Enhancing Non-Motorized Safety by Simulating Non-Motorized Exposure using a Transportation Planning Approach

SAFETY RESEARCH USING SIMULATION

UNIVERSITY TRANSPORTATION CENTER

Dr. Naveen Eluru
Department of Civil, Environmental and Construction Engineering
Associate Professor

Dr. Shamsunnahar Yasmin, Tanmoy Bhowmick, Moshiur Rahman and Dr. Naveen Eluru
Department of Civil, Environmental and Construction Engineering

Enhancing Non-Motorized Safety by Simulating Non-Motorized Exposure using a Transportation Planning Approach

Dr. Shamsunnahar Yasmin

Postdoctoral Associate

Department of Civil, Environmental and Construction Engineering

Tanmoy Bhowmik, BSc

Graduate Student

Department of Civil, Environmental and Construction Engineering

Moshiur Rahman, BSc

Graduate Student

Department of Civil, Environmental and Construction Engineering

Dr. Naveen Eluru

Associate Professor

Department of Civil, Environmental and Construction Engineering

A Report on Research Sponsored by SAFER-SIM University Transportation Center

June 2016

Table of Contents

List of Figures

List of Tables

Abstract

Safety researchers and analysists have employed land use and urban form variables as surrogates for traffic exposure information (pedestrian and bicyclist volumes and vehicular traffic). The quality of these crash prediction models is affected by the lack of "true" non-motorized exposure data. The current research effort is focused on developing a transportation planning simulation framework to generate exposure information for crash prediction models. Specifically, the research effort is focused on evaluating non-motorist exposure measures in terms of demand at a planning level. The evaluated exposure measures are incorporated in examining non-motorist safety, which would allow us to devise more evidence-based policy implications for improving overall safety and activities related to non-motorized modes of travel. The proposed research approach recognizes that non-motorized safety is affected by vehicular volumes and non-motorized activity at a macro-level in the urban region. The vehicular and non-motorized exposure measures are generated to enhance the vulnerable road user crash prediction models. In identifying non-motorist exposure measures, we develop aggregate-level demand models to identify critical factors contributing to non-motorist generators and attractors at a zonal level. In evaluating non-motorist safety, we estimate four different aggregate level models: (1) zonal-level crash count model for examining pedestrian-motor vehicle crash occurrences, (2) zonal-level crash count model for examining bicycle-motor vehicle crash occurrences, (3) zonal-level crash severity model for examining pedestrian crash injury severity by proportions, and (4) zonal-level crash severity model for examining bicycle crash injury severity by proportions. These models are estimated as a function of zonal level sociodemographic characteristics, roadway/traffic attributes, built environment, land-use characteristics, and exposure measures identified from demand models. The formulated demand models are estimated by using 2009 National Household Travel Survey data and the crash models are estimated by using the Signal Four Analytics crash database for the year 2010 for the Central Florida region. Model estimation results are further augmented by a validation exercise. To demonstrate the implication of the estimated models, we also perform policy analysis for ten different scenarios, including changes in traffic volume within the vicinity of central business district, reduction in zonal-level speed limit, increasing walking facilities, and restrictions on the number of traffic lanes. From the policy scenario analysis, we identify beneficial changes to existing infrastructure and traffic operation for improving non-motorized road user safety at a planning level. The research methodology as proposed in our study recognizes that zonal-level attributes are likely to influence non-motorist exposure. At the same time, non-motorist exposure along with the zonal-level attributes are critical factors in developing non-motorist safety models.

1 Introduction

Urban regions in North America are encouraging the adoption of active modes of transportation by proactively developing infrastructure for non-motorized modes. According to data from the 2009 National Household Travel Survey (NHTS), about 37.6% of the trips by private vehicles in the United States (US) are less than 2 miles long. If even a small proportion of the shorter private vehicle trips (around dense urban cores) are substituted with active transportation trips, there are substantial benefits to individuals, cities and the environment. However, a strong impediment to the increasing adoption of active modes of transportation is the risk associated with these modes. The safety risk posed to active transportation users in Florida is exacerbated compared to active transportation users in the rest of the US. While the national average for pedestrian (bicyclist) fatalities per 100,000 population in 2015 is 1.67 (2.50), the corresponding number for the state of Florida is 3.10 (7.40), which clearly demonstrates the challenges faced in Florida (NHTSA 2017a, b). An important tool for determining the critical factors affecting the occurrence of pedestrian and bicycle crashes and identifying vulnerable locations is the application of planning-level crash prediction models.

Traditionally, in developing these models, safety researchers and analysists have employed land use and urban form variables as surrogates for exposure information (pedestrian and bicyclist volumes and vehicular traffic). The quality of these crash prediction models is affected by the lack of "true" non-motorized exposure data. Moreover, to assess the implication of different strategies in improving non-motorized safety, it is important to evaluate and document demand of non-motorized road users. The current research effort is focused on developing a transportation planning simulation framework to generate exposure information for crash prediction models. Specifically, the current research effort is focused on evaluating non-motorist exposure measures in terms of demand at a planning level. The evaluated exposure measures are further incorporated in examining non-motorist safety, which would allow us to devise more evidence-based policy implications for improving overall safety and activities related to non-motorized modes of travel.

1.1 Research Methodology

The proposed research approach recognizes that non-motorized safety is affected by vehicular volumes and non-motorized activity at a macro-level in the urban region. The vehicular and non-motorized exposure measures are generated to enhance the vulnerable road user crash prediction models. The details of the methodology are discussed below.

In identifying non-motorist exposure measures, we develop aggregate-level demand models to identify critical factors contributing to non-motorist's generators and attractors at a zonal level. Specifically, in our current study, we investigate non-motorist demand at a zonal level by using aggregate trip information based on origin and destination locations of trips. Further, we also estimate a pedestrian destination choice model and a bicycle destination choice model, which would allow us to micro-simulate aggregate-level non-motorist travel demand in the form of detailed origin-destination (O-D) matrices. The demand model would allow us to identify the number of non-motorists at a zonal level, while the destination choice models would allow us to identify the location of non-motorist activities. These models are estimated as a function of zonal-level sociodemographic characteristics, roadway/traffic attributes, built environment and land-use characteristics. In evaluating non-motorist exposure, we also generate different zonal-level trip exposure matrices with the numbers of daily trip origin and daily trip destination at the zonal level for both pedestrian and bicycle modes to be considered as non-motorist exposure measures for safety evaluation.

In evaluating non-motorist safety, in this research effort, we estimate four different aggregate level models: (1) a zonal-level crash count model for examining pedestrian-motor vehicle crash occurrences, (2) a zonal-level crash count model for examining bicycle-motor vehicle crash occurrences, (3) a zonal-level crash severity model for examining pedestrian crash injury severity by proportions and (4) a zonal-level crash severity model for examining bicycle crash injury severity by proportions. These models are estimated as a function of zonal-level sociodemographic characteristics, roadway/traffic attributes, built environment, land-use characteristics and exposure measures identified from demand models. The outcomes of these macro-level models can be used to devise safety-conscious decision support tools to facilitate a proactive approach in assessing medium- and long-term policy-based countermeasures. Moreover, the tool plays an important role in safety implications of land use planning initiatives and alternate network-planning initiatives.

The research methodology as proposed in our study recognizes that zonal-level attributes are likely to influence non-motorist exposure. At the same time, non-motorist exposure along with the zonal-level attributes are critical factors in developing non-motorist safety models. This intertwined relationship and our research road map is represented in Figure 1.1.

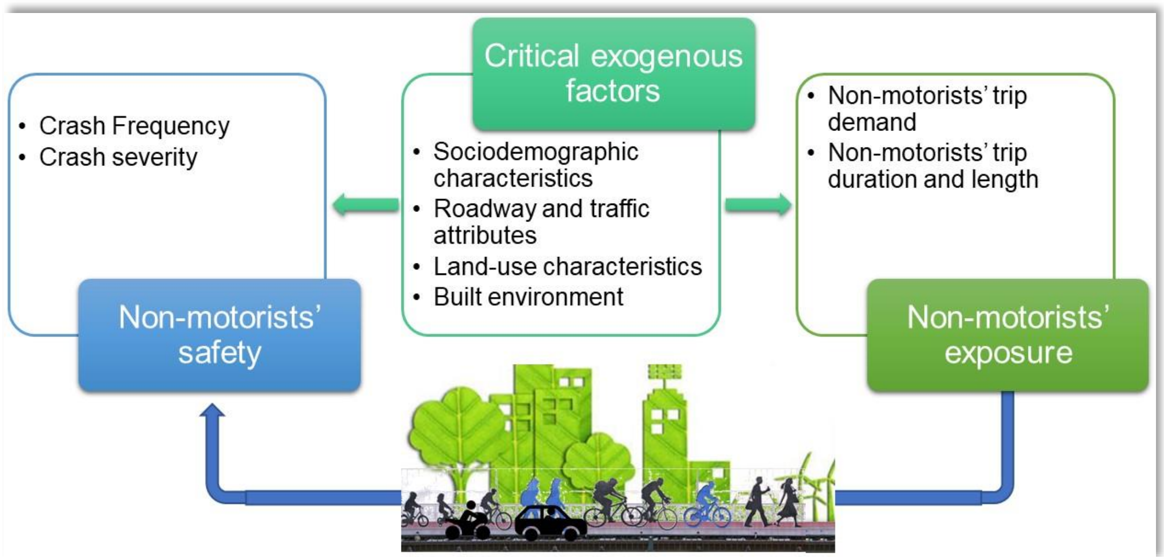


Figure 1.1 – Proposed non-motorist safety evaluation framework

1.2 Study Area

Our study areas include the Central Florida region. Specifically, we consider the region defined for Central Florida Regional Planning Model version 6.0 (CFRPM 6.0). The study area includes 4,747 traffic analysis zones (TAZs). Boundary of the study area encompasses nine counties (Brevard, Flagler, Lake, Marion, Orange, Osceola, Seminole, Sumter and Volusia) within District 5, Polk county within District 1 and part of Indian River county in District 4 of the Florida Department of Transportation (FDOT). The location study area along with the zonal boundaries are shown in Figure 1.2. For estimating models, we consider the year 2010 as the base year.

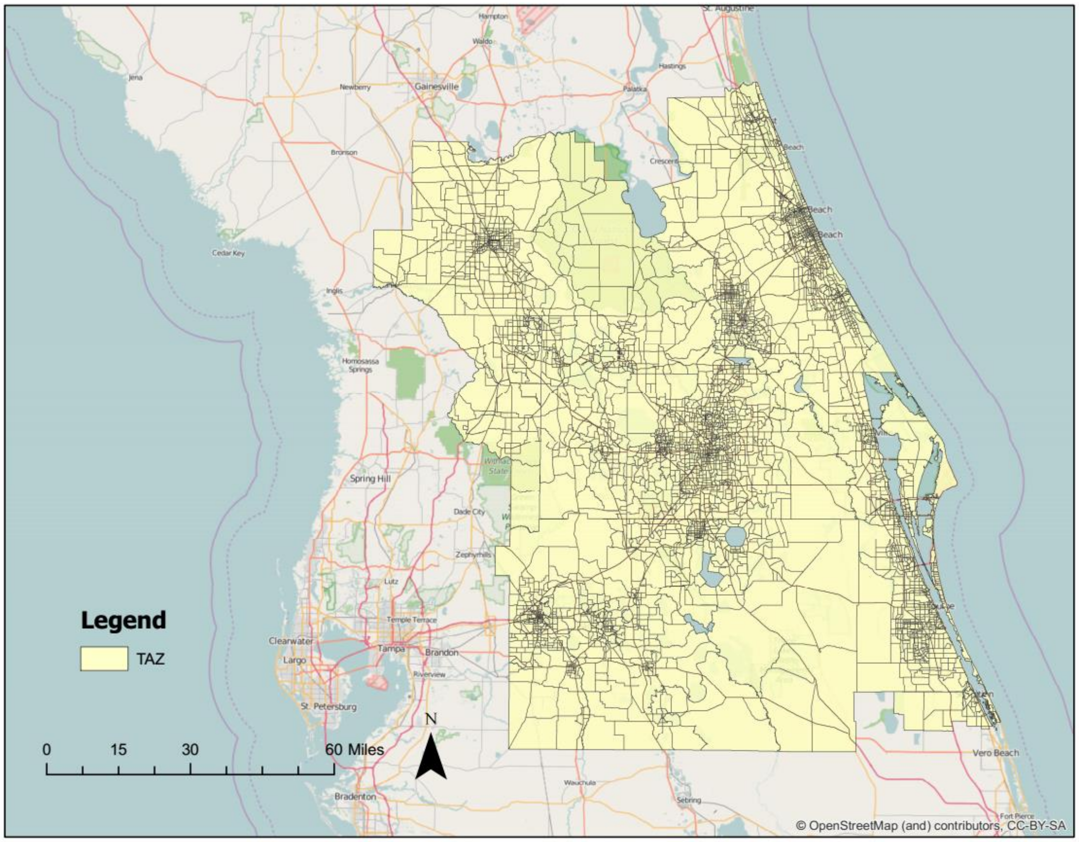


Figure 1.2 – Location of study region

1.3 Outline of the Report

The remainder of the document is structured as follows:

- Section 2 contributes to non-motorized road user exposure evaluation.
- Section 3 presents non-motorized road user safety evaluation.
- Section 4 contributes to policy scenario analysis and recommendations.
- Section 5 concludes the report by summarizing the findings and describing limitations.

2 Non-Motorized Road User Exposure Evaluation

In evaluating non-motorized road user safety, studies often do not consider non-motorized exposure in detail. However, non-motorist exposure is likely to have a significant influence on the safety of this group of road users. Further, with growing emphasis on improving mobility in the Florida region, there is increasing awareness and targeted efforts to enhance non-motorized (pedestrian and bicyclist) mobility. To evaluate the effectiveness of these strategies and to enhance safety, it is useful to develop methods that accommodate the potential adoption of non-motorized modes within the mobility planning process. In order to assess the implication of different strategies in improving non-motorized safety, it is important to evaluate and document the demand of non-motorized road users. Analysts often develop a non-motorist demand model at different local levels, such as regional level, corridor or sub-area level and household/individual level. Among these models, analysis is widely conducted to evaluate non-motorized travel at a zonal level. Several high-resolution modeling frameworks, such as an activity-based or trip-based approach, could be pursued for evaluating planning-level non-motorist demand. However, it is worthwhile to mention here that high-resolution disaggregate-level data of non-motorist activity is still unavailable or available only for few locations at a corridor level. Extrapolating planning-level non-motorist demand from a few corridor-level exposure data would require several assumptions along with a higher level of computational burden. An alternative approach to generating a planning-level non-motorist demand model is to estimate O-D demand at an aggregate level.

The aggregate-level demand models examine critical factors contributing to non-motorist generators and attractors at a zonal level. Outcomes of these studies can be used to devise medium- or long-term area-wide planning and investment policies to encourage and promote non-motorized activities and to improve safety situations for these groups of road users. Moreover, these models can be used as a tool for evaluating non-motorized transportation pilot projects. To that extent, in our current study, we investigate non-motorist demand at a zonal level by using aggregate trip information based on origin and destination locations of trips. Specifically, we develop four non-motorist demand models: (1) <u>pedestrian generator model</u> – based on zonal-level pedestrian origin demand, (2) <u>pedestrian attractor model</u> – based on zonal-level pedestrian destination demand, (3) <u>bicycle generator model</u> – based on zonal-level bicycle origin demand, and (4) <u>bicycle attractor model</u> – based on zonal-level bicycle destination demand. Further, we also estimate the pedestrian destination choice model and the bicycle destination choice model, which would allow us to micro-simulate aggregate-level non-motorist travel demand in the form of detailed O-D matrices. The demand model would allow us to identify the number of non-motorists at a zonal level, while the destination choice models would allow us to identify the location of non-motorist activities. These models are estimated for the study area defined by the CFRPM 6.0 area by using trip records from the 2009 NHTS database. In the following section, we have presented and discussed estimation results of these models along with data compilation procedures. Further, we also present the validation results of these estimated models.

2.1 <u>Data Source</u>

For developing non-motorist demand and destination choice models, the data is sourced from the 2009 NHTS database gathered in the United States. The 2009 NHTS collected detailed information on more than one million trips undertaken by 320,000 individuals from 150,000 households sampled from all over the country. The database includes information on mode taken by trip makers for each trip, trip purpose and trip location, along with the trip maker's characteristics, household characteristics and trip characteristics. The 2009 NHTS database from FDOT with add-ons allowed us to identify trips which were recorded for the Central Florida region. In the 2009 NHTS, there were 2,749 households surveyed in the Central Florida region. It included a total of 5,090 individuals and 22,359 trips. Among these trips, walk and bike trip shares were 8.8 % and 1.3 %, respectively. In the current study context, we incorporate "person-trip weight" – as defined in the NHTS database – to extrapolate the representative number of trips for the Central Florida region.

2.2 Non-Motorist Demand Model

Non-motorist travel demand models are estimated at the zonal level based on information about trip origin and destination. Specifically, we estimate four different models: (1) pedestrian generator model, (2) pedestrian attractor model, (3) bicycle generator model and (4) bicycle attractor model. In generator models, we examine the daily zonal trip origin count (total number of trips originated at zones) to identify critical factors that are likely to generate non-motorist origin demand. On the other hand, in attractor models, we examine the daily zonal trip destination count (total number of trip destined to zones) to identify critical factors that are likely to generate non-motorist destination demand. In the current research effort, we formulated and estimated Hurdle-Negative Binomial (HNB) models for examining non-motorist travel demand to account for the large share of zones with zero non-motorized activity. The HNB models are estimated at the TAZ level for the CFRPM 6.0 area employing a comprehensive set of exogenous variables. Based on the model results, we identify important exogenous variables that influence pedestrian and bicycle O-D demand.

2.2.1 Model Framework

In our current research effort, the non-motorist O-D demands are examined by using the Hurdle count regression approach. The non-motorist demands are represented as the total number of non-motorist trips originated and destined to a zone. Thus, the demands are non-negative integer values. Naturally, these integer counts can be examined by employing count regression approaches, such as the Poisson and Negative Binomial (NB) regression approaches. However, for the zonal-level non-motorist trip counts, more than 84% and 96% TAZs have zero pedestrian and bicycle trip records, respectively. The traditional count models (Poisson and NB models) do not account for such over-representation of zero observations in the data. The Hurdle model is typically used in the presence of such excess zeroes. Cameron and Trivedi [1] presented these models as finite mixture models with a degenerate distribution and probability mass concentrated at zeroes. The Hurdle approach is generally used for modeling excess sampling zeroes. It is usually interpreted as a two-part model [2]: the first part is a binary response structure modeling the probability of crossing the hurdle of zeroes for the response and the second part is a zero-truncated formulation modeled in the form of standard count models (Poisson or NB). Thus, the probability expression for the Hurdle model can be expressed as:

$$\Lambda_i[y_i] = \begin{cases} \pi_i & y_i = 0 \\ \frac{(1-\pi_i)}{(1-e^{-\mu_i})} P_i(y_i) & y_i > 0 \end{cases} \tag{2.1}$$

where i is the index for TAZ ($i = 1,2,3, \dots, N$) and y_i is the index for non-motorist (pedestrian and bicycle) trips occurring daily in a TAZ i.

In Equation 2.1, π_i is the probability of zero trip count and is modeled as a binary logit model as follows:

$$\pi_i = \frac{exp(\gamma \boldsymbol{\eta}_i)}{1+exp(\gamma \boldsymbol{\eta}_i)} \tag{2.2}$$

where $\boldsymbol{\eta}_i$ is a vector of attributes and γ is a conformable parameter vector to be estimated.

$P_i(y_i)$ in Equation 2.1 can be presented as Poisson and NB expressions in forming Hurdle Poisson (HP) and HNB regression models, respectively. Given the setup as presented in Equation 2.1, the probability distribution for Poisson can be written as:

$$P_i(y_i|\mu_i) = \frac{e^{-\mu_i}(\mu_i)^{y_i}}{y_i!}, \mu_i > 0 \tag{2.3}$$

where μ_i is the expected number of daily trips non-motorists are making in TAZ i.

We can express μ_i as a function of explanatory variable ($\boldsymbol{z}_i$) by using a log-link function as $\mu_i = E(y_i|\boldsymbol{z}_i) = exp(\boldsymbol{\delta} \boldsymbol{z}_i)$, where $\boldsymbol{\delta}$ is a vector of parameters to be estimated. However, one of the most restrictive assumptions of Poisson regression, often violated, is that the conditional mean is equal to the conditional variance of the dependent variable.

The variance assumption of Poisson regression is relaxed in NB by adding a Gamma distributed disturbance term to Poisson distributed count data [3]. Given the above setup, the NB probability expression for $\boldsymbol{y_i}$ can be written as:

$$P_i(y_i|, \mu_{i,}\alpha) = \frac{\Gamma(y_i+\alpha^{-1})}{\Gamma(y_i+1)\Gamma(\alpha^{-1})} \left(\frac{1}{1+\alpha\mu_i}\right)^{\frac{1}{\alpha}} \left(1 - \frac{1}{1+\alpha\mu_i}\right)^{y_i} \tag{2.4}$$

where $\Gamma(\cdot)$ is the Gamma function and α is the NB dispersion parameter.

Finally, the weighted log-likelihood function for the Hurdle count model can be written as:

$$LL = w_i * \begin{cases} ln(\pi_i) & y_i = 0 \\ \ln\left(\frac{(1-\pi_i)}{(1-e^{-\mu_i})} P_i(y_i)\right) & y_i > 0 \end{cases} \tag{2.5}$$

The daily trip weight at the zonal level is generated by using the following formulation:

$$w_i = \sum_{j=1}^{J} \frac{Yearly\ person\ trip\ weight}{365} \tag{2.6}$$

where j ($j = 1,2,3, \dots J$) represents the index for trip.

The reader should note that in computing the weighting factor, as presented in Equation 2.6, we divided the yearly person trip factor, as obtained from NHTS data, by 365 to convert the yearly trip count to a daily trip count. Substitution of $\left(P_i(y_i)\right)$ by Equations 2.3 and 2.4 into Equation 2.5 results in HP and HNB models, respectively. The model presented in Equation 2.5 is estimated by using a maximum likelihood approach.

2.2.2 Data Description

The non-motorist demand model is focused on non-motorist O-D demand at the TAZ level. With respect to origin and destination demand, we examine daily zonal trip origin count and daily zonal trip destination count, respectively. Table 2.1 offers summary characteristics of these daily trip counts for pedestrian and bicycle trip activities based on their trip origin and trip destination along with the number of zones with sample characteristics. From Table 2.1, we can see that number of zones with pedestrian demand is much higher than the number of zones with bicycle demand. Locations of zones with pedestrian and bicycle O-D demand are shown in Figure 2.1.

Table 2.1 - Summary characteristics of trip counts

Sample characteristics	Frequency (percentage)
Total number of zones	4747
Zones with zero pedestrian origin trip counts	4007 (84.4011)
Zones with pedestrian origin trip counts	740 (15.589)
Zones with zero pedestrian destination trip counts	4010 (15.53)
Zones with pedestrian destination trip counts	737 (84.47)
Zones with zero bicycle origin trip counts	4574 (3.64)
Zones with bicycle origin trip counts	173 (96.36)
Zones with zero bicycle destination trip counts	4581 (3.50)
Zones with bicycle destination trip counts	166 (96.50)

Variable names	Definition	Zonal (weighted)		
		Minimum	**Maximum**	**Mean**
Dependent variables				
Pedestrian origin trip count	Total daily pedestrian trip origin demand at a zone	0.000	39232.010	265.450
Pedestrian destination trip count	Total daily pedestrian trip destination demand at a zone	0.000	39232.010	261.696
Bicycle origin trip count	Total daily bicycle trip origin demand at a zone	0.000	7012.434	35.022
Bicycle destination trip count	Total daily bicycle trip destination demand at a zone	0.000	7012.434	34.937

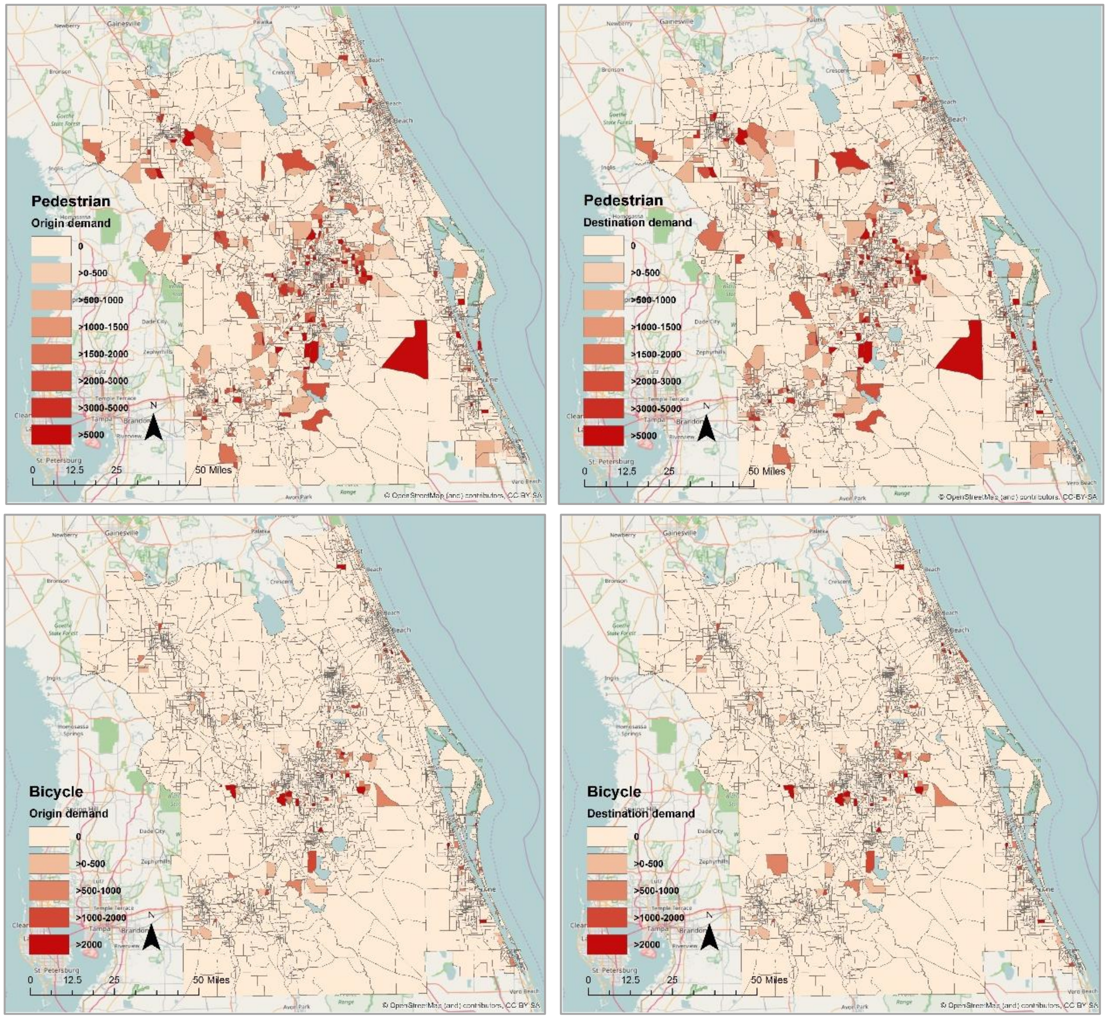


Figure 2.1 - Zones with pedestrian and bicycle O-D demand

In addition to the trip counts, the explanatory attributes considered in the empirical study are also aggregated at the TAZ level accordingly. For the empirical analysis, the selected explanatory variables can be grouped into four broad categories: sociodemographic characteristics, roadway and traffic attributes, built environment and land use characteristics. The sociodemographic characteristics are compiled from the U.S. Census Bureau's Tiger/line data and American Community Survey database. Moreover, roadway and traffic attributes, built environment and land use characteristics are obtained from the Florida Geographic Data Library and the FDOT data repository. Table 2.2 offers a summary of the sample characteristics of the exogenous variables and the definition of variables considered for final model estimation along with the zonal minimum, maximum and average.

Table 2.2 - Summary characteristics for exogenous variables

Variable names	Definition	Zonal

		Minimum	Maximum	Mean
Sociodemographic characteristics				
Population density	Total number of Population of TAZ/ Area of TAZ in acre	0	19.956	2.366
Proportion of male population	Total number of male of TAZ/ Total number of Population of TAZ	0	0.998	0.49
Proportion of 22-29 aged population	Total number of population who are 22 to 29 years old of TAZ/ Total number of Population of TAZ	0	0.397	0.096
Proportion of 65+ aged population	Total number of people above 65 years old of TAZ/ Total number of Population of TAZ	0	0.899	0.182
Roadway and traffic attributes				
Average zonal speed	Average zonal speed in mph	0	70	36.028
AADT	Total Annual Average Daily Traffic (AADT) of TAZ/10000	0	27.550	0.931
Truck AADT	Total Truck AADT of TAZ/10000	0	2.747	0.083
Proportion of arterial road	Total length of arterial road of TAZ/Total roadway length of TAZ	0	1	0.459
Proportion of collector road	Total length of collector road of TAZ/Total roadway length of TAZ	0	1	0.417
Proportion of 3 and more lane road	Total length of through roadway with 3 or more number of lanes of TAZ/Total roadway length of TAZ	0	1	0.096
Length of sidewalk	Total sidewalk length in meter of TAZ	0	36.346	0.280
Availability of bike lane	Presence of bike lane in TAZ	0	1	0.041
Number of flashing beacon sign	Total number of flashing beacon of TAZ	0	2	0.009
Number of school signal	Total number of school signal of TAZ	0	1	0.001
Built environment				
Number of educational center	Total number of educational center of TAZ	0	5	0.275
Number of financial center	Total number of financial center of TAZ	0	17	0.586

Number of park and recreational center	Total number of park and recreational center of TAZ	0	20	0.245
Number of commercial center	Total number of commercial center of TAZ	0	4	0.087
Number of entertainment center	Total number of entertainment center of TAZ	0	3	0.017
Number of restaurant	Total number of restaurant of TAZ	0	36	1.335
Number of shopping center	Total number of shopping center of TAZ	0	78	1.492
Number of transit hub	Total number of transit hub of TAZ	0	11	0.051
Land-use characteristics				
Institutional area	Ln (Institutional area in a TAZ in acre)	-16.417	7.071	0.785
Residential area	Ln (Residential area in a TAZ in acre)	-12.427	8.014	3.596
Industrial area	Ln (Industrial area in a TAZ in acre)	-12.943	6.709	0.671
Recreational area	Ln (Recreational area in a TAZ in acre)	-13.946	10.04	0.388
Retail/office area	Ln (Office/Retail area in a TAZ in acre)	-17.312	6.611	1.744
Urban area	Ln (Urban area in a TAZ in acre)	-9.275	8.491	4.291
Land-use mix	Land use mix = $\left[\frac{-\sum_k (p_k(ln p_k))}{ln N}\right]$, where k is the category of land-use, p is the proportion of the developed land area devoted to a specific land-use, N is the number of land-use categories in a TAZ	0	0.929	0.35496

2.2.3 Model Specification and Overall Measures of Fit

The empirical analysis of non-motorist demand involves the estimation of model using two different econometric frameworks: HP and HNB. Prior to discussing the estimation results, we compare the performance of these models in this section. To compare the performance of estimated models, Bayesian information criterion (BIC) and Akaike information criterion (AIC) measures are used. These measures can be computed as follows:

$$BIC = -2\ln(\mathrm{L}) + \mathrm{K}\ln(\mathrm{Q}) \tag{2.7}$$

$$AIC = 2\mathrm{K} - 2\ln(\mathrm{L})$$

where $ln(L)$ is the log-likelihood value at convergence, K is the number of parameters and Q is the number of observations.

The model with the lower BIC and AIC values is the preferred model. The computed BIC and AIC values along with the log-likelihood at the convergence and number of parameters estimated for all the models are presented in Table 2.3. The BIC (AIC) values for the final specifications of the HP and HNB models clearly indicate that the HNB model shows superior fit compared to the HP models for all four models. Therefore, in explaining the effect of exogenous variable, we will restrict ourselves to the discussion of the HNB models.

Table 2.3 - Fit measures of the estimated demand models

Models	**Econometric framework**	**Log-likehood at convergence**	**Number of parameters**	**BIC**	**AIC**
Pedestrian generator model	**HP**	-933160.513	16	1866456.470	1866353.026
	HNB	-845920.147	17	*1691984.204*	*1691874.294*
Pedestrian attractor model	**HP**	-924530.467	21	1849238.705	1849102.934
	HNB	-835125.469	22	*1670437.174*	*1670294.938*
Bicycle generator model	**HP**	-113462.794	15	227052.567	226955.588
	HNB	-112380.003	16	*224895.451*	*224792.007*
Bicycle attractor model	**HP**	-109786.243	21	219750.256	219614.485
	HNB	-109381.323	22	*218948.883*	*218806.647*

2.2.4 Pedestrian Trip Demand Models

Table 2.4 presents the estimation results of the pedestrian generator and attractor models. The pedestrian generator model results are presented in 2nd and 3rd columns of Table 2.4 and pedestrian attractor model results are presented in 4th and 5th columns of Table 2.4. In the Hurdle model, the positive (negative) coefficient in the probabilistic component corresponds to increased (decreased) propensity of zero trip events. On the other hand, the positive (negative) coefficient in the count component of the Hurdle model corresponds to increased (decreased) non-zero trip count events. The final specification of the model was based on removing the statistically insignificant variables in a systematic process based on statistical significance (90% significance level) and intuitive coefficient effect. In estimating the models, several functional forms and variable specifications are explored. The functional form that provided the best result is used for the final model specifications and, in Table 2.4, the variable definitions are presented based on these final functional forms of variables. The effects of exogenous variables in model specifications for both pedestrian generator and attractor models are discussed in this section by variable groups.

Table 2.4 - Estimation results of pedestrian demand models

Variable name	Pedestrian generator model		Pedestrian attractor model	
	Estimates	t-stat	Estimates	t-stat
Probabilistic component				
Constant	2.346	55.615	2.319	54.774
Land-use mix	0.605	8.143	0.539	7.212
Urban area	0.224	37.315	0.215	35.200
Number of Household	0.212	27.324	0.228	29.528
Count component				
Constant	-0.217	-27.198	-0.422	-57.616
Sociodemographic characteristics				
Proportion of 65+ aged population	0.802	62.096	--	--
Roadway and traffic attributes				
Average zonal speed	-0.008	-59.952	--	--
AADT	-0.035	-31.141	-0.047	-40.822
Proportion of arterial road	0.320	53.077	0.255	43.828
Proportion of 3 and more lane road	-0.316	-32.398	-0.420	-39.923
Length of sidewalk	0.048	48.038	0.030	31.668
Built environment				
Number of business center	--	--	0.158	10.811
Number of entertainment center	--	--	0.194	14.437
Number of financial center	--	--	0.021	17.835
Number of park and recreational center	--	--	0.099	38.188
Number of restaurant	--	--	-0.022	-27.858
Number of shopping center	--	--	0.032	46.627
Number of transit hub	--	--	-0.057	-10.832
Land-use characteristics				
Industrial area	-0.029	-22.989	-0.055	-42.162
Recreational area	0.070	70.274	0.042	38.617

Residential area	0.060	57.244	0.062	55.280
Retail/office area	0.049	40.450	0.037	25.914
Institutional area	0.126	110.646	0.146	124.131
Overdispersion parameter	0.917	116.574	0.826	110.526

Probabilistic Component: In the probabilistic component, land-use mix, urban area and number of households are found to be significant in both pedestrian generator and attractor models. As expected, these variables are positively correlated with the propensity of non-zero pedestrian demand. As these variables serve as surrogates for pedestrian activity, it is expected that TAZs with higher levels of these variables are likely to be associated with pedestrian generator and attractor.

Count Component:

Sociodemographic characteristics: With respect to sociodemographic characteristics, from Table 2.4 we can see that proportion of 65+ aged population is positively associated with pedestrian generator, indicating that TAZs with higher number of population aged 65+ have higher pedestrian origin demand.

Roadway and Traffic Attributes: Zones with higher average speed limit of roadways are likely to generate less pedestrian origin demand. Annual average daily traffic (AADT) is negatively associated with both pedestrian demand components, indicating lower pedestrian activities in the zones with higher vehicular traffic. From Table 2.4, we can see that zones with a higher proportion of arterial roads are likely to have a higher level of pedestrian activities, both in terms of pedestrian activity generation and attraction. A higher proportion of roadways with 3 or more lanes is negatively associated with zonal level pedestrian activities. As expected, zones with higher sidewalk length are likely to have a higher level of pedestrian activities – both generation and attraction.

Built Environment: Built environment attributes are considered only in pedestrian attractor models as these attributes are more likely to attract pedestrians. With respect to built environment, we find that higher numbers of business centers, entertainment centers, financial centers, park/recreational centers and restaurants are positively associated with pedestrian attraction demand. On the other hand, higher numbers of shopping centers and transit hubs are found to be negatively associated with pedestrian destination demand at the zonal level.

Land-Use Characteristics: Land-use characteristics are found to have significant influence in both pedestrian generator and attractor demand models. Among different land-use categories, industrial area is found to be negatively associated with both pedestrian origin and destination demands. All other land-use categories (recreational, residential, retail/office and institutional area) are likely to generate higher levels of pedestrian demands.

2.2.5 Bicycle Trip Demand Model

Table 2.5 presents the estimation results of the bicycle generator and attractor models. The bicycle generator model results are presented in the 2nd and 3rd columns of Table 2.5, and bicycle attractor model results are presented in the 4th and 5th columns of Table 2.5. In the Hurdle model, the positive (negative) coefficient in the probabilistic component corresponds to increased (decreased) propensity of zero trip events. On the other hand, the positive (negative) coefficient in the count component of the Hurdle model corresponds to increased (decreased) non-zero trip count events. The final specification of the model was based on removing the statistically insignificant variables in a systematic process based on statistical significance (90% significance level) and intuitive coefficient effect. In estimating the models, several functional forms and variable specifications are explored. The functional form that provided the best result is used for the final model specifications and, in Table 2.5, the variable definitions are presented based on these final functional forms of variables. The effects of exogenous variables in model specifications for both bicycle generator and attractor models are discussed in this section by variable groups.

Table 2.5 - Estimation results of bicycle demand models

Variable name	Bicycle generator model		Bicycle attractor model	
	Estimates	**t-stat**	**Estimates**	**t-stat**
PROBABILISTIC COMPONENT				
Constant	-0.197	-3.661	-0.339	-6.208
Land-use mix	0.596	8.187	0.719	9.832
Urban area	0.305	38.242	0.300	36.621
Number of household	0.287	25.106	0.304	26.455
COUNT COMPONENT				
Constant	-2.351	-69.340	-1.974	-70.397
Sociodemographic characteristics				
Proportion of 65+ aged population	-0.546	-12.745	--	--
Roadway and traffic attributes				
AADT	-0.028	-8.577	--	--
Proportion of arterial road	0.095	6.921	0.044	3.473
Proportion of 3 and more lane road	-0.740	-33.999	-1.243	-55.656
Length of sidewalk	0.052	16.866	0.049	15.968
Built environment				
Number of commercial center	--	--	-0.416	-29.226
Number of educational center	--	--	0.112	21.645
Number of entertainment center	--	--	2.941	23.494

Number of financial center	--	--	-0.144	-43.018
Number of park and recreational center	--	--	0.339	54.894
Number of restaurant	--	--	0.225	73.716
Number of shopping center	--	--	-0.098	-36.605
Number of transit hub	--	--	0.260	23.207
Land-use characteristics				
Industrial area	0.092	31.510	0.052	17.338
Recreational area	0.016	6.847	-0.057	-23.155
Residential area	0.440	82.309	0.361	74.286
Retail/office area	-0.127	-39.940	-0.191	-53.656
Institutional area	0.041	12.410	0.032	9.903
Overdispersion parameter	3.081	26.618	6.009	20.365

Probabilistic Component: In the probabilistic component, land-use mix, urban area and number of households are found to be significant in both bicycle generator and attractor models. As expected, these variables are positively correlated with the propensity of non-zero bicycle demand. As these variables serve as surrogates for bicycle activity, it is expected that TAZs with higher levels of these variables are likely to be associated with bicycle generator and attractor.

Count Component:

Sociodemographic characteristics: With respect to sociodemographic characteristics, from Table 2.5 we can see that proportion of 65+ aged population is negatively associated with bicycle generator, indicating that TAZs with a higher number of population aged 65+ have lower bicycle origin demand.

Roadway and Traffic Attributes: AADT is negatively associated with bicycle generator demand component, indicating lower bicycle origin demand in the zones with higher vehicular traffic. From Table 2.5, we can see that zones with a higher proportion of arterial roads are likely to have higher level of zonal-level bicycle activities, both in terms of bicycle activity generation and attraction. A higher proportion of roadways with 3 or more lanes is negatively associated with zonal-level bicycle activities. Zones with higher sidewalk lengths are likely to have higher levels of bicycle activities – both generation and attraction, perhaps indicating that in Central Florida bicyclists use sidewalks as well as roads for biking.

Built Environment: Built environment attributes are considered only in bicycle attractor models as these attributes are more likely to attract bicyclists. With respect to built environment, we find that higher numbers of education centers, entertainment centers, park/recreational centers, restaurants and transit hubs are positively associated with bicycle attraction demand. On the other hand, higher numbers of commercial centers, financial centers and shopping centers are found to be negatively associated with bicycle destination demand at the zonal level.

2.2.6 Validation Results of Mobility Component

To demonstrate the predictive performance of the estimated pedestrian and bicycle demand models, a validation experiment is also carried out. The most common approach to perform a validation exercise for an aggregate-level model is to evaluate the in-sample predictive measures. To evaluate the in-sample goodness-of-fit measures, we computed the predicted count events for both zero and non-zero events and compared those with the observed values. These measures are presented in Table 2.6 below. From Table 2.6 we can see that the error between observed and predicted values are marginal, and hence we can argue that the predictive performance of the estimated models is reasonable for all four estimated demand models.

Table 2.6 - Predictive performance evaluation

Models	Events	Observed	Predicted
Pedestrian generator model	**Total zones with zero trip count**	4007.00	4006.80
	Total number of zonal trips	1260090.60	1255479.90
	Average zonal trips	265.45	264.48
Pedestrian attractor model	**Total zones with zero trip count**	4010.00	4010.49
	Total number of zonal trips	1242270.50	1236690.70
	Average zonal trips	261.70	260.52
Bicycle generator model	**Total zones with zero trip count**	4574.00	4573.82
	Total number of zonal trips	166248.45	165671.36
	Average zonal trips	35.02	34.90
Bicycle attractor model	**Total zones with zero trip count**	4581.00	4581.18
	Total number of zonal trips	165845.77	171959.97
	Average zonal trips	34.94	36.22

2.3 Non-Motorist Destination Choice Model

In this section, we examine non-motorist destination choice at a trip level to analyze destination zone preferences of pedestrians and bicyclists. Specifically, we examine the zonal attributes that influence the decision process of non-motorists in identifying destination locations after starting a trip from any particular zone. The decision process formulated and employed in the current study approach for the non-motorist destination choice is analogous to the destination choice evaluation framework of vehicular demand within the traditional travel demand modeling approach. In non-motorist trip destination choice models, given the origin location of trips, a quantitative model framework is employed to identify the possible destination of the trips. In developing these destination choice models, we explore the influence of built environment, roadway and land use characteristics at the potential destinations accessible from the origin on the decision process of choosing a destination in a trip. We estimate two different models: (1) pedestrian destination choice model, and (2) bicycle destination choice model. These models are developed by using a random utility maximization approach where the trip maker chooses the destination that offers the highest utility from the universal choice set of the destination zone and is estimated as a function of destination zone attributes. The random utility framework employed in the current study takes the form of a multinomial logit (MNL) model. The MNL models are estimated at the trip level for the CFRPM 6.0 study region employing a comprehensive set of exogenous variables. Based on the model results, we identify important exogenous variables that influence non-motorist destination preferences.

2.3.1 Model Framework

In the current research effort, we assume a random utility-based framework (MNL model) for modeling the non-motorist destination choice models (following McFadden [4]). The MNL model is widely used in existing transportation literature to study location choice and in related literature [5-8]. In this section, we explain the econometric framework of the MNL model employed in the current study.

Let j $(j = 1,2,3, \dots, J)$ be the index to represent a destination zone among a set of C_i alternatives of trip i. Thus, the destination choice takes the familiar discrete outcome formulation as the linear function as follows:

$$u_{ij}^{*} = \left(\boldsymbol{\delta z}_{ij} + \xi_{ij}\right) \tag{2.8}$$

where u_{ij}^{*} is the latent variable of destination choice for trip i with alternative j.

Within the traditional random utility maximization-based discrete outcome framework as presented in Equation 2.8, trip i will have the possibility of a destination in zone j if $u_{ij}^{*} > \max_{\substack{d=1,2,3,\dots,J \\ d \neq j}} u_{iJ}^{*}$. $\boldsymbol{z}_{ij}$ is a vector of destination zonal attributes corresponding to destination zone j. $\boldsymbol{\delta}$ is a vector of coefficients to be estimated. ξ_{ij} is an idiosyncratic error term assumed to be identically and independently standard logistic distributed across trip i with destination j. thus, the probability of trip i representing the destination choice of trip makers takes the typical MNL form given by:

$$R_{ij} = \frac{exp(\boldsymbol{\delta z}_{ij})}{\sum_{j \in C_i} exp(\boldsymbol{\delta z}_{ij})} \tag{2.9}$$

Finally, the weighted log-likelihood function is:

$$LL = \omega_i * \left(\sum_i Ln(R_{ij})\right) \tag{2.10}$$

The daily trip weight is generated by using the following formulation:

$$\omega_i = \frac{Yearly\ person\ trip\ weight}{365} \tag{2.11}$$

The reader should note that in computing the weighting factor, as presented in Equation 2.11, we divided the yearly person trip factor, as obtained from NHTS data, by 365 to convert the yearly trip count to a daily trip count. All the parameters in the model are estimated by maximizing the logarithmic function LL presented in Equation 2.10. The maximum likelihood model estimation is programmed in the GAUSS matrix programming language.

2.3.2 Dependent Variable and Data Description

The non-motorist trip destination choice model is focused on non-motorist destination zonal choice. We examine two trip-level destination choices for walk and bike modes by employing a random utility-based model. We generate the destination choice set by assuming that people will walk up to 2 miles and bike up to 6 miles in a trip. Thus, the destination choice set is identified based on generating a 2-mile and 6-mile buffer around the trip origin location for pedestrian and bike trips, respectively. However, on any occasion, if non-motorists walk or bike beyond these limits, we accommodate those alternatives in our choice set as well. Thus, the associated data records for pedestrian destination choice model are identified to be 56,223 with 1895 trips (unweighted). At the same time, the associated data records for the bicycle destination choice model are identified to be 39,585 with 276 trips (unweighted).

In the current research effort, the destination choice models for non-motorist trips are examined by considering the zonal attributes of the identified destination zones. Table 2.7 offers a summary of the sample characteristics of the exogenous factors in the final estimation dataset for the pedestrian and bicycle destination models. For the empirical analysis, the selected explanatory variables can be grouped into four broad categories: sociodemographic characteristics, roadway and traffic attributes, built environment and land use characteristics. The table represents the definition of variables considered for final model estimation along with the zonal minimum, maximum and average values. It is worthwhile to mention here that in our destination choice model specifications we did not consider person-level or trip-level attributes. One of the major objectives for developing the destination choice models is to forecast and/or evaluate policy implications for the future year considering real-world changes. However, for such analysis by employing destination choice models with person- and trip-level attributes would require scenario-level person and trip attributes for the future year, which are not readily available. Therefore, we examine the destination choice of non-motorists by employing zonal-level attributes only.

Table 2.7 – Summary characteristics for destination choice models

Variable name	Description	Pedestrian	Bicycle
		Mean	**Mean**
Sociodemographic characteristics			
Population density	Total number of Population of TAZ/ Area of TAZ in acre	3.197	3.301

Proportion of 22-29 aged population	Total number of population who are 22 to 29 years old of TAZ/ Total number of Population of TAZ	0.056	0.108
Proportion of people aged 18 to 21	Total number of population who are 18 to 21 years old of TAZ/ Total number of Population of TAZ	---	---
Proportion of people aged 65+	Total number of people above 65 years old of TAZ/ Total number of Population of TAZ	0.193	---
Roadway and traffic attributes			
Length of bike lane	Total bike length in meter of TAZ	---	0.310
Average zonal speed	Mean maximum speed in mph	34.579	35.646
Traffic signal	Number of traffic signal in TAZ	0.479	---
AADT	Total Annual Average Daily Traffic (AADT) of TAZ/10000	1.095	---
Truck AADT	Total Truck AADT of TAZ/10000	0.089	---
Built environment			
Number of commercial center	Total number of commercial center of TAZ	0.084	0.096
Number of educational center	Total number of educational center of TAZ	0.330	0.387
Number of financial center	Total number of financial center of TAZ	0.771	0.844
Number of restaurant	Total number of restaurant of TAZ	1.690	2.020
Number of shopping center	Total number of shopping center of TAZ	1.783	2.120
Number of transit hub	Total number of transit hub of TAZ	0.056	0.069
Land-use Characteristics			
Urban area	Ln (Urban area in a TAZ in acre)	4.950	4.955
Residential area	Ln (Residential area in a TAZ in acre)	3.590	3.670
Industrial area	Ln (Industrial area in a TAZ in acre)	0.586	0.683
Recreational area	Ln (Recreational area in a TAZ in acre)	0.404	0.344

Institutional area	Ln (Institutional area in a TAZ in acre)	0.813	0.897
Retail/Office area	Ln (Office/Retail area in a TAZ in acre)	1.856	1.970

The final specification of the model development was based on removing the statistically insignificant variables in a systematic process based on statistical significance (90% significance level). The specification process was also guided by prior research and parsimony considerations. In estimating the models, several functional forms and variable specifications were explored. The functional form that provided the best result was used for the final model specifications and, in Table 2.7, the variable definitions are presented based on these final functional forms.

2.3.3 Estimation Results

Table 2.8 presents the estimation results of the pedestrian and bicycle destination choice models. The pedestrian destination choice model results are presented in the 2nd and 3rd columns of Table 2.8, and the bicycle destination choice model results are presented in the 4th and 5th columns of Table 2.5. In the MNL model, the positive (negative) coefficient corresponds to increased (decreased) likelihood of destination choice. The effects of exogenous variables in model specifications for both the pedestrian and bicycle destination choice models are discussed in this section by variable groups.

Table 2.8 - Estimation results of destination choice models

Variable name	Pedestrian		Bicycle	
	Estimates	t-stat	Estimates	t-stat
Sociodemographic characteristics				
Population density	0.116	237.353	-0.043	-32.367
Proportion of 22-29 aged population	---	---	2.506	40.419
Proportion of people aged 18 to 21	-0.148	-7.287	---	---
Proportion of people aged 65+	1.757	182.073	---	---
Roadway and traffic attributes				
Length of bike lane	---	---	0.009	12.570
Average zonal speed	-0.001	-9.327	0.010	56.096
Traffic signal	0.079	66.305	---	---
AADT	-0.019	-14.721	---	---
Truck AADT	-0.715	-55.190	---	---

Built environment				
Number of commercial center	0.299	143.347	0.072	13.207
Number of educational center	0.078	61.863	0.265	91.467
Number of financial center	0.013	23.002	0.130	87.656
Number of restaurant	0.044	122.330	-0.054	-44.092
Number of shopping center	-0.003	-13.470	0.010	11.786
Number of transit hub	0.153	70.265	0.312	80.481
Land-use Characteristics				
Urban area	0.247	230.929	0.217	92.081
Residential area	0.147	187.598	0.402	171.613
Industrial area	-0.022	-35.855	-0.119	-82.827
Recreational area	-0.016	-29.005	0.068	49.466
Institutional area	0.117	178.904	0.025	14.561
Retail/Office area	0.060	84.343	0.009	5.244
Log-likelihood at convergence	-3705303.725		-750448.472	

Sociodemographic Characteristics: From the destination choice model of pedestrian trip, we can see that trip makers tend to choose zones as destinations with higher population density as highlighted by the positive coefficient of the population density variable in the pedestrian destination choice model. On the contrary, as the population density increases at the zonal level, it is less likely that bikers choose those zones as destination of their trip. For bike riders, the proportion of zonal-level population aged 22-29 years has a positive impact on the destination choice utility. For pedestrians, the negative impact associated with the proportion of population aged 18-21 years on the likelihood of choosing a zone indicates lower utility. The estimate for proportion of population aged 65 and over has a positive coefficient, suggesting that presence of more aged population in zones is likely to incur higher utility for pedestrian trip destination.

Roadway and Traffic Attributes: Several roadway and traffic attributes are found to be significant determinants of non-motorist destination choice. As expected, bikers are more likely to choose zones in the proximity of trip origin with higher lengths of bike lane. It is surprising to note that the effect of average zonal speed has a negative impact on pedestrian destination choice, while the variable has a positive impact on the bicycle destination choice model. Pedestrians are likely to choose zones that have higher traffic signal density. The zones with higher AADT and truck AADT tend to be chosen less as destinations by pedestrians.

Built Environment: With respect to built environment, the coefficients of number of commercial centers, educational centers, financial centers, and transit hubs in both pedestrian and bicycle destination choice models demonstrate a higher likelihood of choosing zones with these attributes for non-motorist trips. Number of restaurants has a positive coefficient in the model for pedestrians, while the variable has a negative impact on bicycle destination choice. Pedestrians are likely to choose destination zones with higher numbers of shopping center, while bikers are likely to do the opposite, as indicated by the positive sign of the variable for bikers.

Land-use Characteristics: From Table 2.8, we can see that different land use characteristics have similar effects (other than recreational areas) in both pedestrian and bicycle destination choice models. The likelihood of choosing a zone as a destination for a non-motorized trip is higher when there are more urbanized, residential, institutional and retail/office areas. With respect to industrial area, people are less likely to walk or bike to the zones with more industrial areas. The effect of recreational area has opposing signs in pedestrian and bicycle models, indicating lower and higher destination choice utility for pedestrians and bikers, respectively.

2.4 Non-Motorist Trip Exposure Matrices

In evaluating non-motorist exposure, we also generate different zonal-level trip exposure matrices with the number of daily trip origins and daily trip destinations at the zonal level for both the pedestrian and bicycle modes. Specifically, four different zonal-level exposure matrices are generated: 1) trip O-D demand matrices, 2) trip origin demand matrices, 3) trip destination demand matrices and 4) total trip demand matrices. These matrices are generated for pedestrian and bicycle modes separately for the 4,747 TAZs in the area defined by the Central Florida region. The procedure for generating these matrices along with the summary reports are discussed in this section.

Trip origin-destination (O-D) demand matrices: Zonal-level trip O-D demand matrices are generated by using predictions from destination choice models as presented in Section 2.3. Specifically, we micro-simulate and assign the total number of trips originated in different zones to different destination zones by using the estimated probability shares identified from the destination choice models. Thus, the dimension of the generated O-D demand matrices are $[4747 \times 4747]$ with origin zones in the rows, destination zones in the column and the number of trips destined across different zones in each cell. The O-D demand matrices are generated for pedestrian and bicycle modes separately. For brevity, we present a snapshot of the O-D demand matrices in Table 2.9.

Table 2.9 – Trip O-D matrices example

Pedestrian												
Origin TAZ	**Destination TAZ**											
	1	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**…**	**13**
1	169.52	0	0	0	0	0	0	0	0	0	…	365.06
2	0	0	0	0	0	0	0	0	0	0	…	0
3	8.15	9.35	3.56	7.94	0	0	0	0	0	0	…	17.56
4	0	0	0	0	0	0	0	0	0	0	…	0

5	0	0	0	0	0	0	0	0	0	0	…	0
6	0	0	0	0	0	0	0	0	0	0	…	0
7	0	0	0	0	0	0	0	0	0	0	…	0
8	0	0	0	0	0	0	0	0	0	0	…	0
9	0	0	0	0	0	0	0	0	0	0	…	0
10	0	0	0	0	0	0	0	0	0	0	…	0
…	…	…	…	…	…	…	…	…	…	…	…	0
13	172.17	0	0	0	0	0	0	0	0	0	0	370.77
Bicycle												
Origin TAZ	**Destination TAZ**											
	1	**2**	**3**	**4**	…	**13**	**14**	**15**	**16**	**17**	**18**	**19**
1	0	0	0	0	…	0	0	0	0	0	0	0
…	…	…	…	…	…	…	…	…	…	…	…	**…**
62	18.43	8.29	2.06	9.32	…	26.22	4.28	0.93	0.51	2.22	3.2	1.96
…	…	…	…	…	…	…	…	…	…	…	…	**…**
70	33.54	15.09	0	0	…	47.71	7.79	1.7	0.92	4.04	0	0
…	…	…	…	…	…	…	…	…	…	…	…	**…**
77	0	0	0	0	…	23.82	0	0	0	0	0	0
…	…	…	…	…	…	…	…	…	…	…	…	…
109	0	0	0	0	…	0	0	0	0	0	0	0
…	…	…	…	…	…	…	…	…	…	…	…	…
112	0	0	0	0	…	0	0	0	0	0	0	0

Trip origin demand matrices: Zonal-level trip origin demand matrices are computed by using predictions from non-motorist generator models as presented in Section 2.2, which are further used to generate the trip origin matrices for the pedestrian and bicycle trip modes. Thus, the dimensions of the generated trip origin demand matrices are $[4747 \times 1]$ with origin trip counts across different rows. The origin demand matrices are generated for the pedestrian and bicycle modes separately.

Trip destination demand matrices: Zonal-level trip destination demand matrices are computed by using predictions from non-motorist attractor models as presented in Section 2.2, which are further used to generate the trip destination matrices for the pedestrian and bicycle trip modes. Thus, the dimension of the generated trip destination demand matrices are $[4747 \times 1]$ with destination trip counts across different rows. The destination demand matrices are generated for the pedestrian and bicycle modes separately. It is worthwhile to mention here that we can also generate destination demand matrices by using O-D demand matrices. In doing so, we should follow two steps: 1) summing up trip counts of the O-D demand matrices across columns and 2) transposing the generated row matrices to generate the column matrices, generating trip destination demand matrices with dimension $[4747 \times 1]$ of destination trip counts across different rows. In the current research effort, we resort to generating trip destination demand matrices by using information from trip attractor models as it involves only one step.

<u>Total trip demand matrices:</u> Finally, zonal-level total trip demand matrices are generated by combining the trip origin and destination demand matrices across different zones (total trip demand = trip origin demand + trip destination demand). Thus, the dimensions of the generated total trip demand matrices are $[4747 \times 1]$ with total trip counts across different rows. The total trip demand matrices are generated for the pedestrian and bicycle modes separately.

<u>Summary Report:</u> For representation purposes, the summary report for trip origin, destination and total trip demands are presented at the county level. In Table 2.10 we present the county-level trip origin, trip destination and total trip demand matrices for the pedestrian and bicycle modes. The locations of these counties are presented in Figure 2.2.

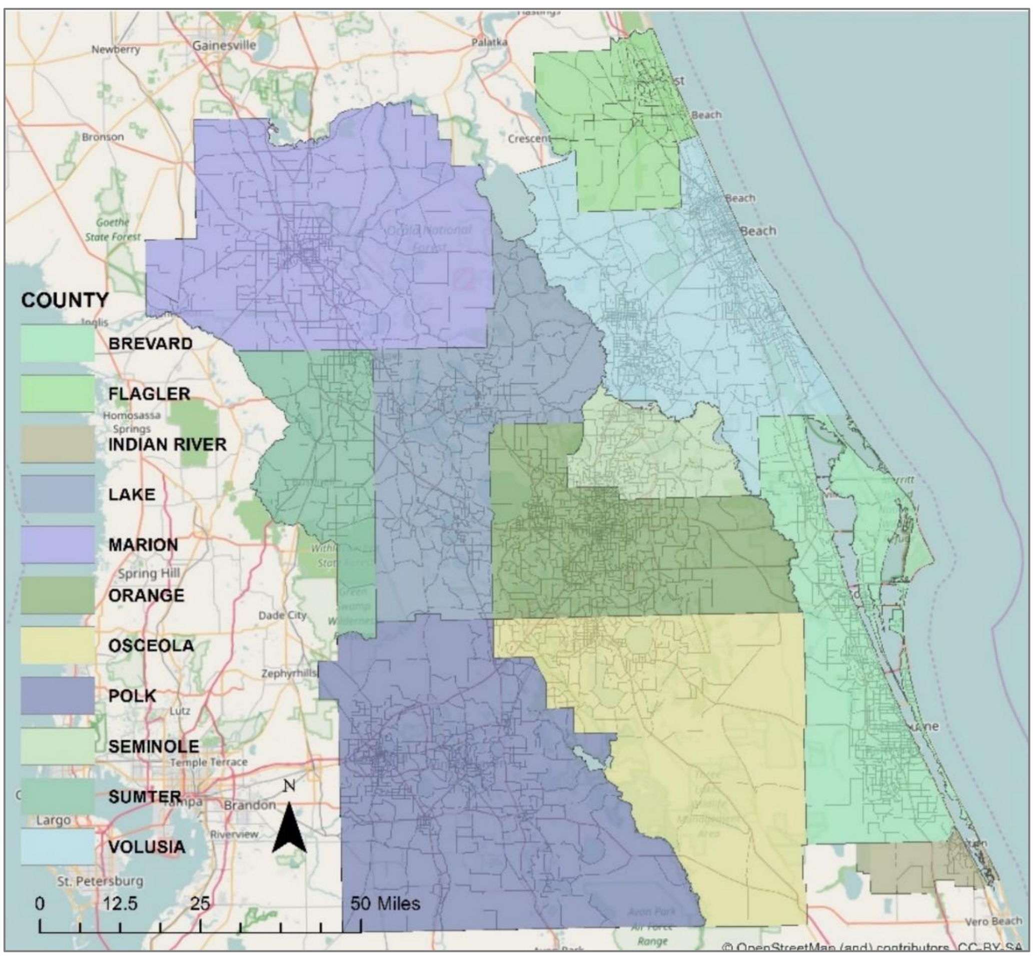


Figure 2.2 – County locations in Central Florida region

Table 2.10 – Trip demand matrices by county level

County	No. of TAZs	Pedestrian			Bicycle		
		Trip origin demand	Trip destination demand	Total trip demand	Trip origin demand	Trip destination demand	Total trip demand
Brevard	692	154936.5	149804.8	304741.3	21663.59	23172.9	44836.49
Flagler	141	26241.46	23153.66	49395.12	2940.338	2634.027	5574.365
Indian River	37	12066.78	11826.16	23892.94	1735.289	999.454	2734.743
Lake	350	67309.28	66545.88	133855.2	10784.29	9977.642	20761.94
Marion	422	95199.85	89602.94	184802.8	5238.246	4226.254	9464.501
Orange	781	348163.9	355169.8	703333.7	57661.94	64084.73	121746.7
Osceola	250	67651.62	65181.71	132833.3	4026.134	3875.623	7901.758
Polk	621	185959.9	195543.4	381503.4	10931.12	10687.68	21618.8
Seminole	230	75690.14	79212.17	154902.3	12179.38	11558.89	23738.27
Sumter	147	32272.77	26598.91	58871.68	553.048	817.907	1370.955
Volusia	1076	189987.7	174051.2	364038.8	37957.98	39924.86	77882.84
Total	4747	1255480	1236691	2492171	165671.4	171960	337631.3

2.5 Non-Motorist Trip Exposure Measures for Safety Evaluation

In the current research effort, our objective is to evaluate non-motorist safety at a planning level. To that extent our focus is on examining pedestrian and bicycle crash risk (in terms of total crashes and crashes by injury severity levels) at the zonal level to evaluate critical planning-level factors. By using the outcome of these models, we can identify medium- and long-term safety improvement strategies to encourage non-motorist travel. As discussed in Section 1, aggregate-level non-motorist safety is likely to be influenced by non-motorist exposure along with other aggregate-level attributes. Hence, we select an exposure measure identified in Section 2.4 to be used as non-motorist exposure measures in examining their safety at the zonal level. Specifically, we consider zonal-level total trip demand matrices, as presented in Table 2.10, as exposure measures for further safety evaluation. The decision process of incorporating exposure measures in examining aggregate-level non-motorists crash and severity risks is presented in Figure 2.3 below.

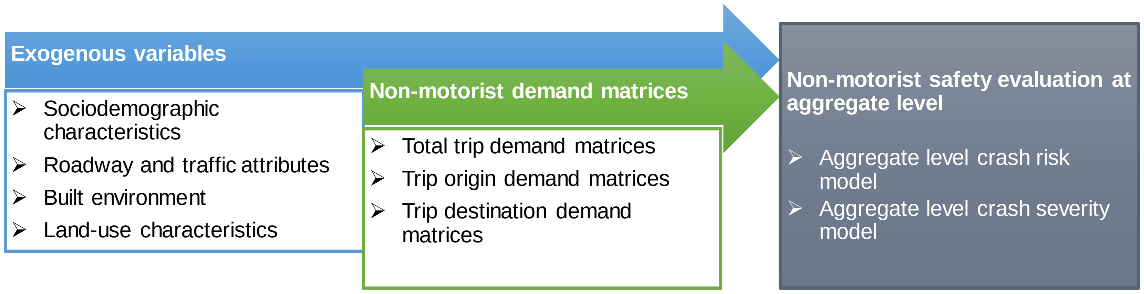


Figure 2.3 – Non-motorist safety evaluation road map

3 Non-Motorized Road User Safety Evaluation

Among the different modes of transportation, active forms such as walking and bicycling are the most sustainable, leaving the lowest carbon footprint on the environment. These modes also contribute to improving the physical health of non-motorists. However, non-motorist safety is a global health concern, and the United States of America is no exception. Several previous studies have revealed that the possibility of being involved in a collision and the risk of injury or fatality is higher for pedestrians and bicyclists than for motorists [9-10]. Thus, the safety concerns remain a detriment for walking or biking. As a consequence, these transport alternatives have the lowest mode shares, specifically in North American communities where personal automobiles are the most common mode of transportation [11]. However, the transportation decision makers of developed countries are proactively encouraging the adoption of these active forms of transportation for short-distance trips given the growing concern of worsening global climate change and increasing obesity among adults. For increasing the adoption of active transportation, there is a need to reduce the risk to pedestrians and bicyclists on roadways. Any effort to reduce the social burden of these crashes and enhance non-motorist safety would necessitate the examination of factors that contribute significantly to crash likelihood and/or severity outcomes in the event of a crash and the implementation of policies that enhance safety for pedestrians and bicyclists. Important tools for identifying and evaluating road safety policies are forecasting and policy evaluation, which are predominantly devised through evidence-based and data-driven safety analysis.

Traditionally, transportation safety analysis by using crash records has evolved along two major streams: crash frequency analysis and crash severity analysis. Crash frequency or crash prediction analysis is focused on identifying attributes that result in traffic crashes and proposing effective countermeasures to improve the roadway design and operational attributes (Lord and Mannering [12] offer a review of these studies). The crash frequency models study aggregate information, such as total number of crashes at an intersection or at a spatial aggregation level (zone or tract level), and are developed by using non-crash-specific data. On the other hand, crash severity analysis is focused on examining crash events, identifying factors that impact the crash outcome and providing recommendations to reduce the consequences (injuries and fatalities) in the unfortunate event of a traffic crash (see references for reviews [13, 14]). The crash severity models are developed by using detailed post-crash data and are quite disaggregate in nature because they consider every crash as a record for model development. In evaluating the impact of a safety measure, crash frequency analysis forecasts the change in crash occurrences, whereas crash severity analysis forecasts the change in crash consequences (injuries and fatalities).

The disaggregate-level crash severity analysis is focused on examining crash events. It is worthwhile to mention here that these models cannot be directly employed to incorporate safety considerations in the transportation planning process. The outcomes of aggregate-level crash count models, specifically macro-level models, can be used to devise safety-conscious decision support tools to facilitate proactive approach in assessing medium- and long-term policy-based countermeasures. Moreover, the tool plays an important role in safety implications of land use planning initiatives and alternate network-planning initiatives. Therefore, aggregate-level crash count models are more feasible for planning-level policy analysis and identifying policy measures.

To that extent, in this research effort, we estimate four different sets of aggregate-level models: (1) zonal-level crash count model for examining pedestrian–motor vehicle crash occurrences, (2) zonal-level crash count model for examining bicycle–motor vehicle crash occurrences (3) zonal-level crash severity model for examining pedestrian crash injury severity by proportions and (4) zonal-level crash severity model for examining bicycle crash injury severity by proportions. These models are estimated for the study area defined by CFRPM 6.0 by using crash records of the base year 2010. In the following sections, we present the outcomes of these models.

3.1 Crash Frequency Analysis

A regional- or zonal-level safety planning tool can be devised by using a macro-level study and hence is useful not only for planners but also for decision-makers. Therefore, it is important to investigate zonal-level pedestrian and bicycle crashes to identify critical factors and propose implications to facilitate proactive safety-conscious planning. In this current research effort, we formulate and estimate count models for examining pedestrian and bicycle crash risks. The count models are estimated at the TAZ level for the CFRPM 6.0 area employing a comprehensive set of exogenous variables. Based on the model results, we identify important exogenous variables that influence pedestrian and bicycle crash counts. The NB model, which offers a closed-form expression while relaxing the mean variance equality constraint of Poisson regression, serves as the workhorse for crash count modeling. Therefore, crash count models for examining pedestrian and bicycle crash events are developed in this study by using the NB modeling approach.

3.1.1 Model Framework

The focus of our study is to model pedestrian crash frequency and bicycle crash frequency at the zonal level by employing the NB modeling framework. The econometric framework for the NB model is presented in this section.

Let i be the index for TAZ $(i = 1,2,3, \dots, N)$ and y_i be the index for crashes occurring over a period of time in a TAZ i. The NB probability expression for random variable y_i can be written as:

$$P_i(y_i|\mu_i, \alpha) = \frac{\Gamma\left(y_i+\frac{1}{\alpha}\right)}{\Gamma(y_i+1)\Gamma\left(\frac{1}{\alpha}\right)}\left(\frac{1}{1+\frac{\mu_i}{\alpha}}\right)^{\frac{1}{\alpha}}\left(1-\frac{1}{1+\frac{\mu_i}{\alpha}}\right)^{y_i} \quad (3.1)$$

where $\Gamma(\cdot)$ is the Gamma function, α is the NB dispersion parameter, and μ_i is the expected number of crashes occurring in TAZ i over a given period of time.

We can express μ_i as a function of explanatory variable (x_i) by using a log-link function as: $\mu_i = E(y_i|x_i) = exp(\beta x_i)$, where β is a vector of parameters to be estimated. Finally, the log-likelihood function for the NB model can be written as:

$$LL = \sum_{i=1}^{N} log(P_i) \quad (3.2)$$

The parameters to be estimated in the model of Equation 3.2 are $\boldsymbol{\beta}$ and $\boldsymbol{\alpha}$. The parameters are estimated using maximum likelihood approaches.

3.1.2 Dependent Variable and Data Description

The crash frequency analysis is focused on pedestrian and bicycle crashes at the TAZ level for 4,747 TAZs in the area defined by the CFRPM 6.0 area. For this report, we have examined the pedestrian and bicycle crash count events for the year 2010 to reflect the base year situation in terms of non-motorized safety. For the year 2010, 1,474 (with 0, 9 and 0.31 zonal minimum, maximum and average, respectively) and 1,012 (with 0, 8 and 0.21 zonal minimum, maximum and average, respectively) crashes were reported involving pedestrians and bicycles, respectively. Spatial representation of these crashes at the zonal level is shown in Figure 3.1.

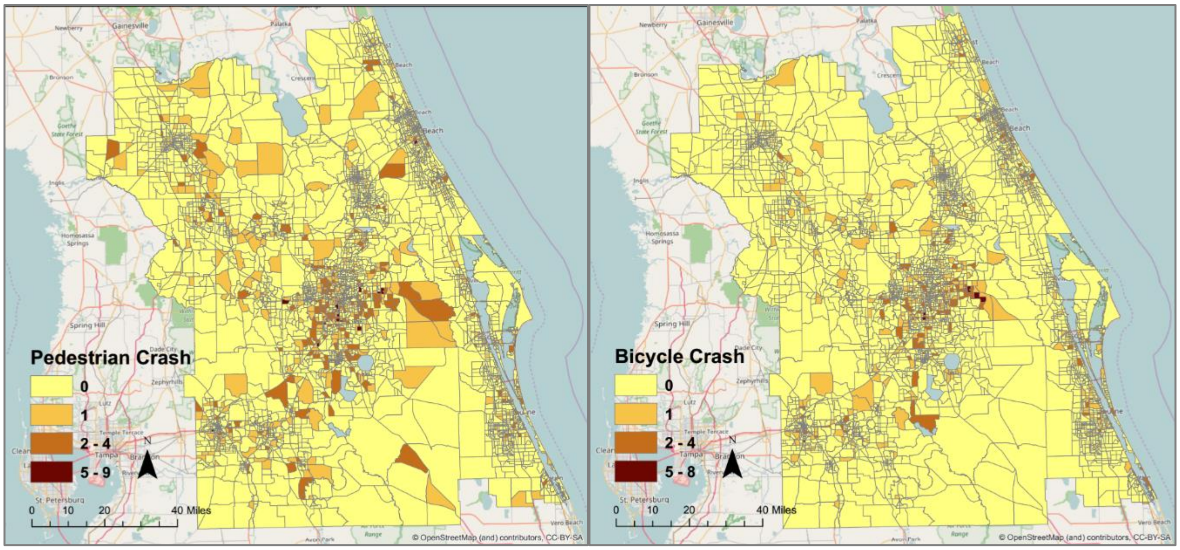


Figure 3.1 - Total number of pedestrian and bicycle crashes for the year 2010

In addition to the crash database, the explanatory attributes considered in the empirical study are also aggregated at the TAZ level accordingly. To reflect the base year characteristics of the analysis zone, all attributes are generated for the year 2010. For the empirical analysis, the selected explanatory variables can be grouped into five broad categories: sociodemographic characteristics, socioeconomic characteristics, roadway attributes, land use characteristics and exposure measures. Table 3.1 offers a summary of the sample characteristics of the exogenous variables and the definition of variables considered for final model estimation along with the zonal minimum, maximum and average.

Table 3.1 – Sample characteristics for crash frequency models

Variable name	Description	Zonal		
		Minimum	Maximum	Mean
Sociodemographic characteristics				
Population density	Total number of population of TAZ/ Area of TAZ in acre	0.000	19.956	2.366

Proportion of people aged 65+	Total number of people above 65 years old of TAZ/ Total number of population of TAZ	0.000	0.899	0.182
Roadway and traffic attributes				
Traffic signal density	Total number of traffic signals in TAZ	0.000	8.000	0.379
Proportion of arterial road	Total length of arterial road of TAZ/Total roadway length of TAZ	0.000	1.000	0.459
Proportion of local road	Total length of local road of TAZ/Total roadway length of TAZ	0.000	1.000	0.040
Length of sidewalk	Total sidewalk length in meter of TAZ	0.000	36.346	0.280
Length of bike lane	Total bike lane length of TAZ in meters	0.000	58.525	0.421
Length of bus lane	Total bus lane length of TAZ in kilometers	0.000	31.161	0.888
AADT	Total annual average daily traffic (AADT) of TAZ/10000	0.000	27.550	0.931
Truck AADT	Total truck AADT of TAZ/10000	0.000	2.747	0.083
Built environment				
Number of commercial centers	Total number of commercial centers of TAZ	0.000	4.000	0.087
Number of financial centers	Total number of financial centers of TAZ	0.000	17.000	0.586
Number of educational centers	Total number of educational center of TAZ	0.000	5.000	0.275
Number of transit hubs	Total number of transit hubs of TAZ	0.000	11.000	0.051
Number of restaurants	Total number of restaurants of TAZ	0.000	36.000	1.335
Number of parks and recreational centers	Total number of parks and recreational centers of TAZ	0.000	20.000	0.245

Number of hospitals	Total number of hospitals of TAZ	0.000	2.000	0.017
Land-use characteristics				
Urban area	Ln (Urban area in a TAZ in acres)	-9.275	8.491	4.291
Residential area	Ln (Residential area in a TAZ in acres)	-12.427	8.014	3.596
Recreational area	Ln (Recreational area in a TAZ in acres)	-13.946	10.040	0.388
Land-use mix	Land use mix = [(-∑k(Pk (lnPk)))/lnN], where k is the category of land-use, p is the proportion of the developed land area devoted to a specific land-use, N is the number of land-use categories in a TAZ	0.000	0.929	0.355
Exposure measures				
Total pedestrian trip demand per household	Total pedestrian daily trip demand in a TAZ/(Total number of household in a TAZ*100)	0.000	948.164	0.321
Total bicycle trip demand	Ln(Total bicycle daily trip demand in a TAZ)	0.000	9.549	0.259

3.1.3 Estimation Results

In this research effort, we estimate two different NB models: one model for pedestrian crash count events at the zonal level and another model for bicycle crash count events at the zonal level. Table 3.2 presents the estimation results of the NB models. The pedestrian crash count model results are presented in 2nd and 3rd columns of Table 3.2, and the bicycle crash count model results are presented in the 4th and 5th columns. The effects of exogenous variables in model specifications for both pedestrian and bicycle crash count models are discussed in this section by variable groups.

In NB models, the positive (negative) coefficient corresponds to increased (decreased) crash risk. The final specification of the model was based on removing the statistically insignificant variables in a systematic process based on statistical significance (90% significance level) and intuitive coefficient effect. In estimating the models, several functional forms and variable specifications are explored. The functional form that provided the best result is used for the final model specifications and, in Table 3.1, the variable definitions are presented based on these final functional forms of variables.

Table 3.2 - Estimation results of negative binomial models

Variable name	Pedestrian		Bike	
	Estimates	t-stat	Estimates	t-stat
Constant	-3.063	-22.318	-3.789	-23.884
Sociodemographic characteristics				
Population density	0.131	10.645	0.130	10.050
Proportion of people aged 65+	-1.401	-4.229	-0.979	-3.019
Roadway and traffic attributes				
Traffic signal density	0.223	6.001	0.146	3.994
Proportion of arterial road	0.325	3.723	0.341	3.619
Proportion of local road	---	---	-0.799	-2.241
Length of sidewalk	0.025	2.090	---	---
Length of bike lane	---	---	0.016	1.681
Length of bus lane	---	---	0.087	5.040
AADT	0.037	2.373	0.090	2.272
Truck AADT	---	---	-1.054	-2.510
Built environment				
Number of commercial centers	---	---	0.182	1.863
Number of financial centers	---	---	0.063	3.204
Number of educational centers	0.085	1.822	---	---
Number of transit hubs	0.254	5.506	---	---
Number of restaurants	0.086	9.055	0.052	5.135
Number of parks and recreational centers	0.123	3.173	---	---
Number of hospitals	---	---	0.307	3.143
Land-use characteristics				
Urban area	0.123	5.098	0.165	5.876
Residential area	0.041	2.076	0.082	3.736
Recreational area	---	---	-0.049	-2.222

Land-use mix	0.810	4.673	0.697	3.719
Exposure measures				
Total pedestrian trip demand per household	-0.277	-1.482	---	---
Total bicycle trip demand	---	---	0.042	2.055
Overdispersion parameter	1.004	9.297	0.641	5.642

Sociodemographic characteristics: With respect to sociodemographic characteristics, the estimates indicate that both pedestrian and bicycle crashes are positively associated with population density. At the same time, the results in Table 3.2 indicate a reduced crash propensity for both pedestrians and bicyclists with a higher proportion of population aged 65 and over.

Roadway and traffic attributes: Several roadway and traffic attributes are found to be significant determinants of pedestrian and bicycle crashes at the zonal level. The results associated with traffic signal density reveal that an increase in traffic signal density in a zone increases the likelihood of both pedestrian and bicycle crashes. A higher proportion of arterial road results in higher pedestrian and bicycle crash risks. At the same time, a higher proportion of local roads is found to have negative impact on bicycle crash risk. From Table 3.2, we can see that the likelihood of a pedestrian crash is higher in the zone with a higher sidewalk length. It is also surprising to note that TAZs with higher bicycle lane lengths have an increased likelihood of bicycle crashes. The result for length of zonal-level bus lanes reveals an increasing likelihood of bicycle crash. An increase in zonal AADT increases the likelihood of both pedestrian and bicycle crashes at the TAZ level. The result in the bicycle crash model suggests that zones with higher truck AADT have a decreased likelihood of bicycle crashes.

Built environment: With respect to built environment, the estimation results of the pedestrian crash risk model reveal that a higher number of educational centers, transit hubs, restaurants and parks/recreational centers results in a higher pedestrian crash risk at the zonal level. From the results of the bicycle crash risk models, we can see that bicycle crash risk is positively associated with a higher number of commercial centers, financial centers, restaurants and hospitals.

Land-use characteristics: Several land-use characteristics are found to be significant determinants of pedestrian and bicycle crash risks. Pedestrian and bicycle crash risks increase with increasing urbanized and residential areas. In the bicycle crash risk model, recreational area is found to decrease the likelihood of zonal-level bicycle crash risk. TAZs with higher land-use mix have increased propensity for both pedestrian and bicycle crashes.

Exposure measures: The non-motorist exposure measures generated from Section 2.4 are used in evaluating zonal-level pedestrian and bicycle crash risk. Specifically, we use the total daily trip demand of pedestrians and bicyclists as exogenous variables in pedestrian and bicycle crash risk models, respectively. We consider different functional forms of pedestrian and bicycle exposure measures in estimating NB models and the functional form that provides the best fit is considered in the final specifications. With respect to the pedestrian crash risk model, pedestrian exposure measures with any of the functional forms are not found to be significant at a 90% confidence level. However, pedestrian trip demand per household at a zonal level provides the best data fit and hence is considered in our final pedestrian crash risk model. From Table 3.2, we can see that a higher number of pedestrians per household decreases the risk of pedestrian–motor vehicle crashes. With respect to bicycle crash risk model, bicycle exposure measures are found to have a significant impact on zonal-level bicycle-motor vehicle crash risk. The estimation result of exposure measure in the bicycle crash risk model reveals that a higher bicyclist trip demand at a zonal level increases the risk of bicycle crashes.

3.1.4 Validation Exercise of Crash Count Models

In order to demonstrate the predictive performance of the estimated crash count models, a validation experiment is also carried out. The most common approach to performing a validation exercise for an aggregate-level model is to evaluate the in-sample predictive measures. To evaluate the in-sample goodness-of-fit measures, we employ different fit measures that are widely used in statistical analysis. For crash frequency models, we compute mean prediction bias (MPB) and mean absolute deviation (MAD). These fit measures quantify the error associated with model predictions, and the model with lower fit measures provides better predictions of the observed data. These measures are computed as:

$$MPB = \frac{\sum_{i=1}^{n}(\hat{y}_i - y_i)}{n} \qquad (3.3)$$

$$MAD = \frac{\sum_{i=1}^{n}|\hat{y}_i - y_i|}{n}$$

where $\hat{y}_i$ and y_i are the predicted and observed values for event i (i is the index for event $(i = 1,2,3, \dots, N)$) and n is the number of events.

Table 3.3 presents the values for these measures for NB models for pedestrian and bicycle crash count models. Further, we also compared the predictive performance of NB models by comparing the observed and predictive counts across different count events, which are presented in Figure 3.2. From Table 3.3 and Figure 3.2, we can argue that the resulting fit measures for comparing the predictive performance clearly indicate that the models' predictive performances are overall reasonable with less error in predictions.

Table 3.3 - Predictive performance evaluation

In-sample predictive fit measures for count models				
Models	**Mean crash**		**MPB**	**MAD**
	Observed	**Predicted**		
Pedestrian	0.31	0.33	-0.81	11.44
Bicycle	0.21	0.22	-0.28	6.41

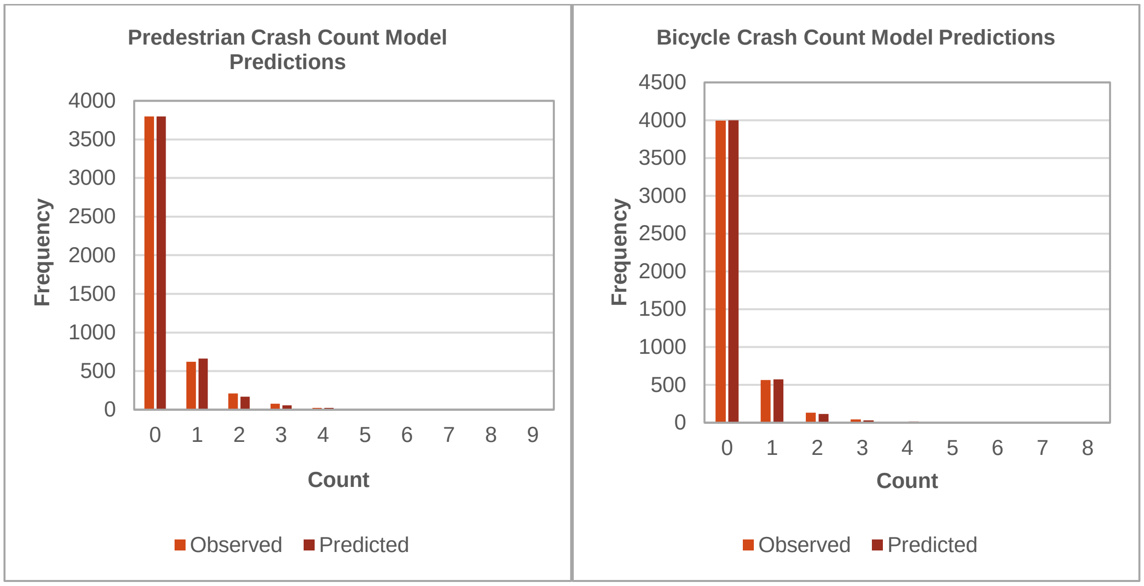


Figure 3.2 - Crash count model predictions

3.2 Crash Severity Analysis

Crash count data are often compiled by injury severity outcomes (for example: no injury, minor injury, major injury and fatal injury crashes). Given the consequences of road traffic crashes, it is important to examine crash frequency by severity level as it would play a significant role in model implications. To that extent, we can develop independent crash prediction models for different injury severity levels. However, for the same observation record, crash frequencies by different severity levels are likely to be dependent. Therefore, it might be beneficial to evaluate the impact of exogenous variables in a framework that directly relates a single exogenous variable to all severity count variables simultaneously, i.e., a framework where the observed propensities for crashes are examined by severity level directly. To that extent, in this current research effort, as opposed to modeling the number of crashes, we adopt a fractional split modeling approach to study the fraction of crashes by each severity level at a TAZ level. Specifically, we formulate and estimate ordered probit fractional split (OPFS) models for examining pedestrian and bicycle crash proportions by severity levels. The fractional split models are estimated at the TAZ level for the CFRPM 6.0 area employing a comprehensive set of exogenous variables. Based on the model results, we identify important exogenous variables that influence pedestrian and bicycle crash severity proportions.

3.2.1 Model Framework

The formulation for the OPFS model for modeling the proportion of crashes by severity is presented in this section. The reader should note that conventional maximum likelihood approaches are not suited for factional proportion models. Hence, we resort to a quasi-likelihood approach. See Yasmin et al. [15] for a detailed description of the modeling approach. Yasmin et al. [15] developed the ordered outcome fractional split model that allows the analysis of proportion for variables with multiple alternatives while also recognizing the inherent ordering in the severity outcomes.

Model Structure

Let q $(q = 1, 2, \ldots, Q)$ be an index to represent TAZ, and let k $(k = 1, 2, \ldots, K)$ be an index to represent severity category. The latent propensity equation for severity category at the q th zone:

$$y_q^* = \alpha' z_q + \xi_q \tag{3.4}$$

This latent propensity y_q^* is mapped to the actual severity category proportion y_{qk} by the ψ thresholds ($\psi_0 = -\infty$ and $\psi_k = \infty$). z_q is an (L x 1) column vector of attributes (not including a constant) that influences the propensity associated with severity category. α is a corresponding (L x 1)-column vector of mean effects. ξ_q is an idiosyncratic random error term assumed to be identically and independently standard normal distributed across zones q.

Model Estimation

The model cannot be estimated using conventional maximum likelihood approaches. Hence, we resort to a quasi-likelihood-based approach for our methodology. The parameters to be estimated in Equation 3.4 are the α and ψ thresholds. To estimate the parameter vector, we assume that

$$E(y_{qk} \mid z_{qk}) = H_{qk}(\alpha, \psi), 0 \le H_{qk} \le 1, \sum_{k=1}^{K} H_{qk} = 1 \tag{3.5}$$

H_{qk} in our model takes the ordered probit probability (P_{qk}) form for severity category *k* defined as

$$P_{qk} = \left\{ G\left[\psi_k - \alpha_q' z_q\right] - G\left[\psi_{k-1} - \alpha_q' z_q\ \right] \right\} \tag{3.6}$$

The proposed model ensures that the proportion for each severity category is between 0 and 1 (including the limits). Then, the quasi-likelihood function, for a given value of δ_q vector, may be written for site *q* as:

$$L_q(\alpha, \psi) = \prod_{k=1}^{K} \left\{ G\left[\psi_k - \alpha_q' z_q\right] - G\left[\psi_{k-1} - \alpha_q' z_q\ \right] \right\}^{d_{qk}} \tag{3.7}$$

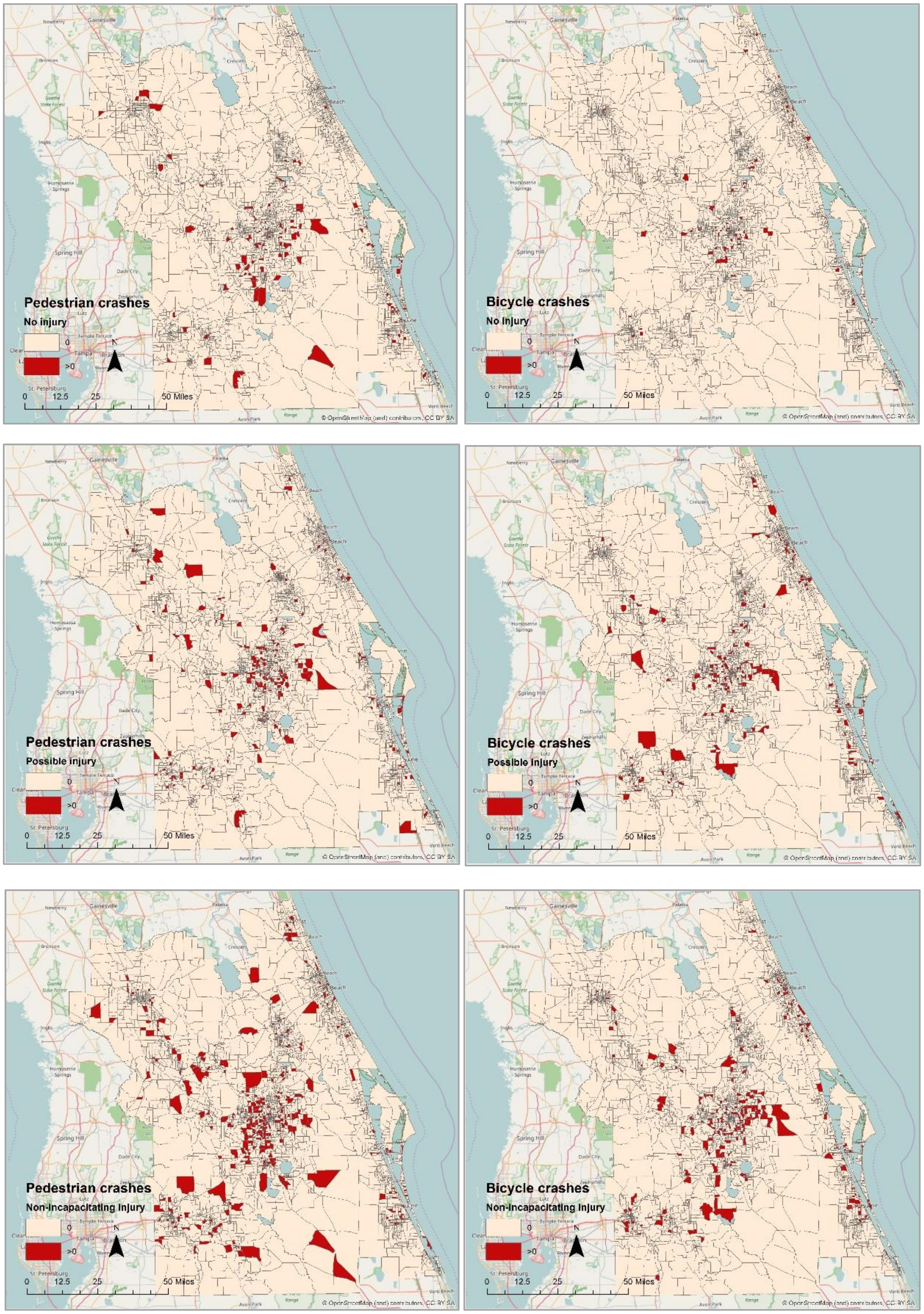
where *G*(.) is the cumulative distribution of the standard normal distribution and d_{qk} is the proportion of crashes in severity category *k*. The model estimation is undertaken using routines programmed in the Gauss matrix programming language.

3.2.2 Dependent Variable and Data Description

The crash proportion analysis is focused on pedestrian and bicycle crashes at the TAZ level for 4,747 TAZs in the area defined by the CFRPM 6.0 model. For this report, we have examined the pedestrian and bicycle crash count by severity levels for the year 2010 to reflect the base year situation in terms of non-motorized safety. Injury severity levels of non-motorist-involved crashes are presented in 5 ordinal scale variables: property damage only, possible injury, non-incapacitating, incapacitating injury and fatal crashes. For the year 2010, sample characteristics of crash injury severity outcome for pedestrian and bicycle crashes are presented in Table 3.4. From Table 3.4, we can see that number of zones with pedestrian crashes is higher than the number of zones with bicycle crashes. Moreover, the number of pedestrians involved in fatal crashes is much higher than the number of bicyclists involved in fatal crashes. Locations of zones with pedestrian and bicycle crashes for different injury severity levels are shown in Figure 3.3.

Table 3.4 – Sample characteristics of crash injury severity outcomes

Crash severity levels	Zonal					
	Pedestrian			Bicycle		
	Minimum	Maximum	Total	Minimum	Maximum	Total
Zone with no crashes	3798			4028		
Zones with crashes	949			719		
Number of property damage only crashes	0	3	168	0	2	104
Number of minor injury crashes	0	6	382	0	5	325
Number of non-incapacitating injury crashes	0	4	580	0	4	416
Number of incapacitating injury crashes	0	4	282	0	3	124
Number of fatal crashes	0	2	129	0	1	15


Pedestrian crashes
No injury
0
>0
0 12.5 25 50 Miles
Bicycle crashes
No injury
0
>0
0 12.5 25 50 Miles
Pedestrian crashes
Possible injury
0
>0
0 12.5 25 50 Miles
Bicycle crashes
Possible injury
0
>0
0 12.5 25 50 Miles
Pedestrian crashes
Non-incapacitating injury
0
>0
0 12.5 25 50 Miles
Bicycle crashes
Non-incapacitating injury
0
>0
0 12.5 25 50 Miles

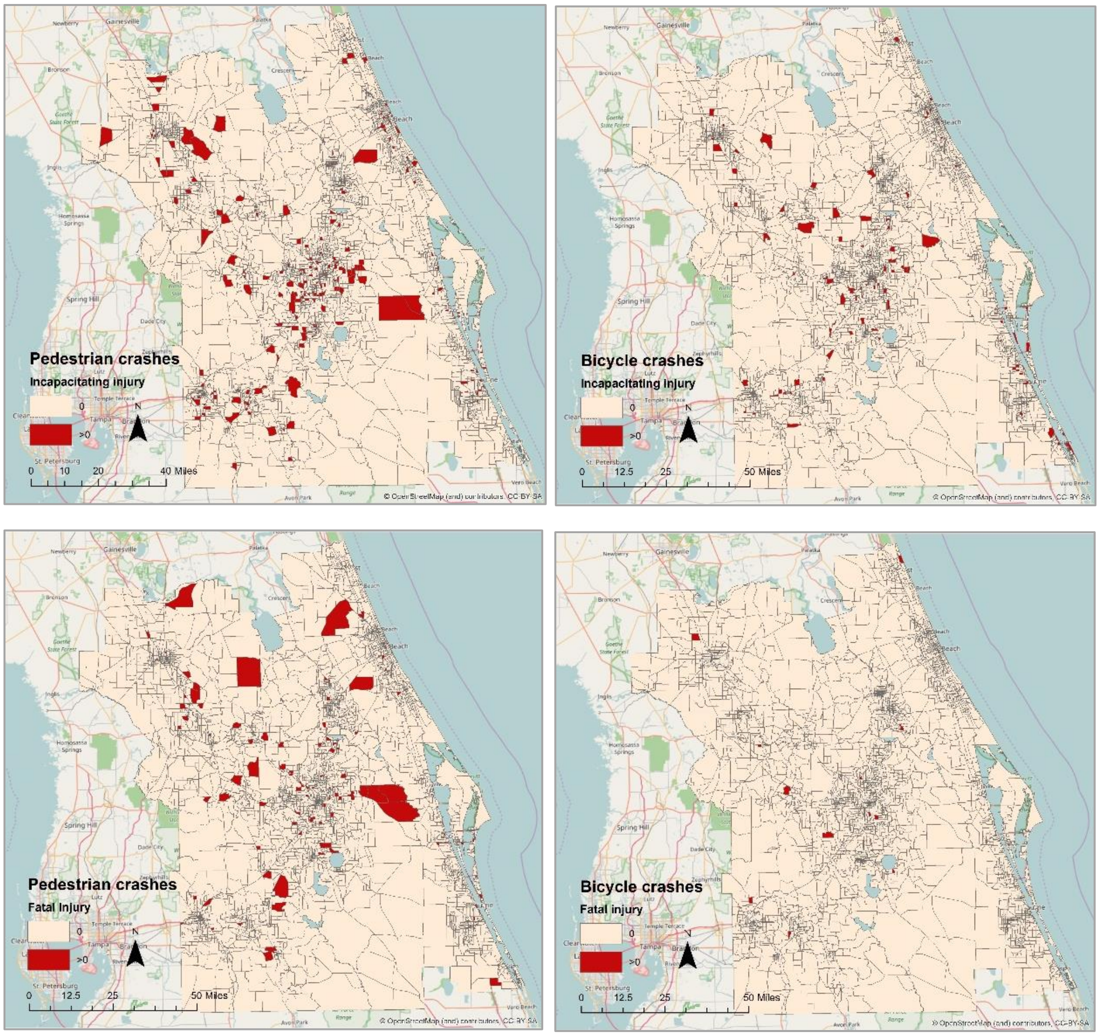


Figure 3.3 – Location of zones with different severity outcomes for pedestrian and bicycle crashes for the year 2010

For examining crash injury severity proportions, we consider zones with non-zero crashes only. Thus, for pedestrian and bicycle crash proportion models, the datasets have 949 and 719 records, respectively. In the case of five severity levels, the dependent variable in this research effort is represented as proportions (number of specific crash level/total number of all crashes) as follows: (1) proportion of property damage only crashes, (2) proportion of minor injury crashes, (3) proportion of non-incapacitating injury crashes, (4) proportion of incapacitating injury crashes and (5) proportion of fatal crashes. The dependent variable proportions and sample size for pedestrian and bicycle crashes are presented in Table 3.5. From Table 3.5, we can observe that fatal crash proportion is higher for pedestrians than for bicycle-involved crashes.

Table 3.5 - Severity proportions

Crash severity levels	Pedestrian	Bicycle

Sample	949	719
Proportion of property damage only crashes	0.113	0.115
Proportion of minor injury crashes	0.237	0.320
Proportion of non-incapacitating injury crashes	0.382	0.407
Proportion of incapacitating injury crashes	0.183	0.141
Proportion of fatal crashes	0.085	0.017

In addition to the crash database, the explanatory attributes considered in the empirical study are also aggregated at the TAZ level accordingly. To reflect the base year characteristics of the analysis zone, all attributes are generated for the year 2010. For the empirical analysis, the selected explanatory variables can be grouped into five broad categories: sociodemographic characteristics, roadway and traffic attributes, built environment characteristics, land-use characteristics and exposure measures. Table 3.6 offers a summary of the sample characteristics of the exogenous variables and the definition of variables considered for final model estimation along with the zonal minimum, maximum and average.

Table 3.6 – Summary characteristics for crash severity models

Variable name	**Description**	**Pedestrian**			**Bike**		
		Zonal			**Zonal**		
		Minimum	**Maximum**	**Mean**	**Minimum**	**Maximum**	**Mean**
Sociodemographic Characteristics							
Population density	Total number of population of TAZ/ Area of TAZ in acres	0.000	19.956	3.362	0.000	19.956	3.622
Proportion of people aged 22 to 29	Total number of population of TAZ who are 22 to 29 years old / Total number of population of TAZ	0.000	0.373	0.111	-	-	-
Roadway and Traffic Attributes							

Number of flashing beacon signs	Total number of flashing beacons of TAZ	-	-	-	0.000	1.000	0.006
Number of school signals	Total number of school signals of TAZ	-	-	-	0.000	1.000	0.003
Availability of bike lanes	Availability of bike lanes in TAZ	-	-	-	0.000	1.000	0.058
VMT	Vehicle miles traveled = Total road length in miles * Average annual daily traffic / 100000	0.000	17.052	0.430	-	-	-
Built Environment							
Number of commercial centers	Total number of commercial centers of TAZ	0.000	3.000	0.113	-	-	-
Number of hospitals	Total number of hospitals of TAZ	-	-	-	0.000	2.000	0.033
Number of parks and recreational centers	Total number of parks and recreational centers of TAZ	-	-	-	0.000	7.000	0.307
Land-use Characteristics							
Urban area	Ln (Urban area in a TAZ in acres)	-6.254	8.384	5.236	-4.661	8.384	5.328
Residential area	Ln (Residential area in a TAZ in acres)	-	-	-	-9.052	7.647	4.070
Exposure measures							
Total pedestrian trip demand per household	Total pedestrian daily trip demand in a TAZ/(Total number of households in a TAZ*100)	0.000	1.316	0.021	-	-	-
Total bicycle trip demand per household	Total bicycle daily trip demand in a TAZ/Total number of households in a TAZ	-	-	-	0.000	134.686	0.498

3.2.3 Estimation Results

In this research effort, we estimate two different OPFS models: one model for pedestrian crash severity proportions at the zonal level and another model for bicycle crash severity proportions at the zonal level. Table 3.7 presents the estimation results of the OPFS models. The pedestrian crash severity proportion results are presented in the 2nd and 3rd columns of Table 3.7, and the bicycle crash severity proportion model component are presented in the 4th and 5th columns. In OPFS models, the positive (negative) coefficient corresponds to increased (decreased) proportions of severe injury categories. The final specification of the model was based on removing the statistically insignificant variables in a systematic process based on statistical significance (90% confidence level) and intuitive coefficient effect. In estimating the models, several functional forms and variable specifications are explored. The functional form that provided the best result is used for the final model specifications. The effects of exogenous variables in model specifications for both pedestrian and bicycle crash severity proportion models are discussed in this section by variable groups.

Table 3.7 - Estimation results of ordered probit fraction split models

Variable name	Pedestrian		Bike	
	Estimates	t-stat	Estimates	t-stat
Threshold 1	-1.708	-13.117	-1.450	-8.330
Threshold 2	-0.870	-6.818	-0.395	-2.309
Threshold 3	0.146	1.148	0.798	4.589
Threshold 4	0.916	7.018	1.954	9.929
Sociodemographic Characteristics				
Population density	-0.022	-1.898	-0.032	-2.061
Proportion of people aged 22 to 29	-1.321	-1.965	---	---
Roadway and Traffic Attributes				
Number of flashing beacon sign	---	---	0.936	2.347
Number of school signal	---	---	0.362	2.474
Availability of bike lane	---	---	-0.288	-1.797
VMT	0.049	1.675	---	---
Built Environment				
Number of commercial center	-0.149	-1.936	---	---
Number of hospital	---	---	-0.189	-1.795
Number of park and recreational center	---	---	0.139	2.802

Land-use Characteristics				
Urban area	-0.046	-2.466	-0.076	-2.079
Residential area	---	---	0.066	2.560
Exposure Measures				
Total pedestrian trip demand per household	-1.063	-2.756	---	---
Total bicycle trip demand per household	---	---	-0.005	-1.040

Sociodemographic characteristics: With respect to sociodemographic characteristics, the estimates indicate that population density results in lower likelihood of severe crash proportions for both pedestrian and bicycle crashes. The proportion of 22- to 29-year-olds in the population has a negative impact on the proportion of pedestrian crash severity outcomes, implying a reduced likelihood of more severe pedestrian crashes.

Roadway and traffic attributes: The OPFS model results for bicycles reveal a higher proportion of severe crash outcomes for zones with a higher number of flash beacon signs and a higher number of school signals. As expected, availability of bike lanes is found to reduce the proportion of less severe bicycle crashes. With respect to traffic attributes, higher VMT is positively associated with more severe crash proportions in the pedestrian crash proportion model.

Built environment: The crash proportion model for pedestrian-involved crashes reveals that the pedestrian crash proportion of severe crashes is lower in TAZs with a higher number of commercial centers. A higher number of hospitals is associated with a lower likelihood of severe crash proportion in the OPFS model for bicycles. At the same time, the OPFS model results reveal that a higher number of parks and recreational centers increases the possibility of higher proportions of severe bicycle crash outcomes.

Land-use characteristics: From both the pedestrian and bicycle models, we find that the possibility of more severe crashes decreases with increasing shares of urbanized areas of a TAZ. Residential area is found to be a significant determinant of bicycle crash proportion by severity outcomes. The estimate for residential area has a positive coefficient in the bicycle crash severity model, suggesting that proportion of severe bicycle crashes increases with increasing zonal-level residential areas.

Exposure measures: The non-motorist exposure measures generated from Section 2.4 are used in evaluating zonal-level pedestrian and bicycle crash severity proportions. In estimating OPFS models, several functional forms and variable specifications are explored. The functional form that provided the best result is used for the final model specifications as presented With respect to the pedestrian crash severity proportion model, pedestrian exposure measures are found to have a significant impact on zonal-level bicycle–motor vehicle crash severity outcome proportions. The estimation result of exposure measures in the pedestrian crash severity proportion model reveals that a higher number of pedestrian trip demand per household at a zonal level decreases the propensity for a higher proportion of severe crashes. With respect to the bicycle crash severity proportion model, bicycle exposure measures with any of the functional forms are not found to be significant at a 90% confidence level. However, bicycle trip demand per household at a zonal level provides the best data fit and hence is considered in our final OPFS model. From Table 3.7, we can see that a higher number of bicyclists per household decreases the risk of a higher proportion of severe bicycle–motor vehicle crashes.

3.2.4 Validation Exercise of Crash Proportion Models

In order to demonstrate the predictive performance of the estimated crash proportion models, a validation experiment is also carried out. The most common approach to perform a validation exercise for an aggregate-level model is to evaluate the in-sample predictive measures. For crash proportion models, we compute mean absolute percentage error (MAPE) and root mean square error (RMSE). These fit measures quantify the error associated with model predictions, and the model with lower fit measures provides better predictions of the observed data. These measures are computed as:

$$MAPE = \sum_{i=1}^{n}\left|\frac{\hat{y}_i - y_i}{y_i}\right| \quad (3.8)$$

$$RMSE = \sqrt{\frac{\sum_{i=1}^{n}(\hat{y}_i - y_i)^2}{n}}$$

where $\hat{y}_i$ and y_i are the predicted and observed values for event i (i is the index for event $(i = 1,2,3, \ldots, N)$) and n is the number of events.

Table 3.8 presents the values for these measures for the OPFS models for the pedestrian and bicycle crash models. From Table 3.8, we can argue that the resulting fit measures for comparing the predictive performance clearly indicate that the models' predictive performances are overall reasonable with less error in predictions.

Table 3.8 - Predictive performance evaluation

In-sample predictive fit measures for fractional split models					
Models	**Mean proportion**			**MAPE**	**RMSE**
	Severity levels	**Observed**	**Predicted**		
Pedestrian	Proportion of property damage only crashes	0.113	0.114	0.003	0.531
	Proportion of minor injury crashes	0.237	0.237		
	Proportion of non-incapacitating injury crashes	0.382	0.381		
	Proportion of incapacitating injury crashes	0.183	0.184		

	Proportion of fatal crashes	0.085	0.085		
	Proportion of property damage only crashes	0.115	0.116		
	Proportion of minor injury crashes	0.320	0.320		
Bicycle	Proportion of non-incapacitating injury crashes	0.407	0.407	0.006	0.292
	Proportion of incapacitating injury crashes	0.141	0.141		
	Proportion of fatal crashes	0.017	0.017		

4 Policy Scenario Analysis and Recommendations

The parameter effects of exogenous variables in Sections 2 and 3 do not directly provide the magnitude of the effects on zonal-level non-motorist demand and crash risks (both in terms of frequency and proportions of severity) and therefore cannot be directly employed for policy scenario analysis. For policy scenario analysis, we compute aggregate-level "elasticity effects" of exogenous variables both in the demand models and safety models (crash frequency and crash severity by proportions) (see work by Eluru and Bhat [16] for a discussion on the methodology for computing elasticities). We investigate the effect as a percentage change in the expected total zonal demand, total zonal crash counts and total proportions by severity levels to the change in exogenous variables for the study region. In the current study context, we perform policy analysis for different scenarios as follows:

- Scenario 1: 50% reduction in traffic volume within 2 miles buffer area of different central business districts (CBD).
- Scenario 2: 30% reduction in traffic volume within 2 miles buffer area of different central business districts (CBD).
- Scenario 3: 15% reduction in traffic volume within 4 miles buffer area of different central business districts (CBD).
- Scenario 4: 5% reduction in traffic volume within 6 miles buffer area of different central business districts (CBD).
- Scenario 5: All zones have sidewalk, and the new proposed $\text{sidewalk length} = \frac{(TAZ\ area)^{0.5}}{2}\ meter$.
- Scenario 6: 50% increase in existing sidewalk length.
- Scenario 7: 15% reduction in zonal average maximum speed.
- Scenario 8: 25% reduction in zonal average maximum speed.
- Scenario 9: 15% reduction in zonal proportion of 3+ lane road.
- Scenario 10: 25% reduction in zonal proportion of 3+ lane road.

These scenarios are evaluated for all zones and for both the pedestrian and bicycle groups of road users separately. Moreover, we also evaluate Scenarios 1, 2, 3 and 4 for the zones within 2 (for Scenarios 1 and 2), 4 (for Scenario 3) and 6 (for Scenario 4) miles of buffer area for multiple CBDs in the Central Florida region, including Orlando, Sanford, Lakeland, Kissemme, Deland, Ocala, Melbourne, Palm Bay, Leesburg, Daytona Beach and Port Orange. In evaluating each scenario, we perform policy scenario analysis for three different components:

1. Component 1: Policy analysis for non-motorist demand - Evaluate change in total demand due to the change considered in the scenario.
2. Component 2: Policy analysis for non-motorist crash frequency - Evaluate the change in total crash frequencies considering the change in the scenario and the change in demand from Component 1 accordingly.
3. Component 3: Policy analysis for non-motorist crash severity proportions - Evaluate the change in total crash proportions by severity considering the change in the scenario and the change in demand from Component 1 accordingly.

By performing policy scenario analysis for these three components, we ensure that the updated demand matrices for each scenario are produced and employed in developing exposure measures for non-motorized travel as well as vehicular volumes on roadways. With these new exposure measures, the safety models are re-run to generate estimates of scenario-based crash and severity rates and the change in safety situation. A comparison across scenarios would allow us to identify beneficial changes to the existing infrastructure for improving non-motorized road user safety. The spatial representation of the considered CBD locations is shown in Figure 4.1. In the following sections, we describe the results from these policy scenario matrices for all three components.

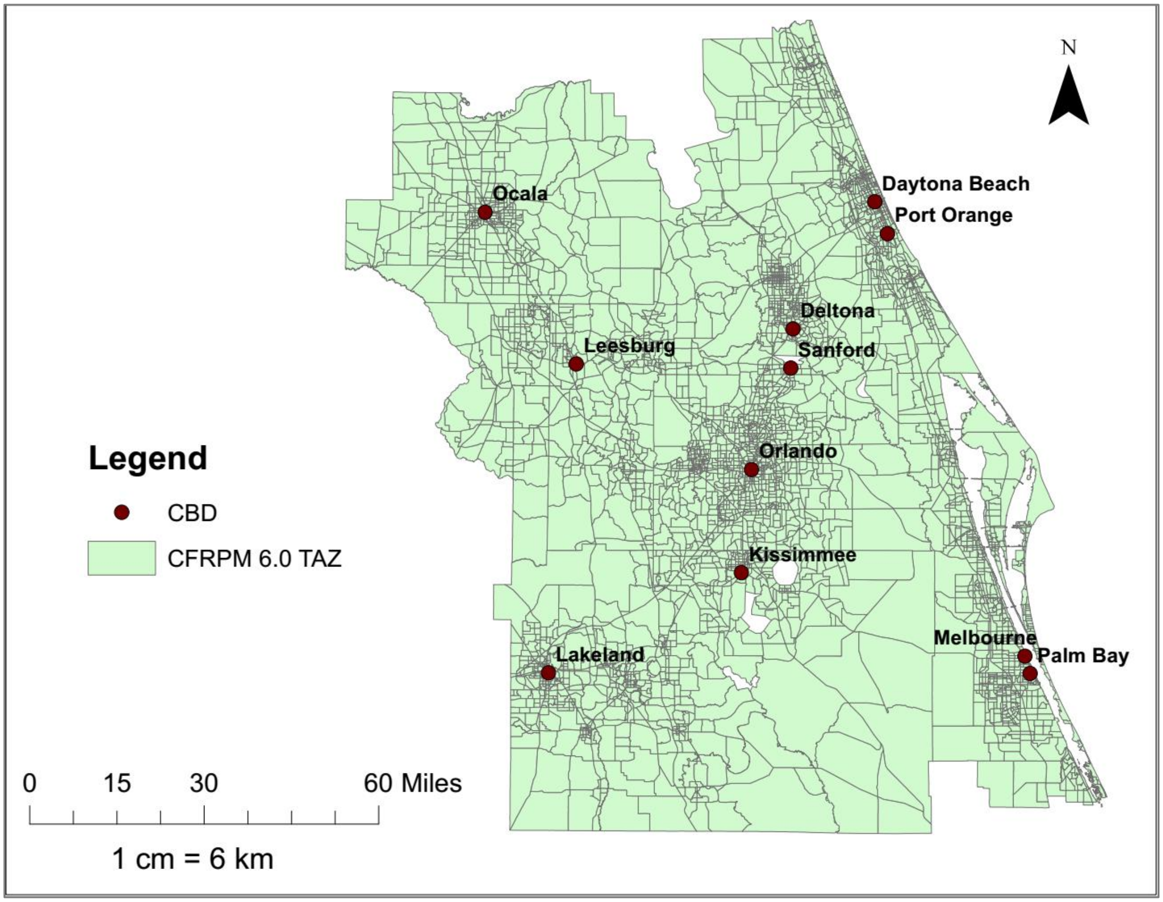


Figure 4.1 – Considered central business district locations

4.1 Policy Analysis for Non-Motorist Demand

Policy scenario analysis for non-motorist travel demand is presented in this section. The change in total demand is evaluated across all scenarios for the pedestrian and bicycle groups of road users separately. The computed elasticities for total change in demand are presented in Table 4.1. The numbers in Table 4.1 may be interpreted as the percentage change in the expected total zonal demand per day due to the change in exogenous variable. The following observations can be made based on the elasticity effects presented in Table 4.1.

First, decreasing vehicular traffic volume near the CBD location have a greater effect on pedestrian demand than on bicycle demand. For both modes, we can observe from the table that a higher level of non-motorist activities can be gained by restricting vehicular traffic; the greater the restrictions, the higher the level of non-motorist demand. Second, increasing sidewalk facilities is likely to attract more non-motorists, but for the hypothetical Scenario 5, the demand for pedestrians is likely to get reduced. Third, the reduction in speed has a greater impact on increasing pedestrian demand. However, for bicycles, the variable has no impact, as it was found insignificant in bicycle demand models. Fourth, a restriction in the number of traffic lanes is likely to have a similar impact; as we can see from Table 4.1, it increases non-motorist demand.

Recommendations: From the policy scenario analysis, it is quite clear that providing more walking- and bicycle-friendly facilities is likely to encourage more people to use non-motorized modes. Thus, we can argue that restricting lanes, reducing speed and reducing/restricting vehicular volume in a certain zone would increase non-motorist volume.

Table 4.1 – Elasticity effects for non-motorist total zonal demand

Scenarios	Study region	Number of zones	Pedestrian	Bicycle
Scenario 1	**All zones**	4747	0.164	0.043
	Zones within 2-mile buffer of CBD	703	1.804	0.389
Scenario 2	**All zones**	4747	0.096	0.026
	Zones within 2-mile buffer of CBD	703	1.060	0.231
Scenario 3	**All zones**	4747	0.125	0.030
	Zones within 4-mile buffer of CBD	1375	0.498	0.090
Scenario 4	**All zones**	4747	0.071	0.013
	Zones within 6-mile buffer of CBD	1985	0.166	0.027
Scenario 5	**All zones**	4747	-0.438	0.108
Scenario 6	**All zones**	4747	0.705	0.289
Scenario 7	**All zones**	4747	1.407	0.000
Scenario 8	**All zones**	4747	2.389	0.000
Scenario 9	**All zones**	4747	0.287	0.576
Scenario 10	**All zones**	4747	0.484	0.337

4.2 Policy Analysis for Non-Motorist Crash Frequency

Policy scenario analysis for non-motorist crash frequency is presented in this section. The change in total crash frequency is evaluated across all scenarios for the pedestrian and bicycle groups of road users separately. The computed elasticities for total change in crash frequency are presented in Table 4.2. To be sure, in evaluating the change in each scenario, the corresponding change in non-motorist demand (as presented in Section 4.1) is also incorporated for evaluating elasticity effects for non-motorist crash frequency. The numbers in Table 4.2 may be interpreted as the percentage change in the expected total zonal crashes per year due to the change in exogenous variable. The following observations can be made based on the elasticity effects presented in Table 4.2.

First, decreasing vehicular traffic volume near CBD locations is likely to reduce pedestrian crashes, with a greater impact within the vicinity of the CBD. However, bicycle crashes are likely to increase by about 3%. However, the number of bicycle–motor vehicle crashes is likely to decrease within the vicinity of CBD with a greater reduction in vehicular volume. Second, the hypothetical scenario of sidewalk length shows that providing walking facilities has the potential to improve pedestrian safety. On the other hand, bicycle crashes are likely to be high for increasing sidewalk length – perhaps indicating greater exposure. Third, reduction in speed and restrictions in traffic lanes decrease pedestrian crashes. On the other hand, restrictions in traffic lanes increase bicycle crashes by about 4%.

Recommendations: It is a well-known fact that non-motorist safety tends to decrease with increasing non-motorist exposure, and only after a certain level of exposure (when traffic becomes familiar with the higher number of non-motorists) does the safety tend to increase. From our policy analysis, we can see that non-motorist-friendly infrastructure has a mixed effect on non-motorist safety. Therefore, it is imperative that policy implications for improving non-motorist safety be identified by considering all known exogenous elements in identifying the appropriate tools. In general, restricting vehicular volume in a targeted zone would improve non-motorist safety.

Table 4.2 – Elasticity effects for non-motorist crash frequency

Scenarios	Study region	Number of zones	Pedestrian	Bicycle
Scenario 1	**All zones**	4747	-0.630	3.144
	Zones within 2-mile buffer of CBD	703	-3.266	-2.889
Scenario 2	**All zones**	4747	-0.437	3.622
	Zones within 2-mile buffer of CBD	703	-2.120	-0.274
Scenario 3	**All zones**	4747	-0.482	3.554

	Zones within 4-mile buffer of CBD	1375	-1.280	1.680
Scenario 4	**All zones**	4747	-0.340	3.935
	Zones within 6-mile buffer of CBD	1985	-0.589	3.281
Scenario 5	**All zones**	4747	-1.360	4.367
Scenario 6	**All zones**	4747	0.985	4.436
Scenario 7	**All zones**	4747	-0.143	0.000
Scenario 8	**All zones**	4747	-0.150	0.000
Scenario 9	**All zones**	4747	-0.138	4.436
Scenario 10	**All zones**	4747	-0.143	4.415

4.3 Policy Analysis for Non-Motorist Crash Severity Proportions

Policy scenario analysis for non-motorist crash severity proportions is presented in this section. The change in total crash severity proportions is evaluated across all scenarios for the pedestrian and bicycle groups of road users separately. The computed elasticities for total change in crash frequency are presented in Table 4.3. To be sure, in evaluating the change in each scenario, the corresponding change in non-motorist demand (as presented in Section 4.1) is also incorporated for evaluating elasticity effects for non-motorist crash severity proportions. The numbers in Table 4.3 may be interpreted as the percentage change in the expected total zonal crash severity proportion across different severity levels due to the change in exogenous variable. The following observations can be made based on the elasticity effects presented in Table 4.3.

First, decreasing vehicular traffic volume near CBD locations is likely to reduce non-motorist crash severity, with greater impact within the vicinity of the CBD. However, the impact on the pedestrian mode is much higher than the impact on the bicycle mode. Second, the decrease in pedestrian fatal crash severity proportions is about 1% for increasing sidewalk length, reducing speed and restricting traffic lanes. The contributions of these measures on bicycle crash severity are less pronounced relative to pedestrian modes.

Recommendations: From our policy analysis, we can see that non-motorist-friendly infrastructure has a positive effect on non-motorist safety by reducing severe crashes. Therefore, we can argue that zonal-level implications of non-motorist-friendly infrastructures and environment should be implemented to reduce the consequences of non-motorist crash severity outcomes.

Table 4.3 – Elasticity effects for non-motorist crash severity proportions

Pedestrian

Scenarios	Study region	Number of zones	O*	C	B	A	K
Scenario 1	All zones	4747	6.082	1.735	-0.505	-2.658	-4.967
	Zones with 2-mile buffer of CBD	703	10.888	0.358	-1.268	-2.922	-4.687
Scenario 2	All zones	4747	6.074	1.733	-0.504	-2.655	-4.963
	Zones with 2-mile buffer of CBD	703	10.851	0.349	-1.263	-2.906	-4.664
Scenario 3	All zones	4747	6.075	1.734	-0.504	-2.655	-4.963
	Zones with 4-mile buffer of CBD	1375	7.432	1.014	-0.807	-2.653	-4.550
Scenario 4	All zones	4747	6.068	1.732	-0.503	-2.653	-4.960
	Zones with 6-mile buffer of CBD	1985	6.688	1.365	-0.690	-2.751	-4.891
Scenario 5	All zones	4747	2.514	0.154	-0.294	-0.675	-1.013
Scenario 6	All zones	4747	2.703	0.191	-0.321	-0.740	-1.111
Scenario 7	All zones	4747	2.685	0.193	-0.318	-0.738	-1.107
Scenario 8	All zones	4747	2.742	0.204	-0.326	-0.758	-1.135
Scenario 9	All zones	4747	2.628	0.181	-0.310	-0.717	-1.077
Scenario 10	All zones	4747	2.644	0.185	-0.312	-0.723	-1.085
			Bicycle				
Scenarios	Study region	Number of zones	PDO	MI	NI-In	I-in	Ft
Scenario 1	All zones	4747	0.142	0.023	-0.034	-0.062	-0.066
	Zones with 2-mile buffer of CBD	703	0.033	0.011	-0.009	-0.028	-0.045
Scenario 2	All zones	4747	0.142	0.023	-0.034	-0.062	-0.066
	Zones with 2-mile buffer of CBD	703	0.033	0.011	-0.009	-0.028	-0.045
Scenario 3	All zones	4747	0.142	0.023	-0.034	-0.062	-0.066
	Zones with 4-mile buffer of CBD	1375	0.024	0.007	-0.007	-0.016	0.003

Scenario 4	**All zones**	4747	0.142	0.023	-0.034	-0.062	-0.066
	Zones with 6-mile buffer of CBD	1985	-0.005	0.000	0.001	0.000	0.015
Scenario 5	**All zones**	4747	0.134	0.023	-0.033	-0.060	-0.063
Scenario 6	**All zones**	4747	0.149	0.024	-0.036	-0.065	-0.071
Scenario 7	**All zones**	4747	0.000	0.000	0.000	0.000	0.000
Scenario 8	**All zones**	4747	0.000	0.000	0.000	0.000	0.000
Scenario 9	**All zones**	4747	0.143	0.024	-0.034	-0.063	-0.068
Scenario 10	**All zones**	4747	0.142	0.024	-0.034	-0.063	-0.066

*O=property damage only, C=minor injury, B=non-incapacitating injury, A=incapacitating injury, K=fatal

4.4 Future Year Demand Predictions

In order to demonstrate the implications from the estimated demand models, we also generate the predicted demand matrices for the year 2015. Specifically, we have estimated predicted origin demand, predicted destination demand and predicted total demand for the year 2015. These matrices are presented in Table 4.4 at the county level. From Table 4.4 we can see that overall bicycle demand has increased from 2010 to 2015, but that total pedestrian demand has decreased over the same period. These generated demand matrices can be used as non-motorist exposure measures for developing a crash prediction model for the year 2015. Similar matrices can be generated for any other year.

Table 4.4 – Trip demand matrices by county level for the years 2010 and 2015

Pedestrian									
County	**Trip origin demand**			**Trip destination demand**			**Total trip demand**		
	2010	**2015**	**% change**	**2010**	**2015**	**% change**	**2010**	**2015**	**% change**
Brevard	154936.5	153610.7	-0.9	149804.8	144628.0	-3.5	304741.3	298238.7	-2.1

Flagler	26241.5	24853.4	-5.3	23153.7	22261.3	-3.9	49395.1	47114.6	-4.6
Indian River	12066.8	12169.7	0.9	11826.2	11663.3	-1.4	23892.9	23833.0	-0.3
Lake	67309.3	68943.5	2.4	66545.9	65799.1	-1.1	133855.2	134742.6	0.7
Marion	95199.9	93593.9	-1.7	89602.9	89575.3	0.0	184802.8	183169.2	-0.9
Orange	348163.9	342918.6	-1.5	355169.8	349371.2	-1.6	703333.7	692289.8	-1.6
Osceola	67651.6	68006.6	0.5	65181.7	64571.8	-0.9	132833.3	132578.4	-0.2
Polk	185959.9	195780.4	5.3	195543.4	205340.1	5.0	381503.4	401120.4	5.1
Seminole	75690.1	79112.2	4.5	79212.2	80228.2	1.3	154902.3	159340.4	2.9
Sumter	32272.8	30488.9	-5.5	26598.9	25489.9	-4.2	58871.7	55978.8	-4.9
Volusia	189987.7	189005.7	-0.5	174051.2	172072.2	-1.1	364038.8	361077.9	-0.8
Total	1255480.0	1258483.6	0.2	1236691.0	1231000.4	-0.5	2492171.0	2489483.9	-0.1
Bicycle									

County	Trip origin demand			Trip destination demand			Total trip demand		
	2010	2015	%change	2010	2015	%change	2010	2015	%change
Brevard	21663.6	21822.8	0.7	23172.9	23344.3	0.7	44836.5	45167.1	0.7
Flagler	2940.3	2964.9	0.8	2634.0	3031.2	15.1	5574.4	5996.1	7.6
Indian River	1735.3	1734.3	-0.1	999.5	998.4	-0.1	2734.7	2732.8	-0.1
Lake	10784.3	10676.6	-1.0	9977.6	9774.7	-2.0	20761.9	20451.2	-1.5
Marion	5238.3	5448.9	4.0	4226.3	4344.1	2.8	9464.5	9793.0	3.5
Orange	57661.9	60551.9	5.0	64084.7	68918.9	7.5	121746.7	129470.8	6.3
Osceola	4026.1	4308.8	7.0	3875.6	3974.1	2.5	7901.8	8282.9	4.8
Polk	10931.1	11589.5	6.0	10687.7	11851.7	10.9	21618.8	23441.2	8.4
Seminole	12179.4	12529.5	2.9	11558.9	11903.0	3.0	23738.3	24432.5	2.9

Sumter	553.1	614.6	11.1	817.9	1019.8	24.7	1371.0	1634.4	19.2
Volusia	37958.0	38199.6	0.6	39924.9	41457.9	3.8	77882.8	79657.5	2.3
Total	165671.4	170441.4	2.9	171960.0	180618.0	5.0	337631.3	351059.4	4.0

5 Conclusions

The current research effort proposed an approach to identify and incorporate non-motorist exposure for developing macro-level crash models both in terms of total crashes and crashes by different severity levels. We developed non-motorist demand models and estimated aggregate-level crash frequency and severity models for non-motorized modes of transportation by incorporating the exposure measures predicted from the estimated demand models. The validation exercise performed provided evidence that the estimated models are reasonable. Further, the implications of the estimated models are demonstrated by analyzing several policy scenario analyses. The research methodology as proposed in our study recognizes that zonal-level attributes are likely to influence non-motorist exposure. At the same time, non-motorist exposure along with the zonal-level attributes are critical factors in developing non-motorist safety models.

5.1 Limitations

Our study is not without limitations. In our study approach, we evaluated non-motorist demand by using the NHTS database at an aggregate level, which is not readily transferable for developing a micro-level model. It might be interesting to generate a micro-level trip demand model to identify non-motorist exposure at a corridor level.

5.2 Future Directions

With respect to future research, it might be useful to forecast non-motorist safety by employing non-motorist demand generated for a future year.

Acknowledgement

The authors would like to thank FDOT for providing additional funds as well as data access to support the research. We also thank S4A for providing the crash data used in the research.

References

1. NHTSA, (2017a). Traffic Safety Facts 2015 Data, Pedestrians, National Highway Traffic Safety Administration, February 2017.
2. NHTSA, (2017b). Traffic Safety Facts 2015 Data, Bicyclists, National Highway Traffic Safety Administration, March 2017.
3. Cameron, A. C., & Trivedi, P. K. (1998). *Regression analysis of count data*. New York: Cambridge University Press.
4. Heilbron, D. C. (1994). Zero-altered and other regression models for count data with added zeros. *Biometrical Journal*, 36(5), 531-547.
5. Jang, T. Y. (2005). Count data models for trip generation. *Journal of Transportation Engineering*, 131(6), 444-450.
6. McFadden, D. (1978). Modeling the choice of residential location. In A. Karlqvist et al. (Eds.), *Spatial interaction theory and planning models*. Amsterdam: North Holland Publishers.
7. Rashidi, T. H., Auld, J., & Mohammadian, A. K. (2012). A behavioral housing search model: Two-stage hazard-based and multinomial logit approach to choice-set formation and location selection. *Transportation Research Part A: Policy and Practice*, 46, 1097-1107.
8. Zolfaghari, A., Sivakumar, A., & Polak, J. (2013). Simplified probabilistic choice set formation models in a residential location choice context. *Journal of Choice Modelling*, 9, 3-13.
9. Chakour, V., & Eluru, N. (2014). Analyzing commuter train user behavior: A decision framework for access mode and station choice. *Transportation*, 41(1), 211-228.
10. Faghih-Imani, A., & Eluru, N. (2015). Analyzing bicycle sharing system user destination choice preferences: An investigation of Chicago's Divvy System. *Journal of Transport Geography,* 44, 53-64.
11. Reynolds, C., Harris, M. A., Teschke, K., Cripton, P. A., & Winters, M. (2009). The impact of transportation infrastructure on bicycling injuries and crashes: a review of the literature. *Environmental Health*, 8(1), 47.
12. Elvik, R. (2009). The Non-linearity of risk and the promotion of environmentally sustainable transport. *Accident Analysis & Prevention*, 41(4), 849-855.
13. Wei, F., & Lovegrove, G. (2013). An empirical tool to evaluate the safety of cyclists: Community based, macro-level collision prediction models using negative binomial regression. *Accident Analysis & Prevention*, 61, 129-137.
14. Lord, D., & Mannering, F. (2010). The statistical analysis of crash-frequency data: a review and assessment of methodological alternatives. *Transportation Research Part A: Policy and Practice*, 44(5), 291-305.
15. Savolainen, P. T., Mannering, F. L., Lord, D., & Quddus, M. A. (2011). The statistical analysis of highway crash-injury severities: A review and assessment of methodological alternatives. *Accident Analysis and Prevention*, 43(5), 1666-1676.
16. Yasmin. S., & Eluru, N. (2013). Evaluating alternate discrete outcome frameworks for modeling crash injury severity, *Accident Analysis & Prevention*, 59(1), 506-521.
17. Yasmin. S., Eluru, N., Lee, J., & Abdel-Aty, M. (2016). Ordered fractional split approach for aggregate injury severity modeling. *Transportation Research Record*, 2583, 119-126.
18. Eluru, N., & Bhat, C. R. (2007). A joint econometric analysis of seat belt use and crash-related injury severity. *Accident Analysis & Prevention*, 39(5), 1037-1049.