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

Shamsunnahar Yasmin*
Senior Lecturer/Senior Research Fellow
Queensland University of Technology (QUT)
Centre for Accident Research & Road Safety – Queensland (CARRS-Q) &
Research Affiliate, Department of Civil, Environmental & Construction Engineering, University of Central Florida
Tel: +61731384677
Email: shams.yasmin@qut.edu.au

Tanmoy Bhowmik
Postdoctoral Scholar
Department of Civil, Environmental & Construction Engineering
University of Central Florida
Tel: 1-407-927-6574; Fax: 1-407-823-3315
Email: tanmoy78@knights.ucf.edu

Moshiur Rahman
Traffic Engineer
Marlin Engineering
Tel: 321-276-7580
Email: moshiur@knights.ucf.edu

Naveen Eluru
Professor
Department of Civil, Environmental & Construction Engineering
University of Central Florida
Tel: 407-823-4815, Fax: 407-823-3315
Email: naveen.eluru@ucf.edu

Date: March, 2021

---

* Corresponding author

ABSTRACT

Traditionally, in developing non-motorized crash prediction models, safety researchers have employed land use and urban form variables as surrogate for exposure information (such as pedestrian, bicyclist volumes and vehicular traffic). The quality of these crash prediction models is affected by the lack of “true” non-motorized exposure data. High-resolution modeling frameworks such as activity-based or trip-based approach could be pursued for evaluating planning level non-motorist demand. However, running a travel demand model system to generate demand inputs for non-motorized safety is cumbersome and resource intensive. The current study is focused on addressing this drawback by developing an integrated non-motorized demand and crash prediction framework for mobility and safety analysis. Towards this end, we propose a three-step framework to evaluate non-motorists safety: (1) develop aggregate level models for non-motorist generation and attraction at a zonal level, (2) develop non-motorists trip exposure matrices for safety evaluation and (3) develop aggregate level non-motorists crash frequency and severity proportion models. The framework is developed for the Central Florida region using non-motorist demand data from National Household Travel Survey (2009) Florida Add-on and non-motorist crash frequency and severity data from Florida. The applicability of the framework is illustrated through extensive policy scenario analysis.

Keywords: *Pedestrian; Bicycle; Active travel; Travel Demand; Safety; Negative Binomial; Fractional Split; Non-motorist*

1. INTRODUCTION

Urban regions in North America are encouraging the adoption of active modes of transportation by proactively developing infrastructure for non-motorist (pedestrians and bicyclists). However, a strong impediment to the increasing adoption of active modes of transportation is the safety 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 is 1.50 (2.35), the corresponding number for the state of Florida is 2.56 (6.80), which clearly demonstrates the safety risk for non-motorists in Florida (NHTSA, 2015). An important tool to determine 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 have employed land use and urban form variables as surrogate for exposure information (pedestrian, bicyclist volumes and vehicular traffic). The quality of these crash prediction models is affected by the lack of "true" non-motorist exposure data. Moreover, to assess how recent investments in pedestrian and bicycle transportation infrastructure are influencing their mobility and safety, it is important to develop demand models. High-resolution modeling frameworks such as activity-based or trip-based approaches could be pursued for evaluating planning level non-motorist demand. However, the current state-of-the-art of travel demand models focus on generating vehicular demand (for automobile and transit). For example, the existing Central Florida Regional Planning Model (CFRPM) is predominantly focused on auto mode and public transit mode. The modeling approach does not consider non-motorized modes in detail. Therefore, travel demand matrices of active transport modes are not often readily available to integrate those in road safety evaluation. Even when non-motorist demand is considered, the models employed for these dimensions are rule-based or simplistic models with very few parameters. Further, running a travel demand model system to generate inputs for non-motorist safety is cumbersome, resource intensive and unlikely to be implemented.

The current study is focused on addressing this drawback by developing an integrated demand and crash prediction framework for active modes (pedestrians and bicyclists) with the objective of using it for mobility and safety analysis. To be sure, analysts often develop non-motorists' demand model at different local levels, such as: regional level (Porter *et al.* 1999), corridor (Matlick 1998) or sub-area level and household/individual level (Pulugurtha and Repaka 2008, Schneider *et al.* 2009a). Extrapolating planning level non-motorists demand from corridor level exposure data is not straightforward. An alternative approach to generating planning level non-motorist demand is to estimate origin-destination (O-D) demand at an aggregate level. The proposed integrated demand and safety framework would allow us to devise evidence-based policy implications for improving mobility and safety of pedestrians and bicyclists. In the following section, non-motorist refers to pedestrians and bicyclists collectively, while non-motorist crash refers to pedestrian and bicycle involved crashes and finally non-motorist demand refers to pedestrian and bicycle trips.

2. BACKGROUND AND CURRENT STUDY IN CONTEXT

2.1 Earlier Research

As the focus of our research is on examining non-motorist demand and non-motorist safety, we organize our literature along these two dimensions.

Non-motorist Demand

Accurate information on non-motorized trip volumes is useful for many studies including public health studies (Cervero and Duncan 2003), non-motorist safety research (Miranda-Moreno *et al.* 2011), and active mode infrastructure improvements (Ercolano *et al.* 1997). Table 1 provides a summary of the literature on non-motorist demand modeling. The table provides information on the unit of analysis, the spatial and temporal aggregation level, (non-motorists counts at what spatial temporal unit), methodological framework employed, and different categories of exogenous variables considered. The studies presented in Table 1 are categorized along two streams: (1) studies examining the non-motorist activity only and (2) studies analyzing non-motorist demand and safety simultaneously.

Several observations can be made from Table 1. First, only a small share of studies had developed an integrated framework for analyzing non-motorist demand and safety. Second, several spatial units were considered for analyzing the non-motorist demand including segments, intersections, census block and household amongst which intersections and segments are the most prevalent one. Third, in terms of temporal aggregation, majority of the earlier research has examined the non-motorist traffic at a daily level or at an hourly level. Finally, the methodological frameworks adopted in these studies include Linear regression, Ordinary least square, Negative binomial, Poisson regression, Hurdle negative binomial, Generalized linear mixed model, Time series model and Space syntax tool.

With respect to exogenous variables, the overall findings from earlier research effort are consistent. The various factors identified as influencing non-motorist demand include: (1) socio-demographic characteristics such as population density and household income, (2) land-use characteristics such as residential land use and land use mix, (3) built environment characteristics such as transport accessibility and nearby educational center (such as universities), (4) roadway characteristics such as sidewalk length and presence of traffic signal, and (5) weather variables such as average temperature and precipitation rate.

Non-motorist Safety

There is a vast body of safety literature examining the factors affecting crash occurrence of active travellers (pedestrians and bicyclists) and the severity of different types of non-motorist crashes with motorized vehicles. It is beyond the scope of the paper to review all the research on non-motorists safety (see (Eluru et al., 2008; Cottrill and Thakuriah, 2010; Ukkusuri et al., 2012, 2011; Siddiqui et al., 2012; Abdel-Aty et al., 2013; Wei and Lovegrove, 2013; Yasmin et al., 2014; Lee et al., 2015; Cai et al., 2016; Nashad et al., 2016) for a detailed review). In general, studies evaluating non-motorist road user safety do not consider non-motorist exposure in detail. In our paper, we focus our attention on studies that attempted to incorporate non-motorist exposure in their studies.

A critical component in the process of analyzing non-motorist safety is the selection of appropriate exposure measure. A number of research efforts have examined several surrogate measures to gain a comprehensive understanding of the individual non-motorist crash risk. In general, total five types of matrices are being developed throughout the years to be used as exposure for the non-motorist safety analysis (see (Jamali & Wang, 2017) for detail review) including: (1) area-based approach such as population density (Chakravarthy et al., 2010; Cottrill & Thakuriah, 2010; Wang et al., 2017; Sze et al., 2019), number of trips (Bouaoun et al., 2015; Kerr et al., 2013; Bao et al., 2017; Su et al., 2020; Kamel and Sayed, 2020) and vehicular traffic (C. Lee & Abdel-Aty, 2005; Wier et al., 2009; Wei and Lovegrove, 2013; Kamel and Sayed, 2020);

(2) point-based approach such as non-motorist volume (Lee et al., 2019, 2018a;; Guo et al., 2018; Xie et al., 2018; Ding et al., 2020; Cai et al., 2020; Heydari et al., 2020; Kwayu et al., 2020); (3) segment-based approach such as pedestrian volume at a segment (Clifton et al., 2008); (4) distance-based approach such as distance (Molino et al., 2012); and (5) trip-based approach such as space time prism (Lam et al., 2013, 2014; Yao et al., 2015).

Over the past few years, a significant debate has emerged among the researchers about the best exposure matrix that can explain the non-motorist crash risk. Several studies investigated the link between non-motorist trip frequency and collision; and concluded that motorists were more likely to exhibit safer driving behavior in the presence of higher volumes of pedestrians and bicyclists (Jacobsen, 2015; Schepers, 2012). However, the relationship is non-linear which implies that with the increase in number of walking or bicycling trips, the absolute number of non-motorist involved crashes might increase, but the individual risk of the non-motorist involved in a crash may decline (for example see (Robinson, 2005; Elvik, 2009; Elvik and Bjørnskau, 2017; Xu et al., 2017). Robinson (2005) investigated the accuracy of the relationship between non-motorist volume and safety by conducting a comparison study in Australia and found that the risk per kilometer for cyclist dropped by around 34% when the volume of the cyclists was doubled. In recent years, researchers have geared towards trip-based approach where the amount of time or distance travelled by walking/cycling are used as exposure for examining the non-motorist involved crash occurrences (Lam et al., 2013, 2014; Yao et al., 2015; Yao & Loo, 2016; Ding et al., 2020;). Yao et al. (2015) investigated the vehicle-pedestrian crashes and proposed two approaches for estimating the pedestrian exposure matrix including: (1) Deterministic approach - space time path (SPT) method using the shortest path algorithm based on the origin-destination of the trip and (2) Probabilistic approach - potential path tree (PPT) method that considers all the possible paths between origin and destination given the link level travel time. The findings from the study indicate that while both methods are useful for measuring the pedestrian exposure, the PPT method is more efficient in explaining the manner of vehicle-pedestrian crashes. A relatively recent study by Ding and colleagues (Ding et al., 2020) considered both bicycle use and time duration (based on the public bicycle rental system) to serve as proxy for exposure in their bicycle crash risk model. The authors concluded that the duration of bicycle serve as a better exposure in capturing the interactions between bicycles and motor vehicles as indicated by the superior performance of the model using duration relative to the model considering bicycle frequency as exposure. Another study by Li et al., 2020 considered three different types of exposure including trip frequency, distance travelled and number of roads crossed in their pedestrian crash model and found that model using trip frequency to serve as a surrogate for exposure provided inferior performance (statistical fit and prediction accuracy) compared to the other models.

2.2 Current Study

It is evident from the literature review that apart from a handful of studies, non-motorist safety literature has not adequately addressed the link between non-motorist demand and safety. With growing emphasis on improving mobility in Florida region there are targeted efforts to enhance non-motorist mobility. To evaluate the effectiveness of these strategies and to enhance safety, it is useful to develop methods that accommodate the potential adoption of active modes within the mobility planning process. Thus, developing an integrated framework of demand and safety would allow a seamless evaluation of various scenarios that influence mobility and/or safety. In transportation research, the concept of demand/exposure is defined based on the underlying research question and the intended use of data. The exposure measures can be identified at aggregate or disaggregate level by considering different unit of analysis (as discussed in the earlier research section for the non-motorists group). The distance travelled and/or time-based exposures, disaggregate exposure measures, allows us to disaggregate trip over space and time. However, generating these exposure measures are often computationally burdensome. Therefore, in our study, we resort to using aggregate level demand in terms of total trip as exposure measure, since generation of these measures are practice-oriented and less expensive[1]. Aligned with "trip generation step" of traditional four-step approach of travel demand modeling, we identify the total number of non-motorist trips generated and attracted in different zones. Based on these aggregate counts, we further develop separate models for trip generation and attraction as a function of zonal level attributes. Finally, we hypothesize that total zonal level exposure (zonal level trip generation count + zonal level trip attraction count) would influence zonal level safety and, hence, considered as the exposure matrix in examining zonal-level safety. Thus, in developing an integrated framework of demand and safety for non-motorist, we propose a three-step approach as follows: (1) develop aggregate level models for non-motorist trip generation and attraction at a zonal level, (2) develop non-motorists trip exposure matrices for safety evaluation and (3) develop aggregate level non-motorists crash frequency and crash severity proportion models.

We investigate non-motorists demand at a zonal level by using aggregate trip information based on origin and destination locations of trips. We develop four models: (1) Pedestrian generation model – based on zonal level pedestrian origin trip demand, (2) Pedestrian attraction model – based on zonal level pedestrian destination trip demand, (3) Bicycle generation model –

[1] Chu (2003) identified that distance-based exposure measure can generate misleading result because of difference in travel speed across different trip mode. Using such disaggregate measure in an aggregate level analysis may mislead the outcome. Beyond all the arguments on which measure should adopt, there is no clear consensus in existing safety literature on which exposure measure is more appropriate and more effective. In fact, different exposure measure can lead to different results. In our study design, we hypothesized that the aggregate level demand is surrogate for aggregate level safety analysis. It is beyond the scope of this study is to examine which exposure measure – aggregate/disaggregate – is better representative of non-motorist exposure matrices.

based on zonal level bicycle origin trip demand, (4) Bicycle attraction model – based on zonal level bicycle destination trip demand. In the second step, predicted origin and destination trip counts are used from the exposure models to generate different zonal level trip exposure matrices for both pedestrian and bicycle modes to be considered as non-motorists exposure measures for safety evaluation. Finally, in the third step, we estimate non-motorist safety models by employing predicted exposure matrices, generated from second step, along with other zonal attributes. Specifically, we estimate four different aggregate level safety 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 Central Florida Region as a function of zonal level sociodemographic characteristics, roadway/traffic attributes, built environment and land-use characteristics. The applicability of the framework is illustrated through extensive policy scenario analysis.

The rest of the paper is organized as follows: The description of the data and exogenous variables adopted in the analysis are describes in Section 3. Model estimation results are presented in Section 4. Section 5 presents the policy scenario analysis and finally, conclusions are presented in section 6.

3. DATA

3.1 Study Area and Data Sources

The study area is the Central Florida Region defined by Central Florida Regional Planning Model version 6.0 (CFRPM 6.0) which includes 4,747 traffic analysis zones (TAZ). Data for developing non-motorist exposure models are sourced from 2009 National Household Travel Survey (NHTS) Add-on database provided by the Florida Department of Transportation (FDOT) that allowed us to geo-tag trips recorded in the Central Florida region. In the dataset, there were 2,749 households, 5,090 individuals and 22,359 trips. Among these trips, walk and bike trip shares were 8.8% and 1.3%, respectively. Data for the non-motorist safety analysis is compiled from FDOT Crash Analysis Reporting System (CARS) and Signal Four Analytics (S4A) databases. CARS and S4A are long and short forms of crash reports in the State of Florida, respectively. The long form crash report includes higher injury severity level or crash related to criminal activities (such as hit-and-run or Driving Under Influence). The Short Form Report is used to report all other types of traffic crashes. Crash data records from short and long form databases are compiled to generate complete information on road crashes and hence are used for the purpose of analysis in the current study context. For this study, we have examined the pedestrian and bicycle crash events for the year 2010 to incorporate the exposure measures in terms of non-motorist safety[2]. For the year 2010, 1,474 and 1,012 crashes were reported involving pedestrian and bicycle, respectively.

3.2 Data Description

The dependent variables for the exposure models are daily zonal origin trip count and daily zonal destination trip count for pedestrians and bicyclists. We incorporate "person-trip weight" – as

[2] The proposed integrated demand-safety approach can be employed by using recent data, if both demand and safety data are available to maintain the same base year condition. In our study, we had access to the demand data for the year 2009 from NHTS data. Therefore, for safety models, we have selected year 2010 to reflect the base year demand condition from 2009 NHTS data.

defined in NHTS database – to extrapolate the non-motorized trips to represent number of trips for the zones in the Central Florida region. Locations of zones with pedestrian and bicycle O-D demand are shown in Figure 2. With respect to safety component, the geo-coded crash data involving non-motorists are aggregated at the level of TAZ for the year 2010. These crashes are further classified by crash severity outcomes (property damage only (PDO), possible injury, non-incapacitating, incapacitating injury and fatal crashes) at the zonal level. Locations of zones with pedestrian and bicycle crashes (total crashes and by crashes by different injury severity levels) are shown in Figures 3 and 4, respectively. The corresponding variables by proportion (number of specific severity level/total number of all crashes) include: (1) proportion of PDO crashes, (2) proportion of possible injury crashes, (3) proportion of non-incapacitating injury crashes, (4) proportion of incapacitating injury crashes and (5) proportion of fatal crashes. The dependent variables and sample size for both exposure and safety models are presented in Table 2. The crash proportion models for pedestrian and bicyclists are estimated only for zones with non-zero crashes.

In addition to the different zonal level dependent variables, 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 characteristics and land use characteristics. To ensure that the exogenous variables considered reflect the analysis year trend, we generate these variables using data from 2010. Table 3 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.

4. EMPIRICAL ANALYSIS

The model estimation results for different components are discussed separately. The final specifications of the models were based on removing the statistically insignificant variables in a systematic process based on statistical significance (90% confidence level). 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 as presented in Table 3.

4.1 Exposure Models

In estimating aggregate level exposure models, the non-motorist trip demand is represented as total number of non-motorist trips originated from and destined to at a zonal resolution. Thus, the demands are non-negative integer values. Naturally, these integer values 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 (1998) 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 interpreted as a two-part model: 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). Therefore, to accommodate for the preponderance of zero trip counts, exposure models are developed using Hurdle Negative Binomial (HNB) regression approach (see Cai *et al.* 2016 for methodological framework)[3]. Table 4 presents the estimation results of the exposure models: pedestrian trip generation (2nd and 3rd columns), pedestrian trip attraction (4th and 5th columns), bicycle trip generation (6th and 7th columns), and bicycle trip attraction (8th and 9th columns) models. In the Hurdle model, the positive (negative) coefficient in the probabilistic component corresponds to increased (decreased) propensity of non-zero trip events. The positive (negative) coefficient in the count component of the Hurdle model corresponds to increased (decreased) non-zero trip count events. Pedestrian and bicycle trip demand models are discussed in the following sections.

Pedestrian Trip Demand Models

<u>Probabilistic Component:</u> In the probabilistic components, land-use mix, urban area and number of households are found to be significant in both pedestrian trip generation and attraction models. 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 trip generation and attraction.

<u>Count Component:</u> With respect to sociodemographic characteristics, from Table 4 we can see that proportion of 65+ aged population is positively associated with pedestrian trip generation. Zones with higher average speed limit on roadways are likely to generate less pedestrian trip origin demand (Pulugurtha and Repaka, 2008). Annual average daily traffic (AADT) is negatively associated with both pedestrian demand components. Higher proportion of arterial roads in zones are likely to increase pedestrian activities (Hankey et al., 2012) for generation and attraction. Higher proportion of roadways with 3 or more lanes are negatively associated with zonal level

[3] The econometric framework of HNB model is presented in APPENDIX A.

pedestrian activities. As expected, zones with higher length of sidewalk are likely to have higher level of pedestrian activities (Lu et al., 2018).

With respect to built environment, we find that higher number of business centers, entertainment centers, financial centers, park/recreational centers and shopping centers are positively associated with pedestrian attraction. On the other hand, higher number of transit hubs and restaurants are found to be negatively associated with pedestrian destination demand. Land-use characteristics are found to have significant influence in both pedestrian trip generation and attraction models. Among different land-use categories, industrial area is found to be negatively associated with pedestrian trip origin and trip destination demands (see (Hankey and Lindsey, 2016; Lu et al., 2018) for similar result). All other land-use categories (recreational, residential, retail/office and institutional area) are likely to generate higher level of pedestrian activities.

Bicycle Trip Demand Models

<u>Probabilistic Component:</u> Land-use mix, urban area and number of households are found to be significant in both bicycle trip generation and attraction models. 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 higher levels of bicycle trip generation and attraction.

<u>Count Component:</u> Proportion of 65+ aged population is negatively associated with bicycle generation indicating that TAZs with higher number of population aged 65+ have lower bicycle origin demand (Guo et al., 2007). AADT is negatively associated with bicycle trip generation component. Furthermore, higher proportion of arterial roads in zones are likely to have higher bicycle activity (see (Nordback et al., 2017)). Higher proportion of roadway with 3 or more lanes

are negatively associated with zonal level bicycle activities. Zones with higher sidewalk length are likely to have higher level of bicycle activities.

Built environment attributes are considered only in bicycle attraction model. The study finds that higher number of education centers, entertainment centers, park/recreational centers, restaurants and transit hubs are positively associated with bicycle attractions. On the other hand, higher number of commercial centers, financial centers and shopping centers impose a negative effect on bicycle destination demand. Among different land-use categories, industrial, residential and institutional areas are found to be positively associated with bicycle activities (Tabeshian and Kattan, 2014). With respect to recreational area, the variable shows positive association in bicycle generation model but has a negative correlation with bicycle attraction model. On the other hand, retail/office area is found to be negatively associated with both bicycle trip origin and destination demand (Chen et al., 2017).

4.2 Non-motorist Trip Exposure Matrices

In evaluating non-motorist exposure, we also generate different zonal level trip exposure matrices with the predicted number of daily trip origin (by using trip generation model results) and daily trip destination (by using trip attraction model results) at zonal level for both pedestrian and bicycle group of road users. Then, 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 [4747×1] with total trip counts across different rows. The total zonal level trip demand matrices are generated for pedestrians and bicyclists separately, which is used as exogenous variables in developing safety models along with other zonal attributes.

4.3 Safety Models

We estimate two crash count models and two crash proportions by severity models for pedestrians and bicyclists. Crash count models are developed by using NB model, while the crash proportions by severity models are developed using Ordered Probit Fractional Split (OPFS) approach (see Bhowmik et al., 2018; Lee et al., 2018b for unordered fractional split structure and Yasmin and Eluru 2018; Yasmin et al. 2016; Bhowmik et al., 2019c for ordered fractional split structure). 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 are developed in this study by using the NB modeling approach[4]. 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 and policy implications, it is a common practice among safety researcher community to develop independent crash prediction models for different injury severity levels. However, for the same observation record, 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. Such a framework would allow us to make inferences based on a single model. 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. Specifically, we employ OPFS models for examining pedestrian and bicycle crash proportions by severity levels[5]. The estimation results of these models are presented in the following sections.

[4] The econometric framework of NB model is presented in APPENDIX B.
[5] The econometric framework of OPFS model is presented in APPENDIX C.

Crash Count Models

Table 5 presents the estimation results of the count models. The pedestrian crash count model results are presented in 2nd and 3rd columns while the bicycle crash count model results are presented in 4th and 5th columns.

The model results indicate that both pedestrian and bicycle crashes are positively associated with population density (see (Bhowmik et al., 2019b) for similar results) i.e. zones with higher population density are likely to experience more pedestrian and bicycle crashes (as expected). The results, surprisingly, indicate a reduced crash risk for both pedestrian and bicyclists with higher proportion of population aged 65 and more. One reasonable explanation can be attributed to the fact that senior people are more experienced which eventually protects them from colliding with the motor vehicles. Similar results are also observed in the study of Saha et al. (2018). Several roadway and traffic attributes are found to be significant determinants of non-motorist 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 (Nashad et al., 2016). The result is expected as the density of traffic intersections increases potential conflicts between vehicles and non-motorist road users are likely to increase. Higher proportion of arterial roads results in higher pedestrian and bicycle crash risks (Bhowmik et al., 2019a; Nashad et al., 2016). At the same time, higher proportion of local roads is found to have negative impact of bicycle crash risk. From Table 5, we can see that the likelihood of pedestrian crash is higher in zones with longer sidewalk length. This is intuitive as sidewalk lengths are reflections of pedestrian access. For instance, zones with higher length of sidewalk are likely to have higher level of pedestrian activities (Lu et al., 2018) which eventually results in more pedestrian crashes (see Cai et al., 2016; Nashad et al; ,2016 for similar results). Similarly, TAZs with longer bicycle lane length have an increased likelihood of bicycle crashes. The length of zonal level bus lane result reveals an increasing likelihood of bicycle crash risk. An increase in zonal AADT increases the likelihood of both pedestrian and bicycle crashes at the TAZ level. The result in bicycle crash model suggests that zones with higher truck AADT have a decreased likelihood of bicycle crashes possibly because bicycling is less prevalent in these zones (see Cai et al., 2016 for similar results). Trucks usually travel on the highways in their majority part of the trips. As a result, zones with higher truck volume are basically the zones with major highways which explain the reduced likelihood of bicycle crashes in those zones.

With respect to built environment, the estimation results of pedestrian crash risk model reveal that higher number of educational centers, transit hubs, restaurants and park/recreational centers result in higher pedestrian crash risk at zonal level. On the other hand, bicycle crash risk is positively associated with higher number of commercial centers, financial centers, restaurants and hospitals. 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 area. 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 increase the propensity of both pedestrian and bicycle crashes.

The major objective of the current study is to integrate the non-motorist trip exposure as exogenous variable in developing aggregate level crash risk models. We use total daily trip demand of pedestrian and bicycle (as explained in section 4.2) 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 are considered in the final specifications. With respect to pedestrian crash risk model, pedestrian 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 5, we can see that higher number of pedestrians per household is likely to decrease the risk of pedestrian-motor vehicle crashes (Miranda-Moreno *et al.* 2011). The result perhaps is indicating that the motorists are more likely to exhibit safer driving behavior in the presence of higher volumes of pedestrians (Jacobsen, 2015; Schepers, 2012). With respect to bicycle crash risk model, bicycle exposure measures are found to have significant impact on zonal level bicycle-motor vehicle crash risk. The estimation result of exposure measure in bicycle crash risk model reveal that higher bicyclists trip demand at a zonal level increases the risk of bicycle crashes. The reader would note that even after controlling for trip exposure variables, several variables from other variable categories still serve as proxies for exposure.

Crash Proportions by Severity Models

Table 6 presents the estimation results of the crash proportions by severity models. The pedestrian crash proportions by severity model results are presented in 2nd and 3rd columns and bicycle crash proportions by severity model results are presented in 4th and 5th columns of Table 6. These models are estimated by using OPFS framework. The effects of exogenous variables in model specifications for both pedestrian and bicycle crash proportions by severity models are discussed in this section. In OPFS models, the positive (negative) coefficient corresponds to increased (decreased) proportion for severe injury categories.

With respect to sociodemographic characteristics, the estimates indicate that population density results in lower likelihood of severe injury proportions for both pedestrian and bicycle crashes. Proportion of 22-29 years old group of population has negative impact on proportion of pedestrian crash severity outcomes implying a reduced likelihood of more severe pedestrian crashes (Yasmin et al., 2014). Relative to older people, young individuals are more flexible in handling any sudden activity which in turn protects them from enduring severe injuries (see Saha et al., 2018 for similar results). The OPFS model results for bicycle reveal a higher proportion of severe crash outcomes for zones with higher number of flashing beacon signs and higher number of school signals. As expected, availability of bike lane is found to reduce the likelihood of less severe bicycle crash proportions. With respect to traffic attributes, higher vehicles miles travelled (VMT) is positively associated with more severe crash proportions in the model for pedestrians.

The pedestrian severity model reveals that the proportion of severe crashes is lower in TAZs with higher number of commercial centers (Moudon et al., 2011; Aziz et al., 2013). Higher number of hospitals is associated with lower likelihood of severe crash proportion in OPFS model for bicycle crashes. At the same time, the OPFS model results reveal that higher number of park and recreational centers increases the possibility of higher proportions of severe bicycle crash outcomes. From both pedestrian and bicycle models, we find that the possibility of more severe crashes decreases with increasing share of urbanized area of a TAZ (Boufous et al., 2012). 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 bicycle crash severity model suggesting that proportion of severe bicycle crashes increases with increasing zonal level residential area. Similar results are also observed in the study of Bhat et al. (2017).

The non-motorist exposure measures generated (as presented in Section 4.2) are used as exogenous variables in evaluating zonal level pedestrian and bicycle crash severity proportions. With respect to, pedestrian crash severity proportion model, higher pedestrian demand per household at a zonal level decreases the propensity of higher proportion of severe crashes. With respect to bicycle crash severity proportion model, increase in bicycle trip demand per household at a zonal level decreases the risk of higher proportion of severe bicycle-motor vehicle crashes. The reader would note that the impact of exposure is contrasting in the count and severity models highlighting how increased exposure is likely to increase the number of crashes but at the same time contributing to reduced proportion of severe crashes.

4.4 Predictive Performance Evaluation

In order to demonstrate the predictive performance of the estimated exposure and safety (count and severity) models, a validation exercise is also carried out. The most common approach of performing validation exercise for aggregate level model is to evaluate the in-sample predictive measures. Therefore, to evaluate the predictive performance of the estimated eight models, we compute the predicted count/proportion events and compared those with the observed values across different zones. For demand models, 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. For crash frequency models, we compute mean prediction bias (MPB) and mean absolute deviation (MAD). 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:

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

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

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

where, $\hat{y}_i$ and y_i are the predicted and observed values for event i (i be the index for event ($i = 1,2,3, \dots, N$)) and n is the number of events. These measures are presented in Table 7. From Table 7 we can see that the error between observed and predicted values across all events of different models are quite small (ranging from -6.452% to 0.451%). Hence, we can conclude that the predictive performance of the estimated models is reasonable for all eight estimated models.

5. IMPLICATIONS

5.1 Policy Scenario Analysis

The parameter effects of exogenous variables as presented in Sections 4 do not directly provide the magnitude of the effects on zonal level non-motorists demand and safety 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 trip demand models and safety models (see Eluru and Bhat 2007 for a discussion on the methodology for computing elasticities). We investigate the effect as percentage change in the expected change in zonal demand, change in zonal crash counts, and change in 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 scenario as follows:

- Scenario 1: 50% reduction in traffic volume within 2 miles buffer area of different central business district (CBD).
- Scenario 2: 30% reduction in traffic volume within 2 miles buffer area of different central business district (CBD).
- Scenario 3: 15% reduction in traffic volume within 4 miles buffer area of different central business district (CBD).
- Scenario 4: 5% reduction in traffic volume within 6 miles buffer area of different central business district (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 pedestrian and bicycle group of road users separately. For the buffer area around CBD scenarios, we consider multiple CBDs in the Central Florida region including Orlando, Sanford, Lakeland, Kissemme, Deland, Ocala, Melbourne, Palm Bay, Leesburg, Daytona Beach and Port Orange of Central Florida region. The spatial representation of the considered CBD locations is shown in Figure 1. By performing policy scenario analysis for exposure and safety components, we ensure that the updated demand matrices for each scenario is produced and employed in generating exposure measures for non-motorist travel as well as vehicular volumes on roadways. With these new exposure measures, the safety models are used to generate estimates of scenario-based crash and severity proportions and the change in safety situation. By following the simulation procedure, it is possible to predict demand matrices for future year and in turn predict safety by incorporating exposure measures. A comparison across scenarios would allow us to identify beneficial changes to existing infrastructure for improving non-motorist road user safety. Policy scenario analysis for non-motorist travel demand and safety components are presented in Table 8. We generated elasticity effects for all severity levels in crash proportion by severity models. However, we present the elasticity effects only for the highest injury severity category (fatal crash proportions). The following observations can be made based on the elasticity effects presented in Table 8.

With respect to demand component, we can observe that - First, decreasing vehicular traffic volume near CBD locations has greater effect on pedestrian demand than bicycle demand. For both modes, we can observe from the table that higher level of non-motorist activities can be attained by restricting vehicular traffic; greater the restrictions, higher the level of non-motorist demand. Second, increasing sidewalk facilities are likely to attract more non-motorists, but for the hypothetical scenario 5, the demand for pedestrian is likely to get reduced. Third, the reduction in speed has greater impact on increasing pedestrian demand. However, for bicycle, the variable has no impact as it was found insignificant in bicycle demand models. Fourth, restriction in number of traffic lanes are likely to have similar impact and as we can see from Table 8, it increases non-motorists demand.

With respect to crash count component, we can observe that - First, decreasing vehicular traffic volume near CBD locations are likely to reduce pedestrian crashes with greater impact within the vicinity of CBD. However, bicycle crashes are likely to increase by about 3%. But number of bicycle-motor vehicle crashes are likely to decrease within the vicinity of CBD with greater reduction in vehicular volume. Second, the hypothetical scenario of sidewalk length shows that providing walk facilities has the potential to improve pedestrian safety. On the other hand, bicycle crashes are likely to be adversely affected by increasing sidewalk length – perhaps indicating greater exposure. Third, reduction in speed and restrictions in traffic lanes decreases pedestrian crashes. On the other hand, restrictions in traffic lanes increases bicycle crashes by about 4%.

With respect to crash severity by proportions component, we can observe that - First, non-motorist friendly facilities are likely to reduce proportion of fatal crashes for both pedestrians and bicyclists. However, the impact on pedestrian mode is much higher than the impact on bicycle mode. Second, the decrease in pedestrian fatal crash severity proportions are about 1% for increase in sidewalk length, reducing speed and restricting traffic lanes. The contribution of these measures on bicycle crash severity are less pronounced relative to pedestrian modes.

It is a well-known fact that non-motorist safety tend to decrease with increasing non-motorist exposure, and only after a certain level of exposure (when traffic become familiar with higher number of non-motorists), the safety tends to improve. From the policy analysis, we can see that non-motorist friendly infrastructure has mixed effect on non-motorist safety in current study context. Therefore, it is imperative that policy implications for improving non-motorist safety should be identified by considering all known exogenous elements in identifying the appropriate tools. In general, providing more walking and bicycle friendly facilities are likely to encourage more people to use non-motorized mode and in targeted zones these measures are likely to improve non-motorist safety.

5.2 Predictions for Future Year

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 trip origin demand, predicted trip destination demand and predicted total trip demand for the year 2015. In generating demand matrices for the year 2015, we consider the increase in weight based on the change in population at a zonal level. These matrices are presented in Table 9 at the county level (the matrices are generated at the zonal-level but are shown at the county level for presentation purposes). From Table 9 we can see that overall bicycle demand has increased from 2010 to 2015, but total pedestrian demand has decreased over the same period. Similar matrices can be generated for any other year. These generated demand matrices can be further be used as non-motorist exposure measures in the safety model predictions for the year 2015. We employ the predicted results for the year 2015 to plot the spatial distribution of predicted crash counts and predicted crash counts by severity levels for both non-motorist road user groups. These plots are presented in Figure 5 (5(a) – 5(c)). From the spatial representation, we can see that high crash risk zones and zones with higher proportion of severe crashes are dispersed throughout the state.

6. CONCLUSION

In developing non-motorist crash prediction models safety researchers have employed land use and urban form variables as surrogate for exposure information (such as pedestrian, bicyclist volumes and vehicular traffic). The quality of these crash prediction models is affected by the lack of "true" non-motorist exposure data. Modeling with high-resolution data frameworks such as activity-based or trip-based approach could be pursued for evaluating planning level non-motorist demand. However, running a travel demand model system to generate demand inputs for non-motorist safety evaluation is cumbersome and resource intensive. The current study focused on addressing this drawback by developing an integrated non-motorist trip demand and crash prediction framework for mobility and safety analysis. Towards this end, we proposed a three-step framework to evaluate non-motorists safety: (1) develop aggregate level models for non-motorist trip generation and attraction at a zonal level, (2) develop non-motorists trip exposure matrices for safety evaluation and (3) develop aggregate level non-motorists crash frequency and severity proportion models.

The three-step approach entailed estimation of eight different models for pedestrian and bicyclist road user groups – two trip attraction models, two trip generation models, two crash count models and two crash proportions by severity models. The model estimation was based on National Household Travel Survey (NHTS) Florida Add-on and Florida Department of Transportation non-motorist crash data. The integrated framework was employed for policy scenario analysis. The results provided useful insights on mobility and safety changes associated with these hypothetical scenarios.

To be sure, our study is not without limitations. We evaluated non-motorist demand by using NHTS database at an aggregate level which is not readily transferable for developing micro-level model. It might be interesting to generate micro-level trip demand model to identify non-motorist exposure at a corridor level. It might also be useful to conduct a pooled model estimation with random effects for pedestrians and bicyclists to improve model estimation efficiency.

ACKNOWLEDGEMENT

The authors would like to thank Florida Department of Transportation (FDOT) for providing additional funds as well as data access to support the research. We also thank SAFER-SIM

University Transportation Center for partially funding the research effort. We are grateful to S4A for providing the crash data used in the research.

REFERENCES

Abdel-Aty, M., Lee, J., Siddiqui, C., & Choi, K., (2013). Geographical unit based analysis in the context of transportation safety planning. Transportation Research Part A: Policy and Practice, 49, 62-75.

Aziz, H. A., Ukkusuri, S. V., & Hasan, S. (2013). Exploring the determinants of pedestrian–vehicle crash severity in New York City. Accident Analysis & Prevention, 50, 1298-1309.

Bao, J., Liu, P., Yu, H., & Xu, C. (2017). Incorporating twitter-based human activity information in spatial analysis of crashes in urban areas. Accident Analysis & Prevention, 106, 358-369.

Bhat, C. R., Astroza, S., & Lavieri, P. S. (2017). A new spatial and flexible multivariate random-coefficients model for the analysis of pedestrian injury counts by severity level. Analytic methods in accident research, 16, 1-22.

Bhowmik, T., Rahman, M., Yasmin, S., & Eluru, N. (2019a). Alternative Model Structures for Multivariate Crash Frequency Analysis: Comparing Simulation-based Multivariate Model with Copula-based Multivariate Model (No. 19-05507). Presented at the 98th Transportation Research Board (TRB) Annual Meeting, Washington D.C., 2019.

Bhowmik, T., Yasmin, S., & Eluru, N. (2018). A joint econometric approach for modeling crash counts by collision type. Analytic methods in accident research, 19, 16-32.

Bhowmik, T., Yasmin, S., & Eluru, N. (2019b). Do we need multivariate modeling approaches to model crash frequency by crash types? A panel mixed approach to modeling crash frequency by crash types. Analytic Methods in Accident Research, 24, 100107.

Bhowmik, T., Yasmin, S., & Eluru, N. (2019c). A multilevel generalized ordered probit fractional split model for analyzing vehicle speed. Analytic methods in accident research, 21, 13-31.

Bouaoun, L., Haddak, M. M., & Amoros, E., (2015). Road crash fatality rates in France: a comparison of road user types, taking account of travel practices. Accident Analysis & Prevention, 75, 217-225.

Boufous, S., de Rome, L., Senserrick, T., & Ivers, R., (2012). Risk factors for severe injury in cyclists involved in traffic crashes in Victoria, Australia. Accident Analysis & Prevention, 49, 404-409.

Cai, Q., Abdel-Aty, M., & Castro, S. (2020). Explore effects of bicycle facilities and exposure on bicycle safety at intersections. International Journal of Sustainable Transportation, 1-12.

Cai, Q., Lee, J., Eluru, N., Abdel-Aty, M., (2016). Macro-level pedestrian and bicycle crash analysis: Incorporating spatial spillover effects in dual state count models. Accident Analysis & Prevention 93, 14-22.

Cameron, A. C., Trivedi, P. K., (1998). Regression analysis of count data. New York: Cambridge University Press.

Cervero, R., Duncan, M., (2003). Walking, bicycling, and urban landscapes: Evidence from the san francisco bay area. American journal of public health 93 (9), 1478-1483.

Chakravarthy, B., Anderson, C. L., Ludlow, J., Lotfipour, S., & Vaca, F. E., (2010). The relationship of pedestrian injuries to socioeconomic characteristics in a large Southern California County. Traffic injury prevention, 11(5), 508-513.
Chen, P., Zhou, J., & Sun, F., (2017). Built environment determinants of bicycle volume: A longitudinal analysis. Journal of Transport and Land Use, 10(1), 655-674.
Chu, X. (2003). The fatality risk of walking in America: A time-based comparative approach. In Walk21 conference: Health equity and the environment. Portland, Oregon.
Clifton, K. J., Singleton, P. A., Muhs, C. D., & Schneider, R. J., (2016). Representing pedestrian activity in travel demand models: Framework and application. Journal of transport geography, 52, 111-122.
Clifton, K., Burnier, C., Schneider, R., Huang, S., & Kang, M. W., (2008). Pedestrian demand model for evaluating pedestrian risk exposure. University of Maryland.
Cottrill, C. D., & Thakuriah, P. V., (2010). Evaluating pedestrian crashes in areas with high low-income or minority populations. Accident Analysis & Prevention, 42(6), 1718-1728.
Dhanani, A., Tarkhanyan, L., & Vaughan, L., (2017). Estimating pedestrian demand for active transport evaluation and planning. Transportation research part A: policy and practice, 103, 54-69.
Ding, H., Sze, N. N., Li, H., & Guo, Y. (2020). Roles of infrastructure and land use in bicycle crash exposure and frequency: a case study using Greater London bike sharing data. Accident Analysis & Prevention, 144, 105652.
Eluru, N., Bhat, C. R., & Hensher, D. A., (2008). A mixed generalized ordered response model for examining pedestrian and bicyclist injury severity level in traffic crashes. Accident Analysis & Prevention, 40(3), 1033-1054.
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.
Elvik, R., & Bjørnskau, T., (2017). Safety-in-numbers: a systematic review and meta-analysis of evidence. Safety science, 92, 274-282.
Elvik, R., (2009). The non-linearity of risk and the promotion of environmentally sustainable transport. Accident Analysis & Prevention, 41(4), 849-855.
Ercolano, J., Olson, J., Spring, D., (1997). Sketch-plan method for estimating pedestrian traffic for central business districts and suburban growth corridors. Transportation Research Record: Journal of the Transportation Research Board (1578), 38-47.
Ermagun, A., Lindsey, G. and Loh, T.H., (2018a). Urban trails and demand response to weather variations. Transportation Research Part D: Transport and Environment, 63,404-420.
Ermagun, A., Lindsey, G. and Loh, T.H., (2018b). Bicycle, pedestrian, and mixed-mode trail traffic: A performance assessment of demand models. Landscape and Urban Planning, 177, 92-102.
Fagnant, D.J. and Kockelman, K., (2016). A direct-demand model for bicycle counts: The impacts of level of service and other factors. Environment and Planning B: Planning and Design, 43(1), 93-107.
Fatality in Collisions with Motor Vehicles, a Social Ecological Study of State Routes and City Streets in King County, Washington. Accident Analysis & Prevention 43 (1), 11-24.
Fournier, N., Christofa, E. and Knodler Jr, M.A., (2017). A sinusoidal model for seasonal bicycle demand estimation. Transportation research part D: transport and environment, 50, 154-169.

Guo, J., Bhat, C., Copperman, R., (2007). Effect of the built environment on motorized and nonmotorized trip making: Substitutive, complementary, or synergistic? Transportation Research Record: Journal of the Transportation Research Board 2010(1), 1-11.

Guo, Y., Osama, A., & Sayed, T. (2018). A cross-comparison of different techniques for modeling macro-level cyclist crashes. Accident Analysis & Prevention, 113, 38-46.

Hankey, S., & Lindsey, G. (2016). Facility-demand models of peak period pedestrian and bicycle traffic: comparison of fully specified and reduced-form models. Transportation research record, 2586(1), 48-58.

Hankey, S., Lindsey, G., & Marshall, J. D. (2017). Population-level exposure to particulate air pollution during active travel: planning for low-exposure, health-promoting cities. Environmental health perspectives, 125(4), 527-534.

Hankey, S., Lindsey, G., Wang, X., Borah, J., Hoff, K., Utecht, B. and Xu, Z., (2012). Estimating use of non-motorized infrastructure: Models of bicycle and pedestrian traffic in Minneapolis, MN. Landscape and Urban Planning, 107(3), 307-316.

Hewawasam, C., Bandara, S., & Wirasinghe, S. C. (2014). Development of walking trip rates for different land uses.

Heydari, S., Miranda-Moreno, L., & Hickford, A. J. (2020). On the causal effect of proximity to school on pedestrian safety at signalized intersections: a heterogeneous endogenous econometric model. Analytic methods in accident research, 26, 100115.

Jacobsen, P. L. (2015). Safety in numbers: more walkers and bicyclists, safer walking and bicycling. Injury prevention, 21(4), 271-275.Jamali, A., & Wang, Y. (2017). Estimating pedestrian exposure for small urban and rural areas. Transportation research record, 2661(1), 84-94.

Jones, M., Ryan, S., Donlon, J., Ledbetter, L., Ragland, D., Arnold, L., (2010). Measuring bicycle and pedestrian activity in san diego county and its relationship to land use, transportation, safety, and facility type. Berkeley: Institute of Transportation Studies-UC Berkeley Traffic Safety Center.

Kamel, M. B., & Sayed, T. (2020). Cyclist-vehicle crash modeling with measurement error in traffic exposure. Accident Analysis & Prevention, 144, 105612.

Kerr, Z.Y., Rodriguez, D.A., Evenson, K.R. and Aytur, S.A., (2013). Pedestrian and bicycle plans and the incidence of crash-related injuries. Accident Analysis & Prevention, 50, 1252-1258.

Kraemer, J.D., Zaccaro, H.N., Roffenbender, J.S., Baig, S.A., Graves, M.E., Hauler, K.J., Hussain, A.N. and Mulroy, F.E., (2015). Assessing the potential for bias in direct observation of adult commuter cycling and helmet use. Injury prevention, 21(1), 42-46.

Kwayu, K. M., Kwigizile, V., & Oh, J. S. (2020). Development of systemwide pedestrian safety performance function using stratified random sampling and a proxy measure of pedestrian exposure. International journal of injury control and safety promotion, 1-12.

Lam, W.W., Loo, B.P. and Yao, S., (2013). Towards exposure-based time-space pedestrian crash analysis in facing the challenges of ageing societies in Asia. Asian Geographer, 30(2), 105-125.

Lam, W.W., Yao, S. and Loo, B.P., (2014). Pedestrian exposure measures: a time-space framework. Travel behaviour and society, 1(1), 22-30.

Lee, C. and Abdel-Aty, M., (2005). Comprehensive analysis of vehicle–pedestrian crashes at intersections in Florida. Accident Analysis & Prevention, 37(4), 775-786.

Lee, J., Abdel-Aty, M., & Jiang, X. (2015). Multivariate crash modeling for motor vehicle and non-motorized modes at the macroscopic level. Accident Analysis & Prevention, 78, 146-154.

Lee, J., Abdel-Aty, M., & Shah, I. (2019). Evaluation of surrogate measures for pedestrian trips at intersections and crash modeling. Accident Analysis & Prevention, 130, 91-98.

Lee, J., Abdel-Aty, M., Cai, Q., Wang, L., & Huang, H. (2018a). Integrated modeling approach for non-motorized mode trips and fatal crashes in the framework of transportation safety planning. Transportation research record, 2672(32), 49-60.

Lee, J., Yasmin, S., Eluru, N., Abdel-Aty, M., & Cai, Q. (2018b). Analysis of crash proportion by vehicle type at traffic analysis zone level: A mixed fractional split multinomial logit modeling approach with spatial effects. Accident Analysis & Prevention, 111, 12-22.

Li, H., Wu, D., Graham, D. J., & Sze, N. N. (2020). Comparison of exposure in pedestrian crash analyses: A study based on zonal origin-destination survey data. Safety science, 131, 104926.

Lu, T., Mondschein, A., Buehler, R. and Hankey, S., (2018). Adding temporal information to direct-demand models: Hourly estimation of bicycle and pedestrian traffic in Blacksburg, VA. Transportation Research Part D: Transport and Environment, 63, 244-260.

Matlick, J. M. (1998). If we build it, will they come?. Washington State Department of Transportation.

Miranda-Moreno, L. F., & Fernandes, D. (2011). Modeling of pedestrian activity at signalized intersections: land use, urban form, weather, and spatiotemporal patterns. Transportation research record, 2264(1), 74-82.

Miranda-Moreno, L. F., Morency, P., & El-Geneidy, A. M. (2011). The link between built environment, pedestrian activity and pedestrian–vehicle collision occurrence at signalized intersections. Accident Analysis & Prevention, 43(5), 1624-1634.

Molino, J. A., Kennedy, J. F., Inge, P., Bertola, M. A., Beuse, P. A., Fowler, N. L., ... & Do, A. (2012). A Distance-Based Method to Estimate Annual Pedestrian and Bicyclist Exposure in an Urban Environment (No. FHWA-HRT-11-043 de). United States. Federal Highway Administration. Office of Safety Research and Development.

Moudon, A. V., Lin, L., Jiao, J., Hurvitz, P., Reeves, P., (2011). The Risk of Pedestrian Injury and

Nashad, T., Yasmin, S., Eluru, N., Lee, J., & Abdel-Aty, M. A. (2016). Joint modeling of pedestrian and bicycle crashes: copula-based approach. Transportation Research Record: Journal of the Transportation Research Board, (2601), 119-127.

NHTSA, 2015. Traffic Safety Facts 2015: A Compilation of Motor Vehicle Crash Data from the Fatality Analysis Reporting System and the General Estimates System. Publication DOT HS 812 139. National Highway Traffic Safety Administration (NHTSA), U.S. Department of Transportation.

Nordback, K., Sellinger, M., Phillips, T., (2017). Estimating Walking and Bicycling at the State Level. doi:10.15760/trec.163

Porter, C., Suhrbier, J., & Schwartz, W. (1999). Forecasting bicycle and pedestrian travel: State of the practice and research needs. Transportation Research Record, 1674(1), 94-101.

Pulugurtha, S. S., & Repaka, S. R. (2008). Assessment of models to measure pedestrian activity at signalized intersections. Transportation Research Record, 2073(1), 39-48.

Raford, N., Ragland, D.R., (2004). Space Syntax : An Innovative Pedestrian Volume Modeling Tool for Pedestrian Safety. Transportation Research Record: Journal of the Transportation Research Board 1878, 66–74.

Reardon, T. G., Wallace, E., & Brown, C. M. (2017). Network Utility Scores: A New Approach to Estimating Bicycle and Pedestrian Latent Demand (No. 17-06644).

Robinson, D.L., (2005). Safety in numbers in australia: More walkers and bicyclists, safer walking and bicycling. Health promotion journal of Australia 16 (1), 47-51.

Saha, D., Alluri, P., Gan, A., & Wu, W. (2018). Spatial analysis of macro-level bicycle crashes using the class of conditional autoregressive models. Accident Analysis & Prevention, 118, 166-177.)

Schepers, P., (2012). Does more cycling also reduce the risk of single-bicycle crashes? Injury prevention 18 (4), 240-245.

Schneider, R. J., Arnold, L. S., & Ragland, D. R. (2009a). Methodology for counting pedestrians at intersections: use of automated counters to extrapolate weekly volumes from short manual counts. Transportation research record, 2140(1), 1-12.

Schneider, R. J., Arnold, L. S., & Ragland, D. R. (2009b). Pilot model for estimating pedestrian intersection crossing volumes. Transportation research record, 2140(1), 13-26.

Schneider, R. J., Henry, T., Mitman, M. F., Stonehill, L., & Koehler, J. (2012). Development and application of volume model for pedestrian intersections in San Francisco, California. Transportation research record, 2299(1), 65-78.

Siddiqui, C., Abdel-Aty, M., & Choi, K. (2012). Macroscopic spatial analysis of pedestrian and bicycle crashes. Accident Analysis & Prevention, 45, 382-391.

Strauss, J., Miranda-Moreno, L.F., Morency, P., (2013). Cyclist activity and injury risk analysis at signalized intersections: A Bayesian modelling approach. Accident Analysis & Prevention.59, 9–17.

Strauss, J., Miranda-Moreno, L.F., Morency, P., (2015). Mapping cyclist activity and injury risk in a network combining smartphone GPS data and bicycle counts. Accident Analysis & Prevention.83, 132–142.

Su, J., Sze, N. N., & Bai, L. (2020). A joint probability model for pedestrian crashes at macroscopic level: roles of environment, traffic, and population characteristics. Accident Analysis & Prevention, 150, 105898.

Sze, N. N., Su, J., & Bai, L. (2019). Exposure to pedestrian crash based on household survey data: effect of trip purpose. Accident Analysis & Prevention, 128, 17-24.

Tabeshian, M., & Kattan, L. (2014). Modeling nonmotorized travel demand at intersections in Calgary, Canada: use of traffic counts and Geographic Information System data. Transportation Research Record, 2430(1), 38-46.

Tian, G. and Ewing, R., (2017). A walk trip generation model for Portland, OR. Transportation Research Part D: Transport and Environment, 52, 340-353.

Ukkusuri, S., Hasan, S., Aziz, H., (2011). Random Parameter Model Used to Explain Effects of Built-Environment Characteristics on Pedestrian Crash Frequency. Transportation Research Record: Journal of the Transportation Research Board 2237, 98–106.

Ukkusuri, S., Miranda-Moreno, L.F., Ramadurai, G. and Isa-Tavarez, J., (2012). The role of built environment on pedestrian crash frequency. Safety science, 50(4), 1141-1151.

Wang, J., Hankey, S., Wu, X., & Lindsey, G. (2016). Monitoring and modeling of urban trail traffic: validation of direct demand models in Minneapolis, Minnesota, and Columbus, Ohio. Transportation Research Record, 2593(1), 47-59.

Wang, J., Huang, H., & Zeng, Q. (2017). The effect of zonal factors in estimating crash risks by transportation modes: Motor vehicle, bicycle and pedestrian. Accident Analysis & Prevention, 98, 223-231.

Wang, X., Lindsey, G., Hankey, S., & Hoff, K. (2014). Estimating mixed-mode urban trail traffic using negative binomial regression models. Journal of Urban Planning and Development, 140(1), 04013006.

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.

Wier, M., Weintraub, J., Humphreys, E.H., Seto, E. and Bhatia, R., (2009). An area-level model of vehicle-pedestrian injury collisions with implications for land use and transportation planning. Accident Analysis & Prevention, 41(1), 137-145.

Xie, S. Q., Dong, N., Wong, S. C., Huang, H., & Xu, P. (2018). Bayesian approach to model pedestrian crashes at signalized intersections with measurement errors in exposure. Accident Analysis & Prevention, 121, 285-294.

Xu, P., Xie, S., Dong, N., Wong, S. C., & Huang, H. (2019). Rethinking safety in numbers: are intersections with more crossing pedestrians really safer?. Injury prevention, 25(1), 20-25.

Yao, S., Loo, B.P.Y., (2016). Safety in numbers for cyclists beyond national-level and city-level data: A study on the non-linearity of risk within the city of Hong Kong. Injury Prevention. 22 (6) , 379–385.

Yao, S., Loo, B.P.Y., Lam, W.W.Y., (2015). Measures of activity-based pedestrian exposure to the risk of vehicle-pedestrian collisions: Space-time path vs. potential path tree methods. Accident Analysis & Prevention.75, 320–332.

Yasmin, S., & Eluru, N. (2018). A joint econometric framework for modeling crash counts by severity. Transportmetrica A: transport science, 14(3), 230-255.

Yasmin, S., Eluru, N., & Ukkusuri, S. V., (2014). Alternative ordered response frameworks for examining pedestrian injury severity in New York City. Journal of Transportation Safety & Security, 6(4), 275-300.

Yasmin, S., Eluru, N., Lee, J., & Abdel-Aty, M. (2016). Ordered fractional split approach for aggregate injury severity modeling. Transportation Research Record, 2583(1), 119-126.

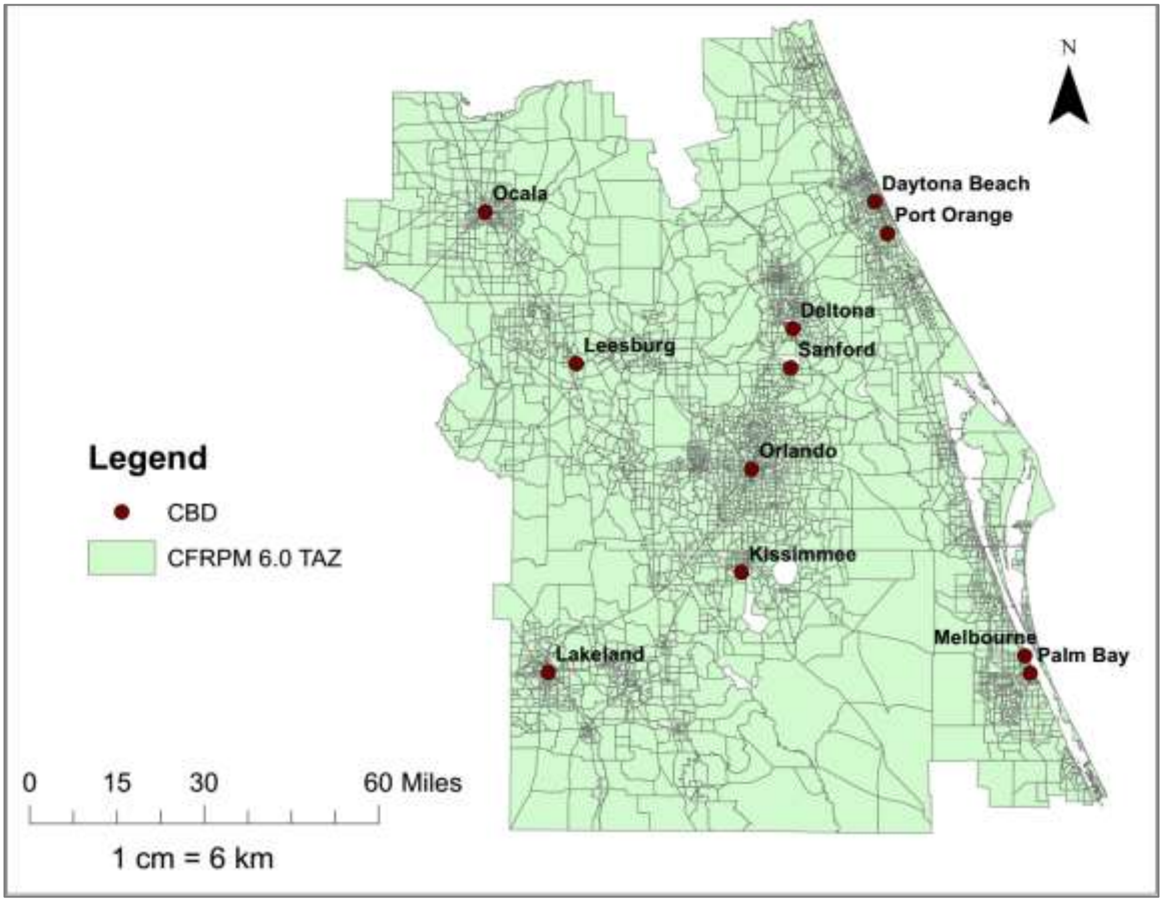

FIGURE 1 Considered Central Business District (CBD) Locations

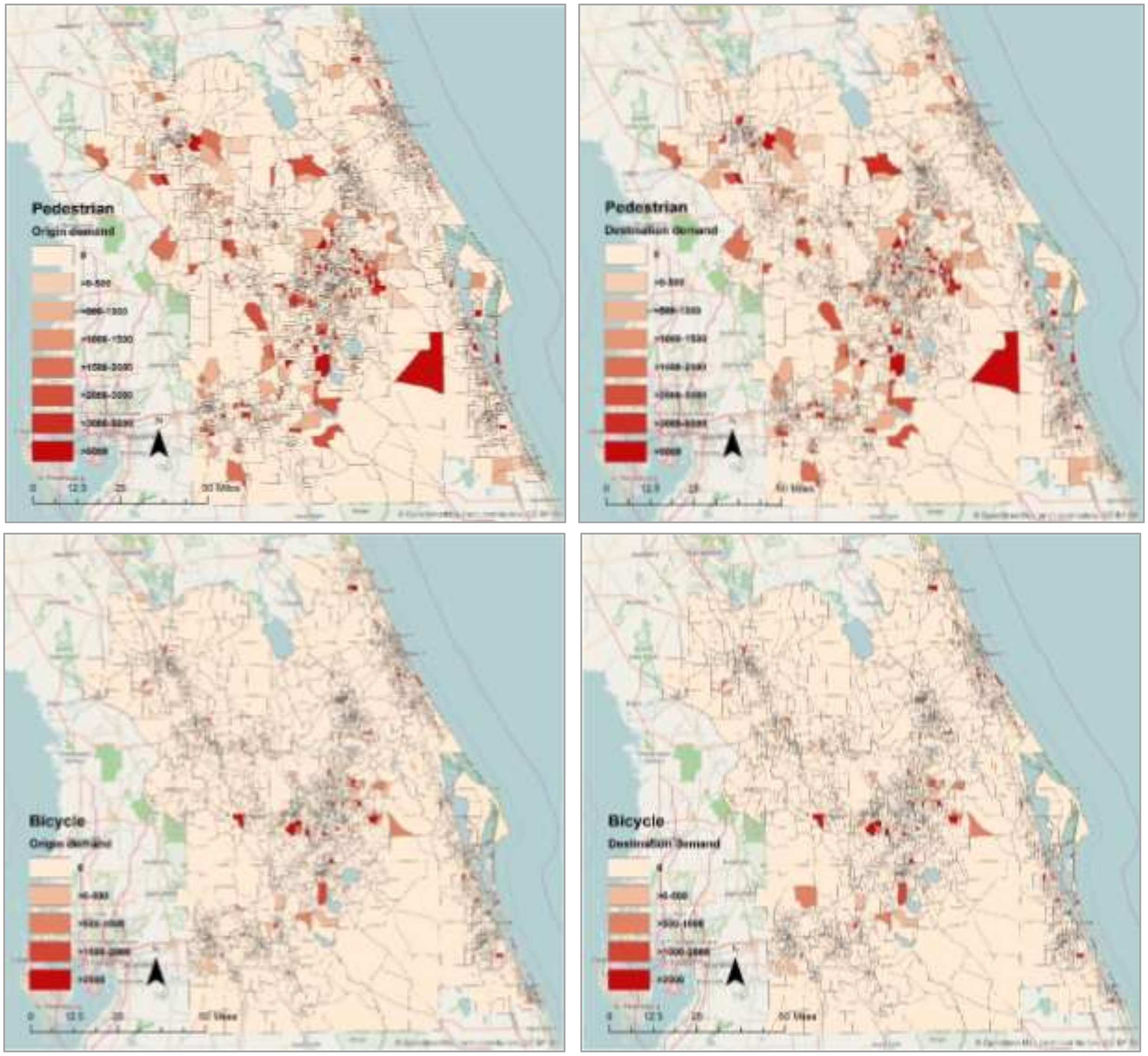

FIGURE 2 Zones with pedestrian and bicycle O-D demand for the year 2009

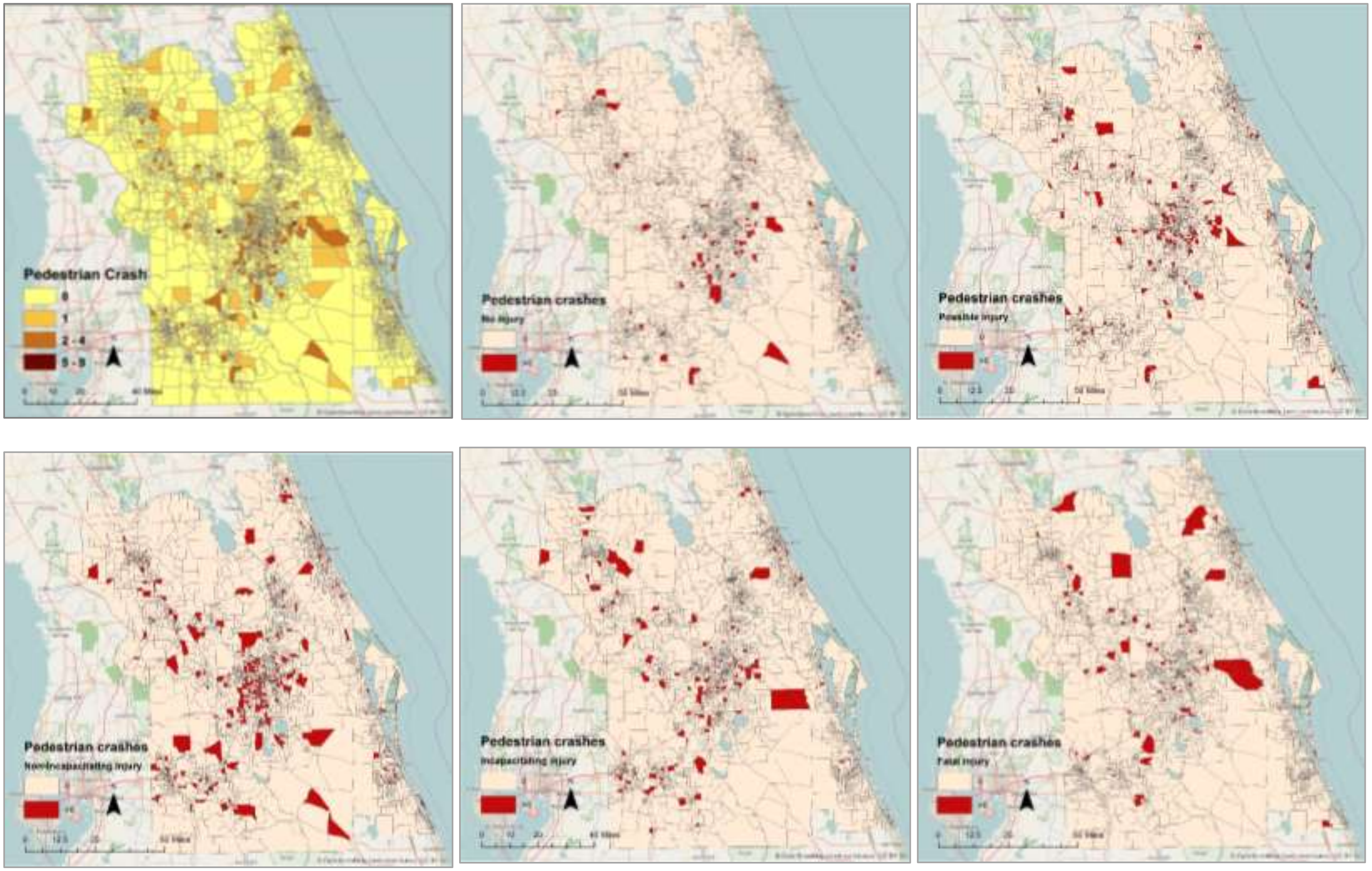

FIGURE 3 Pedestrian crashes (total and by different injury severity) for the year 2010

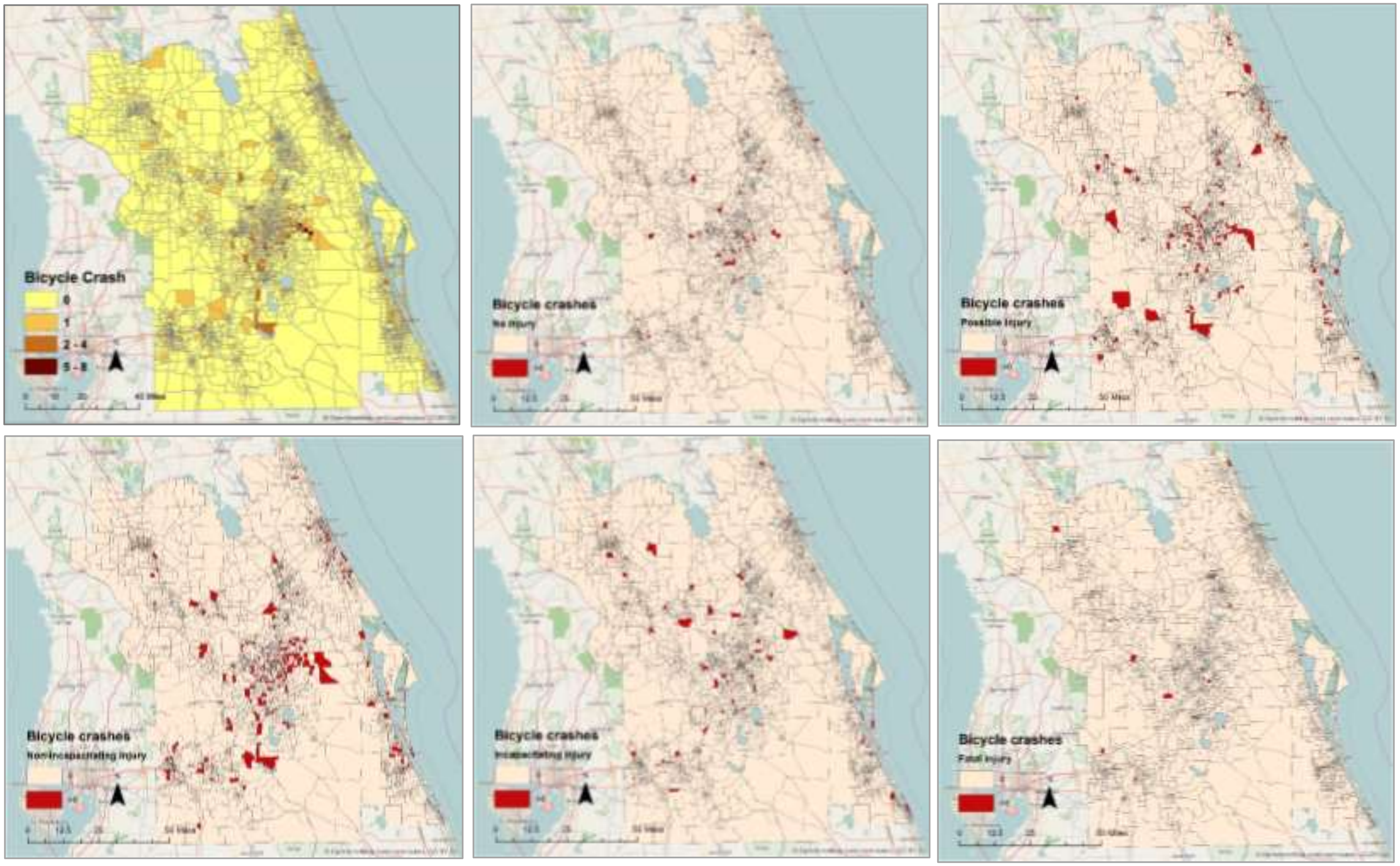

FIGURE 4 Bicycle crashes (total and by different injury severity) for the year 2010

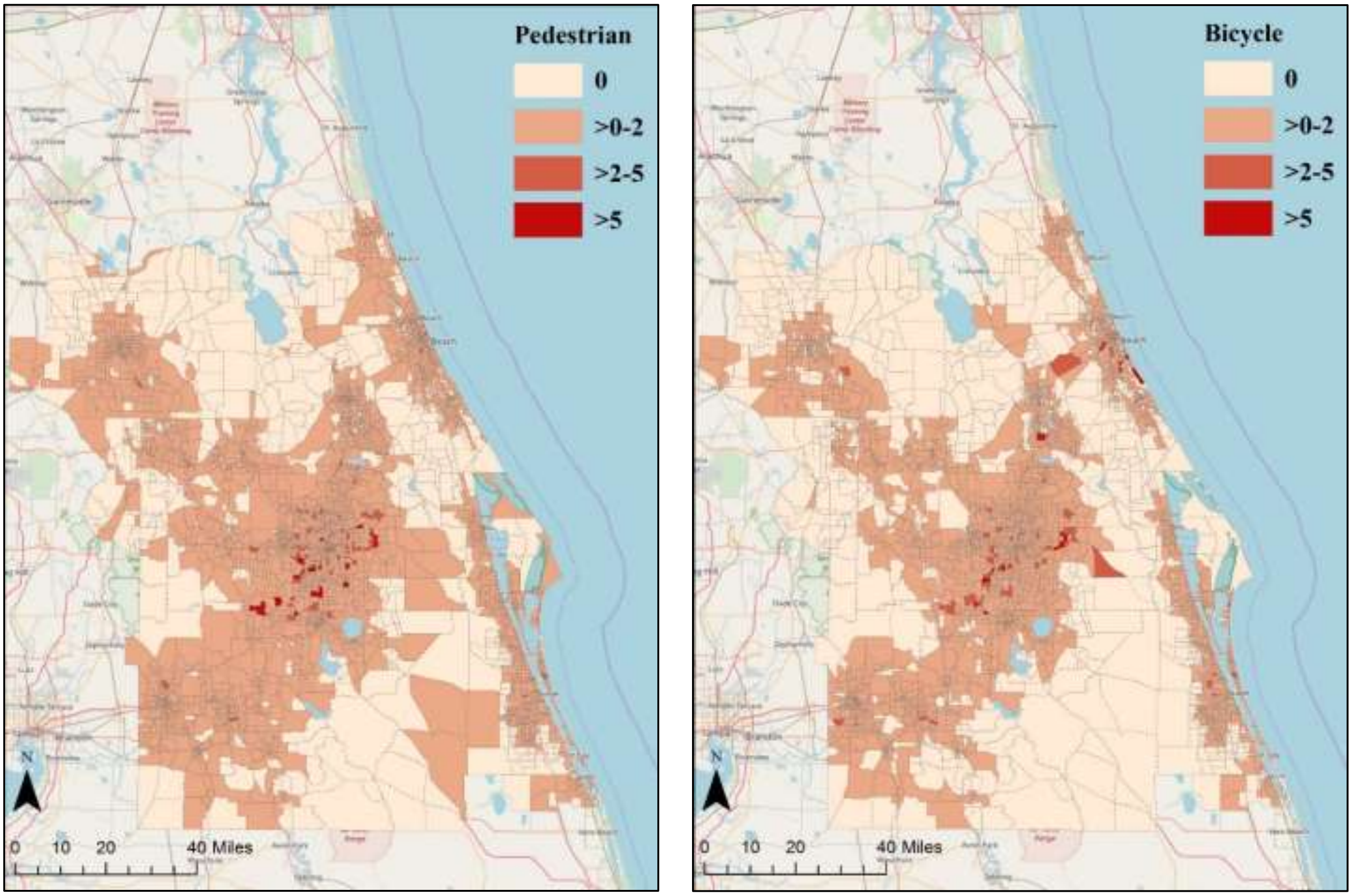

FIGURE 5 (a) Spatial Distribution of Expected Pedestrian and Bicycle Crash Counts for the year 2015

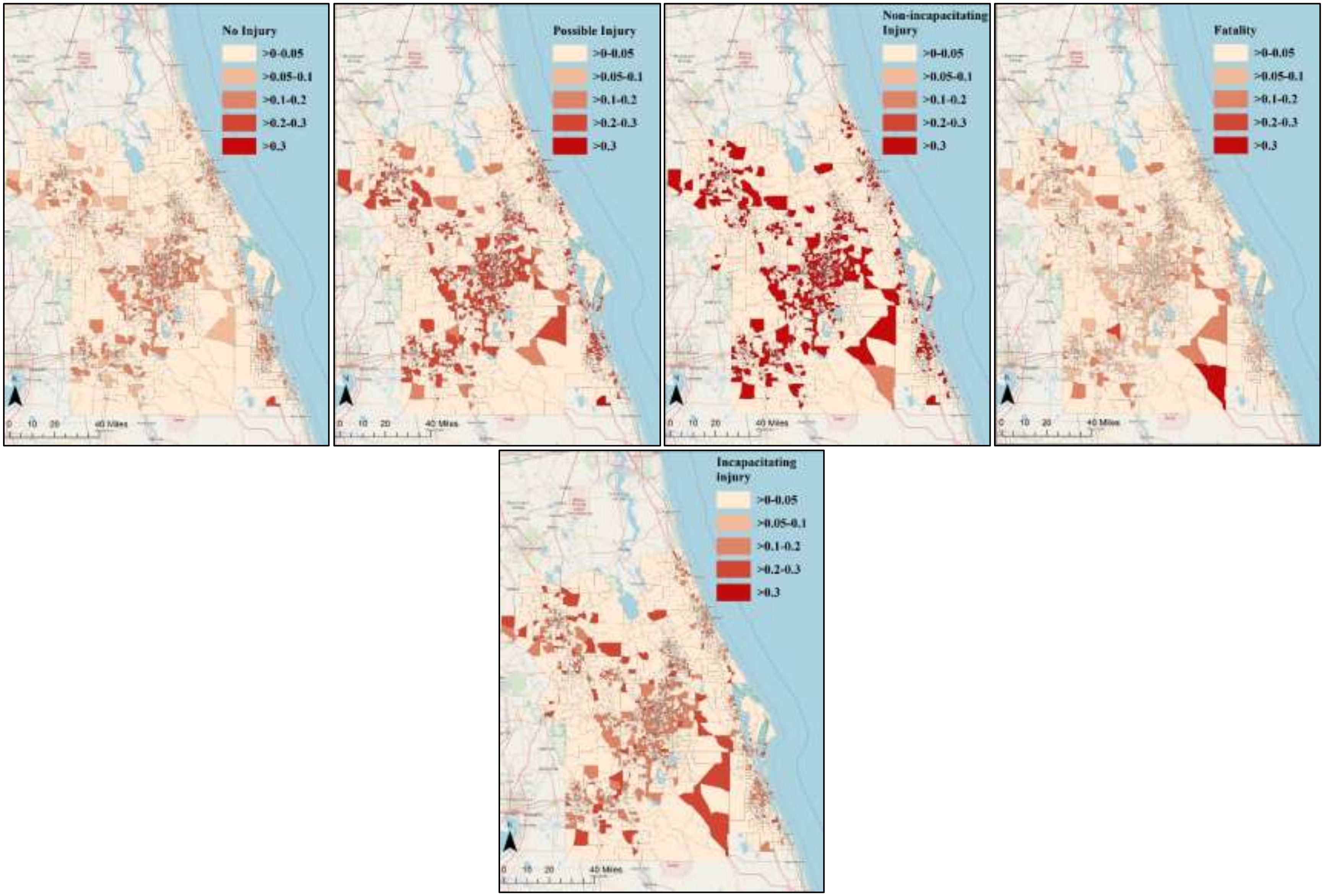

FIGURE 5 (b) Spatial Distribution of Predicted Fraction of Pedestrian Crashes by Severity levels for the year 2015

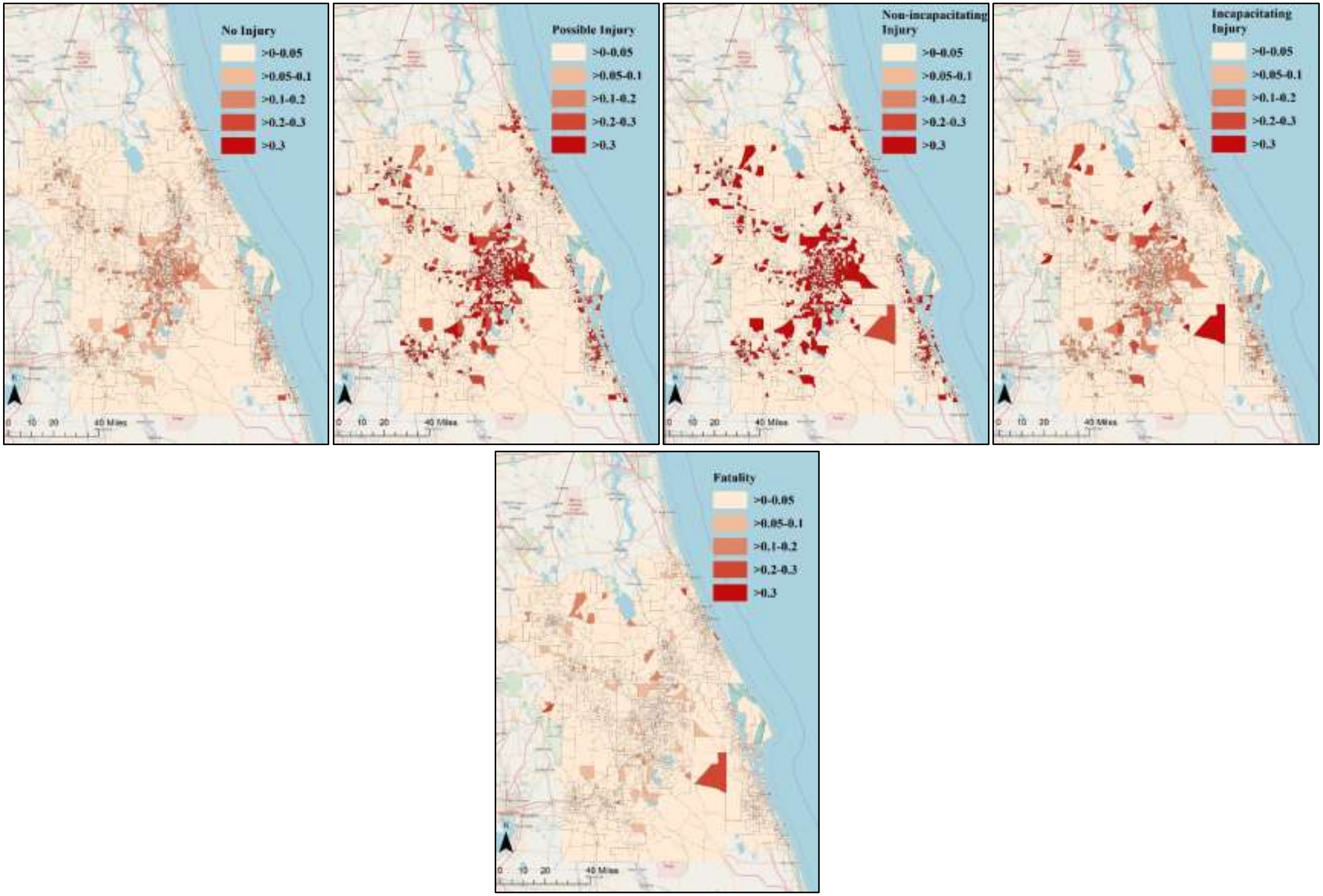

FIGURE 5 (c) Spatial Distribution of Predicted Fraction of Bicycle Crashes by Severity levels for the year 2015

1 **TABLE 1 Summary of Earlier Research on Non-motorized Demand Model**

Studies	Unit of Analysis	Spatial Unit	Temporal unit	Methodological Approach	Independent Variables Considered				
					Socio-demographic	*Land-use*	*Built Environment*	*Roadway/ Infrastructure*	*Weather*
Traditional Demand Model									
(Pulugurtha and Repaka, 2008)	Pedestrian	Intersection	1-hour	Multiple regression	Yes	Yes	Yes	Yes	--
(Schneider et al., 2009b)	Pedestrian	Intersection	2-hour	Ordinary least square	Yes	Yes	Yes	Yes	--
(Jones et al., 2010)	Bicycle, Pedestrian	Different locations	1-hour, 2-hour, daily	Ordinary least square	Yes	Yes	Yes	Yes	--
(Miranda-Moreno and Fernandes, 2011)	Pedestrian	intersections	8-hour	Log-linear, Negative binomial	Yes	Yes	Yes	Yes	Yes
(Hankey et al., 2012)	Bicycle, Pedestrian	Street segment	12-hour	Ordinary least square, Negative binomial	Yes	--	Yes	Yes	Yes
(Schneider et al., 2012)	Pedestrian	Intersections	Annual	Log-linear	--	--	Yes	Yes	Yes
(Hewawasam et al., 2014)	Pedestrian	Household	Daily	Multivariate regression	Yes	Yes	--	--	--
(Wang et al., 2014)	Bicycle, Pedestrian	Multiuse trails	Daily	Linear and negative binomial	Yes	--	Yes	--	Yes
(Tabeshian and Kattan, 2014)	Bicycle, Pedestrian	Intersection	2-hour	Multiple linear and Poisson regression	Yes	Yes	--	Yes	--
(Kraemer et al., 2015)	Bicycle	Sites such as corridors	1-hour	Multiple linear regression	--	--	--	--	Yes
(Wang et al., 2016)	Bicycle, Pedestrian	Trail segment	Annual	Negative binomial	Yes	--	Yes	--	--
(Hankey and Lindsey, 2016)	Bicycle, Pedestrian	Block level	3-hour	Stepwise linear regression, Reduced form core and time average model	--	Yes	--	Yes	Yes

(Fagnant and Kockelman, 2016)	Bicycle	Segments, intersections	3-hour	Poisson and Negative binomial	Yes	--	--	Yes	Yes
(Clifton et al., 2016)	Pedestrian	Pedestrian and traffic analysis zones	Daily	Cross classification	Yes	--	Yes	Yes	--
(Tian and Ewing, 2017)	Pedestrian	Household	Daily	Hurdle negative binomial	Yes	--	Yes	Yes	--
(Dhanani et al., 2017)	Pedestrian	hexagons (diameter 350 m)	Six year	Poisson regression	--	Yes	Yes	--	--
(Reardon et al., 2017)	Bicycle, Pedestrian	census blocks	Daily	Four step model	Yes	--	Yes	Yes	--
(Fournier et al., 2017)	Bicycle	Continuous counters	Daily, monthly, annual	Time series model	--	--	--	--	Yes
(Chen et al., 2017)	Bicycle	Bicycle count site buffer (0.25,0.5, 1-mile)	2-hour	A generalized linear mixed model	Yes	Yes	Yes	Yes	--
(Nordback et al., 2017)	Bicycle, Pedestrian	Count stations	2-hour	Survey-based, count based, and a sketch planning tool	Yes	--	--	Yes	--
(Hankey et al., 2017)	Bicycle, Pedestrian	Block level	3-hour	Facility demand model, land-use regression	Yes	--	Yes	Yes	--
(Ermagun et al., 2018b)	Bicycle, Pedestrian	Multiuse trails	Daily	Negative binomial	--	--	--	--	Yes
(Ermagun et al., 2018a)	Bicycle, Pedestrian	Infrared-inductive loop counters	1-hour, Daily	Generalized linear model with gamma distribution	Yes	--	Yes	--	Yes
(Lu et al., 2018)	Bicycle, Pedestrian	Traffic monitoring station	1-hour	Stepwise linear regression	--	Yes	--	Yes	--
Studies Incorporate Demand in Safety									
(Raford and Ragland, 2004)	Pedestrian	Segment, intersection	2-hour, annual	Space syntax tool	Yes	--	--	--	--
(Miranda-Moreno et al., 2011)	Pedestrian	Intersection	3-hour	Log-linear and count regression model	--	--	Yes	Yes	--

(Strauss et al., 2013)	Bicycle	Intersection	8-hour	Bivariate mixed Poisson model	--	Yes	Yes	Yes	Yes
(Strauss et al., 2015)	Bicycle	Segments, Intersection	8-hour	Linear regression model	--	--	Yes	Yes	--
(Lee et al., 2018a)	Bicycle, Pedestrian	American household survey metropolitan area	Annual	Bayesian integrated bivariate probit regression	Yes	Yes	Yes	Yes	--
(Lee et al., 2019)	Pedestrian	Intersections	Annual	Generalized linear model, Tobit regression	Yes	Yes	--	Yes	--

TABLE 2 Sample Statistics of Dependent Variables

EXPOSURE MODELS						
Models	**Dependent variables**	**Definitions**	**Sample size**	**Zonal (weighted)**		
				Minimum	**Maximum**	**Mean**
Pedestrian generation model	Pedestrian origin trip count	Total number of daily pedestrian trips originated in TAZs	4747	0.00	39232.01	265.45
Pedestrian attraction model	Pedestrian destination trip count	Total number of daily pedestrian trips destined in TAZs	4747	0.00	39232.01	261.70
Bicycle generation model	Bicycle origin trip count	Total number of bicycle trips originated in TAZs	4747	0.00	7012.43	35.02
Bicycle attraction model	Bicycle destination trip count	total number of bicycle trips destined in TAZs	4747	0.00	7012.43	34.94
SAFETY MODELS						
Models	**Dependent variables**	**Definitions**	**Sample size**	**Zonal**		
				Minimum	**Maximum**	**Mean**
Pedestrian crash count model	Pedestrian crash counts	Total number of pedestrian crashes in TAZs	4747	0.00	9.00	0.31
Bicycle crash count model	Bicycle crash counts	Total number of bicycle crashes in TAZs	4747	0.00	8.00	0.21
Pedestrian crash proportion by severity model						
Proportion of PDO pedestrian crashes		Total number of PDO pedestrian crashes in TAZs/ Total number of pedestrian crashes in TAZs		0.00	1.00	0.11
Proportion of possible injury pedestrian crashes		Total number of possible injury pedestrian crashes in TAZs/ Total number of pedestrian crashes in TAZs		0.00	1.00	0.24
Proportion of non-incapacitating injury pedestrian crashes		Total number of non-incapacitating injury pedestrian crashes in TAZs/ Total number of pedestrian crashes in TAZs	949	0.00	1.00	0.38
Proportion of incapacitating injury pedestrian crashes		Total number of incapacitating injury pedestrian crashes in TAZs/ Total number of pedestrian crashes in TAZs		0.00	1.00	0.18
Proportion of fatal pedestrian crashes		Total number of fatal pedestrian crashes in		0.00	1.00	0.09

	TAZs/ Total number of pedestrian crashes in TAZs				
Bicycle crash proportion by severity model					
Proportion of PDO bicycle crashes	Total number of PDO bicycle crashes in TAZs/ Total number of bicycle crashes in TAZs		0.00	1.00	0.12
Proportion of possible injury bicycle crashes	Total number of possible injury bicycle crashes in TAZs/ Total number of bicycle crashes in TAZs		0.00	1.00	0.32
Proportion of non-incapacitating injury bicycle crashes	Total number of non-incapacitating injury bicycle crashes in TAZs/ Total number of bicycle crashes in TAZs	719	0.00	1.00	0.41
Proportion of incapacitating injury bicycle crashes	Total number of incapacitating injury bicycle crashes in TAZs/ Total number of bicycle crashes in TAZs		0.00	1.00	0.14
Proportion of fatal bicycle crashes	Total number of fatal bicycle crashes in TAZs/ Total number of bicycle crashes in TAZs		0.00	1.00	0.02

1

TABLE 3 Summary Characteristics for Exogenous Variables

Variable names	Definitions	Zonal		
		Minimu	Maximu	Mean
Sociodemographic characteristics				
Population density	Total number of Population of TAZ/ Area of TAZ in acre	0.000	19.956	2.366
Proportion of male population	Total number of male of TAZ/ Total number of Population of TAZ	0.000	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	0.000	0.397	0.096
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 signal in TAZ	0.000	8.000	0.379
Proportion of arterial roads	Total length of arterial road of TAZ/Total roadway length of TAZ	0.000	1.000	0.459
Proportion of local roads	Total length of local road of TAZ/Total roadway length of TAZ	0.000	1.000	0.040
Length of sidewalks	Total sidewalk length in meter of TAZ	0.000	36.346	0.280
Length of bike lane	Total bike lane length in meter of TAZ	0.000	58.525	0.421
Availability of bike lane	Presence of bike lane in TAZ	0.000	1.000	0.041
Length of bus lanes	Total bus lane length in kilometer of TAZ	0.000	31.161	0.888
Average zonal speed	Average zonal speed in mph	0.000	70.000	36.028
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
VMT	Vehicle Miles Travel (VMT) = Total road length in miles * Average annual daily traffic / 100000	0.000	29.928	0.225
Number of flashing beacon sign	Total number of flashing beacon of TAZ	0.000	2.000	0.009
Number of school signal	Total number of school signal of TAZ	0.000	1.000	0.001
Built environment characteristics				
Number of commercial centers	Total number of commercial center of TAZ	0.000	4.000	0.087
Number of financial centers	Total number of financial center 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 hub of TAZ	0.000	11.000	0.051

Number of restaurants	Total number of restaurant of TAZ	0.000	36.000	1.335
Number of park and recreational centers	Total number of park and recreational center of TAZ	0.000	20.000	0.245
Number of hospitals	Total number of hospital of TAZ	0.000	2.000	0.017
Number of entertainment centers	Total number of entertainment center of TAZ	0.000	3.000	0.017
Number of shopping centers	Total number of shopping center of TAZ	0.000	78.000	1.492
Land-use characteristics				
Urban area	Ln (Urban area in a TAZ in acre)	-9.275	8.491	4.291
Institutional area	Ln (Institutional area in a TAZ in acre)	-16.417	7.071	0.785
Industrial area	Ln (Industrial area in a TAZ in acre)	-12.943	6.709	0.671
Retail/Office area	Ln (Office/Retail area in a TAZ in acre)	-17.312	6.611	1.744
Residential area	Ln (Residential area in a TAZ in acre)	-12.427	8.014	3.596
Recreational area	Ln (Recreational area in a TAZ in acre)	-13.946	10.040	0.388
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_k 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

TABLE 4 Estimation Results of Exposure Models – Hurdle-Negative Binomial Models

Variable names	Pedestrian demand models				Bicycle demand models			
	Pedestrian generation model		Pedestrian attraction model		Bicycle generation model		Bicycle attraction model	
	Estimates	t-stat	Estimates	t-stat	Estimates	t-stat	Estimates	t-stat
Probabilistic component								
Constant	2.346	55.615	2.319	54.774	-0.197	-3.661	-0.339	-6.208
Land-use mix	0.605	8.143	0.539	7.212	0.596	8.187	0.719	9.832
Urban area	0.224	37.315	0.215	35.200	0.305	38.242	0.300	36.621
Number of Household	0.212	27.324	0.228	29.528	0.287	25.106	0.304	26.455
Count component								
Constant	-0.217	-27.198	-0.422	-57.616	-2.351	-69.340	-1.974	-70.397
Sociodemographic characteristics								
Proportion of 65+ aged population	0.802	62.096	--*	--	-0.546	-12.745	--	--
Roadway and traffic attributes								
Average zonal speed	-0.008	-59.952	--	--				
AADT	-0.035	-31.141	-0.047	-40.822	-0.028	-8.577	--	--
Proportion of arterial roads	0.320	53.077	0.255	43.828	0.095	6.921	0.044	3.473
Proportion of 3 and more lane roads	-0.316	-32.398	-0.420	-39.923	-0.740	-33.999	-1.243	-55.656
Length of sidewalk	0.048	48.038	0.030	31.668	0.052	16.866	0.049	15.968
Built environment								
Number of commercial centers	--	--	--	--	--	--	-0.416	-29.226
Number of educational centers	--	--	--	--	--	--	0.112	21.645
Number of business centers	--	--	0.158	10.811	--	--		
Number of entertainment centers	--	--	0.194	14.437	--	--	2.941	23.494
Number of financial centers	--	--	0.021	17.835	--	--	-0.144	-43.018
Number of park and recreational centers	--	--	0.099	38.188	--	--	0.339	54.894
Number of restaurants	--	--	-0.022	-27.858	--	--	0.225	73.716
Number of shopping centers	--	--	0.032	46.627	--	--	-0.098	-36.605
Number of transit hubs	--	--	-0.057	-10.832	--	--	0.260	23.207
Land-use characteristics								
Industrial area	-0.029	-22.989	-0.055	-42.162	0.092	31.510	0.052	17.338
Recreational area	0.070	70.274	0.042	38.617	0.016	6.847	-0.057	-23.155

Residential area	0.060	57.244	0.062	55.280	0.440	82.309	0.361	74.286
Retail/office area	0.049	40.450	0.037	25.914	-0.127	-39.940	-0.191	-53.656
Institutional area	0.126	110.646	0.146	124.131	0.041	12.410	0.032	9.903
Over dispersion parameter	0.917	116.574	0.826	110.526	3.081	26.618	6.009	20.365

*variable insignificant at 90% significance level

TABLE 5 Estimation Result of Crash Count Models – Negative Binomial Models

Variable names	Pedestrian crash count model		Bicycle crash count model	
	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 roads	0.325	3.723	0.341	3.619
Proportion of local roads	---*	---	-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 park 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 [Total pedestrian daily trip demand in a TAZ/(Total number of household in a TAZ*100)]	-0.277	-1.482	---	---
Total bicycle trip demand [Ln(Total bicycle daily trip demand in a TAZ]	---	---	0.042	2.055
Over-dispersion parameter	1.004	9.297	0.641	5.642

*variable insignificant at 90% significance level

TABLE 6 Estimation Results of Crash Proportions by Severity Models – Ordered Probit Fractional Split Models

Variable name	Pedestrian crash proportions by severity model		Bike crash proportions by severity model	
	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 signals	---	---	0.362	2.474
Availability of bike lane	---	---	-0.288	-1.797
VMT	0.049	1.675	---	---
Built Environment				
Number of commercial centers	-0.149	-1.936	---	---
Number of hospitals	---	---	-0.189	-1.795
Number of park and recreational centers	---	---	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 [Total pedestrian daily trip demand in a TAZ/(Total number of household in a TAZ*100)]	-1.063	-2.756	---	---
Total bicycle trip demand per household [Total bicycle daily trip demand in a TAZ/Total number of household in a TAZ]	---	---	-0.005	-1.040

*variable insignificant at 90% significance level

TABLE 7 Predictive Performance Evaluation

In sample predictive fit measures for Demand Models				
Models	**Events**	**Observed**	**Predicted**	**Percentage Error**
Pedestrian generator model	Total Zones with zero trip count	4007.00	4006.80	0.005
	Total number of zonal trips	1260090.6	1255479.9	0.366
	Average zonal trips	265.45	264.48	0.365
Pedestrian attractor model	Total Zones with zero trip count	4010.00	4010.49	-0.012
	Total number of zonal trips	1242270.5	1236690.7	0.449
	Average zonal trips	261.70	260.52	0.451
Bicycle generator model	Total Zones with zero trip count	4574.00	4573.82	0.004
	Total number of zonal trips	166248.45	165671.36	0.347
	Average zonal trips	35.02	34.90	0.343
Bicycle attractor model	Total Zones with zero trip count	4581.00	4581.18	-0.004
	Total number of zonal trips	165845.77	171959.97	-3.687
	Average zonal trips	34.94	36.22	-3.663

In sample predictive fit measures for Count Models				
Models	**Mean crash**		**MPB**	**MAD**
	Observed	**Predicted**		
Pedestrian	0.31	0.33	-0.80	11.49
Bicycle	0.21	0.22	-0.26	6.39

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	0.113	0.003	0.526
	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	0.084		
Bicycle	Proportion of property damage only crashes	0.115	0.115	0.005	0.2912
	Proportion of minor injury crashes	0.320	0.320		

	Proportion of non-incapacitating injury crashes	0.407	0.407		
	Proportion of incapacitating injury crashes	0.141	0.141		
	Proportion of fatal crashes	0.017	0.017		

TABLE 8 Policy Scenarios

Scenarios	Description of scenarios	Study region	Number of zones	Change in zonal demand		Change in crash count		Change in crash severity proportions — Fatal Crash	
				Pedestrian	Bicycle	Pedestrian	Bicycle	Pedestrian	Bicycle
Scenario 1	50% reduction in traffic volume with 2 miles buffer area of different central business district (CBD)	**All zones**	4747	0.164	0.043	-0.63	3.144	-4.967	-0.066
		Zones within 2 miles buffer of CBD	703	1.804	0.389	-3.266	-2.889	-4.687	-0.045
Scenario 2	30% reduction in traffic volume with 2 miles buffer area of different central business district (CBD)	**All zones**	4747	0.096	0.026	-0.437	3.622	-4.963	-0.066
		Zones within 2 miles buffer of CBD	703	1.060	0.231	-2.120	-0.274	-4.664	-0.045
Scenario 3	15% reduction in traffic volume with 4 miles buffer area of different central business district (CBD)	**All zones**	4747	0.125	0.030	-0.482	3.554	-4.963	-0.066
		Zones within 4 miles buffer of CBD	1375	0.498	0.090	-1.280	1.680	-4.55	0.003
Scenario 4	5% reduction in traffic volume with 6 miles buffer area of different central business district (CBD)	**All zones**	4747	0.071	0.013	-0.34	3.935	-4.96	-0.066
		Zones within 6 miles buffer of CBD	1985	0.166	0.027	-0.589	3.281	-4.891	0.015
Scenario 5	All zones have sidewalk and the new proposed sidewalk length = $\frac{(TAZ\ area)^{0.5}}{2}\ meter$	**All zones**	4747	-0.438	0.108	-1.360	4.367	-1.013	-0.063
Scenario 6	50% increase in existing sidewalk length	**All zones**	4747	0.705	0.289	0.985	4.436	-1.111	-0.071
Scenario 7	15% reduction in zonal average maximum speed	**All zones**	4747	1.407	0.000	-0.143	0.000	-1.107	0.000
Scenario 8	25% reduction in zonal average maximum speed	**All zones**	4747	2.389	0.000	-0.150	0.000	-1.135	0.000
Scenario 9	15% reduction in zonal proportion of 3+lane road	**All zones**	4747	0.287	0.576	-0.138	4.436	-1.077	-0.068
Scenario 10	25% reduction in zonal proportion of 3+lane road	**All zones**	4747	0.484	0.337	-0.143	4.415	-1.085	-0.066

TABLE 9 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

APPENDIX A: Hurdle Negative Binomial (HNB) Model Framework

The Hurdle approach is generally used for modeling excess sampling zeroes. It is usually interpreted as a two-part model: 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:

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

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

In Equation 1, π_i is the probability of zero trip count and is modeled as a binary logit model as follows:

$$\pi_i = \frac{exp(\gamma \boldsymbol{\eta}_i)}{1+exp(\gamma \boldsymbol{\eta}_i)} \tag{2}$$

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 1 can be presented as NB expression in forming Hurdle NB (HNB) regression model. Given the setup as presented in Equation 1, the probability distribution for NB can be written as:

$$P_i(y_i|, \mu_{i,} \alpha) = \frac{\Gamma(y_i+\alpha^{-1})}{\Gamma(y_i+1)\Gamma(\alpha^{-1})} \left(\frac{1}{1+\alpha\mu_i}\right)^{\frac{1}{\alpha}} \left(1 - \frac{1}{1+\alpha\mu_i}\right)^{y_i} \tag{3}$$

where $\Gamma(\cdot)$ is the Gamma function and α is the NB dispersion parameter. μ_i is the expected number of daily trips non-motorists are making in TAZ i where α represents the overdispersion parameter. We can express μ_i as a function of explanatory variable ($\boldsymbol{z}_i$) by using a log-link function as $\mu_i = E(y_i|\boldsymbol{z}_i) = exp(\boldsymbol{\delta z}_i)$, where $\boldsymbol{\delta}$ is a vector of parameters to be estimated.

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

$$LL = w_i * \begin{cases} ln\ (\pi_i) & y_i = 0 \\ \ln\left(\frac{(1-\pi_i)}{(1-e^{-\mu_i})} P_i(y_i)\right) & y_i > 0 \end{cases} \tag{4}$$

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

$$w_i = \sum_{j=1}^{J} \frac{Yearly\ person\ trip\ weight}{365} \tag{5}$$

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

The reader should note that in computing the weighting factor, as presented in Equation 5, 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 Equation 3 into Equation 4 results HNB model. The model presented in Equation 4 is estimated by using a maximum likelihood approach.

APPENDIX B: Negative Binomial (NB) Model Framework

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

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

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

where, $\boldsymbol{\Gamma}(\cdot)$ is the Gamma function, $\boldsymbol{\alpha}$ is the NB dispersion parameter and $\boldsymbol{\mu_i}$ is the expected number of crashes occurring in TAZ $\boldsymbol{i}$ over a given period of time. We can express $\boldsymbol{\mu_i}$ as a function of explanatory variable $(\boldsymbol{x_i})$ by using a log-link function as: $\boldsymbol{\mu_i} = \boldsymbol{E}(\boldsymbol{y_i}|\boldsymbol{x_i}) = \boldsymbol{exp}(\boldsymbol{\beta x_i})$, where $\boldsymbol{\beta}$ is a vector of parameters to be estimated. Finally, the log-likelihood function for the NB model can be written as:

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

The parameters to be estimated in the model of equation 2 are: $\boldsymbol{\beta}$ and $\boldsymbol{\alpha}$. The parameters are estimated using maximum likelihood approaches.

APPENDIX C: Ordered Probit Fractional Split (OPFS) Model Framework

The formulation for the OPFS model for modeling the proportion of crashes by severity is presented in this section. The reader would note that conventional maximum likelihood approaches are not suited for factional proportion models. Hence, we resort to a quasi-likelihood approach. Let q (q = 1, 2, …, Q) be an index to represent TAZ, and let k (k = 1, 2, 3, …, K) be an index to represent severity category. The latent propensity equation for severity category at the q th zone:

$$y_q^* = \alpha' z_q + \xi_q, \qquad (1)$$

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 quasi-likelihood based approach for our methodology. The parameters to be estimated in the Equation (2) are α, and ψ thresholds. To estimate the parameter vector, we assume that

$$E(y_{qk} \mid z_{qk}) = H_{qk}(\alpha,\psi), 0 \le H_{qk} \le 1, \sum_{k=1}^{K} H_{qk} = 1 \qquad (2)$$

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

$$P_{qk} = \left\{ G\left[\psi_k - \alpha_q' z_q\right] - G\left[\psi_{k-1} - \alpha_q' z_q\right] \right\} \tag{3}$$

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:

$$L_q(\alpha, \psi) = \prod_{k=1}^{K} \left\{ G\left[\psi_k - \alpha_q' z_q\right] - G\left[\psi_{k-1} - \alpha_q' z_q\right] \right\}^{d_{qk}} \tag{4}$$

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 Gauss matrix programming language.