Patterning Demographic and Socioeconomic Characteristics Affecting Pedestrian and Bicycle Crash Frequency

***Deo Chimba, P.E., PTOE**
Assistant Professor
Civil Engineering Department
Tennessee State University
3500 John A. Merritt Blvd, Nashville, TN 37209
Phone: 615-963-5430
Email: dchimba@tnstate.edu

Daniel Emaasit
Graduate Research Assistant
Civil Engineering Department
Tennessee State University

Christopher R. Cherry
Associate Professor
Department of Civil and Environmental Engineering
University of Tennessee

Zane Pannell
Graduate Research Assistant
Department of Civil and Environmental Engineering
University of Tennessee

***Corresponding Author**

Prepared for Publication Consideration and Presentation at the 93rd Annual meeting of the Transportation Research Board, Washington, D.C. January, 2014

3568 + 250 x 3 Tables + 250x4 Figure = 5318 Words
Abstract = 216 Words

ABSTRACT

The objective of this study was to investigate factors influencing occurrence of pedestrian and bicycle crashes in Tennessee. Of interest were demographic and socio-economic, roadway geometry, traffic, and land use factors that could influence pedestrian crash rates on specific infrastructure. Geographic Information System (GIS) and statistical modeling were applied to study the crash patterns with respect to these factors. GIS was used to geo-locate and cluster the crash locations onto the roadway network, joined with background data of the crash locations. Negative Binomial (NB) regression was used to model the relationship between contributing factors and the crashes to detect any positive or negative correlations with the crashes. The following factors were found to have significant correlation with pedestrian and bicycle crash occurrences; percentage distribution of population by race, age groups, mean household income, percentage in the labor force, poverty level, and vehicle ownership. Land use, number of lanes crossed by pedestrians or bicyclists, posted speed limit and the presence of special speed zones, all were found to influence the occurrence of these crashes significantly. The findings were used to identify patterns of pedestrian and bicycle high crash locations in Tennessee and flagged combination of demographic, socioeconomic and geometry variables which if present are good indicators to TDOT as areas likely to experience pedestrian and bicycle crashes.

INTRODUCTION

Bicyclists and pedestrians are a class of vulnerable road users that are often over-represented in fatal or incapacitating injury crash statistics. Many cities in Tennessee are seeing large increases in bicycling, and walking because these non-motorized trips are vital to Tennessee residents, it is essential to identify factors affecting safety of these modes of transportation. A cluster analysis with a combination of GIS and statistical analysis on demographic, socio-economic, land use and roadway data within respective census tracts was conducted in order to verify, if any, clustering of the accident points per identified locations, zones or neighborhoods. GIS cluster analysis was conducted to find patterns of crashes starting with all of the observations in one cluster and then proceeds to split them into smaller clusters. The combination of neighborhood attributes, socio-economic and demographic characteristics revealed factors associated with bicycle and pedestrian crashes. With the initial finding from the GIS mapping and cluster analysis and by using available disaggregate historical crash data, a comparative crash pattern was performed followed by development of statistical crash models. The models examined relationships between bicycle and pedestrian crash frequency with respect to influencing variables. Through these models, impact and significance of each variable was examined.

Many studies have found a number of demographic and roadway factors to be associated with pedestrian crashes *(1-17)*. For example, pedestrians over the age of 65 are more prone to motor vehicles crashes as a result of slower reaction times (*1*). Traffic speeds also increase the frequency of pedestrian and bicycle crashes *(2)*. Higher speeds reduce the reaction time available for drivers and the non-motorized road user, while increasing the distance and time needed for drivers to successfully brake, greatly increasing the likelihood of a collision. Other studies have found the number of lanes to significantly increase the pedestrian crash frequency because wider roadway may be related to higher vehicle speeds (*3*). Some studies have examined the relationship between pedestrian/bicycle crashes and the social and physical characteristics of neighborhoods. Indeed, an area's socioeconomic composition has a relationship with pedestrian and bicycle crash frequency (*4*). LaScala *(5)* found child pedestrian crashes in four communities in California are related to environmental factors. For example, higher rates of crashes are associated with higher youth population densities, higher unemployment, lower household income, and higher traffic flow *(5)*. The topic of race and ethnicity has been linked greatly with pedestrian and bicycle safety studies. A study in Arizona found that Latino were 60 percent more likely to be involved in pedestrian-vehicle crashes *(6)*. However, some studies have recommended a deeper analysis when studying the relationship between race and pedestrian/bicycle crashes. This is because some of the minority groups live in lower income areas, have low vehicle ownerships, live in high-density areas with high traffic volumes, and walk more *(7)*. Land use is another important factor when studying pedestrian/bicycle crash occurrence (*4, 8, 5, 9, and 10*). For example retail and entertainment areas (especially liquor stores, bars and restaurants that serve alcohol, and retail stores) tend to be associated with high pedestrian crashes (*5*). In addition, dense urban areas usually have high levels of pedestrian activity and therefore tend to have higher frequencies of pedestrian-vehicle collisions *(9)*. Commercial and multifamily housing areas are frequently found to be linked to high collision frequency *(10)*.

There is a vast body of statistical safety literature examining the factors affecting crash occurrence of non-motorized road users and the frequency of different types of non-motorized crashes with motorized vehicles (*11-17*). For example, Ukkusuri et al. (*11*) developed a random parameter negative binomial model for predicting pedestrian crash frequencies using the census tract level crash data in New York City. They found significant positive correlation between pedestrian crash frequencies and black or Latino neighborhoods, and areas with higher proportion of median-aged population and uneducated population. Greater commercial and industrial land uses are more prone to pedestrian crashes. Abdel-Aty et al. (*12*) applied a log linear model to study the safety of students around schools. They found that middle and high school children were involved in crashes more frequently than younger children, particularly on high speed multi-lane roads. Other pedestrian safety factors were also found to be correlated with the frequency of crashes including: the driver's age, gender, and alcohol use; the pedestrian's/bicyclist's age; and the number of lanes, median type, speed limits, and speed ratio. Garder (*13*) examined pedestrian crash data from Maine, and found that pedestrian crashes are more prevalent on Saturdays, in the afternoons between 4 and 7 pm, at times of clear weather, on level, straight, roads, and at locations without any traffic control devices or signage (this study did not control for exposure). Tomoyuki (*14*) used the negative binomial model to understand the contributing factors that cause pedestrian accidents in Hillsborough County in the State of Florida. He found that commercial and service land use, residential land use, and the density of state roadways increase the number of pedestrian accidents. He also found that an increase in average household income and a high percentage of elderly residents lower the number of pedestrian accidents. Mohamed and Sai (*15*) used log-linear models to relate pedestrian and bicyclist crashes involving school-aged children with various factors. They found that more crashes occurred around middle and high schools than elementary schools. Middle-aged, alcohol-impaired and male drivers are more likely to be involved in school-aged children crashes. Chris and Mohamed (*16*) classified pedestrian crashes into two types: crashes at driver's fault and crashes at pedestrian's fault; and then modelled using negative binomial. Their findings revealed that middle age (25-64 years) and male drivers are more involved in crashes as causers than other driver groups. To understand how the built environment affects both pedestrian activity and collision frequency pedestrian activity at intersections, Miranda-Moreno et al. (*17*) proposed a modeling framework based on the standard Negative Binomial (NB) model, the Generalized Negative Binomial (GNB) model and the latent class NB model. Among other results, they found that the built environment in the proximity of an intersection has a powerful association with pedestrian activity but a small direct effect on pedestrian-vehicle collision frequency

STUDY DATA

Pedestrian and bicycle crash data used in this study was obtained from Tennessee Department of Transportation (TDOT) and Tennessee Department of Safety (TDOS) databases. The crash data was downloaded for the year 2003 to 2009 which consisted of 5,360 pedestrian crashes and 2,558 bicyclist crashes. In order to locate individual crashes on a map, geospatial data is required. The geospatial data provided by TDOT consisted of shape files of the entire road network and link and surrounding land use attributes in Tennessee. In addition to the roadway data, 2010 US census demographic and socioeconomic data was drawn from the US census website (*18-19*) at census tract level. The census data consist of population, housing, race, and age distribution while the socioeconomic data consists of income, employment, commuting to work, occupations, and poverty status. Pedestrian and Bicycle crashes were geocoded using GIS. Approximately 75% of pedestrian and bicycle crashes were accurately mapped. Some of the crashes had unrecognizable route numbers or occurred in locations not monitored by TDOT and therefore they were not mapped (e.g., private driveways or streets). This is a limitation of our approach.

DISTRIBUTION BY SOCIOECONOMIC, DEMOGRAPHIC AND ECONOMIC DATA

The general distribution of the pedestrian and bicycle crashes were found to be in urban and suburban areas, consistent with earlier findings. The crash clusters are patterned in major cities such as Memphis (Shelby County), Nashville (Davidson County), Chattanooga (Hamilton County), and Knoxville (Knox County). Overall, the GIS cluster and mapping indicate these types of crash are more problematic in urban vicinity than in rural area (a fact which is obvious). Figures 1 to 4 show the distribution of crashes in two major counties, Shelby and Davidson Counties.

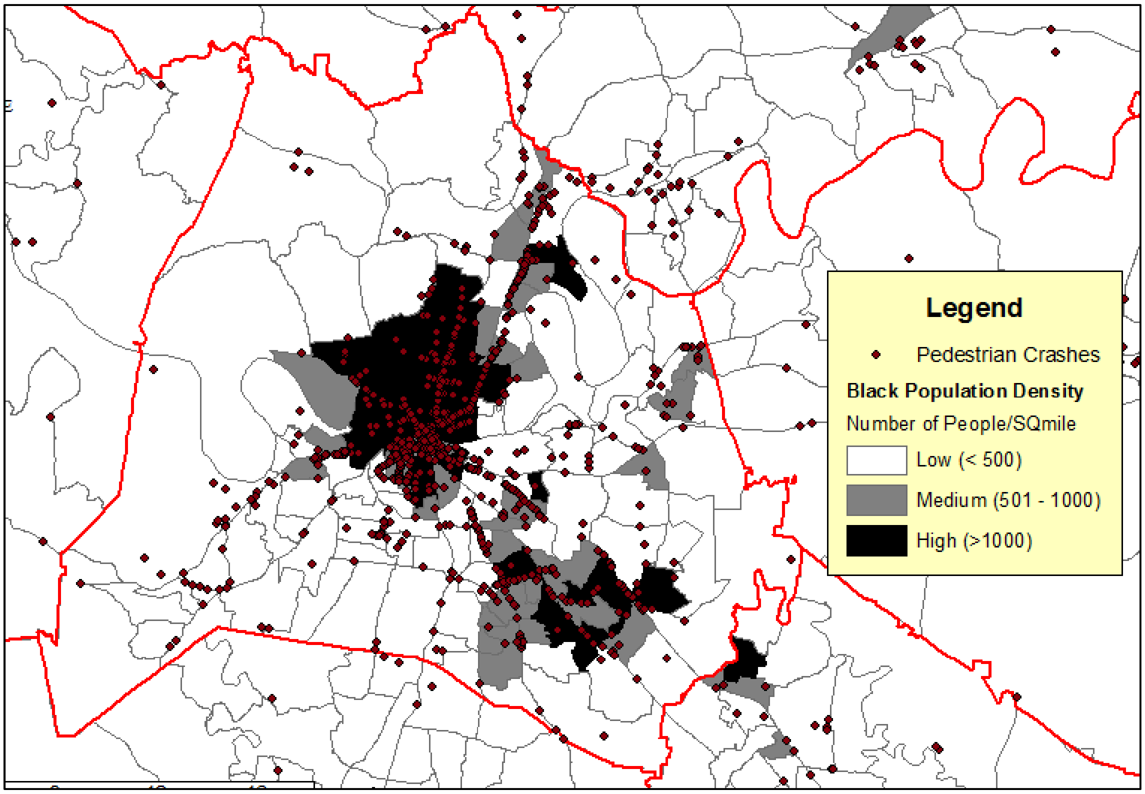

FIGURE 1 Distribution of Pedestrian Crashes with Black Population in Davidson County

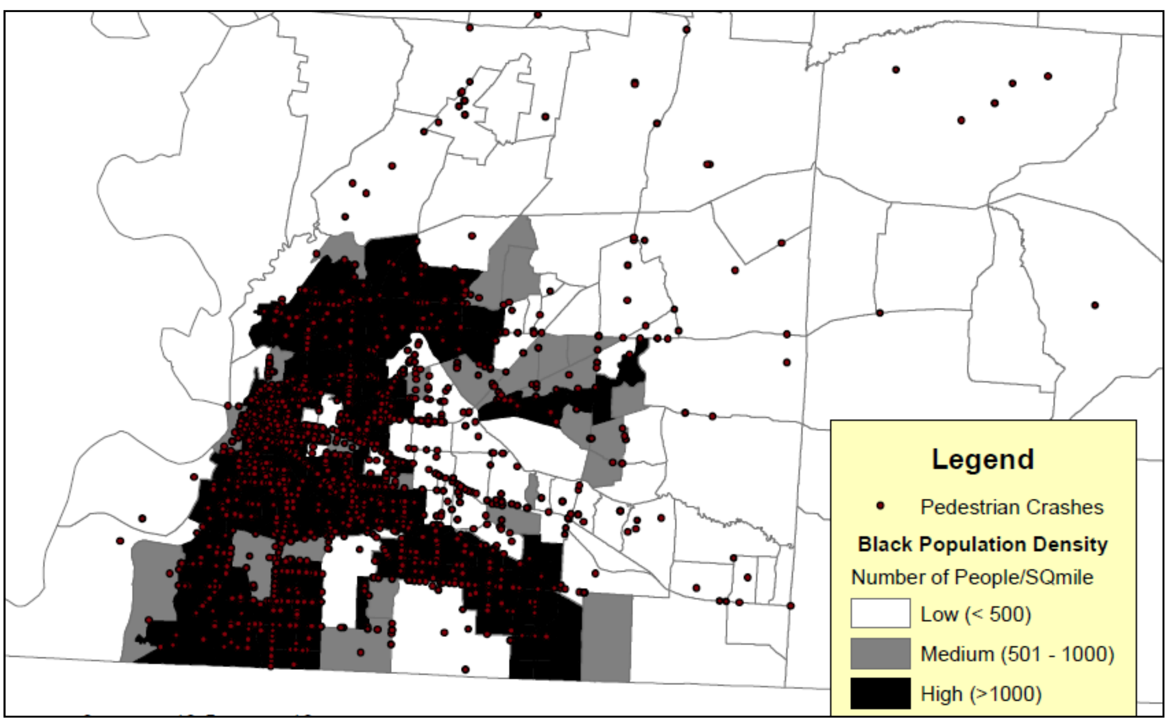

FIGURE 2 Distributions of Pedestrian Crashes with Black Population in Shelby County

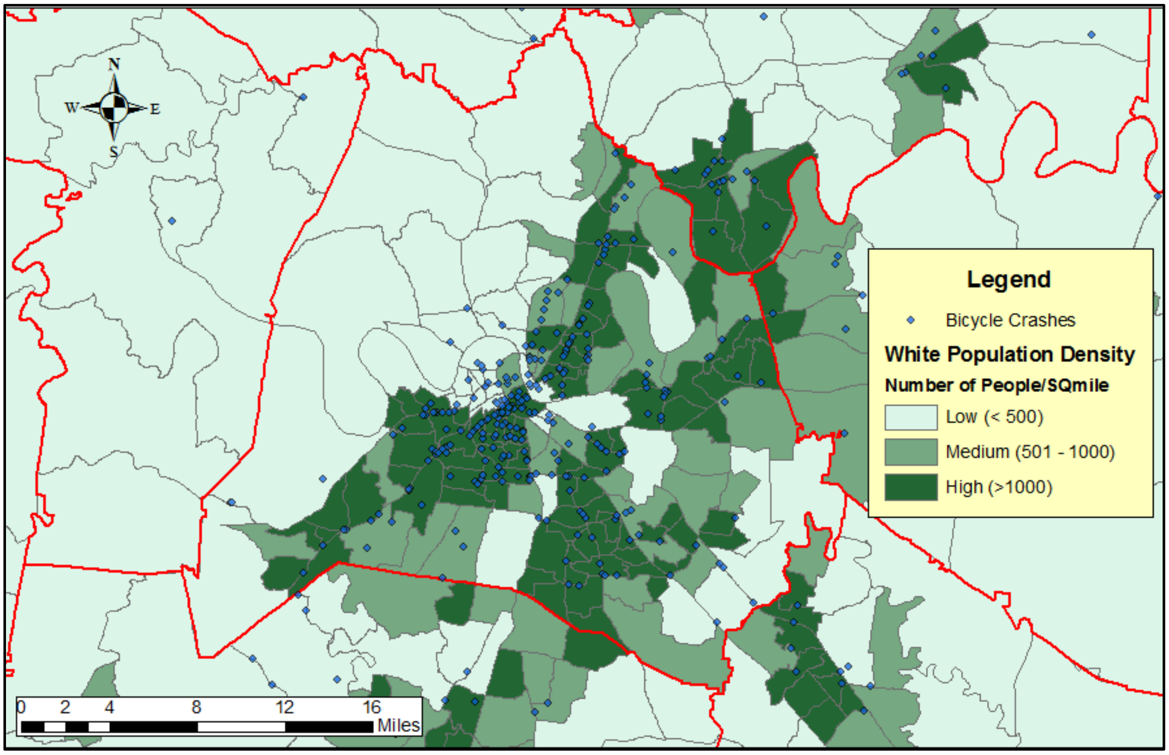

FIGURE 3 Distributions of Bicycle Crashes with White Population in Davidson County

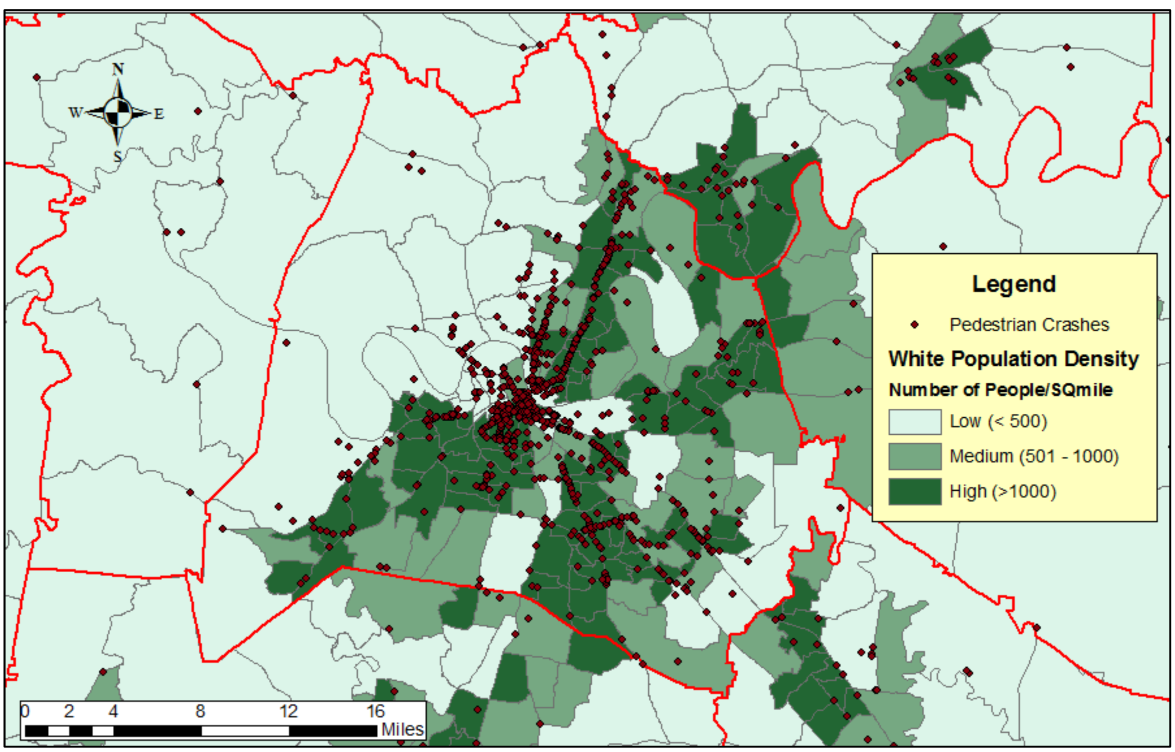

FIGURE 4 Distributions of Pedestrian Crashes with White Population in Davidson County

STATISTICAL ANALYSIS

A comparative crash pattern analysis performed in GIS was followed by statistical analysis. Summarized in Table 1 are the mean, standard deviation, and range of crash, socio-economic, and demographic data used either as a dependent or independent variables in the model. The statistical analysis examined the relationships between bicycle and pedestrian crashes with respect to influencing variables such as demographic characteristics, population, socio-economic characteristics, age groups, neighbourhood and land use characteristics, roadway geometry, and speed characteristics. Some variables were modeled as categorical and not as continuous variables. Posted speed limit was coded and modeled as "0" for speed limits <30 mph, "1" for speed limits >= 30 but <=35 mph and "2" for speed limit> 35 mph. Land use was grouped as "0" for rural, "1" for CBD, commercial, fringe and industrial land uses and "2" for residential, public use and parks. Terrain was coded as "0" for all others and "1" for rolling. Presence of school zone speed was coded as "1" and non-school zone sections was coded as "0".The crashes were merged on 168,920 uniform non-freeway roadway segments throughout the state with each segment averaging 0.5 miles, Table 1.

TABLE 1 Summary Statistics of Data Used

Variable	Mean	Std. Dev.	Min	Max
By Road Segment				
Number of Pedestrian Crashes	0.034	0.41	0	48
Number of Bicycle Crashes	0.006	0.12	0	18
Segment Length	0.50	0.90	0.001	29.4
Number of Lanes	1.94	0.68	1	10
By Census tract				
Total Population (n)	4838	1962.76	0	21763
White Population (%)	86.03	18.59	0	100
Black Population (%)	9.09	16.98	0	100
Latino Population (%)	3.69	4.43	0	52.8
Black and Latino Population (%)	12.78	18.39	0	100
Population Under 20years of Age (%)	25.34	4.54	0	61.18
Population from 20 to 64years of Age (%)	59.50	5.19	0	100
Population Above 65years of Age (%)	15.01	5.26	0	56.65
Housing units with No vehicles (%)	5.74	5.88	0	73.3
Housing units with 1 vehicle (%)	29.99	10.14	0	100
Housing units with 2 or more vehicles (%)	64.02	13.97	0	96.2
Population below Poverty Level (%)	14.90	9.05	0	79.6
Population in Labor Force (%)	60.53	9.57	0	88.8
Mean Household Income ($)	56058	24714.76	0	247239
Households with Income Below $25000 (%)	30.56	13.58	0	96.5
Households with Income: $25000 to $49999 (%)	27.87	7.21	0	63.2
Households with Income: $50000 & Above (%)	38.97	14.09	0	84.2

The literature review of crash modeling revealed that the crash rate and crash frequency are commonly used as response variables. The literature review also showed that Poisson and Negative Binomial (NB) distributions are often more appropriate for modeling discrete counts of events such as crashes which are likely to be zero or a small integer. The testing procedures was used to determine which distribution among the two (Poisson and negative binomial) was appropriate for the available crash data. Both pedestrian and bicycle crashes were found to be highly overdispersed, a fact that favors Negative Binomial over Poisson distribution. Therefore, NB model was considered more relevant for modeling these pedestrian and bicycle crashes. The probability mass function of the Negative Binomial (NB) model is expressed as, Chimba et al., (*20*):

$$P(y) = \frac{\tau(y+\alpha^{-1})}{\tau(\alpha^{-1})\tau(y+1)}\left[\frac{1}{1+\alpha\mu}\right]^{1/\alpha}\left[\frac{\alpha\mu}{1+\alpha\mu}\right]^{y} \quad (1)$$

Where;

$\mu = E(y) = \exp^{(X_i\beta)}$

X_i= the value of the independent variable related to the occurrence of crash

β=the coefficient of the independent variable

A stepwise model specification was performed, focusing on maintaining variables that met >90% significance levels. Although some few variables did not meet the significance levels, they were retained in the model to show how they affect the response variable and to help control for possible correlation effects.

Statistical Analysis Results

The primary objective of this statistical modeling effort was to evaluate the impact of different variables to pedestrian and bicycle crash frequency, the NB model was used to complete this objective. The frequency here is defined as the number of pedestrian/bicycle crashes along a road segment. As expected, not all variables were statistically significant in the models; hence the presented results show only the significant variables that were retained (in exception to % of white population in pedestrian model, and % population from 20 to 64years of age and housing units with 1 vehicle in bicycle model which were not significant). The impact and significance of each of the variables retained in the model to crash frequencies are summarized in Tables 2 & 3. As shown in Table 2, all the retained variables (except % of white population) were found to significantly influence pedestrian crashes at or above 90% significance level. Most of the variables were found to have positive coefficients while some had negative coefficients. The variables that have a positive correlation with pedestrian crash frequency include; number of lanes, speed limit, presence of a school zone, category 1 and 2 of the land use (CBD, commercial, fringe & industrial land use, residential, public use & parks land uses), percentage population of blacks and Latino, percentage population 20 to 64 years of age, mean household income, percentage population of 16 years and over in civilian labour force, percentage population of households below poverty level, percentage of households with no vehicle and those with one vehicle only. The positive coefficient variables indicate as the measure or presence of those variables increases, the likelihood of pedestrian crashes increases too. That means, from the model results, pedestrians crossing a roadway with multilanes are more likely to be involved in a crash than those crossing two lanes or fewer lanes. As shown in the GIS crash maps, there is very high correlation between pedestrian crashes and black and Latino populations. Model results in Table 2 show that strong significance of the two variables well above a 90% level. No vehicle ownership is significantly positively associated with pedestrian crashes indicating that the probability of pedestrian crashes increases with increase in the number of occupied housing units with no vehicles in a geographic area. This shows that households without vehicles have to use other means of transportation such as walking or public transportation which increases the likelihood of being involved in a pedestrian crash.

The variables that have negative correlation with pedestrian crash frequency include; percentage population of whites, households with income and benefits below $25000, and percentage of households with two or more vehicles. This can be interpreted that areas resided in predominantly by white populations have lower probability of pedestrian crashes compared those resided in mainly by black and Latino populations. In addition, as the number of occupied housing units with two or more vehicles increases, the lower the probability of pedestrian crashes. The finding that the increase in the percentage of households with income below $25000 lowers the likelihood of pedestrian crashes is contrary to common perception which needs further investigation. Household income and percentage in labor force both have positive coefficients. Positive coefficients for these two variables are also considered to be contradictory to intuition where reduction in pedestrian crashes could have been expected. However, close look to the data showed some of these pedestrian crashes occur at the intersections located in large shopping centers in wealthy neighborhoods, which could influence the model results.

TABLE 2 Pedestrian Crash Frequency Model Results

	Coefficient	P-Value
Number of Lanes	0.585	0.000
Speed Limit: 30 to 35mph	2.063	0.000
Speed Limit: 40 to 55mph	2.052	0.000
Presence of a School Zone	0.365	0.000
CBD, Commercial, Fringe & Industrial Land Use	2.057	0.000
Residential, Public Use & Parks Land Use	1.368	0.000
White Population (%)	-2.6E-04	0.963
Black Population (%)	0.010	0.062
Latino Population (%)	0.023	0.000
Population from 20 to 64years of Age (%)	0.014	0.000
Mean Household Income ($)	5.3E-06	0.000
Households with Income below $25000 (%)	-0.025	0.000
Population in Labor Force (%)	0.006	0.037
Population below Poverty Level (%)	0.024	0.000
Housing units with No vehicles (%)	0.015	0.009
Housing units with 1 vehicle (%)	0.014	0.002
Housing units with 2 or more vehicles (%)	-0.029	0.000
Constant	-8.620	0.000
Segment Length	Offset	

Note: Number of observations = 168920, Model P-Value = 0.000 and Pseudo R^2 = 0.279

Overall, these model results shows that the number of pedestrian crashes tends to increase with an increase in the number of families below the poverty level. Roadway Central Business District (CBD) areas, commercial areas, fringe areas and industrial land use types are positively associated with pedestrian crashes. In this study, fringe land use is a location that has a mix of both residential and commercial activities. Pedestrians are more prone to crashes in such areas. This observation is also seen in residential, public use and parks land use areas whereby the exposure rate to crashes is much higher. Posted speed limit was found to be significant in increasing pedestrian crashes. As the speed increases, crossing the street becomes more dangerous which increases the likelihood of a pedestrian to be involved in a crash. The positive influence of number of lanes influencing crash rate may reflect the longer the time it takes pedestrian to cross the road and the higher the likelihood of a crash. In addition, roads with more lanes are associated with a high traffic volume that is also positively associated with high crash frequency. The presence of a school zone is positively associated with pedestrian crashes, likely because of higher exposure when school is in session. Higher middle aged population (20-64 years) is positively associated with pedestrian crashes.

A similar model was estimated for bicycle crash frequency (Table 3). With model coefficients interpreted as those for pedestrian crash model, only two variables were found to have negative correlation with bicycle crash frequency, black population and the number of housing units with two or more vehicles. The rest of the variables had positive coefficients meaning that they tend to increase the probability of bicycle crash occurrences.

TABLE 3 Bicycle Crash Frequency Model Results

	Coefficient	P-Value
Number of Lanes	0.675	0.000
Speed Limit: 30 to 35mph	3.184	0.000
Speed Limit: 40 to 55mph	3.692	0.000
Rolling Terrain	0.371	0.06
CBD, Commercial, Fringe & Industrial Land Use	3.068	0.000
Residential, Public Use & Parks Land Use	1.555	0.000
Black Population (%)	-0.006	0.035
Latino Population (%)	0.019	0.048
Population from 20 to 64years of Age (%)	0.010	0.207
Mean Household Income ($)	1.3E-05	0.000
Households with Income from $50000 and above (%)	0.011	0.121
Population below Poverty Level (%)	0.027	0.000
Population in Labor Force (%)	0.012	0.074
Housing units with 1 vehicle (%)	0.007	0.357
Housing units with 2 or more vehicles (%)	-0.028	0.000
Constant	-14.224	0.000
Segment Length	Offset	

Note: Number of observations = 168920, Model P-Value = 0.000 and Pseudo R^2 = 0.295

CONCLUSION

This paper is part of the study focused primarily on identifying pedestrian and bicycle high crash locations in Tennessee. The study analyzed factors considered to influence an individual or household to use transit, walk or ride a bike which eventually can lead to increase in the likelihood of a crash. Roadway features were also included, i.e., number of lanes which have to be crossed by a pedestrian or bicyclist as well as land use and the posted speed limit on a particular segment or intersection. The study findings are expected to help decision makers at the state and local level develop road safety audits where officials can rank pedestrian and bicycle high crash locations experiencing high risk of bicycle and pedestrian crashes and make engineering recommendations and improvements.

Geographic Information System (GIS) and Negative Binomial (NB) statistical methods were applied in this research to investigate factors influencing crash frequency. Pedestrian and bicycle crash data from 2003 to 2009 was used together with geometry, census, socio-economic and demographic factors associated with these crashes within respective census tracts. GIS mapping found the pedestrian patterns to be associated with areas having higher percentage of minority populations (Black and Latino), urban areas, areas with higher percentage of households below poverty level, and those where a lower population is in labor force. However, GIS showed areas with high percentage of white population density to have low frequency of pedestrian and bicycle crashes. Negative Binomial distribution was used to linearize and detect which factors had positive or negative correlation with crash frequency. The statistical models revealed the following factors to have positive correlation with pedestrian crash occurrence; low vehicle ownership, high below poverty levels, high minority populations, residential and fringe areas, presence of a rolling terrain, high speed limits, high number of lanes, and presence of a school zones. In addition, the following factors were found to have positive correlation with bicycle crash occurrence; number of lanes, speed limit, presence of a rolling terrain, commercial, residential fringe and industrial land uses, Latino population, mean household income, households with income & benefits from $50000 and above, and low vehicle ownership.

The GIS mapping and statistical analysis findings are incorporated in planning framework of identifying pedestrian and bicycle high crash locations. Connecting patterns generated from the GIS and the magnitude and significance of variables in the statistical analysis helped in laying out clues which TDOT should focus on to improve bicycle and pedestrian safety. The study flagged combination of demographic, socioeconomic and road geometry variables which are good indicators of areas prone to pedestrian and bicycle crashes and therefore TDOT can direct their improvement funds to such areas.

REFERENCES

1. Campbell, B. J., C. V. Zegeer, H. H. Huang, and M. J. Cynecki. *A Review of Pedestrian Safety Research in the United States and Abroad*. Publication FHWA-RD-03-042, 2004.
2. Garder, P. E. The Impact of Speed and Other Variables on Pedestrian Safety in Maine. *Accident Analysis and Prevention*, Vol. 36, 2004, No. 533-542.
3. Zegeer, C. V., R. Stewart, H. H. Huang, P. A. Lagerway, K. Todd, and P. Jacobsen. Safety Effects of Marked vs. Unmarked Crosswalks at Uncontrolled Locations: Analysis of Pedestrian Crashes in 30 cities. *Journal of the Transportation Research Board*, No. 1773, 2001, pp. 56–58.
4. Shin, H. S., G. Chen, K. Majka, A. Shorris, and S. Ukkusuri. *Pedestrian Fatality and Severe Injury Accidents in New York City*. New York City Department of Transportation, 2010.
5. LaScala, E. A., P. J. Gruenewald, and F. W. Johnson. An Ecological Study of the Locations of Schools and Child Pedestrians Injury Collisions. *Accident Analysis and Prevention*, Vol. 36, 2004, pp. 569-576.
6. Campos-Outcalt, D., C. Bay, A. Dellapenna, and M. K. Cota. Pedestrian Fatalities by Race/Ethnicity in Arizona, 1990-1996. *American Journal of Preventive Medicine*, Vol. 23, No. 2, 2002, pp. 129-135.
7. Voas, R. B. A., S. Tipptts, and D. A. Fisher. *Ethnicity and Alcohol-related Fatalities: 1990 to 1994*. Publication DOT HS 809 068. National Highway Traffic Safety Administration, USDOT, 2000.
8. Ukkusuri, S. V., L. F. Miranda, G. Ramadurai, and J. Isa. A Frequency Model of Pedestrian Accidents Accounting for Built Environment Characteristics. Purdue University, 2010.
9. Loukaitou-Sideris, A. Death on the Crosswalk: A Study of Pedestrian-Automobile Collisions in Los Angeles. *Journal of Planning Education and Research*, Vol. 26, 2007, pp. 338- 351.
10. Levine, N., K. E. Kim, and L. H. Nitz. Spatial Analysis of Honolulu Motor Vehicle Crashes: I. Spatial patterns. *Accident Analysis and Prevention*, Vol. 27, 1995, pp. 663-674.
11. Ukkusuri, S., H. Samiul, H. M. Abdul-Aziz. A Random-parameter Model to Explain the Effects of Built Environment Characteristics on Pedestrian Crash Frequency. Presented at TRB Annual Meeting, 2011.
12. Abdel-Aty, M., S. Chundi, and C. Lee. Geo-spatial and log-linear analysis of pedestrian and bicyclist crashes involving school-aged children. *Journal of Safety Research*, Vol 38, 2007, pp. 571–579.
13. Garder, P.E. The impact of speed and other variables on pedestrian safety in Maine. *Accident Analysis and Prevention,* Vol. 36 (4), 2004, pp 533–542.
14. Tomoyuki, H. *Spatial Analysis of Pedestrian Accidents*. M.S. thesis, Department of Civil Engineering, University of South Florida, 2005.
15. Mohamed, A. A., S. C. Sai. Geo-spatial and log-linear analysis of pedestrian and bicyclist crashes involving school-aged children. *The Journal of Safety Research*, Vol. 38, 2007, pp 571–579.
16. Chris, L., A. A. Mohamed. Comprehensive analysis of vehicle-pedestrian crashes at intersections in Florida. *Accident Analysis and Prevention*, Vol. 37, 2005, pp 775–786.
17. Miranda-Moreno, L., P. Morency, and A. El-Geneidy. The link between built environment, pedestrian activity and pedestrian-vehicle collision occurrence at signalized intersections. *Accident Analysis and Prevention*, Vol. 43(5), 2011, 1624–1634.
18. Topologically Integrated Geographic Encoding and Referencing (TIGER), http://www.census.gov/geo/www/tiger/shp.html. Accessed 2/2012.
19. United States Census Bureau (USCB), http://factfinder2.census.gov/faces/nav/jsf/pages/index.xhtml. Accessed 2/2012.
20. Chimba D., Sando T., and Kwigizile V. 2010. Effect of Bus Sizes and Operations to Crash Occurrences. *Accident Analysis & Prevention* Volume 42, Issue 6, pp 2063-2067.