An Investigation of Crash Proximity to the Residential Locations of Crash Occupants

Center for Accessibility and Safety for an Aging Population
Florida State University
In Partnership with Florida A&M University and University of North Florida
2525 Pottsdamer Street, Rm. 129, Tallahassee, FL 32310-6046
Phone: (850) 410-6153 Fax: (850) 410-6142
http://utc.fsu.edu/

Mehmet Baran Ulak[1], Ayberk Kocatepe[1], Eren Erman Ozguven[1], Mark W. Horner[2]

[1] Department of Civil and Environmental Engineering, Florida A&M University - Florida State University College of Engineering

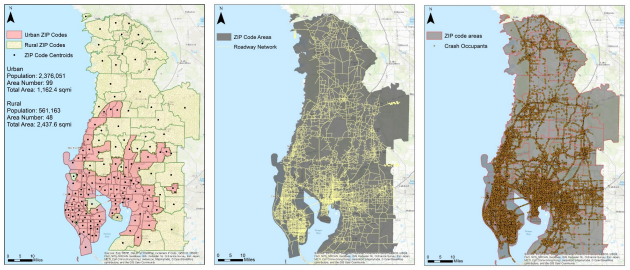
[2] Department of Geography, Florida State University

Motivation

This study investigates the proximity of crashes to the residential locations of the crash occupants through a statistical analysis-based approach. To this end, two years of crash data was disaggregated by the crash occupants' ZIP codes, for a study area in Southwest Florida. This is done in order to calculate the roadway network distances between their residential ZIP code area centroids (origins) and crash spots (destinations). These distances are then used to create separate O-D vectors allowing for several different groups to be analyzed: different non-motorist types (e.g. pedestrians, cyclists), rural vs. urban ZIP codes, different levels of crash severity, DUI involvement, and different age group occupants. Finally, the best-fitting statistical distributions were identified for each group to assess the proximity of crash spots to the residences of crash occupants. Results indicate clear differences in crash involvement among the groups with respect to varying urban densities, people's ages and modes of travel. Such findings can help in the development of more accurate crash prediction methods, as most current approaches only implement variables associated with traffic and roadway geometry.

Demographics & Available Data

- Study area is Tampa Bay region of Florida, also identified as District 7 by the Florida Department of Transportation (FDOT), with more than 3 million inhabitants.
- Areas with a population density higher than 800 (people/mi2) were considered as urban ZIP codes.
- Crash data includes 270,003 crash occupants and was obtained from FDOT for 2013 and 2014.
- It is found that only 7% of crash occupants living in District 7 involve in crashes outside of District 7 boundaries.



ZIP Code Area Type	Number of Areas	Total Population	Average Population	Total Area (mi^2)	Average Area (mi^2)	Average Population Density (Population/mi^2)
All	147	2,937,214	19,981	3,600.0	24.5	2,157.9
Urban	99	2,376,051	24,001	1,162.4	11.7	3,056.0
Rural	48	561,163	11,690	2,437.6	50.8	305.4

Objectives

- **To understand the spatial nature behind residence-crash spot relationship:**

Where and how far away, in Florida, do people live from the crashes they are involved in? What are the statistical properties of the proximity of crash spots to the residence locations of crash occupants?

- **Prospective input to crash prediction:**

To establish a foundation for future applications which can help in developing methodologies that can integrate population into crash frequency prediction methods based on the distances between crash occupant residences and their crash locations.

Methodology

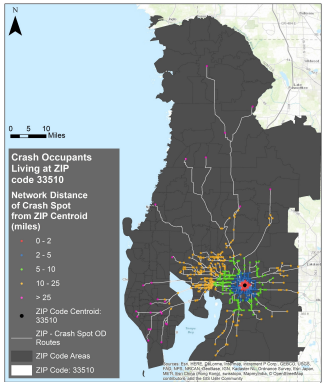
The crash data includes several attributes and these **attributes were used to disaggregate the dataset** in order to conduct the crash spot–residence location proximity analysis **for following analysis groups: (a)** *rural and urban ZIP codes*, **(b)** *different non-motorist types*, **(c)** *different levels of crash severity, and DUI involvement*, and **(d)** *different age group crash occupants*. Then, the analysis was conducted for to determine **crash spot–residence location proximity** for all types of crashes. Three main components were identified in this analysis in order **to obtain the O-D distance vector**:

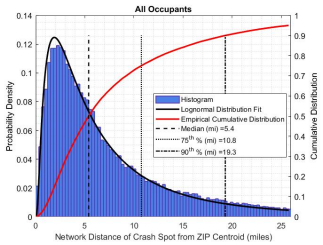
- The ZIP code centroids (origins),
- Crash locations (destinations),
- The roadway network.

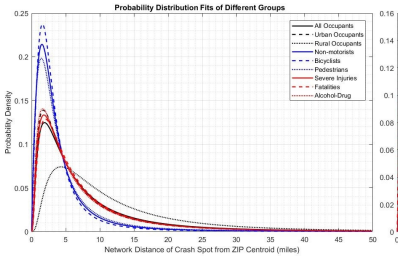
After obtaining O-D distance vectors, **descriptive statistical analyses** were conducted including: plotting histograms and cumulative distributions. Following this analysis, empirical cumulative distribution plots were provided.

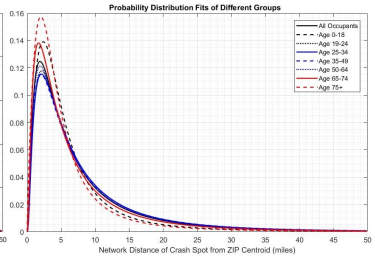
Then **the best-fitting statistical distributions** for each analyzed group were identified from the following distributions: Exponential, Gamma, Generalized Pareto, Inverse Gaussian, Log-normal, Log-logistic, Birnbaum-Saunders, Nakagami and Weibull. Note that, **Log-normal and Log-logistic distributions** were found to **be the best fitting distributions** for all analysis groups.

Results









Crash Occupants	*Obs.*	*Crash/Pop.*	*mean*	*std. dev.*	*median*	*25th %ile*	*75th %ile*	*90th %ile*	*95th %ile*
All	231,834	7.9 %	8.4	8.7	5.4	2.7	10.8	19.3	25.8
Urban	193,336	8.1 %	7.5	7.8	4.8	2.5	9.5	17.5	23.4
Rural	38,498	6.9 %	12.8	11.5	9.3	5.1	16.7	27.0	35.4
Non-motorists	4,441	0.15 %	4.9	6.4	2.9	1.6	5.3	10.6	16.2
Bicyclists	2,338	0.08 %	4.9	6.4	2.9	1.6	5.3	8.3	12.1
Pedestrians	1,880	0.06 %	5.4	6.9	3.1	1.6	6.0	12.6	19.9
Severe Injury	26,162	0.89 %	7.8	8.5	5.0	2.6	9.8	18.1	24.6
Fatality	620	0.02 %	8.1	9.3	5.3	2.7	9.5	18.1	25.1
DUI	4,943	0.17 %	7.6	8.5	4.8	2.4	9.4	17.9	24.7
Age Groups									
<18	41,426	6.8 %	6.8	7.2	4.5	2.4	8.5	15.2	20.8
19-24	33,886	11.9 %	8.7	8.9	5.7	2.9	11.2	19.9	26.1
25-34	42,531	12.1 %	9.0	8.9	5.9	2.9	11.9	20.7	26.7
35-49	49,263	8.3 %	9.1	9.1	5.9	2.9	12.1	21.0	27.4
50-64	39,087	6.6 %	8.7	9.2	5.5	2.7	11.3	20.2	27.1
65-74	13,466	4.9 %	7.8	8.9	4.8	2.4	9.6	18.5	25.6
>75	9,204	3.7 %	6.2	7.3	4.0	2.2	7.2	13.6	19.6

Conclusions

- Results indicate **clear differences in crash involvement** for many of the analysis groups with respect to varying urban densities, people's ages and modes of transportation.
- The **spatial crash distribution** of individuals living in a region can be identified through the proposed evaluation of distances between the crash location and crash occupant residences, it may be **possible to determine the effect of that population on the crash occurrences on nearby roadways**.
- The information gained in this study can **help in developing methodologies that can integrate population into crash frequency prediction methods** based on the distances between crash occupant residences and their crash locations.
- For example, **the statistical distribution parameters** for urban residents **can be helpful to refine the crash prediction** for urban roadways using statistical distribution-based decay functions. The distance-based population weight at every roadway section in a region can be used to develop more accurate crash prediction methods and functions (e.g. safety performance functions).
- As **a limitation, the centroid of a ZIP code** might be further away or closer to a crash spot than the crash occupant's **actual residence** (house) is. However, residence information is difficult to obtain due to privacy concerns and therefore not readily available in crash data, which **compels use of Zip codes**.

For information please contact: mbul4@my.fsu.edu