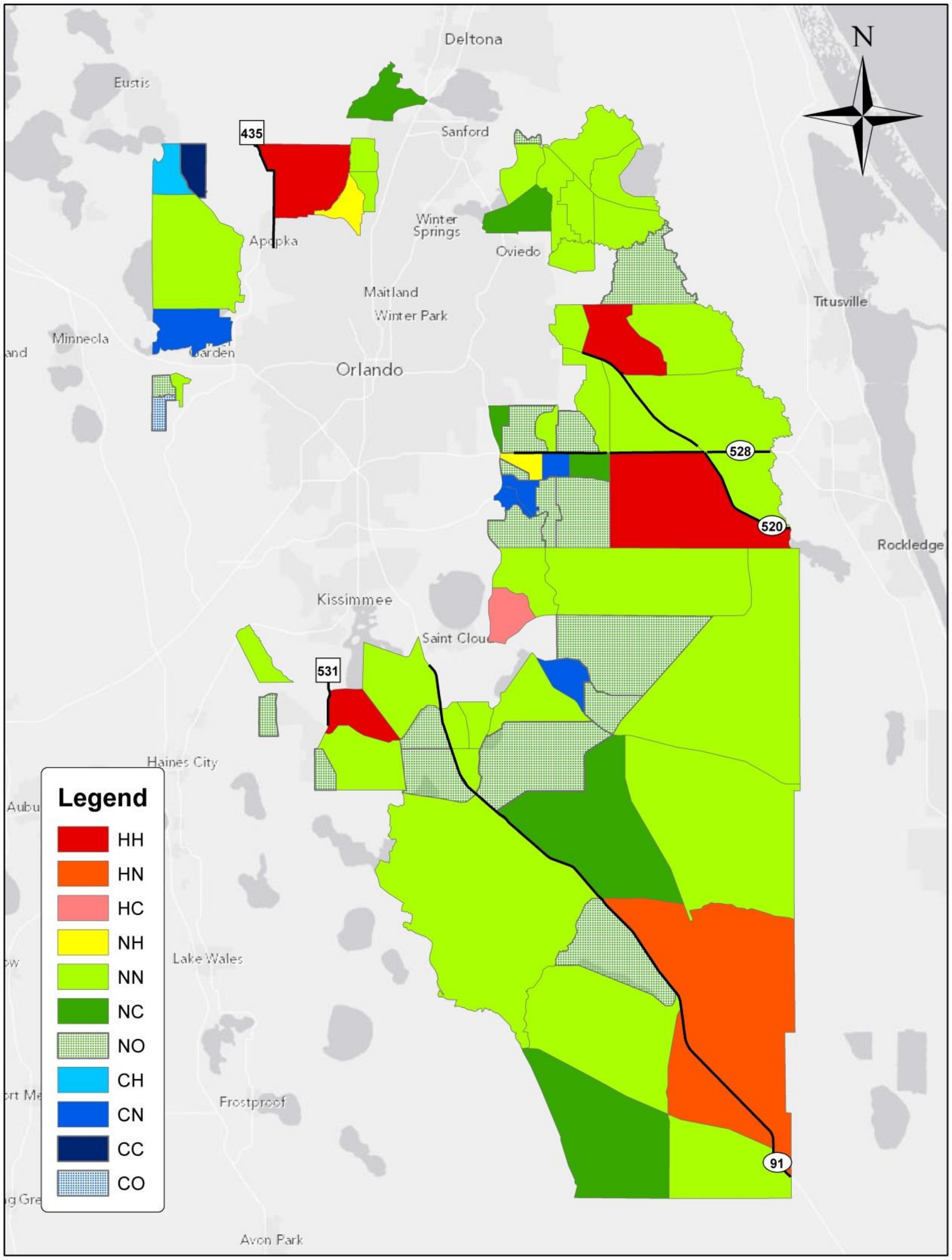


Figure 9-4 Distribution of zones by hot zone classification in rural areas (total crashes)

Table 9-3 compares the local features of HH and CC zones in urban and rural areas. For HH zones in urban areas, it was shown that the 'Population density' values in HH zones are more than three times larger than in the entire urban area. Both 'Proportion of Hispanics' and 'Number of hotel, motel, and timeshare rooms per square mile' in HH zones are also larger, as compared to the overall urban area. Moreover, 'Proportion of roadways with 55 mph or higher speed limits' is higher in HH zones, as well. This suggests that zones containing more high speed roadways are more vulnerable to traffic crash occurrences.

CC zones in urban areas also have larger 'Population density' levels than average. However, the 'Proportion of Hispanics' in CC zones is slightly lower than the average. Likewise, the 'Number of hotel, motel, and timeshare rooms per square mile' and the 'Proportion of roadways with 55 mph or higher speed limits' in CC zones are also smaller than the average.

As compared to the average, HH, and CC zonal features in rural areas (Table 2-2), the 'Population density' in HH zones is much larger than average, whereas in CC zones it is only half the average. 'Proportion of Hispanics' in HH zones is also higher than the average, whereas in CC zones it is slightly lower.

Table 9-3 Comparison of zonal features between the average, HH, and CC zones (total crashes)

Zonal factors	Urban			Rural		
	Average	HH	CC	Average	HH	CC
Population density	410.0	1258.0	1297.5	124.4	551.9	62.4
Proportion of Hispanics	0.274	0.340	0.240	0.279	0.399	0.238
Number of hotel, motel, and timeshare rooms per square mile	139.7	590.9	65.86	51.05	1.138	0.000
Proportion of roadways with 55 mph or higher speed limits	0.052	0.077	0.045	0.075	0.000	0.000

9.2.2. Fatal-and-Injury Crashes

In the previous sub-section, the total crash hot zone screening results showed the general crash distributions. However, it is also necessary to examine where more severe crashes occur, and their corresponding features. Thus, in this sub-section the results of a fatal-and-injury crash hot zone identification are described and compared with the results of the total crash hot zone identification.

Table 9-4 summarizes the number of zones by hot zone classification for fatal-and-injury crashes. In the first section, we start from the zones that are consistent at both the macro- and microscopic levels. There are only 12 HH and two CC zones identified. Considering that there are 26 HH and eight CC zones for total crashes, the number of HH/CC zones for fatal-and-injury crashes is quite small when compared to the total crash case. It seems that consistency in the hot zone/cold zone classifications between the macro- and microscopic levels is reduced in the case of fatal-and-injury crashes. It could thus be concluded that fatal-and-injury crashes are more significantly influenced by network-level characteristics than zonal factors, as compared to total crashes.

Furthermore, it was observed that there is little difference in the percentages of each category in urban and rural areas. The proportion of HH zones in urban areas is 2.1% and in rural areas 4.2%. Similarly, the proportion of CC zones in urban areas is only 0.2%, while in the rural areas 1.4%. This shows that the hot zone classifications from the two levels are more consistent in the rural area than in the urban area.

Table 9-4 Number of zones by hot zone classification (fatal-and-injury crashes)

Classification	Urban		Rural		Sum	
	Zones	%	Zones	%	Zones	%
HH	9	2.1%	3	4.2%	12	2.4%
HN	26	6.1%	4	5.6%	30	6.0%
HC	7	1.6%	0	0.0%	7	1.4%
HO	0	0.0%	0	0.0%	0	0.0%
NH	31	7.2%	4	5.6%	35	7.0%
NN	253	59.1%	30	41.7%	283	56.6%
NC	35	8.2%	7	9.7%	42	8.4%
NO	24	5.6%	16	22.2%	40	8.0%
CH	2	0.5%	0	0.0%	2	0.4%
CN	34	7.9%	5	6.9%	39	7.8%
CC	1	0.2%	1	1.4%	2	0.4%
CO	6	1.4%	2	2.8%	8	1.6%
Sum	428	100.0%	72	100.0%	500	100.0%

As seen in Figure 9-5, the majority of the HH/HC zones in urban areas are located along State Road 50 and State Road 408. However, the HH/HN zones near Interstate 4 showed a considerably reduced number as compared to total crash hot zones. As mentioned earlier, NH zones for total crashes are concentrated in the downtown Orlando area. However, the NH zones for FI crashes are dispersed from the center of Orlando and most are located in suburban areas. This implies that more severe crashes are more likely in suburban areas than in urban areas. It can be concluded, then, that the total crash risk is higher in urban areas because such areas have more significant exposure to traffic; at the same time, driving speeds in urban areas are slower than in suburban areas.

It was also observed that HH zones form two clusters. The first cluster is located between State Road 435 (Kirkman Road) and US 17/92/441 (Orange Blossom Trail) near Interstate 4 in the center of Orlando. The second cluster is located in East Orlando along State Road 50 (Colonial Drive) and surrounds the University of Central Florida.

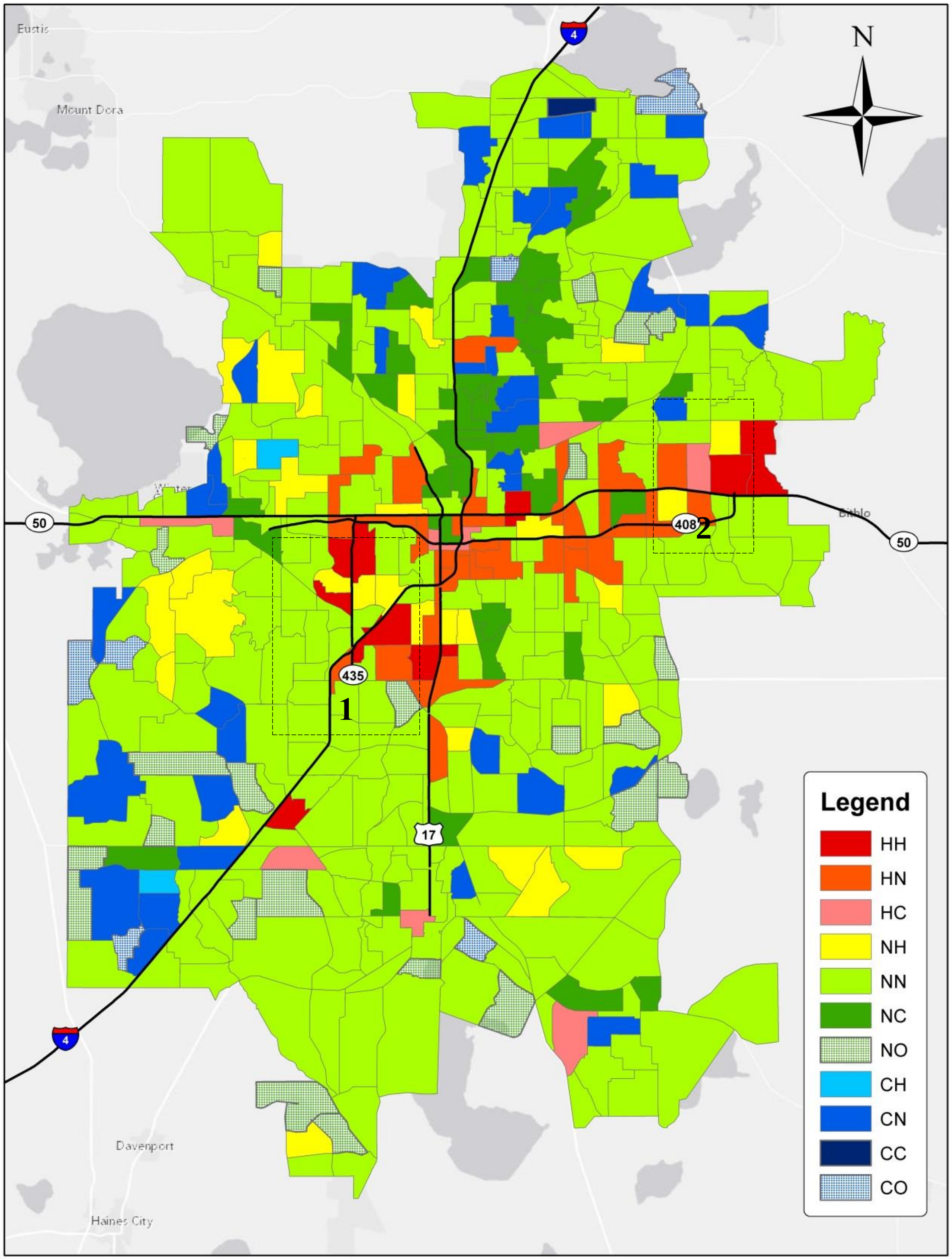


Figure 9-5 Distribution of zones by hot zone classification in urban areas (fatal-and-injury crashes)

As for rural areas, Figure 9-6 shows the spatial distribution of the TSAZs by hot zone classification. Similarly, in the total crash case the majority of HH/HN zones of FI crashes are located near main arterial roadways such as State Road 528, State Road 520, and State Road 91. Only a few HH/HN zones are close to urban areas.

One HH zone in the northwest (which was also classified as an HH zone for total crashes) has a mixed land use of residential and commercial. It was the only zone in a rural area that was classified as a CC zone for fatal-and-injury crashes; it is in a residential area. It was found that most HH/HN zones for fatal-and-injury crashes can also be categorized into HH/HN zones. This indicates that zones that are vulnerable to total crashes are also likely to see fatal-and-injury crashes. This may be because crashes occurring in rural areas tend to be more severe than those in urban areas.

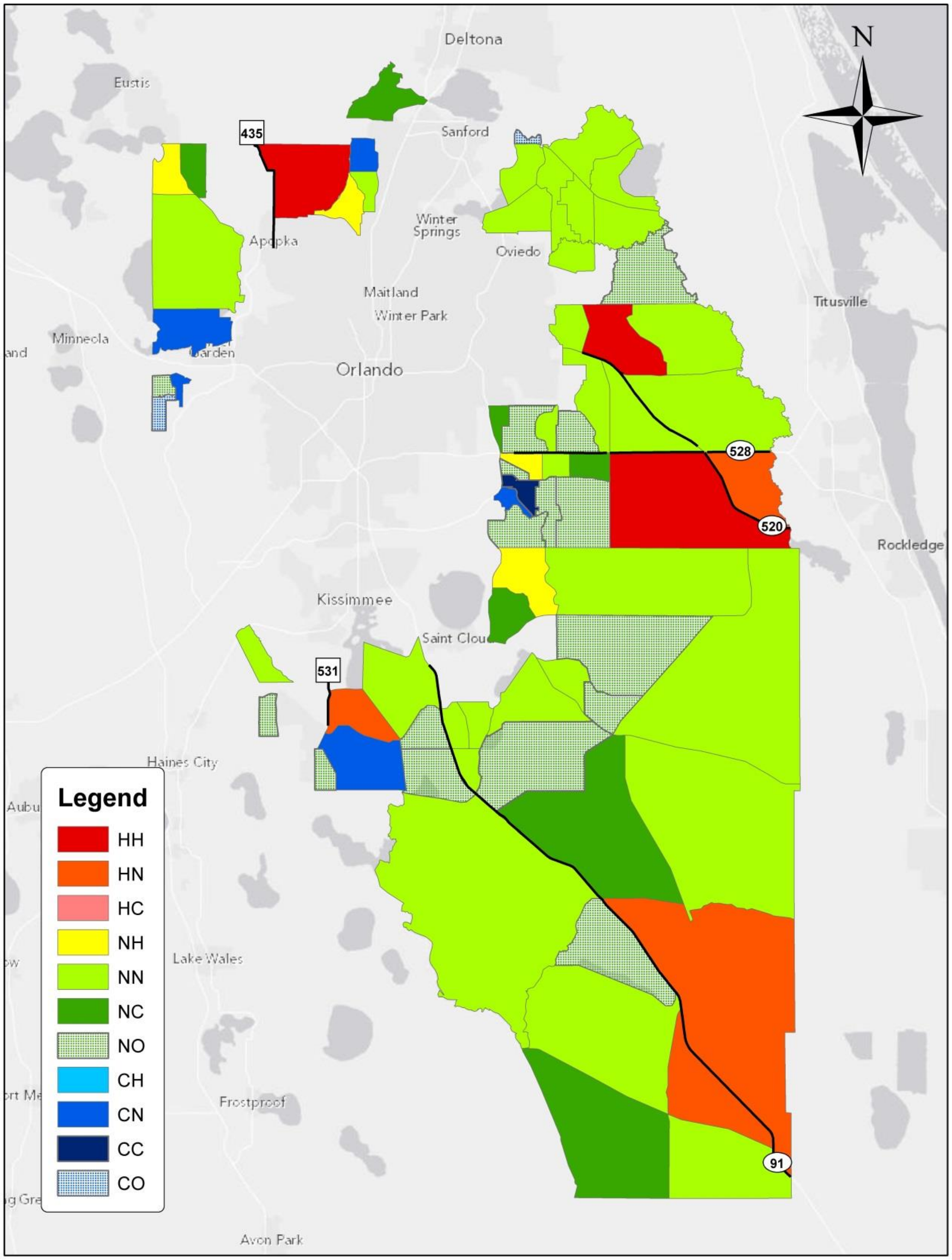


Figure 9-6 Distribution of zones by hot zone classification in rural areas (fatal-and-injury crashes)

Table 9-5 compares zonal features between the average values of all areas, HH, and CC zones for fatal-and-injury crashes. In urban areas, the 'Population density' values for both HH and CC zones are larger than average. However, 'Number of hotel, motel, and timeshare rooms per square mile' in HH zones is nearly triple of the average. In contrast, 'Number of hotel, motel, and timeshare rooms per square mile' in CC zones is only half that of the average. 'Proportion of roadways with 55 mph or higher speed limits' in HH zones is 8.0%, which is higher than the average (5.2%).

Table 9-5 Comparison of zonal features between the average, HH, and CC zones (fatal-and-injury crashes)

Zonal factors	**Urban**			**Rural**		
	Average	**HH**	**CC**	**Average**	**HH**	**CC**
Population density	410.0	733.4	1612.2	124.4	287.2	18.6
Proportion of Hispanics	0.274	0.221	0.213	0.279	0.466	0.233
Number of hotel, motel, and timeshare rooms per square mile	139.7	338.5	61.83	51.05	1.227	0.000
Proportion of roadways with 55 mph or higher speed limits	0.052	0.080	0.000	0.075	0.000	0.000

As for rural areas, the 'Population density' values in HH zones are much larger than the average, but quite a bit smaller than the average in CC zones. This implies that population-dense areas (i.e., residential areas) are more dangerous in terms of fatal-and-injury crashes in rural areas. Also, there is a significant gap in the 'Proportion of Hispanics' between HH zones and the average. Hispanics in HH zones in fatal-and-injury crashes make up 46.6% of the total population, whereas the average for Hispanics is 27.9% in all rural areas.

9.3. Summary of Integration

A novel screening methodology for integrating two levels was developed and used in this research for hotspot/hot zone determination. TSAZs were classified into twelve categories with considerations made for both macro- and micro-level results. It is recommended that different strategies for each hot zone classification be developed because each category has distinctive traffic safety risks at each of the different levels. For HH zones, both macro-level treatments (i.e., education, campaigns, enforcement, etc.) and micro-level treatments (i.e., engineering solutions) are required to improve the traffic safety of the entire area. For example, assuming that one zone has a high safety risk related to bicycle crashes at both the macro- and microscopic levels, only applying engineering treatments at the network level (i.e., adding bike lanes) might not be effective or efficient because the zone also has zonal level factors that contribute to bicycle crashes. Therefore it would be ideal to begin bicycle safety campaigns and education programs at bike facilities.

On the other hand, HN and HC zones might need a greater level of focus on macro-level treatments because no specific safety problems emerge at the microscopic level. For CH zones, applying micro-level treatments for specific hotspots could alleviate traffic risks more efficiently than other types of measures. As seen in the results of this research, no HO zones were identified by our case study. However, they might be observed in other study areas. If HO zones exist, it would mean that such zones do not have major roadways or intersections, but rather only local residential roads with high traffic crash risk. Thus, we would need to screen residential areas and provide macro-level solutions to prevent local traffic crashes (such as installing a traffic-calming zone). Admittedly, NC, NO, CC, and CO zones are not priority zones for safety treatments because they are safe for now. Nevertheless, it is necessary to keep monitoring these areas because traffic crash patterns are unstable and traffic crash risks can be transferred to these zones from other adjusted zones, especially for NC and NO zones.

10. CONCLUSION

Many studies have analyzed at the microscopic level the sites with high traffic safety risk (e.g., segments, intersections, etc.), including the HSM Part B (AASHTO, 2010). Recently, several studies have begun to focus on zonal-based network screening at the macroscopic level. Compared to microscopic safety studies, macroscopic-focused research is more efficient at integrating zonal-level features into crash prediction models and identifying hot zones. However, macroscopic screening has accuracy limitations because it cannot identify and separate hot spots from other sites within a single zone. Thus, this study developed a new integrated screening approach to overcome the above-mentioned shortcomings of current screening techniques and to achieve a balance between efforts towards accuracy and efficiency.

For conducting macro level safety analyses, the research team faced several challenges. First, using current Traffic Analysis Zones (TAZs) as basic geographic units caused a high percentage of boundary crashes because TAZs were delineated for transportation planning but not for traffic crash analysis. In order to solve this problem, the research team used regionalization to develop a new study unit: Traffic Safety Analysis Zones (TSAZs) systems. In other words, this regionalization can alleviate limitations of the TAZ system by aggregating TAZs into a sufficiently large and homogenous zonal system. The research team used the Brown-Forsythe test to select the optimal scale (500 zones as the new zone system for overall crashes) since it minimizes boundary crashes and zones without including rare types of crashes. Approximately 10% of boundary crashes have been eliminated after the regionalization but more than 60% of crashes still occur on the boundary of TSAZs. Hence, a nested structure was proposed to estimate safety performance models separately for boundary and interior crashes. This nested structure allows different contributing factors for different crash types, so this model can provide more accurate and predictable results than a single model. The six types of crashes in each model are varied based on their locations (boundary or interior) and roadways (FACR, other state roads or non-state roads). They are FSB (FACR State road Boundary crashes), FSI (FACR State road Interior crashes), OSB (Other State road Boundary crashes), OSI (Other State road Interior crashes), NSB (Non-state road Boundary crashes) and NSI (Non-state road Interior crashes). In addition, a Bayesian Poisson Lognormal Spatial Error Model (BPLSEM) was adopted for the SPF analysis in this nested structure. The BPLSEM contains a disturbance term for handling the over-dispersion problem, and its spatial error term can control for the spatial autocorrelation of crash data. In addition, the PSI (Potential for Safety Improvements), the difference between the expected crash count and the predicted crash count, was used as our measurement to define hot-zones.

As for the micro level analysis, the research team developed SPFs based on the major function classes of roads in our study area (Osceola, Seminole and Orange counties). For these segments, there are rural 2 lanes undivided, rural 2 or 4 lanes divided, urban 2 lanes divided, urban 4 lanes divided, urban 2 or 4 lanes undivided, six or more lanes interrupted roads, one way roads, and 3 lane with Two Way Left Turn Lane (TWLTL). For the intersection, there are 4 Leg Intersections and 3 Leg Intersections. Overall, these road classes are consistent with the HSM road classification. Moreover, this study includes some new roadway types that are even not presented in the HSM, such as six or more lanes interrupted roads. Because there is no existing SPF or reference group data available, a Full Bayesian model was used to estimate the PSI value for different roadway types in the study area. A Poisson log-normal model with random effect was employed in this project. For the segment, the independent variables were Average Annual Daily Traffic (AADT) and segment length. For the intersection, the model fitting procedure was similar as with the segments. The research team still used the Full Bayesian Poisson Lognormal models to predict crash frequency but tried four different variable combinations to identify the best model.

After identifying hot spot areas at both macro- and microscopic levels, the research team integrated these macroscopic and microscopic screening results. However, this integration task was challenging because we needed to (1) combine various SPFs from different scales, areas, and roadway types; (2) determine an appropriate weight for each group; and (3) choose a measurement for our final results.

In order to solve the above mentioned problems, this study then developed a new criterion to identify whether a zone has safety issues at the macro- and/or microscopic levels. All TSAZs were classified into twelve categories that include two scale groups (macro or micro) and four safety levels (hot, normal, cold, or no data). These categories are: HH, HN, HC, HO, NH, NN, NC, NO, CH, CN, CC, and CO. The first character of the classification represents the macroscopic safety risk, and the second character illustrates the microscopic safety risk.

Then, the research team defined weights for different scales and roadway types. At the macroscopic level, TSAZs were ranked by their zonal PSIs; at the microscopic level, the calculation of average PSI was more complicated because each TSAZ had several intersections and segments. The PSIs of the intersections in each TSAZ were averaged by the number of intersections, and the zones were ranked by their averaged intersection PSI. Simultaneously, the PSIs of segments in each zone were averaged by the total length of the segments in the zone, and zones were ranked by their averaged segment PSI. After that, both the intersection and segment PSI ranks were averaged; the TSAZs were ranked by the final averaged intersection and segment

PSIs. As was the case at the macroscopic level, TSAZs with top 10% micro-level PSIs were categorized as "Hot" zones at the microscopic level.

Finally, the percentile ranks of the PSIs were used in the integration (instead of the original PSIs) because the units of PSI intersections and PSI segments were different. The research team analyzed hot TSAZs for both total crashes and fatal-and-injury crashes in order to be consistent with the HSM. Moreover, by doing so the results also allowed an examination of whether there are any differences with regards to hot zone locations among various crash severity levels. The total crash hot zone screening results display the overall crash distributions within the study area, whereas the fatal-and-injury crash hot zone screening results represent the more severe crash distributions.

In summary, this study presents an integrated screening method that can be used to overcome the shortcomings of macro- and micro-level approaches. In particular, our results provide a comprehensive perspective on appropriate safety treatments by balancing the accuracy and efficiency of screening. Also, it is recommended that different strategies for each hot zone classification be developed because each category has distinctive traffic safety risks at each of the different levels. However, it should be noted that there are some limitations to this study. First, the research team only used data for three counties to estimate the SPFs. It is suggested that future research evaluate the transferability of the SPFs developed in this research for other areas in Florida. Second, the research team did not include the variance of PSIs when calculating the average PSI in each zone. As a result, even two zones could both be classified as hot zones at the microscopic level. It is possible that one zone could have consistently high PSI segments/intersections with low variances, whereas the other zone could have segments/intersections with high variances in their PSIs. In the former case, the zone would be uniformly risky at the microscopic level, so area-wide engineering treatments should be considered. In contrast, the zone in the latter case would have some extremely high risk segments/intersections but other segments/intersections would not be that dangerous. In this case, it is recommended that countermeasures be applied only for the specific sites. Also, if the highway agencies are more concerned about the crash cost, PSIs could be replaced with other hotspot identification methods such as the equivalent-property-damage-only crash frequency method. Lastly, only two types of crashes (total crashes and fatal-and-injury crashes) were analyzed in this research. It would be useful if hot zones for other various types of crashes could be identified, considering both macroscopic and microscopic levels, so practitioners could comprehensively recognize the hot zone locations of specific crash types and apply appropriate safety treatments.

This report includes also 2 Spreadsheets (one for total crashes and the other for Fatal and Injury crashes) to help practitioners to implement the methods developed in this study. Please refer to Appendix D for a short user guide to use these tools.

REFERENCES

Abbess, C., Jarret, D., and Wright, C.C., "Accidents at Blackspots: Estimating the Effectiveness of Remedial Treatment, with Special Reference to the "Regression-to-the-Mean" Effect". Traffic Engineering and Control 22(10), 1981, pp. 535-542.

Abdel-Aty, M. A., Lee, J., Siddiqui, C., and Choi, K., "Geographical Unit Based Analysis in the Context of Transportation Safety Planning", Research Part A: Policy and Practice. 49, 2013, pp. 62-75.

Abdel-Aty, M. A., Siddiqui, C., and Huang, H., 2011. "Integrating Trip and Roadway Characteristics in Managing Safety at Traffic Analysis Zones". Transportation Research Record 2213, pp. 20-28.

Aguero-Valverde, J., and Jovanis, P.P., "Bayesian Multivariate Poisson Lognormal Models for Crash Severity Modeling and Site Ranking". Transportation Research Record 2136, 2009, pp. 82-91

Aguero-Valverde, J., Jovanis, P. P., "Spatial Analysis of Fatal and Injury Crashes in Pennsylvania", Accident Analysis and Prevention, 38(3), 2006, pp. 618-625.

Akaike, H. (1974), "A New Look at the Statistical Model Identification", IEEE Transactions on Automatic Control 19 (6), pp. 716–723

American Association of State Highway and Transportation Officials (AASHTO), "Highway Safety Manual", 2010, AASHTO, Washington, D.C.

Amoros, E., Martin, J. L., and Laumon, B., "Comparison of Road Crashes Incidence and Severity between Some French Counties". Accident Analysis and Prevention, 35(4), 2003, pp. 537-547.

Baass, K. G., 1981. “Design of Zonal Systems for Aggregate Transportation Models”. Transportation Research Record 807, pp. 1-6.

Blatt, J., and Furman, S. M., “Residence location of drivers involved in fatal crashes”. Accident Analysis and Prevention, 30(6), 1998. pp. 705-711.

Chung K., and Ragland D., “A Method for Generating a Continuous Risk Profile for Highway Collisions”, UC Berkeley Traffic Safety Center, 2007.

Clark, D. E., “Effect of Population Density on Mortality after Motor Vehicle Collisions”, Accident Analysis and Prevention, 35(6), 2003, pp. 965-971.

Cottrill, C. D., and Thakuriah P. V. "Evaluating Pedestrian Crashes in Areas with High Low-income or Minority Populations." Accident Analysis & Prevention 42 (6), 2010, pp. 1718-1728.

De Guevara, F. L., Simon, P. W., and Oh, J., “Forecasting Crashes at the Planning Level: Simultaneous Negative Binomial Crash Model Applied in Tucson, Arizona”, Transportation Research Record 1897, 2004, pp. 191-199.

Gehlke, C. E., and Biehl, K., “Certain Effects of Grouping upon the Size of the Correlation Coefficient in Census Tract Material”, Journal of the American Statistical Association 29 (185), Supplement: Proceeding of the American Statistical Journal, 1934, pp. 169-170.

Guo, D., and Wang, H., “Automatic Region Building for Spatial Analysis”, Transactions in GIS, 15(s1), 2011, pp. 29-45.

Hadayeghi, A., Shalaby, A. S., and Persaud, B. N., “Macrolevel Accident Prediction Models for Evaluating Safety of Urban Transportation Systems”, In Transportation Research Board 82th Annual Meeting, Transportation Research Board of the National Academics, Washington, D.C., 2003.

Hadayeghi, A., Shalaby, A. S., and Persaud, B. N., “Development of Planning-Level Transportation Safety Models using Full Bayesian Semiparametric Additive Techniques”, Journal of Transportation Safety and Security 2 (1), 2010a, pp. 45-68.

Hadayeghi, A., Shalaby, A. S., and Persaud, B. N., “Development of Planning-Level Transportation Safety Tools using Geographically Weighted Poisson Regression”, Accident Analysis and Prevention 42, 2010b, pp. 676-688.

Hadayeghi, A., Shalaby, A. S., Persaud, B. N., and Cheung, C., “Temporal Transferability and Updating of Zonal Level Accident Prediction Models”, Accident Analysis and Prevention, 38(3), 2006, pp. 579-589.

Hauer, E. “Detection of Safety Deterioration in a Series of Accident Counts”, Transportation Research Record 1542, 1996, pp. 38-43.

Heydecker, B. J., and Wu, J., “Using the Information in Road Accident Records Proc.”, 19th PTRC Summer Annual Meeting, 1991, London.

Highway Safety Manual, AASHTO, 2010.

Huang, H., Abdel-Aty, M. A., and Darwiche, A. L., “County-level Crash Risk Analysis in Florida”, Transportation Research Record: Journal of the Transportation Research Board, 2148(1), 2010, pp. 27-37.

Kim, K., Brunner, I. M., and Yamashita, E. Y., “Influence of Land Use, Population, Employment, and Economic Activity on Accidents”, Transportation Research Record 1953, 2007, pp.56-64.

Kononov, J., and Allery, B. "Level of Service of Safety: Conceptual Blueprint and Analytical Framework." Transportation Research Record: Journal of the Transportation Research Board 1840, 2003, pp. 57-66.

Lai, P. C., So, F. M., and Chan, K. W. “Spatial Epidemiological Approaches in Disease Mapping and Analysis”. CRC Press, 2010.

LaScala, E. A., Gerber, D., and Gruenewald, P. J. “Demographic and Environmental Correlates of Pedestrian Injury Collisions: A Spatial Analysis”. Accident Analysis and Prevention 32, 2000, pp. 651-658.

Levine, N., Kim, K. E., and Nitz, L. H., “Spatial analysis of Honolulu motor vehicle crashes: II. Zonal generators”. Accident Analysis and Prevention, 27(5), 1995. pp. 675-685.

Loukaitou-Sideris, A., Liggett, R., and Sung, H-G., “Death on the Crosswalk: A Study of Pedestrian-Automobile Collisions in Los Angeles”, Journal of Education and Research, 26(3), 2007, pp. 338-351.

MacNab, Y. C., “Bayesian Spatial and Ecological Models for Small-Area Accident and Injury Analysis”, Accident Analysis and Prevention, 36(6), 2004, pp. 1019-1028.

Moran, P. A. P., “Notes on Continuous Stochastic Phenomena”, Biometrika 37(1), 1950, pp. 17-23.

Naderan, A., and Shahi, J., “Aggregate Crash Prediction Models: Introducing Crash Generation Concept”, Accident Analysis and Prevention 42, 2010, pp. 339-346.

Ng, K-S., Hung, W-T., and Wong, W-G., “An Algorithm for Assessing the Risk of Traffic Accident”, Accident Analysis and Prevention, 33(3), 2002, pp. 387-410.

Noland, R. B., and Oh, L., “The Effect of Infrastructure and Demographic Change on Traffic-related Fatalities and Crashes: A Case Study of Illinois County-Level Data” Transportation Research Record 1897, 2004, pp. 28-33.

Noland, R. B., and Quddus, M. A., "Analysis of Pedestrian and Bicycle Casualties with Regional Panel Data", Transportation Research Record 1897, 2004, pp. 28-33.

Noland, R. B., and Quddus, M. A., “A Spatially Disaggregate Analysis of Road Casualties in England”, Accident Analysis and Prevention, 36(6), 2004, pp. 973-984.

O’Sullivan, D., and Unwin, D. J., “Geographic Information Analysis”, Wiley, 2003.

Openshaw, S., “Ecological Fallacies and the Analysis of Areal Census Data”, Environment and Planning A 16(1), 1984, pp. 17-31.

Pirdavani, A., Brijs, T., Bellemans, T., Kochan, B., and Wets, G., Evaluating the Road Safety Effects of a Fuel Cost Increase Measure by Means of Zonal Crash Prediction Modeling. Accident Analysis and Prevention 50, 2013, pp. 186-195.

Quddus M. A., “Modelling Area-Wide Count Outcomes with Spatial Correlation and Heterogeneity: An Analysis of London Crash Data”, Accident Analysis and Prevention, 40(4), 2008, pp. 1486-1497.

Ragland, D.R., and Chan, C-Y, “High Collision Concentration Location: Table C Evaluation and Recommendations”, 2007, UC Berkeley Traffic Safety Center.

Romano, E. O., Tippetts, A. S., and Voas, R. B., “Language, Income, Education, and Alcohol-related Fatal Motor Vehicle Crashes”, Journal of Ethnicity in Substance Abuse, 5(2), 2006, pp. 119-137.

Root, E. D., “Moving Neighborhoods and Health Research Forward: Using Geographic Methods to Examine the Role of Spatial Scale in Neighborhood Effects on Health”, Annals of the Association of American Geographers, 102 (5), 2012, 986-995.

Root, E. D., Meyer, R. E., & Emch, M. (2011). “Socioeconomic Context and Gastroschisis: Exploring Associations at Various Geographic Scales”, Social Science & Medicine 72 (4), 2010, pp. 625-633.

Siddiqui C., and Abdel-Aty, M., "On The Nature of Modeling Boundary Pedestrian Crashes at Zone", Transportation Research Board 91st Annual Meeting, Transportation Research Board of the National Academics, Washington, D.C., 2012.

Siddiqui, C., "Macroscopic Crash Analysis and Its Implications for Transportation Safety Planning" (Doctoral Dissertation), 2012, University of Central Florida, Florida.

Siddiqui, C., Abdel-Aty, M. A., and Choi, K., "Macroscopic Spatial Analysis of Pedestrian and Bicycle Crashes", Accident Analysis and Prevention 45, 2012, pp. 382-391.

Siddiqui, C., and Abdel-Aty, M. A., "On the Nature of Modeling Boundary Pedestrian Crashes at Zones". Presented in Transportation Research Board 91st Annual Meeting, Transportation Research Board of the National Academics, Washington, D.C., 2012.

Spiegelhalter, D., Thomas, A., Best, N., and Lunn, D., "WinBUGS User Manual", MRC Biostatistics Unit, Cambridge, 2013.

Spiegelhalter, D.J., Best, N.G., Carlin, B.P., and Van der Linde, A., "Bayesian Measures of Model Complexity and Fit", Journal of the Royal Statistical Society B 64, 2003, pp. 1–34.

Ukkusuri, S., Hasan S., and Aziz, M. H. A., "Random Parameter Model Used to Explain Effects of Built-Environment Characteristics on Pedestrian Crash Frequency", Transportation Research Record 2237, 2011, pp. 98-106.

Wier, M., Weintraub, J, Humphreys, E. H., Seto, E., and Bhatia, R., "An Area-Level Model of Vehicle-Pedestrian Injury Collisions with Implications for Land Use and Transportation Planning", Accident Analysis and Prevention, 41(1), 2009, pp. 137-145.

APPENDIX A

MODELING FORMULATION FOR MACRO-LEVEL ANALYSIS

The Nested Bayesian Poisson Lognormal Spatial Error Model (NBPLSEM) was adopted for the SPF analysis because this model has a disturbance term for handling the over-dispersion problem, and its spatial error term can control for the spatial autocorrelation of crash data. The model is specified as follows:

$$y_i \sim Poisson(\mu_i) \quad \text{(A-1)}$$

$$\lambda_i = \exp(\beta_0 + \beta X_i + \theta_i + \varphi_i) \quad \text{(A-2)}$$

$$\theta_i = Normal\ (0, \tau_\theta) \quad \text{(A-3)}$$

where

y_i: aggregated total (or fatal-and-injury) crashes of the i^{th} TSAZ,

β_0: intercept,

βs are the coefficient estimates of covariates (X_i),

θ_i is the random effect term,

φ_i is the spatial effect term, and

τ_θ is the precision parameter, which is the inverse of the variance and a given prior gamma distribution (0.5, 0.005).

The Bayesian model is fit with vague prior distributions, Normal (0, 10^{-6}) for βs. The spatial effect term (φ_i) is included to account for the heterogeneity caused by the spatial correlation. The spatial pattern uses an intrinsic Gaussian Conditional Autoregressive (CAR) with a prior distribution, $Normal\ (0, \tau_\varphi)$. The mean of φ_i is defined by:

$$\bar{\varphi}_i = {\sum_{i \neq j} \varphi_j \times w_{ij}} \Big/ {\sum_{i \neq j} w_{ij}} \quad \text{(A-4)}$$

where

w_{ij} = 1, if zones i and j are adjacent based on the 1st order contiguity, otherwise w_{ij} = 0.

In order to evaluate the contribution of spatial autocorrelations in the error component, the apportionments of spatial variability in the error due to the spatial autocorrelation in each sub-model are calculated using the following formula:

$$\Psi = \text{s.d. of } \varphi_i / (\text{s.d. of } \varphi_i + \text{s.d. of } \theta_i) \quad \text{(A-5)}$$

where
s.d. is standard deviation,
θ_i is the random effect term, and

φ_i is the spatial effect term.

Thus, a high Ψ value means that crashes are affected by the spatial effect rather than the random effect.

APPENDIX B

MODELING FORMULATION FOR MICRO-LEVEL ANALYSIS

A Poisson log-normal model with random effects was employed for each of the crash types (total crashes and fatal-and-injury crashes). The regression model is derived from the Poisson model by assuming that the same intersections share one error term over two years. The framework of the regression model is expressed as follows:

$$y_{it} = Poisson\ (\lambda_{it})$$

$$\lambda_{it} = Exp\ (\chi'_{it}\ \beta + \ \varepsilon_{it}) = \ \mu_{it} * Exp(\varepsilon_{it}) \qquad \text{(B.1)}$$

$$\varepsilon_{it} = Normal\ (0.0, \tau^2)$$

$$\tau = Gamma\ (0.5, 0.005)$$

where *y*_is the observed crash frequency of site *i* in time period *t*,

λ is the mean predicted crash frequency for site *i* in time period *t*,

x is a vector of the independent variables, including the log(AADT) on major roads (AADT_mj) and the log(AADT) on minor roads (AADT_mn),

β is a vector of the coefficients for each independent variable and the intercept term, and

δ^2 is the variance of the normal distribution for ε.

APPENDIX C

THE PSIS FOR ALL TSAZ

Table C-1 Ranking TSAZs using PSIs (urban areas)

Rank	Rank percentile	Total crash		Fatal-and-injury crash	
		TSAZ ID	PSI	TSAZ ID	PSI
1	0.2%	56	1127.880	202	334.644
2	0.5%	15	971.440	8	272.738
3	0.7%	202	791.730	196	255.596
4	0.9%	8	651.180	2	234.250
5	1.2%	9	648.000	56	233.255
6	1.4%	196	625.459	15	204.740
7	1.6%	192	620.349	89	188.698
8	1.9%	89	595.207	207	179.557
9	2.1%	69	549.530	5	178.469
10	2.3%	104	510.150	43	175.275
11	2.6%	382	498.175	69	171.608
12	2.8%	130	492.320	3	156.202
13	3.0%	224	470.914	12	151.874
14	3.3%	0	433.720	192	150.154
15	3.5%	92	429.485	67	138.363
16	3.7%	67	428.796	62	137.494
17	4.0%	62	413.550	130	134.979
18	4.2%	6	411.870	18	133.330
19	4.4%	43	402.370	104	131.018
20	4.7%	66	385.870	9	129.910
21	4.9%	146	384.350	0	125.090
22	5.1%	178	381.803	66	124.134
23	5.4%	18	376.160	58	118.026
24	5.6%	42	361.726	101	111.759
25	5.8%	212	354.540	65	111.366
26	6.1%	195	350.338	93	110.636
27	6.3%	29	345.127	212	110.133
28	6.5%	35	330.897	16	109.178
29	6.8%	180	327.315	180	104.254
30	7.0%	19	318.380	86	96.362
31	7.2%	207	318.163	57	96.124
32	7.5%	60	302.947	6	94.408
33	7.7%	14	293.278	224	94.395
34	7.9%	2	287.020	38	91.203
35	8.2%	28	280.342	105	87.838
36	8.4%	57	268.780	382	87.799
37	8.6%	3	257.644	195	86.882
38	8.9%	250	253.911	250	82.205
39	9.1%	98	252.000	19	80.906
40	9.3%	5	250.610	233	79.481
41	9.6%	38	248.007	345	79.408

42	9.8%	22	247.428	42	78.342
43	10.0%	93	235.027	333	76.341
44	10.3%	295	232.700	161	75.700
45	10.5%	155	231.279	91	75.388
46	10.7%	65	228.712	172	75.240
47	11.0%	218	222.240	146	74.688
48	11.2%	292	220.560	200	74.208
49	11.4%	172	219.390	121	73.262
50	11.7%	121	217.267	160	72.618
51	11.9%	52	216.758	151	72.402
52	12.1%	16	213.589	52	72.197
53	12.4%	399	211.038	78	68.422
54	12.6%	12	197.873	208	67.546
55	12.9%	80	195.390	32	66.875
56	13.1%	50	194.917	178	65.854
57	13.3%	99	194.685	92	65.710
58	13.6%	45	192.949	292	63.797
59	13.8%	107	190.951	117	63.661
60	14.0%	333	190.680	305	63.121
61	14.3%	77	189.101	11	62.572
62	14.5%	78	186.910	50	62.409
63	14.7%	144	185.992	280	61.649
64	15.0%	7	183.210	17	61.047
65	15.2%	74	180.490	126	58.846
66	15.4%	297	180.270	26	57.957
67	15.7%	345	180.270	20	57.839
68	15.9%	375	178.819	29	57.059
69	16.1%	160	178.285	22	56.378
70	16.4%	20	176.797	165	55.250
71	16.6%	405	174.015	375	54.958
72	16.8%	105	166.435	137	54.901
73	17.1%	13	164.186	73	54.421
74	17.3%	94	163.361	94	53.906
75	17.5%	117	162.874	295	51.866
76	17.8%	197	160.625	218	51.827
77	18.0%	126	159.799	187	51.611
78	18.2%	305	159.440	107	51.294
79	18.5%	249	157.915	149	50.325
80	18.7%	404	155.897	271	48.850
81	18.9%	90	154.480	98	48.289
82	19.2%	182	152.978	229	47.793
83	19.4%	111	152.543	276	47.685
84	19.6%	359	148.271	405	47.543
85	19.9%	280	148.107	341	46.986
86	20.1%	177	143.234	156	46.744
87	20.3%	129	139.367	68	46.516
88	20.6%	230	138.655	37	46.407
89	20.8%	203	135.401	125	46.067
90	21.0%	161	135.080	46	45.860

91	21.3%	58	134.929	7	45.587
92	21.5%	101	134.060	129	45.214
93	21.7%	374	130.267	75	44.929
94	22.0%	364	129.627	154	44.620
95	22.2%	23	127.527	322	44.429
96	22.4%	103	127.174	339	44.011
97	22.7%	142	126.915	205	43.977
98	22.9%	366	122.990	188	43.876
99	23.1%	322	119.860	297	43.777
100	23.4%	200	118.584	302	43.685
101	23.6%	229	118.537	404	43.663
102	23.8%	302	113.629	74	43.130
103	24.1%	233	112.898	366	42.956
104	24.3%	339	111.555	256	42.537
105	24.5%	32	110.656	111	41.469
106	24.8%	86	110.280	21	41.382
107	25.0%	100	109.111	289	40.889
108	25.2%	346	107.389	182	40.425
109	25.5%	30	106.505	142	39.762
110	25.7%	187	105.500	364	39.592
111	25.9%	298	105.421	103	39.127
112	26.2%	68	105.025	463	38.790
113	26.4%	175	103.346	213	38.681
114	26.6%	438	101.793	127	38.580
115	26.9%	26	100.800	13	38.123
116	27.1%	226	99.346	230	36.904
117	27.3%	72	98.984	110	36.741
118	27.6%	54	98.622	48	36.681
119	27.8%	289	98.484	25	36.657
120	28.0%	357	92.948	359	36.557
121	28.3%	256	91.000	155	36.039
122	28.5%	271	89.975	45	35.828
123	28.7%	179	88.970	112	35.766
124	29.0%	350	85.277	60	35.615
125	29.2%	213	84.685	141	35.598
126	29.4%	153	84.680	374	35.346
127	29.7%	194	84.092	261	35.042
128	29.9%	391	82.887	88	34.977
129	30.1%	227	82.188	227	34.947
130	30.4%	201	79.026	395	34.901
131	30.6%	381	77.428	132	34.840
132	30.8%	53	76.521	399	34.252
133	31.1%	276	76.044	47	33.814
134	31.3%	395	74.510	222	33.026
135	31.5%	467	70.820	385	32.695
136	31.8%	266	70.665	242	32.563
137	32.0%	88	70.528	90	31.427
138	32.2%	96	69.675	173	31.324
139	32.5%	398	68.654	304	31.240

140	32.7%	136	68.362	35	30.965
141	32.9%	110	67.958	71	30.779
142	33.2%	304	66.850	72	30.050
143	33.4%	51	65.910	83	28.975
144	33.6%	171	65.150	40	28.575
145	33.9%	164	63.417	53	28.526
146	34.1%	27	62.901	122	28.358
147	34.3%	154	62.871	14	28.239
148	34.6%	46	62.630	313	27.617
149	34.8%	132	62.010	438	27.508
150	35.0%	21	61.071	36	27.465
151	35.3%	267	60.941	97	27.448
152	35.5%	120	60.787	376	27.007
153	35.7%	231	58.249	85	26.913
154	36.0%	208	57.530	398	26.824
155	36.2%	173	56.418	153	26.585
156	36.4%	455	56.310	194	26.046
157	36.7%	445	56.220	162	26.034
158	36.9%	351	55.015	119	25.847
159	37.1%	165	54.815	197	25.828
160	37.4%	162	54.799	350	25.718
161	37.6%	264	54.745	163	25.688
162	37.9%	369	52.810	123	25.642
163	38.1%	336	52.325	168	25.263
164	38.3%	373	52.188	175	24.748
165	38.6%	83	50.854	80	24.727
166	38.8%	138	50.647	381	24.633
167	39.0%	294	49.688	113	24.500
168	39.3%	122	49.290	55	24.306
169	39.5%	39	46.524	63	23.957
170	39.7%	198	46.049	368	23.947
171	40.0%	184	45.300	278	23.698
172	40.2%	463	45.211	77	23.623
173	40.4%	137	44.319	177	23.394
174	40.7%	272	44.010	264	23.109
175	40.9%	75	43.659	102	23.016
176	41.1%	84	42.914	144	22.826
177	41.4%	24	42.237	30	22.298
178	41.6%	474	42.179	174	22.084
179	41.8%	157	41.642	100	21.771
180	42.1%	352	40.490	346	21.731
181	42.3%	341	40.133	23	21.614
182	42.5%	106	40.086	99	21.065
183	42.8%	273	38.380	171	21.051
184	43.0%	220	37.860	136	20.981
185	43.2%	152	37.500	120	20.573
186	43.5%	331	36.653	148	20.406
187	43.7%	452	36.023	252	20.118
188	43.9%	91	35.993	203	20.100

189	44.2%	343	35.253	336	19.972
190	44.4%	174	33.750	54	19.713
191	44.6%	321	33.301	28	19.591
192	44.9%	376	33.078	184	18.850
193	45.1%	306	32.962	284	18.630
194	45.3%	219	32.726	84	18.185
195	45.6%	284	32.719	467	18.160
196	45.8%	55	31.890	429	18.151
197	46.0%	246	31.580	201	18.077
198	46.3%	139	31.530	139	17.838
199	46.5%	211	31.437	414	17.801
200	46.7%	261	31.138	358	17.471
201	47.0%	411	30.720	124	17.318
202	47.2%	428	29.902	266	16.909
203	47.4%	167	29.880	179	16.747
204	47.7%	156	29.551	369	16.700
205	47.9%	470	29.460	231	16.282
206	48.1%	31	29.390	198	15.941
207	48.4%	47	28.850	186	15.940
208	48.6%	396	28.440	455	15.770
209	48.8%	326	28.229	428	15.282
210	49.1%	344	25.680	211	15.243
211	49.3%	414	25.625	263	14.804
212	49.5%	400	25.170	357	14.739
213	49.8%	270	24.530	226	14.699
214	50.0%	260	24.350	199	14.555
215	50.2%	252	24.303	109	14.320
216	50.5%	113	24.232	24	13.939
217	50.7%	119	24.154	167	13.836
218	50.9%	296	23.439	157	13.694
219	51.2%	205	23.300	270	13.499
220	51.4%	393	23.150	267	13.208
221	51.6%	349	22.243	330	13.175
222	51.9%	403	21.726	243	12.884
223	52.1%	222	21.567	176	12.799
224	52.3%	293	21.439	403	12.634
225	52.6%	259	20.765	294	12.539
226	52.8%	311	20.660	251	12.489
227	53.0%	199	20.630	27	12.303
228	53.3%	263	20.416	249	12.237
229	53.5%	151	20.087	298	12.140
230	53.7%	116	20.007	351	12.054
231	54.0%	123	19.750	319	11.949
232	54.2%	378	19.263	303	11.859
233	54.4%	465	18.866	452	11.822
234	54.7%	432	18.835	349	11.479
235	54.9%	245	18.687	445	11.362
236	55.1%	97	18.487	372	11.194
237	55.4%	189	18.408	209	10.978

238	55.6%	59	18.050	389	10.836
239	55.8%	303	17.600	79	10.781
240	56.1%	209	17.442	418	10.766
241	56.3%	265	17.293	220	10.433
242	56.5%	429	16.825	245	10.324
243	56.8%	176	16.600	343	10.278
244	57.0%	392	16.117	460	10.226
245	57.2%	163	14.100	311	10.220
246	57.5%	131	13.900	411	10.164
247	57.7%	316	13.840	474	10.100
248	57.9%	389	13.590	246	9.925
249	58.2%	191	12.621	299	9.722
250	58.4%	494	12.484	138	9.665
251	58.6%	325	12.370	492	9.630
252	58.9%	17	11.719	210	9.466
253	59.1%	490	11.014	391	9.179
254	59.3%	149	10.500	51	9.135
255	59.6%	308	10.031	321	9.110
256	59.8%	372	9.817	279	9.107
257	60.0%	275	9.750	44	8.931
258	60.3%	358	9.670	432	8.813
259	60.5%	282	8.869	326	8.661
260	60.7%	228	8.712	96	8.329
261	61.0%	456	8.360	494	8.312
262	61.2%	254	8.348	259	8.191
263	61.4%	423	8.189	308	8.159
264	61.7%	186	7.750	434	7.832
265	61.9%	387	7.250	215	7.726
266	62.1%	124	7.247	254	7.349
267	62.4%	383	6.705	393	7.327
268	62.6%	498	6.240	317	7.220
269	62.9%	112	6.057	116	7.151
270	63.1%	317	5.921	240	6.988
271	63.3%	466	5.552	159	6.592
272	63.6%	63	5.501	312	6.528
273	63.8%	61	4.535	325	6.509
274	64.0%	477	4.107	296	6.478
275	64.3%	193	4.084	378	6.444
276	64.5%	159	3.929	272	6.381
277	64.7%	329	3.926	352	6.370
278	65.0%	114	3.821	293	6.071
279	65.2%	319	3.608	387	5.888
280	65.4%	133	3.038	164	5.747
281	65.7%	268	2.610	400	5.742
282	65.9%	158	2.453	396	5.712
283	66.1%	115	2.011	373	5.479
284	66.4%	11	1.845	490	5.381
285	66.6%	433	1.701	316	5.318
286	66.8%	225	1.555	189	5.289

287	67.1%	460	1.544	106	5.217
288	67.3%	299	1.357	108	5.115
289	67.5%	188	1.245	31	4.741
290	67.8%	288	0.923	291	4.715
291	68.0%	356	0.861	265	4.684
292	68.2%	269	0.450	446	4.677
293	68.5%	215	0.407	487	4.664
294	68.7%	425	0.230	81	4.457
295	68.9%	243	0.071	61	4.431
296	69.2%	87	0.000	344	4.214
297	69.4%	236	0.000	275	4.061
298	69.6%	48	-0.045	409	4.007
299	69.9%	440	-0.223	360	3.911
300	70.1%	320	-0.403	300	3.880
301	70.3%	206	-0.836	301	3.643
302	70.6%	441	-1.539	260	3.464
303	70.8%	473	-1.596	219	3.461
304	71.0%	479	-1.662	431	3.413
305	71.3%	279	-1.664	59	3.357
306	71.5%	338	-1.729	310	3.340
307	71.7%	420	-1.869	150	3.328
308	72.0%	379	-2.187	441	3.288
309	72.2%	190	-2.256	76	3.286
310	72.4%	443	-2.760	314	3.274
311	72.7%	409	-2.787	10	3.240
312	72.9%	487	-3.422	443	3.098
313	73.1%	310	-3.568	118	3.013
314	73.4%	413	-3.686	204	2.894
315	73.6%	489	-3.833	268	2.714
316	73.8%	488	-3.885	423	2.634
317	74.1%	472	-4.164	371	2.451
318	74.3%	145	-4.207	425	2.213
319	74.5%	397	-4.370	413	2.200
320	74.8%	214	-4.560	143	2.136
321	75.0%	290	-4.690	465	2.089
322	75.2%	223	-4.891	140	2.040
323	75.5%	248	-5.049	41	2.031
324	75.7%	439	-5.329	448	1.968
325	75.9%	431	-5.470	477	1.967
326	76.2%	278	-5.630	320	1.751
327	76.4%	386	-5.647	440	1.707
328	76.6%	287	-5.680	158	1.693
329	76.9%	459	-5.994	228	1.655
330	77.1%	324	-6.202	383	1.597
331	77.3%	73	-6.555	221	1.544
332	77.6%	421	-6.567	449	1.467
333	77.8%	457	-6.917	473	1.408
334	78.0%	408	-7.228	356	1.272
335	78.3%	448	-7.280	386	1.229

336	78.5%	434	-7.464	324	1.203
337	78.7%	307	-7.739	379	1.158
338	79.0%	370	-8.436	34	1.031
339	79.2%	85	-9.510	181	1.007
340	79.4%	4	-9.822	288	0.906
341	79.7%	327	-9.827	470	0.855
342	79.9%	486	-9.837	479	0.781
343	80.1%	242	-10.613	327	0.763
344	80.4%	328	-11.279	459	0.716
345	80.6%	426	-13.123	420	0.620
346	80.8%	150	-13.337	115	0.601
347	81.1%	301	-13.490	488	0.495
348	81.3%	368	-13.526	408	0.359
349	81.5%	412	-13.722	329	0.350
350	81.8%	424	-14.125	419	0.273
351	82.0%	79	-14.212	466	0.248
352	82.2%	217	-14.650	217	0.231
353	82.5%	418	-14.875	334	0.230
354	82.7%	300	-15.015	426	0.153
355	82.9%	143	-15.580	306	0.150
356	83.2%	334	-15.857	206	0.139
357	83.4%	102	-15.865	338	0.126
358	83.6%	232	-15.996	451	0.047
359	83.9%	25	-17.000	87	0.000
360	84.1%	312	-17.421	236	0.000
361	84.3%	437	-18.060	133	-0.058
362	84.6%	436	-18.830	39	-0.100
363	84.8%	170	-19.663	95	-0.130
364	85.0%	147	-20.085	331	-0.161
365	85.3%	385	-20.252	4	-0.192
366	85.5%	314	-20.709	397	-0.295
367	85.7%	109	-21.035	273	-0.410
368	86.0%	141	-21.476	269	-0.449
369	86.2%	168	-21.529	476	-0.476
370	86.4%	1	-22.423	135	-0.547
371	86.7%	384	-22.423	433	-0.582
372	86.9%	451	-22.942	498	-0.685
373	87.1%	34	-23.115	214	-0.834
374	87.4%	221	-23.287	145	-0.934
375	87.6%	410	-24.832	82	-0.951
376	87.9%	183	-25.012	223	-0.961
377	88.1%	360	-25.112	384	-1.001
378	88.3%	108	-25.209	421	-1.078
379	88.6%	291	-25.539	457	-1.117
380	88.8%	118	-25.594	392	-1.130
381	89.0%	491	-26.241	274	-1.202
382	89.3%	204	-26.392	486	-1.209
383	89.5%	81	-26.394	472	-1.211
384	89.7%	449	-26.407	489	-1.250

385	90.0%	181	-26.987	328	-1.258
386	90.2%	274	-27.750	290	-1.318
387	90.4%	240	-28.410	248	-1.387
388	90.7%	185	-28.578	232	-1.388
389	90.9%	255	-28.862	437	-1.542
390	91.1%	446	-30.016	439	-1.620
391	91.4%	419	-30.080	491	-1.653
392	91.6%	323	-31.631	424	-1.700
393	91.8%	64	-31.790	262	-1.817
394	92.1%	390	-32.200	185	-1.847
395	92.3%	82	-32.218	412	-2.266
396	92.5%	262	-34.044	456	-2.701
397	92.8%	371	-37.527	436	-2.730
398	93.0%	342	-37.990	114	-2.973
399	93.2%	365	-39.920	282	-3.052
400	93.5%	95	-40.146	191	-3.351
401	93.7%	36	-40.930	370	-3.388
402	93.9%	33	-42.030	1	-3.495
403	94.2%	492	-48.260	225	-3.789
404	94.4%	210	-49.187	152	-3.875
405	94.6%	447	-50.465	447	-4.128
406	94.9%	340	-51.667	253	-4.295
407	95.1%	71	-52.129	410	-4.400
408	95.3%	430	-52.138	499	-4.622
409	95.6%	76	-52.681	307	-4.630
410	95.8%	499	-53.144	390	-4.889
411	96.0%	253	-56.292	147	-4.947
412	96.3%	125	-58.332	323	-5.370
413	96.5%	127	-63.242	170	-5.400
414	96.7%	148	-65.267	131	-6.233
415	97.0%	313	-70.233	193	-6.760
416	97.2%	234	-71.196	255	-6.805
417	97.4%	330	-74.121	235	-7.296
418	97.7%	476	-76.092	234	-7.690
419	97.9%	285	-76.288	287	-8.502
420	98.1%	37	-79.092	33	-8.588
421	98.4%	235	-79.688	285	-8.646
422	98.6%	135	-82.130	365	-9.912
423	98.8%	10	-88.070	190	-10.026
424	99.1%	41	-95.304	183	-11.077
425	99.3%	140	-111.370	430	-11.298
426	99.5%	251	-132.326	342	-11.989
427	99.8%	40	-134.222	340	-16.493
428	100.0%	44	-147.783	64	-16.772

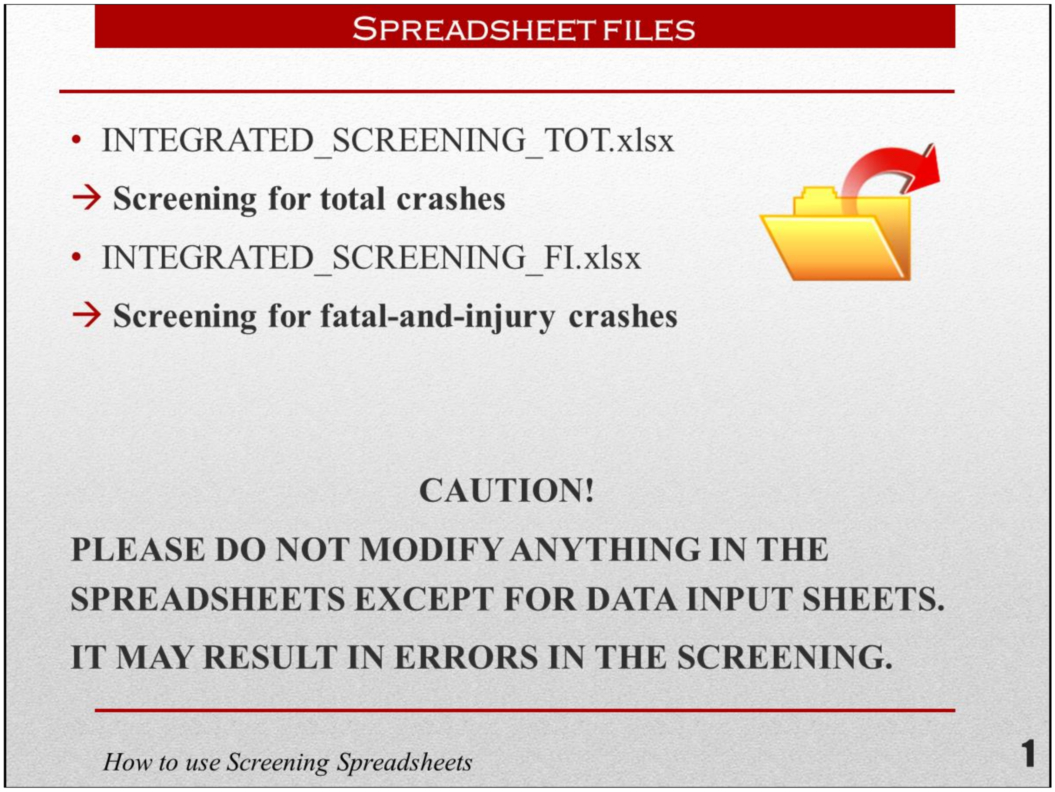
Table C-2 Ranking TSAZs using PSIs (rural areas)

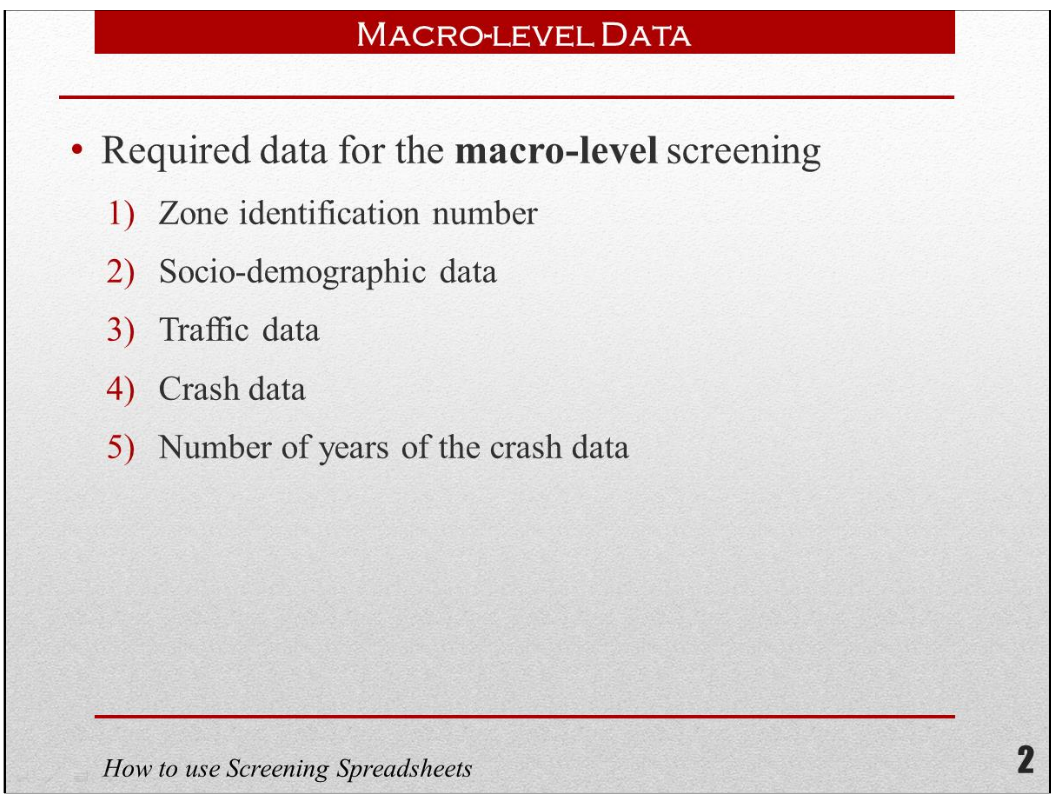
Rank	Rank percentile	Total crash		Fatal-and-injury crash	
		TSAZ ID	PSI	TSAZ ID	PSI
1	1.4%	367	215.548	367	79.229
2	2.8%	337	152.669	337	70.096
3	4.2%	347	145.548	347	51.083
4	5.6%	406	130.475	281	48.928
5	6.9%	281	118.346	406	45.225
6	8.3%	49	103.374	464	31.660
7	9.7%	361	70.069	49	31.319
8	11.1%	247	61.156	394	26.761
9	12.5%	464	47.260	348	23.264
10	13.9%	257	43.886	166	22.284
11	15.3%	318	41.339	361	21.781
12	16.7%	239	34.702	435	21.245
13	18.1%	444	31.853	362	20.253
14	19.4%	497	30.445	238	18.150
15	20.8%	286	28.953	239	15.822
16	22.2%	493	28.502	484	15.514
17	23.6%	475	24.063	332	15.476
18	25.0%	380	23.717	480	15.034
19	26.4%	353	22.464	444	14.607
20	27.8%	332	22.377	247	13.488
21	29.2%	216	19.430	497	12.882
22	30.6%	415	19.150	475	12.396
23	31.9%	496	16.836	416	12.044
24	33.3%	394	15.992	493	10.148
25	34.7%	442	12.340	415	9.658
26	36.1%	238	12.186	258	9.347
27	37.5%	416	11.967	277	9.286
28	38.9%	480	11.695	353	8.122
29	40.3%	435	10.548	496	7.991
30	41.7%	468	10.068	169	7.587
31	43.1%	485	9.824	442	6.624
32	44.4%	355	9.699	257	6.145
33	45.8%	277	8.756	468	5.885
34	47.2%	481	7.505	354	5.636
35	48.6%	362	6.924	309	5.493
36	50.0%	484	6.658	407	5.354
37	51.4%	407	6.079	216	5.145
38	52.8%	169	5.188	485	4.900
39	54.2%	417	5.001	244	4.758
40	55.6%	241	3.252	286	4.266
41	56.9%	363	2.031	417	3.587
42	58.3%	478	1.891	462	3.510
43	59.7%	244	1.455	478	3.247
44	61.1%	458	1.038	481	3.119
45	62.5%	335	1.034	128	1.977

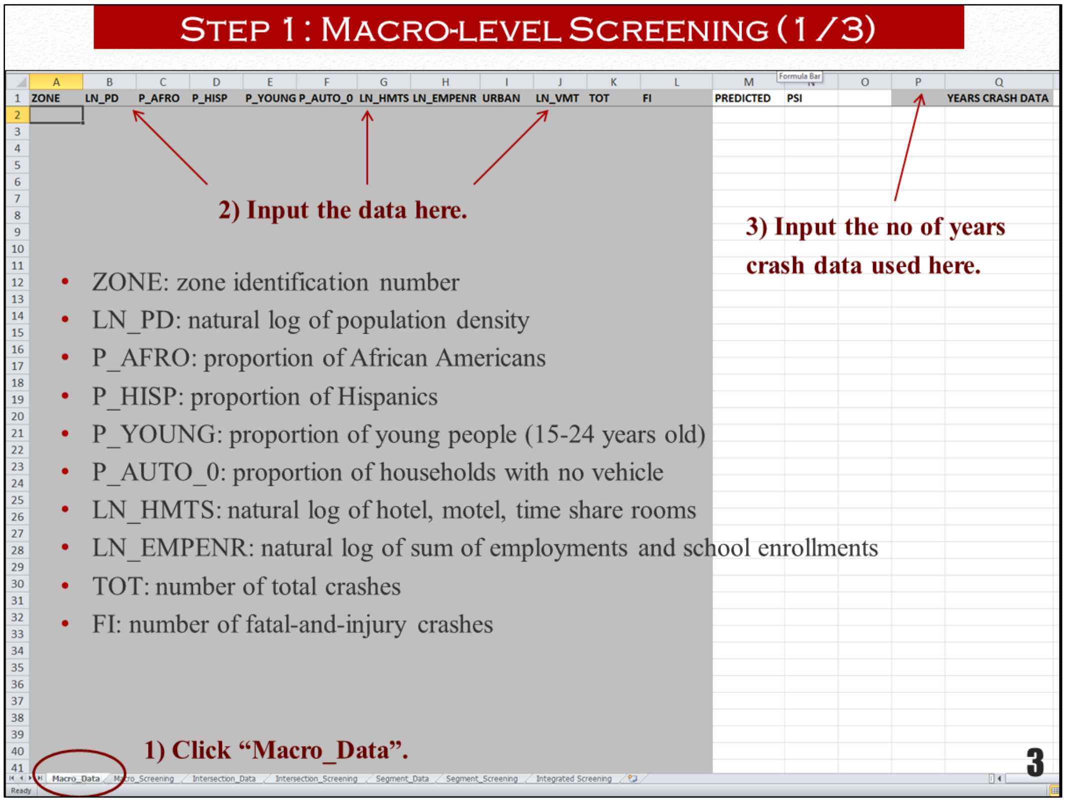
46	63.9%	401	0.544	471	1.858
47	65.3%	70	0.000	377	1.636
48	66.7%	237	0.000	363	1.526
49	68.1%	495	-0.948	241	1.119
50	69.4%	453	-0.976	453	1.064
51	70.8%	462	-1.352	318	1.016
52	72.2%	483	-1.784	458	0.870
53	73.6%	422	-2.248	483	0.658
54	75.0%	482	-3.845	335	0.498
55	76.4%	471	-4.491	401	0.483
56	77.8%	454	-4.863	355	0.452
57	79.2%	128	-5.889	422	0.295
58	80.6%	402	-6.167	380	0.281
59	81.9%	166	-7.550	402	0.069
60	83.3%	388	-8.736	70	0.000
61	84.7%	469	-10.557	237	0.000
62	86.1%	450	-10.641	482	-0.145
63	87.5%	283	-11.850	454	-0.504
64	88.9%	354	-13.836	450	-0.844
65	90.3%	348	-15.137	283	-1.249
66	91.7%	461	-15.552	461	-1.279
67	93.1%	134	-19.530	315	-1.369
68	94.4%	258	-21.533	388	-1.713
69	95.8%	377	-24.869	495	-2.063
70	97.2%	309	-29.978	469	-2.585
71	98.6%	315	-32.029	427	-7.765
72	100.0%	427	-48.191	134	-14.103

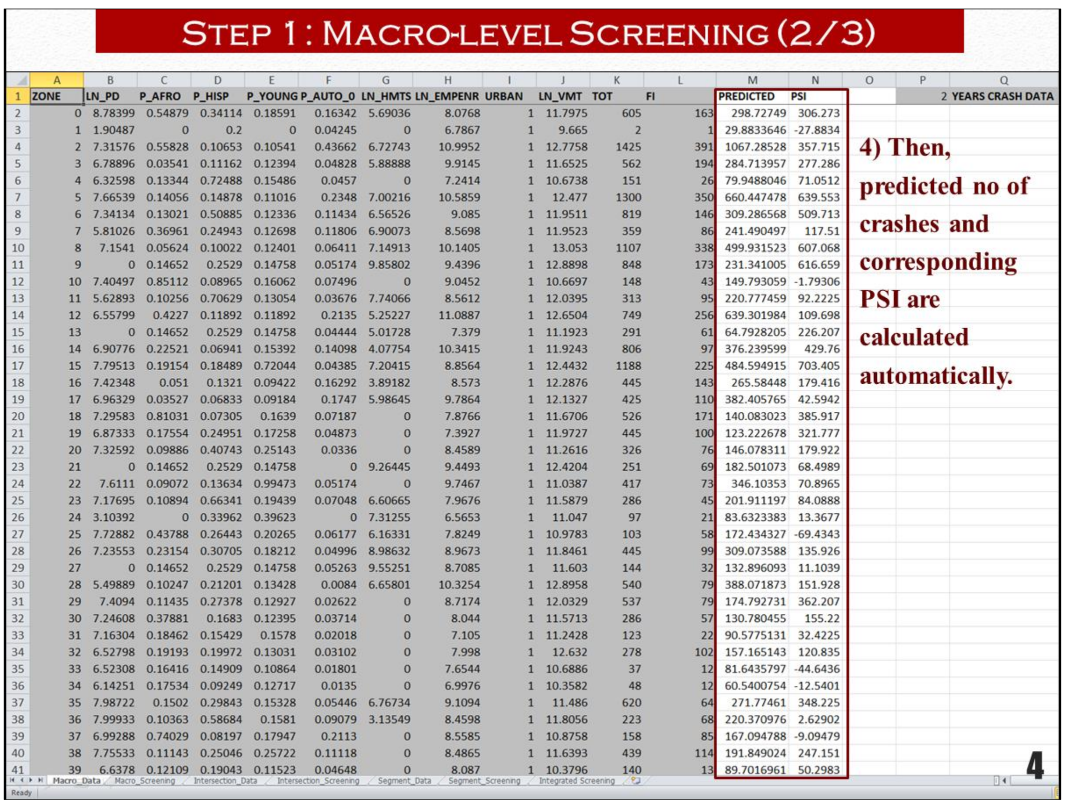
APPENDIX D

A GUIDE TO SCREENING SPREADSHEET



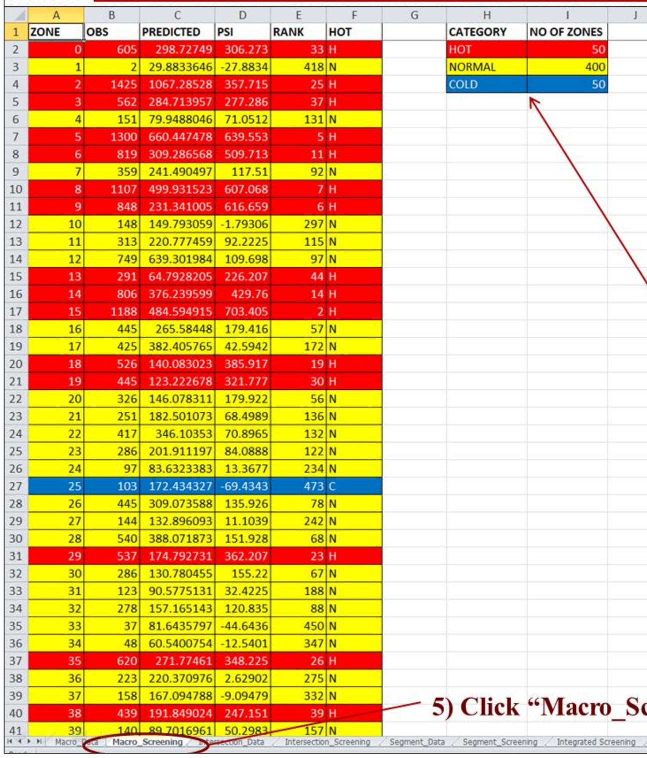






ZONE	LN_PD	P_AFRO	P_HISP	P_YOUNG	P_AUTO_0	LN_HMTS	LN_EMPENR	URBAN	LN_VMT	TOT	FI	PREDICTED	PSI
0	8.78399	0.54879	0.34114	0.18591	0.16342	5.69036	8.0768	1	11.7975	605	163	298.72749	306.273
1	1.90487	0	0.2	0	0.04245	0	6.7867	1	9.665	2	1	29.8833646	-27.8834
2	7.31576	0.55828	0.10653	0.10541	0.43662	6.72743	10.9952	1	12.7758	1425	391	1067.28528	357.715
3	6.78896	0.03541	0.11162	0.12394	0.04828	5.88888	9.9145	1	11.6525	562	194	284.713957	277.286
4	6.32598	0.13344	0.72488	0.15486	0.0457	0	7.2414	1	10.6738	151	26	79.9488046	71.0512
5	7.66539	0.14056	0.14878	0.11016	0.2348	7.00216	10.5859	1	12.477	1300	350	660.447478	639.553
6	7.34134	0.13021	0.50885	0.12336	0.11434	6.56526	9.085	1	11.9511	819	146	309.286568	509.713
7	5.81026	0.36961	0.24943	0.12698	0.11806	6.90073	8.5698	1	11.9523	359	86	241.490497	117.51
8	7.1541	0.05624	0.10022	0.12401	0.06411	7.14913	10.1405	1	13.053	1107	338	499.931523	607.068
9	0	0.14652	0.2529	0.14758	0.05174	9.85802	9.4396	1	12.8898	848	173	231.341005	616.659
10	7.40497	0.85112	0.08965	0.16062	0.07496	0	9.0452	1	10.6697	148	43	149.793059	-1.79306
11	5.62893	0.10256	0.70629	0.13054	0.03676	7.74066	8.5612	1	12.0395	313	95	220.777459	92.2225
12	6.55799	0.4227	0.11892	0.11892	0.2135	5.25227	11.0887	1	12.6504	749	256	639.301984	109.698
13	0	0.14652	0.2529	0.14758	0.04444	5.01728	7.379	1	11.1923	291	61	64.7928205	226.207
14	6.90776	0.22521	0.06941	0.15392	0.14098	4.07754	10.3415	1	11.9243	806	97	376.239599	429.76
15	7.79513	0.19154	0.18489	0.72044	0.04385	7.20415	8.8564	1	12.4432	1188	225	484.594915	703.405
16	7.42348	0.051	0.1321	0.09422	0.16292	3.89182	8.573	1	12.2876	445	143	265.58448	179.416
17	6.96329	0.03527	0.06833	0.09184	0.1747	5.98645	9.7864	1	12.1327	425	110	382.405765	42.5942
18	7.29583	0.81031	0.07305	0.1639	0.07187	0	7.8766	1	11.6706	526	171	140.083023	385.917
19	6.87333	0.17554	0.24951	0.17258	0.04873	0	7.3927	1	11.9727	445	100	123.222678	321.777
20	7.32592	0.09886	0.40743	0.25143	0.0336	0	8.4589	1	11.2616	326	76	146.078311	179.922
21	0	0.14652	0.2529	0.14758	0	9.26445	9.4493	1	12.4204	251	69	182.501073	68.4989
22	7.6111	0.09072	0.13634	0.99473	0.05174	0	9.7467	1	11.0387	417	73	346.10353	70.8965
23	7.17695	0.10894	0.66341	0.19439	0.07048	6.60665	7.9676	1	11.5879	286	45	201.911197	84.0888
24	3.10392	0	0.33962	0.39623	0	7.31255	6.5653	1	11.047	97	21	83.6323383	13.3677
25	7.72882	0.43788	0.26443	0.20265	0.06177	6.16331	7.8249	1	10.9783	103	58	172.434327	-69.4343
26	7.23553	0.23154	0.30705	0.18212	0.04996	8.98632	8.9673	1	11.8461	445	99	309.073588	135.926
27	0	0.14652	0.2529	0.14758	0.05263	9.55251	8.7085	1	11.603	144	32	132.896093	11.1039
28	5.49889	0.10247	0.21201	0.13428	0.0084	6.65801	10.3254	1	12.8958	540	79	388.071873	151.928
29	7.4094	0.11435	0.27378	0.12927	0.02622	0	8.7174	1	12.0329	537	79	174.792731	362.207
30	7.24608	0.37881	0.1683	0.12395	0.03714	0	8.044	1	11.5713	286	57	130.780455	155.22
31	7.16304	0.18462	0.15429	0.1578	0.02018	0	7.105	1	11.2428	123	22	90.5775131	32.4225
32	6.52798	0.19193	0.19972	0.13031	0.03102	0	7.998	1	12.632	278	102	157.165143	120.835
33	6.52308	0.16416	0.14909	0.10864	0.01801	0	7.6544	1	10.6886	37	12	81.6435797	-44.6436
34	6.14251	0.17534	0.09249	0.12717	0.0135	0	6.9976	1	10.3582	48	12	60.5400754	-12.5401
35	7.98722	0.1502	0.29843	0.15328	0.05446	6.76734	9.1094	1	11.486	620	64	271.77461	348.225
36	7.99933	0.10363	0.58684	0.1581	0.09079	3.13549	8.4598	1	11.8056	223	68	220.370976	2.62902
37	6.99288	0.74029	0.08197	0.17947	0.2113	0	8.5585	1	10.8758	158	85	167.094788	-9.09479
38	7.75533	0.11143	0.25046	0.25722	0.11118	0	8.4865	1	11.6393	439	114	191.849024	247.151
39	6.6378	0.12109	0.19043	0.11523	0.04648	0	8.087	1	10.3796	140	13	89.7016961	50.2983

Step 1 : Macro-level Screening (3/3)



ZONE	OBS	PREDICTED	PSI	RANK	HOT
0	605	298.72749	306.273	33	H
1	2	29.8833646	-27.8834	418	N
2	1425	1067.28528	357.715	25	H
3	562	284.713957	277.286	37	H
4	151	79.9488046	71.0512	131	N
5	1300	660.447478	639.553	5	H
6	819	309.286568	509.713	11	H
7	359	241.490497	117.51	92	N
8	1107	499.931523	607.068	7	H
9	848	231.341005	616.659	6	H
10	148	149.793059	-1.79306	297	N
11	313	220.777459	92.2225	115	N
12	749	639.301984	109.698	97	N
13	291	64.7928205	226.207	44	H
14	806	376.239599	429.76	14	H
15	1188	484.594915	703.405	2	H
16	445	265.58448	179.416	57	N
17	425	382.405765	42.5942	172	N
18	526	140.083023	385.917	19	H
19	445	123.222678	321.777	30	H
20	326	146.078311	179.922	56	N
21	251	182.501073	68.4989	136	N
22	417	346.10353	70.8965	132	N
23	286	201.911197	84.0888	122	N
24	97	83.6323383	13.3677	234	N
25	103	172.434327	-69.4343	473	C
26	445	309.073588	135.926	78	N
27	144	132.896093	11.1039	242	N
28	540	388.071873	151.928	68	N
29	537	174.792731	362.207	23	H
30	286	130.780455	155.22	67	N
31	123	90.5775131	32.4225	188	N
32	278	157.165143	120.835	88	N
33	37	81.6435797	-44.6436	450	N
34	48	60.5400754	-12.5401	347	N
35	620	271.77461	348.225	26	H
36	223	220.370976	2.62902	275	N
37	158	167.094788	-9.09479	332	N
38	439	191.849024	247.151	39	H
39	140	89.7016961	50.2983	157	N

CATEGORY	NO OF ZONES
HOT	50
NORMAL	400
COLD	50

6) All zones are classifed as Hot (top 10% PSI), Cold (bottom 10% PSI), and Normal (other) zones.

7) Number of zones by the hot zone category is summarized here.

5) Click "Macro_Screening".

5

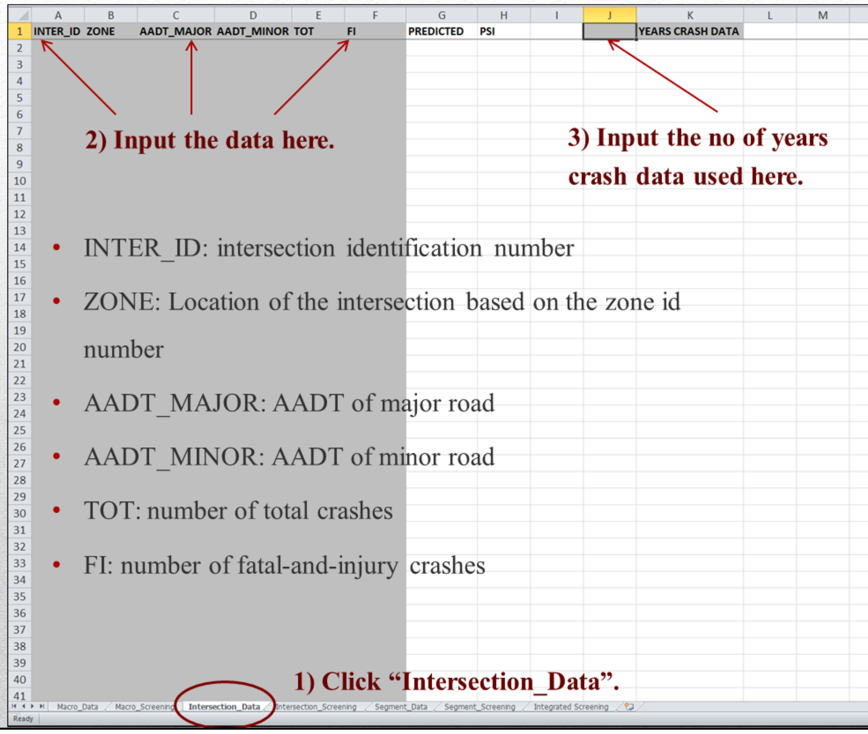
Intersection Data

- Required data for the **intersection** screening
  1) Intersection identification number
  2) Location of the intersection based on the zone identification number.
  3) Traffic data: AADT from major & minor roads
  4) Crash data
  5) Number of years of the crash data

How to use Screening Spreadsheets

6

STEP 2: INTERSECTION SCREENING (1/3)



- INTER_ID: intersection identification number
- ZONE: Location of the intersection based on the zone id number
- AADT_MAJOR: AADT of major road
- AADT_MINOR: AADT of minor road
- TOT: number of total crashes
- FI: number of fatal-and-injury crashes

7

STEP 2: INTERSECTION SCREENING (2/3)

	INTER_ID	ZONE	AADT_MAJOR	AADT_MINOR	TOT	FI	PREDICTED	PSI
2	0	22	108500	107400	121	17	120.373098	0.6269
3	1	152	70500	11000	52	9	30.1353528	21.8646
4	2	8	74500	31200	20	6	49.8671605	-29.8672
5	3	385	134000	104750	62	16	144.433525	-82.4335
6	4	8	88000	37300	85	28	62.7602152	22.2398
7	5	270	73250	35800	43	9	52.1262863	-9.12629
8	6	97	15150	11000	16	6	7.38348046	8.61652
9	7	57	81750	72350	107	31	78.2518802	28.7481
10	8	135	19800	18400	6	1	11.7956058	-5.79561
11	10	69	77500	28546	56	16	49.7409513	6.25905
12	11	6	79500	4800	43	6	23.4555528	19.5444
13	12	60	83243	17000	48	6	42.3899109	5.61009
14	14	14	62750	58750	41	4	56.1189188	-15.1189
15	15	368	69000	19800	29	12	38.1504721	-9.15047
16	17	47	33550	21500	23	7	20.4459006	2.5541
17	18	196	89000	63350	106	17	79.8308198	26.1692
18	19	322	75000	43200	51	12	57.7973088	-6.79731
19	21	233	86500	8800	23	4	32.9761193	-9.97612
20	22	43	92000	17030	28	5	46.4868402	-18.4868
21	23	188	81500	5000	26	13	24.4243166	1.57568
22	25	2	68500	27340	59	16	43.6040444	15.396
23	26	163	63500	50600	64	15	53.1662037	10.8338
24	27	110	117500	27800	41	10	71.9505642	-30.9506
25	29	149	77500	75750	25	1	76.024959	-51.025
26	32	93	95250	46200	80	30	74.0511765	5.94882
27	33	38	106000	19280	61	9	55.8505681	5.14943
28	34	89	106000	10400	43	9	42.7066949	0.29331
29	36	161	63500	9000	24	11	25.098666	-1.09867
30	37	405	50000	19600	11	5	28.2901905	-17.2902
31	38	58	32600	22280	11	0	20.226613	-9.22661
32	39	57	79500	16000	23	9	39.5859493	-16.5859
33	40	18	76750	68700	145	31	72.2188529	72.7811
34	41	8	23200	7108	20	3	9.01801093	10.982
35	42	20	100750	99250	156	36	108.690429	47.3096
36	44	25	116000	47600	119	36	89.8379348	29.1621
37	45	71	47500	12700	15	3	22.3531852	-7.35319
38	46	85	19800	6000	1	1	7.24713402	-6.24713
39	47	2	55300	31500	22	10	38.1268188	-16.1268
40	48	151	152000	20600	19	10	79.9309004	-60.9309
41	49	151	125000	30000	3	3	78.7032185	-75.7032

2 YEARS CRASH DATA

Macro_Data Macro_Screening Intersection_Data Intersection_Screening Segment_Data Segment_Screening Integrated Screening

4) Then, predicted no of crashes and corresponding PSI are calculated automatically.

8

Step 2: Intersection Screening (3/3)

	INTER_ID	OBS	PREDICTED	PSI	RANK	HOT	ZONE
2	0	121	120.3731	0.6269	104	N	22
3	1	52	30.13535	21.8646	45	N	152
4	2	20	49.86716	-29.8672	229	C	8
5	3	62	144.4335	-82.4335	249	C	385
6	4	85	62.76022	22.2398	43	N	8
7	5	43	52.12629	-9.12629	159	N	270
8	6	16	7.38348	8.61652	75	N	97
9	7	107	78.25188	28.7481	34	N	57
10	8	6	11.79561	-5.79561	138	N	135
11	10	56	49.74095	6.25905	82	N	69
12	11	43	23.45555	19.5444	50	N	6
13	12	48	42.38991	5.61009	85	N	60
14	14	41	56.11892	-15.1189	196	N	14
15	15	29	38.15047	-9.15047	160	N	368
16	17	23	20.4459	2.5541	95	N	47
17	18	106	79.83082	26.1692	37	N	196
18	19	51	57.79731	-6.79731	144	N	322
19	21	23	32.97612	-9.97612	167	N	233
20	22	28	46.48684	-18.4868	207	N	43
21	23	26	24.42432	1.57568	99	N	188
22	25	59	43.60404	15.396	60	N	2
23	26	64	53.1662	10.8338	68	N	163
24	27	41	71.95056	-30.9506	231	C	110
25	29	25	76.02496	-51.025	244	C	149
26	32	80	74.05118	5.94882	83	N	93
27	33	61	55.85057	5.14943	87	N	38
28	34	43	42.70669	0.29331	107	N	89
29	36	24	25.09867	-1.09867	114	N	161
30	37	11	28.29019	-17.2902	205	N	405
31	38	11	20.22661	-9.22661	161	N	58
32	39	23	39.58595	-16.5859	202	N	57
33	40	145	72.21885	72.7811	5	H	18
34	41	20	9.018011	10.982	67	N	8
35	42	156	108.6904	47.3096	12	H	20
36	44	119	89.83793	29.1621	32	N	25
37	45	15	22.35319	-7.35319	149	N	71
38	46	1	7.247134	-6.24713	141	N	85
39	47	22	38.12682	-16.1268	198	N	2
40	48	19	79.9309	-60.9309	246	C	151
41	49	3	78.70322	-75.7032	248	C	151

CATEGORY	NO OF ZONES
HOT	25
NORMAL	200
COLD	25

6) All intersections are classifed as Hot (top 10% PSI), Cold (bottom 10% PSI), and Normal (other) intersections.

7) Number of intersections by the hotspot category is summarized here.

5) Click "Intersection_Screening".

9

Segment Data

- Required data for the **segment** screening
  1) Segment identification number
  2) Location of the segment based on the zone identification number.
  3) Traffic data: AADT
  4) Segment length
  5) Crash data
  6) Number of years of the crash data

How to use Screening Spreadsheets

10

STEP 3: SEGMENT SCREENING (1/3)

SEG_ID	ZONE	AADT	LENGTH	TOT	FI	PREDICTED	PSI		YEARS CRASH DATA

1) Click "Segment_Data".

2) Input the data here.

3) Input the no of years crash data used here.

- SEG_ID: segment identification number
- ZONE: Location of the segment based on the zone id number
- AADT: AADT of the segment
- LENGTH: Length of the segment
- TOT: number of total crashes
- FI: number of fatal-and-injury crashes

Macro_Data | Macro_Screening | Intersection_Data | Intersection_Screening | Segment_Data | Segment_Screening | Integrated Screening

11

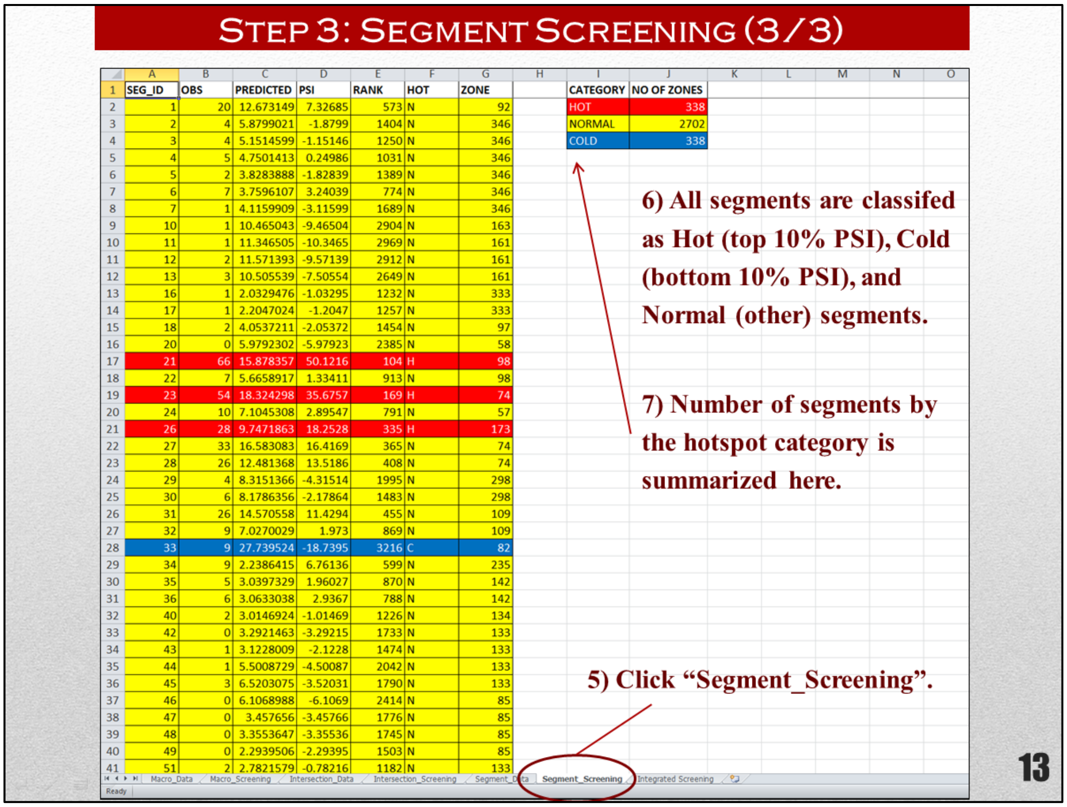
STEP 3: SEGMENT SCREENING (2/3)

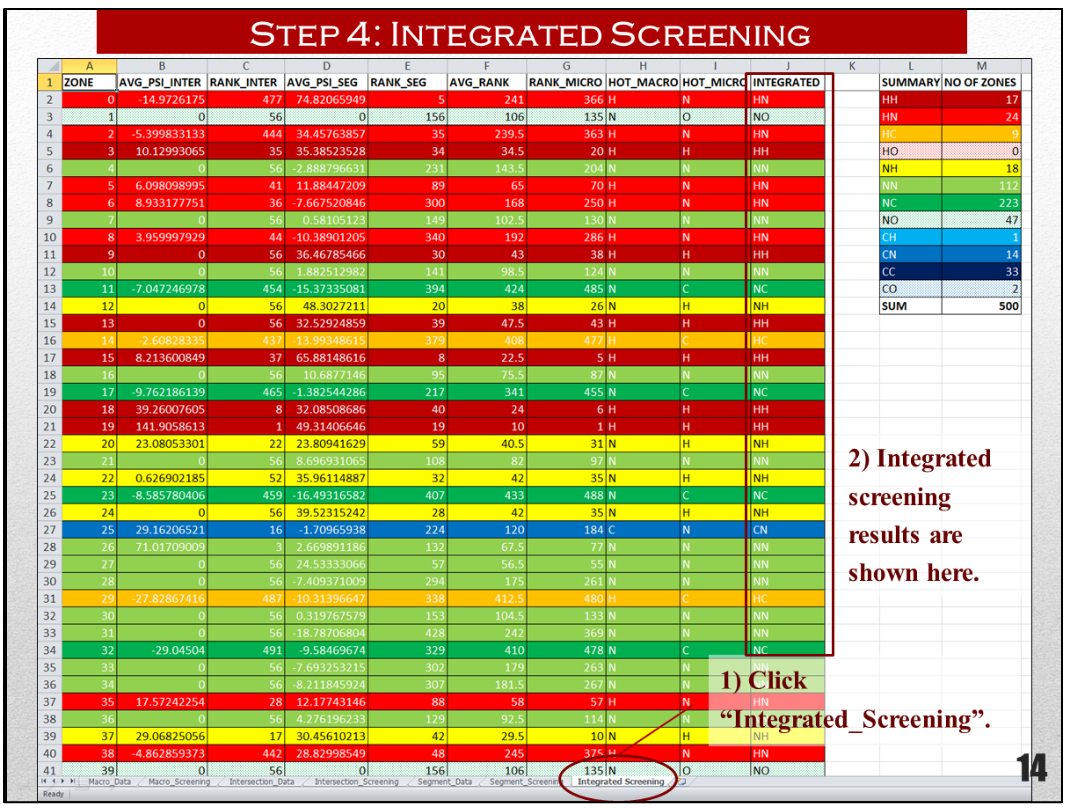
SEG_ID	ZONE	AADT	LENGTH	TOT	FI	PREDICTED	PSI
1	92	20800	0.3939	20	11	12.67315	7.32685
2	346	3900	0.6943	4	3	5.879902	-1.8799
3	346	3900	0.5265	4	1	5.15146	-1.15146
4	346	3900	0.4236	5	1	4.750141	0.24986
5	346	3900	0.1499	2	0	3.828389	-1.82839
6	346	3900	0.1269	7	5	3.759611	3.24039
7	346	3900	0.2418	1	0	4.115991	-3.11599
10	163	22000	0.1083	1	0	10.46504	-9.46504
11	161	22000	0.2109	1	0	11.34651	-10.3465
12	161	22000	0.2358	2	0	11.57139	-9.57139
13	161	22000	0.1132	3	1	10.50554	-7.50554
16	333	1400	0.127	1	0	2.032948	-1.03295
17	333	1400	0.2299	1	0	2.204702	-1.2047
18	97	3200	0.3731	2	1	4.053721	-2.05372
20	58	7500	0.2176	0	0	5.97923	-5.97923
21	98	6700	1.5426	66	22	15.87836	50.1216
22	98	6700	0.2352	7	5	5.665892	1.33411
23	74	8300	1.5613	54	22	18.3243	35.6757
24	57	8300	0.3592	10	4	7.104531	2.89547
26	173	14500	0.3356	28	15	9.747186	18.2528
27	74	14500	1.0098	33	13	16.58308	16.4169
28	74	14500	0.6493	26	7	12.48137	13.5186
29	298	14500	0.134	4	2	8.315137	-4.31514
30	298	14500	0.113	6	3	8.178636	-2.17864
31	109	6600	1.445	26	12	14.57056	11.4294
32	109	6600	0.5198	9	7	7.027003	1.973
33	82	1750	3.2727	9	7	27.73952	-18.7395
34	235	1600	0.1476	9	0	2.238642	6.76136
35	142	1600	0.5357	5	3	3.039733	1.96027
36	142	1600	0.5455	6	2	3.063304	2.9367
40	134	2600	0.1555	2	0	3.014692	-1.01469
42	133	2600	0.2672	0	0	3.292146	-3.29215
43	133	2600	0.2002	1	1	3.122801	-2.1228
44	133	6700	0.1977	1	1	5.500873	-4.50087
45	133	6700	0.4134	3	1	6.520307	-3.52031
46	85	8100	0.1858	0	0	6.106899	-6.1069
47	85	2800	0.273	0	0	3.457656	-3.45766
48	85	2800	0.2349	0	0	3.355365	-3.35536
49	85	1700	0.1324	0	0	2.293951	-2.29395
51	133	1700	0.3772	2	1	2.782158	-0.78216

2 YEARS CRASH DATA

4) Then, predicted no of crashes and corresponding PSI are calculated automatically.

12


Step 3: Segment Screening (3/3)
CATEGORY | NO OF ZONES
HOT 338
NORMAL 2702
COLD 338
6) All segments are classifed as Hot (top 10% PSI), Cold (bottom 10% PSI), and Normal (other) segments.
7) Number of segments by the hotspot category is summarized here.
5) Click "Segment_Screening".
13


Step 4: Integrated Screening
SUMMARY | NO OF ZONES
HH 17
HN 24
HC 9
HO 0
NH 18
NN 112
NC 223
NO 47
CH 1
CN 14
CC 33
CO 2
SUM 500
2) Integrated screening results are shown here.
1) Click "Integrated_Screening".
14