Accident Analysis and Prevention 81 (2015) 14–23

Contents lists available at ScienceDirect

Accident Analysis and Prevention

journal homepage: www.elsevier.com/locate/aap

Analyzing pedestrian crash injury severity at signalized and non-signalized locations

Kirolos Haleem [a,*], Priyanka Alluri [b,1], Albert Gan [c,2]

[a] Transportation Safety Consultant, AgileAssets, Inc., 3001 Bee Caves Road, Suite 200, Austin, TX 78746, USA
[b] Lehman Center for Transportation Center, Department of Civil and Environmental Engineering, Florida International University, 10555 West Flagler Street, EC 3680, Miami, FL 33174, USA
[c] Department of Civil and Environmental Engineering, Florida International University, 10555 West Flagler Street, EC 3680, Miami, FL 33174, USA

A R T I C L E I N F O

Article history:
Received 13 August 2014
Received in revised form 21 April 2015
Accepted 22 April 2015
Available online 2 May 2015

Keywords:
Pedestrian safety
Injury severity
Mixed logit model
Countermeasures

A B S T R A C T

This study identifies and compares the significant factors affecting pedestrian crash injury severity at signalized and unsignalized intersections. The factors explored include geometric predictors (e.g., presence and type of crosswalk and presence of pedestrian refuge area), traffic predictors (e.g., annual average daily traffic (AADT), speed limit, and percentage of trucks), road user variables (e.g., pedestrian age and pedestrian maneuver before crash), environmental predictors (e.g., weather and lighting conditions), and vehicle-related predictors (e.g., vehicle type). The analysis was conducted using the mixed logit model, which allows the parameter estimates to randomly vary across the observations. The study used three years of pedestrian crash data from Florida. Police reports were reviewed in detail to have a better understanding of how each pedestrian crash occurred. Additionally, information that is unavailable in the crash records, such as at-fault road user and pedestrian maneuver, was collected. At signalized intersections, higher AADT, speed limit, and percentage of trucks; very old pedestrians; at-fault pedestrians; rainy weather; and dark lighting condition were associated with higher pedestrian severity risk. For example, a one-percent higher truck percentage increases the probability of severe injuries by 1.37%. A one-mile-per-hour higher speed limit increases the probability of severe injuries by 1.22%. At unsignalized intersections, pedestrian walking along roadway, middle and very old pedestrians, at-fault pedestrians, vans, dark lighting condition, and higher speed limit were associated with higher pedestrian severity risk. On the other hand, standard crosswalks were associated with 1.36% reduction in pedestrian severe injuries. Several countermeasures to reduce pedestrian injury severity are recommended.

1. Introduction

Pedestrian safety is of particular concern to Florida as one in every five traffic-related fatalities in the state is a pedestrian (Fatality Analysis Reporting System “FARS”, 2012). According to the National Highway Traffic Safety Administration (NHTSA, 2009), Florida had the highest pedestrian fatalities per capita in the United States (U.S.) based on the 2009 statistics of 2.51 pedestrian fatalities per 100,000 population. Another recent study by Transportation for America (T4A, 2011) has ranked Florida as the most dangerous state in the U.S. for pedestrians. The same study ranked 52 large metropolitan areas with over 1 million population. In this ranking, the top four regions were located in Florida. This ranking was given based on the Pedestrian Danger Index (PDI), which computes the rate of pedestrian deaths relative to the amount of walk-to-work trips in an area.

As pedestrian crashes typically involve injury, one way to analyze a pedestrian safety issue is to identify the significant factors affecting pedestrian crash injury severity in order to select the appropriate countermeasures. This paper identifies significant factors affecting pedestrian injury severity at signalized and unsignalized intersections using three years of pedestrian crash data (2008–2010) from Florida. The paper focuses on the relative risk of pedestrian injury severity rather than the absolute risk of a vehicle–pedestrian crash. The analysis was performed separately for signalized and unsignalized intersections to identify and compare the significant factors affecting pedestrian injury severity.

* Corresponding author. Tel.: +1 321 276 7889.
E-mail addresses: kirolos60@hotmail.com (K. Haleem), palluri@fiu.edu (P. Alluri), gana@fiu.edu (A. Gan).
[1] Present address: Department of Civil and Environmental Engineering, Florida International University, Miami, FL 33174, USA. Tel.: +1 305 348 1896.
[2] Tel.: +1 305 348 3116.

http://dx.doi.org/10.1016/j.aap.2015.04.025
0001-4575/

The mixed logit (or random parameters logit) modeling approach was applied, which accounts for the influence of unobserved factors, such as pedestrian physical health and driver behavior. This approach allows the parameter estimates to randomly vary across the observations to yield more reliable parameter estimates.

2. Prior research

This section reviews the literature on pedestrian injury severity analysis. It specifically focuses on studies that explored the risk factors affecting severity of pedestrian crashes and those that applied the mixed logit approach.

Zajac and Ivan (2003) used the ordered probit model to evaluate the effect of roadway features on pedestrian crash injury severity in rural Connecticut. The significant variables included roadway width, vehicle type, alcohol involvement, and pedestrian age. Using the same approach, Mohamed et al. (2013) used two pedestrian injury severity datasets from New York City, U.S. (2002–2006) and Montreal, Canada (2003–2006) and applied the ordered probit and multinomial logit models to analyze severity of pedestrian crashes. Several common variables, such as presence of heavy vehicles, absence of lighting, and prevalence of mixed land use, were found to increase the probability of fatal pedestrian crashes in both cities.

Oh et al. (2005) identified the significant factors affecting the probability of pedestrian fatalities in Korea using a logistic regression model. They found that the collision speed was the most significant factor, where a higher speed was associated with a pedestrian fatality increase. This result was consistent with Garder (2004), and was confirmed by Strandroth et al. (2011) and Zhao et al. (2013). Sarkar et al. (2011) also developed binary logistic regression models to identify pedestrian fatality risk factors along Bangladesh's roadways using crash data from 1998 to 2006. The authors found an increased likelihood of a fatality risk among elderly pedestrians (i.e., older than 55 years) and young pedestrians (i.e., younger than 15 years). A higher risk of fatality was observed for pedestrians who crossed the road compared to those who walked along the road. Pedestrian crashes involving trucks, buses, and tractors had a higher fatality risk compared to cars. Furthermore, pedestrian crashes occurring at locations with no traffic control or stop control had a higher fatality risk than those occurring at signalized intersections.

Tarko and Azam (2011) linked both police and hospital crash injury data to identify significant injury risk predictors by applying the bivariate probit model. The authors found that male and older pedestrians were more prone to severe injuries compared to other groups. Rural and high-speed urban roadways were found to be more dangerous for pedestrians, especially for pedestrians crossing the roadways. The most dangerous pedestrian behavior was identified to be crossing a road between intersections (i.e., at midblock locations). In addition, the size and weight of the vehicle involved in a pedestrian crash were significant predictors of pedestrian injury level. Similar findings were observed by Al-Shammari et al. (2009) who investigated risk factors of pedestrian injury severity in Riyadh, Saudi Arabia over three-year period. Results showed that men were at a significantly greater risk than women in pedestrian crash involvement and two-thirds of pedestrians were struck while crossing the road.

Nasar and Troyer (2013) hypothesized that pedestrians could experience reduced awareness of surroundings, distraction, and engage in unsafe behavior while talking or texting on their mobile phones. Using data from the U.S. Consumer Product Safety Commission on injuries in hospital emergency rooms from 2004 to 2010, they found that mobile phone-related injuries among pedestrians increased relative to total pedestrian injuries. Moreover, pedestrian injuries related to mobile phone use were higher for males and for people under 31 years of age. Similarly, Byington and Schwebel (2013) concluded that pedestrian behavior was considered riskier while simultaneously using mobile internet and crossing the street than when crossing the street with no distraction.

Several studies, including Lee and Abdel-Aty (2005) and Jang et al. (2013), analyzed both frequency and severity of pedestrian crashes. Lee and Abdel-Aty (2005) analyzed the frequency and injury severity of vehicle–pedestrian crashes at intersections in Florida using four years of data from 1999 to 2002. Some of the significant factors affecting crash injury severity included pedestrian age, weather and lighting conditions, and vehicle size. For example, the authors found that pedestrian injuries involving a large vehicle were more severe than those involving a passenger car.

Jang et al. (2013) used six years of pedestrian crashes from 2002 to 2007 in San Francisco, California to identify risk factors that affect the frequency and severity of pedestrian crashes. They used an ordered probit model and found that alcohol involvement, cell phone use, and age (either below 15 years or above 65 years) increased pedestrian injury severity. Environmental characteristics that were associated with high pedestrian severity included nighttime, weekends, and rainy weather. The authors also found that larger vehicles such as pickups, trucks, and buses were associated with greater pedestrian injury severities compared to passenger cars.

Roudsari et al. (2005) evaluated the association between the manner of collision and the severity of pedestrian injuries. For passenger cars, they found that colliding with the hood surface and windshield in pedestrian–vehicle crashes is the major scenario contributing to pedestrian injuries. A similar study was performed by Roudsari et al. (2006), who evaluated the impact of pre-crash maneuver on pedestrian severity. They found that vehicles going straight and striking pedestrians were associated with relatively more pedestrian fatalities.

To more accurately identify significant risk predictors of crash injury severity, the mixed logit models have been implemented in recent injury severity studies. The mixed logit model is characterized by its capability to account for unobserved predictors in severity studies due to the difficulty in quantifying some features, such as the pedestrian and driver behavior at the time of the crash (Kim et al., 2011). The application of the mixed logit model in injury severity studies can be found in Gkritza and Mannering (2008), Milton et al. (2008), Pai et al. (2009), Moore et al. (2011), Kim et al. (2011), Haleem and Gan (2013), and Shaheed et al. (2013). One study that applied the mixed logit approach to analyze pedestrian injury severity was performed by Kim et al. (2010). The authors used four-year police-reported crashes in North Carolina from 1997 to 2000. They found several predictors to the probability of fatal (or severe) injuries. Examples of these predictors include darkness without streetlights, trucks and sport utility vehicles, speeding involvement, freeway sections, and increase in pedestrian age.

The above-discussed review suggests that there is a relatively large amount of pedestrian risk literature. However, there is no study that emphasizes many of the contributing factors included in this study that affect pedestrian crash injury severity at signalized and unsignalized intersections. Some of these factors, such as the at-fault road user and crosswalk type, have not been studied. One reason may be because these data are not typically available in the crash and roadway databases. This study was made possible through a major data collection effort including detailed review of police reports and collection of roadway conditions through satellite and roadway images and Google Street View. The police sketches and illustrations were reviewed to have a better understanding of how each pedestrian crash occurred.

3. Mixed logit model specification

As documented in Kim et al. (2010), the traditional approaches of modeling injury severity such as the multinomial logit and nested logit models assume that the effect of each variable is constant across the observations. While these approaches have been applied in many severity studies (e.g., Ulfarsson and Mannering, 2004; Abdel-Aty and Abdelwahab, 2004), they cannot model the influence of unobserved predictors, such as pedestrian behavior, driver behavior, and variables that can capture what happens just prior to the crash or during the crash (for example, microscopic traffic flow variables that represent the traffic volume in very small periods, e.g., in 15-min intervals). Failure to account for these predictors can introduce unobserved heterogeneity and can yield biased and inefficient parameter estimates (Shaheed et al., 2013).

To better account for the influence of unobserved predictors, unobserved heterogeneity, and the spatiotemporal variability of the observed predictors, the mixed logit model is applied in this study. This model is characterized by a random error term that allows the parameter estimates to randomly vary across the crash observations for more reliable parameter estimates (McFadden and Train, 2000; Train, 2009; Malyshkina and Mannering, 2010).

This study follows the methodology discussed by Train (2009). The random utility function (U_{jn}) that defines the crash injury severity category j probability (severe injury or non-severe injury) for crash n is as shown in Eq. (1). Non-severe injury includes property damage only (PDO), possible injury, and non-incapacitating injury, while severe injury includes incapacitating and fatal injuries.

$$U_{jn} = \beta_j' X_{jn} + \varepsilon_{jn} \tag{1}$$

where

β_j' = vector of parameters to be estimated;

X_{jn} = vector of explored variables (geometric, traffic, environmental, pedestrian/driver-related, and vehicle-related); and

ε_{jn} = random error term that is iid (independently and identically-distributed) generalized extreme value (see McFadden, 1981).

Based on Train (2009), the choice probability would be the standard multinomial logit since ε_{jn} is iid extreme value. For this, the logit probability conditioned on β_j [i.e., $L_{jn}(\beta_j)$] from both injury severity categories J (for severe and non-severe injuries) is:

$$L_{jn}(\beta_j) = P_{jn}\left(\frac{X_{jn}}{\beta_j}\right) = \frac{\exp(\beta_j' X_{jn})}{\sum_J \exp(\beta_j' X_{jn})} \tag{2}$$

However, β_j is unknown and the choice probability cannot be conditioned on β_j. Thus, the unconditional choice probability, also known as the mixed logit probability, is the integral of P_{jn} over all possible values of β_j, as follows:

$$P_{jn} = \int \frac{\exp(\beta_j' X_{jn})}{\sum_J \exp(\beta_j' X_{jn})} f\left(\frac{\beta_j}{\theta}\right) d\beta_j \tag{3}$$

where $f(\beta_j/\theta)$ is the density function of β_j, and θ is the vector of parameters for the assumed distribution (e.g., the mean and variance for the normal distribution).

The mixed logit model is a generalization (or a more flexible pattern) of the multinomial logit model to allow for parameter variations across the observations (see Train, 2009; McFadden, 1981). This model allows for the variation within the data by varying β_j across the crash observations. In this model, some parameters are held fixed and some are randomly-distributed. Note that the mixed logit model reduces to the traditional multinomial logit model if the random parameters in the mixed logit model are held fixed (Shaheed et al., 2013). The multinomial logit model suffers from the independence of irrelevant alternatives (IIA) constraint, while the mixed logit model does not (Washington et al., 2011; Shaheed et al., 2013). The IIA constraint means that the ratio of probabilities of choosing any two alternatives (e.g., i and j) from the entire set of alternatives is independent on the availability of a third alternative (h).

The choice probability in Eq. (3) cannot be easily calculated because the integral does not have a closed form (Hensher and Greene, 2002). The integral is thus approximated through simulation, as follows (Train, 2009):

a. Draw a value of β_j from $f(\beta_j/\theta)$ and label it β^r, with r denoting the first draw.
b. Calculate the logit probability $L_{jn}(\beta^r)$ for this draw using Eq. (2).
c. Repeat steps (a) and (b) multiple times and average the results of $L_{jn}(\beta^r)$. The average unbiased simulated probability estimator ($\hat{P}_{jn}$) is thus:

$$\hat{P}_{jn} = \frac{1}{R}\sum_{r=1}^{R} L_{jn}(\beta^r) \tag{4}$$

where R is the total number of draws (simulations).

The log-likelihood (LL) of the mixed logit model can be estimated as shown below:

$$LL = \sum_{n=1}^{N}\sum_{j=1}^{J} \ln(P_{jn}) \tag{5}$$

Furthermore, the simulated log-likelihood (SLL) is estimated as follows:

$$SLL = \sum_{n=1}^{N}\sum_{j=1}^{J} d_{jn} \ln(\hat{P}_{jn}) \tag{6}$$

where d_{jn} = 1 if a crash n has an injury severity j, and 0 otherwise. The maximum simulated likelihood estimator (MSLE) is the value of θ that maximizes SLL in Eq. (6) (Train, 2009).

The simulation procedure in the mixed logit model is based on Halton draws. The Halton sequence is a deterministic method that produces equally-spaced draws in the unit interval and uses a prime number as its base. For example, to generate the sequence for prime number 3, the unit interval (0,1) is divided in thirds, then ninths, twenty-sevenths, etc. A Halton sequence is created for each dimension of the mixing distribution (Train, 1999). For example, if the mixing distribution includes two random parameters, two Halton sequences are constructed. The length of each sequence is determined by the number of observations and the numbers of draws (R) specified. For n crash observations, R draws are created per observation, yielding a sequence of length ($n \times R$). The random parameters are then estimated by taking the inverse of the cumulative normal distribution of each element in the sequence (e.g., 1/3, 1/9, etc.).

The mixed logit model was fitted while considering the normal distribution of the parameters as random parameters since this yielded better-fitted parameter estimates than the uniform and triangular distributions. The parameter estimates of a mixed logit model were computed using 200 Halton draws for building the model. However, 1000 Halton draws were used to fit the final model. The mixed logit model was fitted using the LIMDEP software package (Econometric Software, Inc, 2013). A 10% level of significance was used to keep as many significant variables as possible in the model (since the objective is to identify the significant factors of crash injury severity at signalized and unsignalized intersections).

For the normal distribution of the random parameters, the probability that the distribution is less than zero can be calculated using the common standard normal distribution (Z-formula):

$$Z = \frac{\text{Zero} - \mu}{\sigma} \tag{7}$$

where μ is the mean parameter estimate and σ is the standard deviation of the estimate.

The mean parameter estimate in Eq. (7) represents the estimated value of the random parameter in the mixed logit model. The standard deviation denotes the deviation from the mean (assuming that the parameter follows a normal distribution as applied in this study). A standard normal distribution curve is shown in Fig. 1. After calculating the Z-value, the percentage of the distribution being less than zero can be determined from the Z-tables.

Since one of the objectives is to compare the impact of common variables in the signalized and unsignalized intersections models on the severe injury probability, elasticities (or marginal effects) are used. The elasticities depict the effect of change in an independent variable on the probability of severe injuries. The elasticities are the partial derivatives of the probability of crash injury severity with respect to the vector of independent variables (Zhang, 2010). For continuous variables, the elasticity measures the influence of a unit change in an independent variable on the probability of specific injury (e.g., severe injuries). For categorical or dummy variables, the elasticity represents the percentage the variable is associated with severe injuries.

Note that a major issue while fitting the mixed logit model is to determine random parameters, as described by Moore et al. (2011) and Haleem and Gan (2013). With relatively large number of observations and independent variables, the selection of random variables is not an easy task. A good way to help select the random variables is to rank the list of independent variables to efficiently determine the random parameters. The random forest technique was therefore adopted in this study to screen the severity predictors before fitting the mixed logit model. It is a recent machine learning technique proposed by Breiman (2001). In this technique, a number of decision trees are grown (without pruning) by randomly selecting some observations from the original dataset with replacement, then searching over a randomly selected subset of variables at each split (Grimm et al., 2008). Note that only one prediction is generated from the ensemble of trees. In the case of classification trees (for categorical response as in this study), the prediction is determined by the majority of correct classifications. As indicated by Grimm et al. (2008), random forest can handle noise in the covariates, yield high classification accuracy, and resist overfitting. For more details, refer to Breiman (2001) and Grimm et al. (2008).

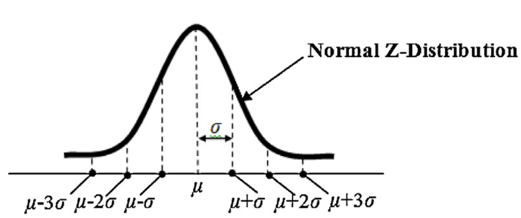

Fig. 1. Standard normal distribution.

4. Data processing and explored variables

Three years of crash data from 2008 to 2010 were used to identify pedestrian crashes on state roads in Florida. In total, 7630 pedestrian crashes were identified from the Florida Department of Transportation (FDOT) Crash Analysis Reporting (CAR) system. Police reports of these 7630 crashes were downloaded and were reviewed in detail to: (1) obtain information from police sketches and illustrations to have a better understanding of how each pedestrian crash occurred; (2) verify the accuracy of the coded crash data including the crash type and crash injury severity; (3) accurately determine the crash location; and (4) collect information that is unavailable in the crash records, such as pedestrian age, at-fault road user, and pedestrian maneuver. For example, in Florida's CAR system, the age is generally recorded for the at-fault road user (can be either a driver or a pedestrian). Since this study investigates the impact of the age of pedestrians on pedestrian injury severity, the birth year of the pedestrian involved in the crash was recorded from the police reports to calculate the pedestrian age.

Examples of the information that is not available in Florida's crash records and was collected from the review of the police reports include:

- birth year of pedestrian,
- injury severity of pedestrian,
- at-fault road user,
- crash location (i.e., at signalized or unsignalized intersection),
- presence of pedestrian signals,
- presence of pedestrian refuge area,
- crosswalk type (as shown in Fig. 2), and
- pedestrian maneuver (i.e., walking along the roadway or crossing the street).

To accurately determine the crash locations, it was vital to review the illustrative sketches and descriptions in the police reports. Note that the geographic location of crashes was broadly categorized into signalized and unsignalized intersections. The five levels of injury severity (fatal, incapacitating injury, non-incapacitating injury, possible injury, and PDO) were collected from the police reports. PDO, possible injury, and non-incapacitating injury were then categorized as non-severe injuries, while incapacitating and fatal injuries were categorized as severe injuries. To facilitate and speed-up the process of reviewing police reports, an in-house web-based tool was customized, as shown in Fig. 3.

Since many exploratory variables were considered in this study, about 35% of crashes had a missing value for at least one of these variables. Although each police report was reviewed in detail, some information was difficult to collect, possibly due to missing data and insufficient information in the police reports. For example, some pedestrian birth year and pedestrian injury severity records were empty in the police reports, some police reports did

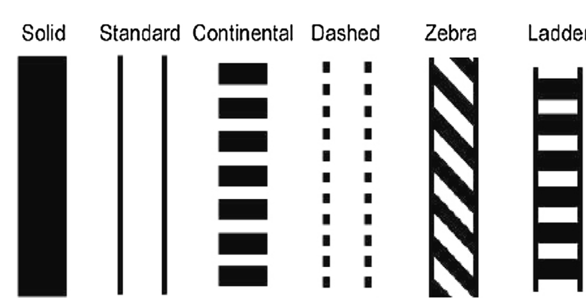

Fig. 2. Crosswalk types (Harkey and Zegeer, 2004).

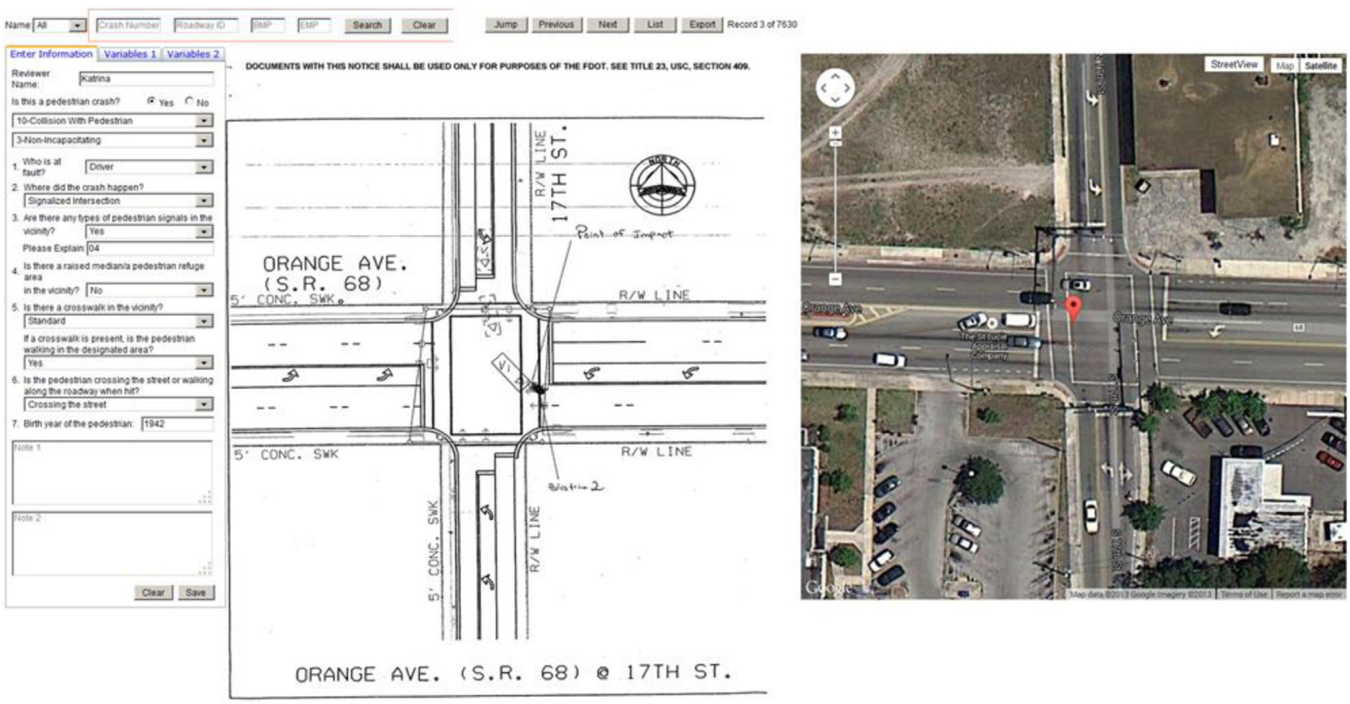

Fig. 3. A web-based tool customized to review police reports.

not include an illustrative sketch, etc. As mixed logit models cannot be fitted for crashes with missing observations, only those with complete information (i.e., no missing values) were considered. As such, data on a total of 4923 pedestrian crashes (out of 7630) were assembled. Of these crashes, 2360 occurred at signalized intersections, 678 occurred at unsignalized intersections, and the remaining 1885 crashes had unknown locations or occurred at midblock segments (not part of this study). A total of 3038 pedestrian crashes were finally included in the analysis. This sample data is considered a decent sample size for pedestrian

Table 1
Summary statistics of explored variables at signalized intersections.

Category	Variable name	Summary statistics (for categorical variables: frequency of crashes and *Italicized* % are displayed)
Continuous independent variables		
Traffic	Ln (AADT)	*Mean* = 10.42, *St. Dev.* = 0.49
	Speed limit (mph)	*Mean* = 40.35, *St. Dev.* = 6.07
	Percentage of trucks	*Mean* = 4.56, *St. Dev.* = 2.58
Categorical variables		
Response	Pedestrian injury severity	Non-severe injury = 1555 (*65.89%*); severe injury = 805 (*34.11%*)
Environmental	Lighting condition	Daylight = 1289 (*54.62%*); dusk = 56 (*2.37%*); dawn = 29 (*1.22%*); Dark street light = 839 (*35.55%*); dark no street light = 134 (*5.67%*); unknown = 13 (*0.55%*)
	Weather condition	Clear = 1821 (*77.16%*); cloudy = 373 (*15.81%*); rainy = 152 (*6.44%*); foggy = 2 (*0.08%*); other = 12 (*0.51%*)
Traffic	Presence of pedestrian signals?	Yes = 2166 (*91.78%*); no = 160 (*6.78%*); not sure = 34 (*1.44%*)
	Hour of crash	Morning peak = 384 (*16.27%*); morning off-peak = 558 (*23.64%*); Afternoon peak = 609 (*25.81%*); night/dawn off-peak = 809 (*34.28%*)
Geometric	Road surface condition	Dry = 2102 (*89.07%*); wet/slippery = 243 (*10.29%*); other = 15 (*0.64%*)
	Land use type	Urban/suburban = 2345 (*99.36%*); rural = 15 (*0.64%*)
	Presence of pedestrian refuge area?	Yes = 1420 (*60.17%*); no = 914 (*38.73%*); not sure = 26 (*1.10%*)
	Crosswalk type	No crosswalk = 166 (*7.03%*); solid = 7 (*0.29%*); standard = 1042 (*44.15%*); ladder = 215 (*9.11%*); Continental = 641 (*27.16%*); zebra = 8 (*0.34%*); dashed = 4 (*0.17%*); other = 277 (*11.74%*)
Pedestrian/driver-related	Pedestrian age	Very young (age $\leq$ 19 yrs) = 303 (*12.84%*); Young (20 yrs $\leq$ age $\leq$ 24 yrs) = 225 (*9.53%*); Middle (25 yrs $\leq$ age $\leq$ 64 yrs) = 1491 (*63.18%*); Old (65 yrs $\leq$ age $\leq$ 79 yrs) = 266 (*11.27%*); Very old (age $\geq$ 80 yrs) = 75 (*3.18%*)
	Who is at-fault?	Pedestrian = 1372 (*58.14%*); driver = 605 (*25.64%*); both = 29 (*1.23%*); not sure = 354 (*15%*)
	Pedestrian maneuver before crash	Crossing street = 2175 (*92.16%*); walking along roadway = 48 (*2.03%*); not sure = 137 (*5.81%*)
Vehicle-related	Driver's vehicle type	Passenger cars = 802 (*33.98%*); vans = 89 (*3.77%*); SUVs and pick-ups = 194 (*8.22%*); Medium trucks = 7 (*0.29%*); heavy trucks = 6 (*0.25%*); buses = 12 (*0.51%*); Bicycles = 1 (*0.04%*); motorcycles = 6 (*0.25%*); other = 1243 (*52.67%*)

Table 2
Summary statistics of explored variables at unsignalized intersections.

Category	Variable name	Summary statistics (for categorical variables: frequency of crashes and *Italicized* % are displayed)
Continuous independent variables		
Traffic	Ln (AADT)	*Mean* = 10.26, *St. Dev.* = 0.60
	Speed limit (mph)	*Mean* = 40.88, *St. Dev.* = 6.84
	Percentage of trucks	*Mean* = 4.92, *St. Dev.* = 3.56
Categorical variables		
Response	Pedestrian injury severity	Non-severe injury = 487 (*71.83%*); severe injury = 191 (*28.17%*)
Environmental	Lighting condition	Daylight = 442 (*65.19%*); dusk = 20 (*2.95%*); dawn = 10 (*1.47%*); Dark street light = 138 (*20.35%*); dark no street light = 65 (*9.58%*); unknown = 3 (*0.44%*)
	Weather condition	Clear = 537 (*79.20%*); cloudy = 99 (*14.60%*); rainy = 33 (*4.87%*); foggy = 3 (*0.44%*); other = 6 (*0.88%*)
Traffic	Hour of crash	Morning peak = 114 (*16.81%*); morning off-peak = 217 (*32%*); Afternoon peak = 183 (*26.99%*); night/dawn off-peak = 164 (*24.19%*)
Geometric	Road surface condition	Dry = 610 (*89.97%*); wet/slippery = 63 (*9.29%*); other = 5 (*0.74%*)
	Land use type	Urban/suburban = 653 (*96.31%*); rural = 25 (*3.69%*)
	Presence of pedestrian refuge area?	Yes = 168 (*24.78%*); no = 485 (*71.53%*); not sure = 25 (*3.69%*)
	Crosswalk type	No crosswalk = 456 (*67.26%*); solid = 2 (*0.29%*); standard = 125 (*18.44%*); ladder = 5 (*0.74%*); Continental = 26 (*3.83%*); other = 64 (*9.44%*)
Pedestrian/ Driver-related	Pedestrian age	Very young (age ≤ 19 yrs) = 79 (*11.65%*); Young (20 yrs ≤ age ≤ 24 yrs) = 70 (*10.32%*); Middle (25 yrs ≤ age ≤ 64 yrs) = 414 (*61.06%*); Old (65 yrs ≤ age ≤ 79 yrs) = 76 (*11.21%*); Very old (age ≥ 80 yrs) = 39 (*5.75%*)
	Who is at-fault?	Pedestrian = 189 (*27.88%*); driver = 357 (*52.65%*); both = 12 (*1.77%*); not sure = 120 (*17.70%*)
	Pedestrian maneuver before crash	Crossing street = 572 (*84.37%*); walking along roadway = 40 (*5.90%*); not sure = 66 (*9.73%*)
Vehicle-related	Driver's vehicle type	Passenger cars = 321 (*47.35%*); Vans = 39 (*5.75%*); SUVs and pick-ups = 108 (*15.93%*); Medium trucks = 5 (*0.74%*); heavy trucks = 4 (*0.59%*); buses = 3 (*0.44%*); Motorcycles = 5 (*0.74%*); other = 193 (*28.47%*)

safety study. Note that it was not a feasible task to separate the influence of crashes at unsignalized intersections from nearby midblock segments in an urban environment.

In addition to the data collected from reviewing the police reports, additional information on roadway characteristics (e.g., road surface condition and land use type) and traffic (e.g., average annual daily traffic (AADT), speed limit, and percentage of trucks) was retrieved from FDOT's Roadway Characteristics Inventory (RCI) database. Tables 1 and 2 describe the dependent and independent variables and their summary statistics at signalized and unsignalized intersections, respectively. Note that this study explores relatively new variables that were not extensively used in previous pedestrian severity studies, such as the at-fault road user, type of crosswalk, and pedestrian maneuver prior to the crash.

At signalized intersections, with the current mix of all variables (i.e., not controlled for), it was found that 6-lane roads experienced the highest percentage of pedestrian severe injuries (51.55%), followed by 4-lane roads (40.87%), 8-lane roads (4.6%), and finally 2-lane roads (2.98%). At unsignalized intersections, 4-lane roads experienced the highest percentage of pedestrian severe injuries (54.97%), followed by 6-lane roads (32.98%), 8-lane roads (8.38%), and finally 2-lane roads (3.66%). This shows that signalized and unsignalized intersections on 4- and 6-lane roadways experienced the highest percentage of pedestrian severe injuries, while both intersections on 2-lane roadways had the least percentage of severe injuries.

To rank the importance of the independent variables before fitting the mixed logit models, the random forest technique was used. As previously indicated, this was performed to facilitate identifying the random parameters in the mixed logit model. Fig. 4 shows sample results from the random forest technique used to screen the variables at signalized intersections. At both signalized and unsignalized intersections, 500 trees were used to grow the forest using the "*randomForest*" library in the R package (R Software, 2013). Using the Gini index measure to indicate the variable purity, the variables were ranked in descending order from the most to the least important.

The cut-off value of 30 for the Gini index in Fig. 4 was purposefully chosen to identify the important variables that yield the best modeling results (i.e., higher McFadden's Pseudo R^2 and lower Akaike Information Criterion "AIC" values), as well as meaningful parameter estimates. As shown in Fig. 4, the cut-off value of 30 yielded ten variables to be included in the model for signalized intersections.

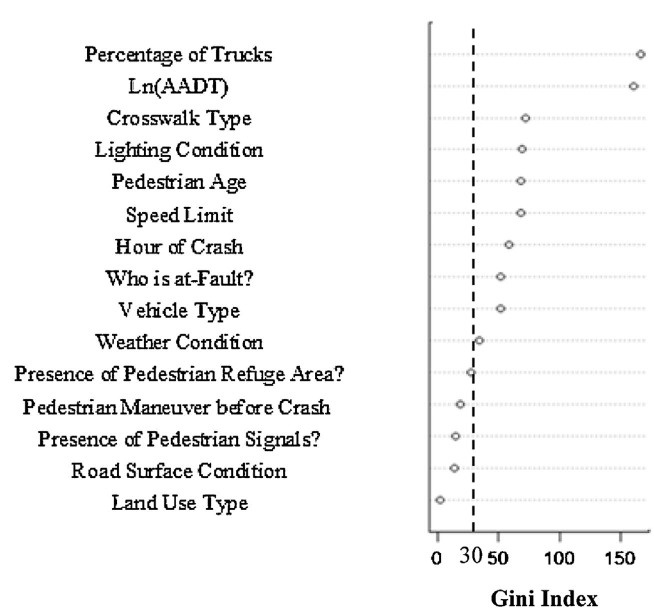

Fig. 4. Sample variable importance ranking using random forest technique at signalized intersections.

To confirm that the pruned 500 trees at both signalized and unsignalized intersections would lead to consistent and unbiased results, the out-of-bag (OOB) error was plotted against the number of trees. The minimum OOB error rates were achieved using 500 trees. The use of 500 trees was thus sufficient to yield reliable variable ranking.

5. Results

Table 3 shows the fitted mixed logit model for pedestrian injury severity at each of the signalized and unsignalized intersections using 1000 simulated Halton draws. Separate models have been generated at signalized and unsignalized intersections since the geometric and traffic characteristics of both locations are different. For example, unsignalized intersections do not include pedestrian push buttons. Table 3 provides the goodness-of-fit statistics for the model including log-likelihood at convergence, log-likelihood at zero, McFadden's Pseudo R^2, and AIC. Both mixed logit models are considered to have a good fit as the McFadden's Pseudo R^2 of the models is greater than 0.1 (Ulfarsson et al., 2010).

The parameters were considered to be random across the observations if they yielded statistically significant standard deviations for the normal distribution. On the other hand, if the parameters' estimated standard deviations were not statistically different from zero, the parameters were held fixed. As shown in the mixed logit model for signalized intersections, percentage of trucks, speed limit, and very young pedestrians were found to be statistically significant random parameters. For the unsignalized intersections model, old pedestrians and crashes occurring in dark lighting conditions with street lights were significant random parameters. The last column in Table 3 shows the elasticities of the severe injury probabilities.

5.1. Random parameters

The percentage of trucks is normally distributed with a mean of 0.381 and a standard deviation of 1.484. The Z-value calculated from Eq. (7) is −0.25, which corresponds to 40.13% from the Z-tables. In other words, 40.13% of the distribution is less than 0 and the remaining 59.87% is greater than 0. This indicates that the percentage of trucks is associated with higher probability of severe injuries at signalized intersections. The severe injury elasticity is 1.37%, which means that a one-percent higher truck percentage increases the probability of severe injuries by 1.37%.

Speed limit is also normally distributed with a mean of 0.67 and a standard deviation of 1.12, which corresponds to a Z-value of −0.6. For this, 27.43% of the distribution is less than 0 and the remaining 72.57% is greater than 0. Based on this, in 72.57% of the pedestrian crash observations, higher speed limits are associated with greater severe injury probability. Furthermore, a one-mile-per-hour higher speed limit increases the probability of severe injuries by 1.22%. In other words, a five-mile-per-hour higher speed limit could increase the probability of severe injuries by 6.1% (i.e., $5 \times 1.22\%$). This result is similar to several other

Table 3
Mixed logit model estimates using 1000 Halton draws.

Variable	Signalized intersections				Unsignalized intersections			
	Estimate	St. error	*P*-value	Elast. (%)	Estimate	St. error	*P*-value	Elast. (%)
Random parameters								
Percentage of trucks	0.381 (1.484)[a]	0.183 (0.463)[a]	0.037 (0.001)[a]	1.37	–	–	–	–
Speed limit (mph)	0.670 (1.120)[a]	0.177 (0.591)[a]	0.000 (0.058)[a]	1.22	–	–	–	–
Very young pedestrians[b]	−0.308 (1.946)[a]	0.186 (0.535)[a]	0.097 (0.000)[a]	−0.03	–	–	–	–
Dark with street light[b]	–	–	–	–	0.692 (1.240)[a]	0.240 (0.516)[a]	0.004 (0.016)[a]	2.21
Old pedestrians[b]	–	–	–	–	0.457 (1.473)[a]	0.347 (0.630)[a]	0.188 (0.019)[a]	1.18
Fixed parameters								
Intercept for severe injury	−0.797	0.063	0.000	–	−1.122	0.142	0.000	–
Ln(AADT)	0.023	0.004	0.000	15.68	–	–	–	–
Speed limit on major road (mph)	–	–	–	–	0.059	0.015	0.000	30.32
Standard crosswalk type[b]	–	–	–	–	−0.743	0.274	0.006	−1.36
Pedestrian age:								
Middle[b]	–	–	–	–	0.832	0.211	0.000	6.49
Very old[b]	0.049	0.007	0.000	30.85	2.383	0.431	0.000	2.26
Pedestrian at-fault[b]	0.367	0.089	0.000	3.44	1.346	0.233	0.000	5.72
Weather condition:								
Clear[b]	−0.263	0.112	0.019	−3.06	–	–	–	–
Rainy[b]	0.353	0.184	0.055	0.39	–	–	–	–
Van vehicle type[b]	–	–	–	–	0.511	0.321	0.100	0.38
Pedestrians walking along roadway[b]	–	–	–	–	−0.655	0.371	0.077	−0.45
Road surface condition: dry[b]	–	–	–	–	0.666	0.294	0.023	7.53
Lighting condition:								
Dark with street light[b]	0.732	0.118	0.000	4.67	–	–	–	–
Dark with no street light[b]	1.637	0.215	0.000	1.52	1.760	0.315	0.000	2.53
Night/dawn off-peak hour (8:00 pm–6:59 am)[b]	0.441	0.118	0.000	2.65	–	–	–	–
Number of observations	2360				678			
Log-likelihood at convergence	−1287.55				−294.70			
Log-likelihood at fitting the intercept	−1635.83				−403.11			
McFadden's pseudo R^2	0.21				0.27			
AIC	2605.10				617.40			

[a] Standard deviation.
[b] Binary categorical variables inserted as "1" if true; "0" otherwise.

studies, e.g., Oh et al. (2005), Milton et al. (2008), Kim et al. (2010), Haleem and Abdel-Aty (2010), and Obeng and Rokonuzzaman (2013).

On the other hand, the speed limit is a fixed parameter in the unsignalized intersections model. A one-mile-per-hour higher speed limit on major roads at unsignalized intersections is associated with 30.32% higher severe injury probability. It can be found that higher speed limits at unsignalized intersections pose greater pedestrian severity risk compared to signalized intersections. From the review of police reports, it was observed that there was an increased probability of jaywalking at unsignalized intersections, especially when crossing the major road with no dedicated pedestrian crosswalks. Thus, higher speed limits (or higher speeds) can increase the pedestrian severity at unsignalized intersections. Another reason is that pedestrians mostly rely on the gap acceptance concept and their best judgment to cross at unsignalized intersections (since no push buttons are available). At signalized intersections, the presence of pedestrian push buttons might provide safer pedestrian crossings.

The last random parameter in the signalized intersections model, very young pedestrians, has a mean of −0.308 and a standard deviation of 1.946 (i.e., Z-value of 0.16). Thus, in 56% of the crash observations, very young pedestrians are associated with lesser probability of severe injuries compared to other age groups. This result is consistent with previous studies, e.g., Lee and Abdel-Aty (2005) and Kim et al. (2011).

In the unsignalized intersections model, the random parameter for old pedestrian age has a mean of 0.457 and a standard deviation of 1.473 (i.e., Z-value of −0.31). From the Z-tables, 62.17% of the distribution is greater than 0. Thus, in 62.17% of pedestrian crashes, older pedestrians are associated with higher probability of severe injuries compared to other age groups. This is consistent with the study by Lee and Abdel-Aty (2005).

The second random parameter in the model for unsignalized intersections includes dark conditions with street lights. The mean is 0.692 and the standard deviation is 1.24. This corresponds to a Z-value of −0.56. From the Z-tables, 28.77% of the distribution is less than 0 and the remaining 71.23% is greater than 0. Thus, in 71.23% of the pedestrian crash observations, dark lighting conditions (with street lights) are associated with an increase in the probability of severe injuries. The severe injury elasticity is 2.21%, which indicates that dark conditions with street lights are associated with an increase in the probability of severe injuries by around 2%. A similar result is also obtained from the signalized intersections model.

5.2. Fixed parameters—environmental and vehicle factors

The parameter for dark lighting conditions (with no street lights) is found significant in both models. This predictor is associated with an increase in the probability of severe injuries at both signalized and unsignalized intersections. Moreover, dark lighting conditions (with no street lights) are associated with slightly higher increase in pedestrian severity at unsignalized intersections (around 2.5%) compared to signalized intersections (around 1.5%).

Clear weather is associated with 3.06% reduction in the severe injury probability compared to other weather conditions at signalized intersections. On the other hand, rainy weather is associated with 0.39% increase in the severe injury probability compared to other weather conditions. This could mainly be due to visibility constraints that impact pedestrian walking patterns. This is consistent with the findings of Sarkar et al. (2011) and Jang et al. (2013); however, it differs from the finding of Kim et al. (2010). Kim et al. (2010) concluded that inclement weather conditions reduced the probability of severe injuries while analyzing pedestrian injury severity in North Carolina. One possible reason for this discrepancy could be the difference in weather conditions between the states of Florida and North Carolina, and the associated differences in pedestrian behavior in these two states.

A significant vehicle-related variable at unsignalized intersections is vans. Vans are associated with 0.38% increase in the probability of severe injuries compared to other vehicle types. A possible explanation could be that vans are relatively large vehicles and when involved in a crash with pedestrians, regardless of who is at-fault, is expected to result in severe injuries to pedestrians. This is consistent with the findings of Zajac and Ivan (2003) and Chu (2006).

5.3. Fixed parameters—geometric factors

Two significant geometric predictors at unsignalized intersections are crosswalk type and road surface conditions. Standard crosswalks are associated with 1.36% reduction in pedestrian severe injuries. This indicates that the presence of crosswalks at unsignalized intersections is essential to alert motorists passing at these locations. Dry surface conditions are associated with 7.53% increase in pedestrian severe injuries compared to other surface conditions. This is possibly due to the increase in the likelihood of speeding and risky maneuvers taken by drivers at locations with good surface conditions. This finding is consistent with the finding of Shaheed et al. (2013).

5.4. Fixed parameters—traffic factors

Increasing the logarithm of AADT at signalized intersections significantly increases the probability of severe injuries by 15.68%, mainly due to the increase in vehicle–pedestrian conflicts. This result is consistent with the study by Obeng and Rokonuzzaman (2013). The probability of severe injury is significantly increased by 2.65% at signalized intersections during the night and dawn off-peak periods, mainly due to increased vehicle speeds and poor visibility.

5.5. Fixed parameters—pedestrian/driver-related factors

At signalized and unsignalized intersections, very old pedestrians (80 years and older) are associated with higher probability of severe injuries compared to other age groups, mainly due to their weak physical conditions. This result is consistent with the findings of other studies, e.g., Tarko and Azam (2011) and Sarkar et al. (2011). However, very old pedestrians experience a greater severity risk at signalized intersections compared to unsignalized intersections. The crossing time provided by the pedestrian signal may not have been sufficient for the slow walking speed of very old pedestrians, especially at wider intersections, making them more vulnerable to severe injuries. Very old pedestrians can experience additional severity risk when crossing multiple approaches at signalized intersections. Middle-aged pedestrians are associated with an increased likelihood of severe injuries at unsignalized intersections. Middle-aged pedestrians are more likely to cross the major approach with no dedicated crosswalks at unsignalized intersections than very old pedestrians, making them more prone to severe injuries.

At both signalized and unsignalized intersections, pedestrian crashes involving at-fault pedestrians are associated with an increase in severe injuries compared to crashes when drivers were at-fault or both pedestrians and drivers were at-fault. However, at-fault pedestrians were more vulnerable of severe injuries at unsignalized intersections. From the review of police reports for crashes at unsignalized intersections, it was found that pedestrians were mostly at-fault when they jaywalked and crossed the major road with no dedicated crosswalks.

At unsignalized intersections, crashes where pedestrians were walking along the roadway are associated with lower severe injuries compared to crashes involving pedestrians crossing the roadway. This might be due to the reduced odds of pedestrian–vehicle conflicts while pedestrians were walking at unsignalized intersections as opposed to crossing at unsignalized intersections. This result is consistent with the findings of Tarko and Azam (2011) and Sarkar et al. (2011).

6. Conclusions and study applications

This study identified the possible geometric, traffic, road user, environmental, and vehicle predictors of pedestrian injury severity at both signalized and unsignalized intersections in Florida using the mixed logit model. These predictors were either retrieved from FDOT's roadway and crash databases or were collected manually. The study focused on the relative risk of pedestrian crash injury severity (e.g., severe injuries) rather than the absolute risk of a vehicle–pedestrian crash. Two separate models were fitted to identify the significant factors affecting injury severity at signalized and unsignalized intersections, respectively. The study also compared the influence of the common significant predictors affecting severe pedestrian injuries at signalized and unsignalized intersections. The analysis was based on 3038 pedestrian crashes that occurred on state roads from 2008 to 2010. The random forest technique was used to rank the importance of independent variables affecting pedestrian injuries.

Significant severity predictors at signalized intersections from the model included percentage of trucks, speed limit, pedestrian age, AADT, at-fault pedestrians, weather and lighting conditions, and hour of crash. At unsignalized intersections, the significant predictors were pedestrian maneuver before crash, pedestrian age, speed limit on major road, at-fault pedestrians, crosswalk type, large-size vehicle types (i.e., vans), dark lighting condition, and dry road surface condition.

The results from this study could help recommend appropriate countermeasures to reduce the severity of pedestrian crashes. These countermeasures should be organized through the coordination of law enforcement officers, safety engineers, and the public to integrate the components of the four E's: engineering, education, enforcement, and emergency response. Of the engineering countermeasures, since pedestrian crashes at night and dawn off-peak periods were associated with an increase in pedestrian injury severities, it is recommended to improve lighting on urban corridors to reduce severe pedestrian injuries at night and early morning. In addition, placing standard crosswalks at unsignalized intersections is recommended since a reduction in pedestrian severity was observed in their vicinity. Since the increase in trucks percentage was associated with severe pedestrian injury increase at signalized intersections, it is recommended to limit truck access at intersections with a high pedestrian concentration.

Since at-fault pedestrians were associated with higher pedestrian injury severities, it is recommended to conduct safety awareness and education campaigns targeting both pedestrians and drivers on the laws and pedestrian right-of-way at both signalized and unsignalized intersections. The education campaigns aim to motivate individuals to alter their behavior and reduce risky actions. As indicated in Harkey and Zegeer (2004), examples of the education programs could include increasing public awareness on pedestrian safety via media and outreach material, highlighting pedestrian facilities when installing new infrastructure, conducting internal meetings within the organization to improve staff support for pedestrian safety programs, and developing working relationships with related state agencies and stakeholders on pedestrian safety. It should be noted that the education campaigns should have long-term goals since changing pedestrian and driver behavior could take long time. Since very old pedestrians were associated with a greater probability of severe injuries at both signalized and unsignalized intersections, older population could be one of the main target age groups in the education programs.

In this study, higher speed limits increased the pedestrian injury severity at both signalized and unsignalized intersections. Speed limits influence speeds and speeds influence traffic safety. Thus, stricter enforcement of speeding by law enforcement officers could help improve pedestrian safety. One example of enforcement could include increased police presence around areas with high pedestrian activity, e.g., school zones and residential neighborhoods. Furthermore, emergency response, one of the components of the four E's, could be improved by enhancing travel time of emergency vehicles and ambulances through traffic pre-emption. Other emergency services (such as the police or fire department) can help shorten the response time by giving the medical services team accurate information about the location and the situation of the crash victims (Road Safety Toolkit, 2014).

Although this study identified several significant predictors of pedestrian injury severity at signalized and unsignalized intersections, some geometric, traffic, and driver behavior predictors could not be explored because the data were not available. However, the mixed logit model used in this study helped to compensate for this deficiency by accounting for the impact from some of these potential predictors. Depending on data availability, future studies could consider using more geometric variables such as the number of exclusive right-turn or left-turn lanes. Furthermore, this study is limited to exploring the AADT traffic flow variable due to unavailability of other traffic flow predictors in the roadway database. Depending on data availability, one suggestion in future crash injury severity studies is to consider microscopic traffic flow variables that can capture what happens just prior to the crash (e.g., traffic volume in 15-min intervals). Another suggestion is to include variables related to the driver behavior (e.g., alcohol/drug use) in future severity studies, given that the data are more complete.

Since this research focused on pedestrian crashes that occurred on state roads alone, one potential expansion of this study is to analyze pedestrian crashes that occurred on local roads. Furthermore, fitting separate models for pedestrian crashes on state roads and non-state roads might highlight differences in the significant predictors of pedestrian severity on both roads.

Acknowledgement

The authors would like to thank the Florida Department of Transportation (FDOT) Research Center for funding this study.

References

Abdel-Aty, M., Abdelwahab, H., 2004. Modeling rear-end collisions including the role of driver's visibility and light truck vehicles using a nested logit structure. Accid. Anal. Prev. 36, 447–456.
Al-Shammari, N., Bendak, S., Al-Gadhi, S., 2009. In-depth analysis of pedestrian crashes in Riyadh. Traffic Inj. Prev. 1, 552–559.
Breiman, L., 2001. Random forests. Mach. Learn. 45 (1), 5–32.
Byington, K., Schwebel, D., 2013. Effects of mobile internet use on college student pedestrian injury risk. Accid. Anal. Prev. 51, 78–83.
Chu, X., 2006. Pedestrian Safety at Midblock Locations. Final Report #BD544-16. Florida Department of Transportation, Tallahassee, FL.
Econometric Software, Inc., 2013. LIMDEP Software Package. http://limdep.com (accessed September 2013).
Fatality Analysis Reporting System (FARS), 2012. http://www-fars.nhtsa.dot.gov/Main/index.aspx (accessed May 2012).
Garder, P., 2004. The impact of speed and other variables on pedestrian safety in Maine. Accid. Anal. Prev. 36, 533–542.
Gkritza, K., Mannering, F., 2008. Mixed logit analysis of safety-belt use in single- and multi-occupant vehicles. Accid. Anal. Prev. 40, 443–451.

Grimm, R., Behrens, T., Marker, M., Elsenbeer, H., 2008. Soil organic carbon concentrations and stocks on Barro Colorado Island—digital soil mapping using random forests analysis. Geoderma 146, 102–113.
Haleem, K., Abdel-Aty, M., 2010. Examining traffic crash injury severity at unsignalized intersections. J. Saf. Res. 41, 347–357.
Haleem, K., Gan, A., 2013. Effect of driver's age and side of impact on crash severity along urban freeways: a mixed logit approach. J. Saf. Res. 46, 67–76.
Harkey, D., Zegeer, C., 2004. PEDSAFE: Pedestrian Safety Guide and Countermeasures Selection System. FHWA-SA-04-003, Federal Highway Administration, U.S. Department of Transportation, Washington, D.C.
Hensher, D., Greene, W., 2002. The Mixed Logit Model: The State of Practice and Warnings for the Unwary. Institute of Transport Studies, the University of Sydney and Monash University.
Jang, K., Park, S., Kang, S., Song, K., Kang, S., Chung, S., 2013. Evaluation of pedestrian safety: geographical identification of pedestrian crash hotspots and evaluating risk factors for injury severity. Proceedings of the 92nd Annual Meeting of the Transportation Research Board, Washington, D.C..
Kim, J., Ulfarsson, G., Kim, S., Shankar, V., 2011. A mixed logit model approach to investigate effects of age on driver-injury severity in single-vehicle accidents. Proceedings of the 90th Annual Meeting of the Transportation Research Board, Washington, D.C..
Kim, J., Ulfarsson, G., Shankar, V., Mannering, F., 2010. A note on modeling pedestrian-injury severity in motor-vehicle crashes with the mixed logit model. Accid. Anal. Prev. 42, 1751–1758.
Lee, C., Abdel-Aty, M., 2005. Comprehensive analysis of vehicle–pedestrian crashes at intersections in Florida. Accid. Anal. Prev. 37, 775–786.
Malyshkina, N., Mannering, F., 2010. Empirical assessment of the impact of highway design exceptions on the frequency and severity of vehicle accidents. Accid. Anal. Prev. 42, 131–139.
McFadden, D., 1981. Econometric models of probabilistic choice. In: Manski, C., McFadden, D. (Eds.), Structural Analysis of Discrete Data with Econometric Applications. MIT Press, Cambridge, MA.
McFadden, D., Train, K., 2000. Mixed MNL models for discrete response. J. Appl. Econometrics 15, 447–470.
Milton, J., Shankar, V., Mannering, F., 2008. Highway accident severities and the mixed logit model: an exploratory empirical analysis. Accid. Anal. Prev. 40, 260–266.
Mohamed, M., Saunier, N., Miranda-Moreno, L., Ukkusuri, S., 2013. A clustering regression approach: a comprehensive injury severity analysis of pedestrian–vehicle crashes in New York, US and Montreal, Canada. Saf. Sci. 54, 27–37.
Moore, D., Schneider IV, W., Savolainen, P., Farzaneh, M., 2011. Mixed logit analysis of bicyclist injury severity resulting from motor vehicle crashes at intersection and non-intersection locations. Accid. Anal. Prev. 43, 621–630.
Nasar, J., Troyer, D., 2013. Pedestrian injuries due to mobile phone use in public places. Accid. Anal. Prev. 57, 91–95.
National Highway Traffic Safety Administration (NHTSA), 2009. Traffic Safety Facts. http://www-nrd.nhtsa.dot.gov/Pubs/811394.pdf (accessed December 2011).
Obeng, K., Rokonuzzaman, M., 2013. Pedestrian injury severity in automobile crashes. Open J. Saf. Sci. Technol. 3, 9–17.
Oh, C., Kang, Y., Kim, B., Kim, W., 2005. Analysis of pedestrian–vehicle crashes in Korea: focused on developing probabilistic pedestrian fatality model. Proceedings of the 84th Annual Meeting of the Transportation Research Board, Washington, D.C..
Pai, C., Hwang, K., Saleh, W., 2009. A mixed logit analysis of motorists' right-of-way violation in motorcycle accidents at priority T-junctions. Accid. Anal. Prev. 41, 565–573.
R Software, 2013. http://www.r-project.org (accessed September 2013).
Road Safety Toolkit, 2014. http://toolkit.irap.org/default.asp?page=treatment&id=54 (accessed June 2014).
Roudsari, B., Kaufman, R., Koepsell, T., 2006. Turning at intersections and pedestrian injuries. Traffic Inj. Prev. 7, 283–289.
Roudsari, B., Mock, C., Kaufman, R., 2005. An evaluation of the association between vehicle type and the source and severity of pedestrian injuries. Traffic Inj. Prev. 6, 185–192.
Sarkar, S., Richard, T., Hunt, J., 2011. Logistic regression model of risk of fatality in vehicle–pedestrian crashes on national highways in Bangladesh. Transportation Research Record No. 2264. Transportation Research Board of the National Academies, Washington D.C, pp. 128–137.
Shaheed, M., Gkritza, K., Zhang, W., Hans, Z., 2013. A mixed logit analysis of two-vehicle crash severities involving a motorcycle. Accid. Anal. Prev. 61, 119–128.
Strandroth, J., Rizzi, M., Sternlund, S., Lie, A., Tingvall, C., 2011. The correlation between pedestrian injury severity in real-life crashes and Euro NCAP pedestrian test results. Traffic Inj. Prev. 12, 604–613.
Tarko, A., Azam, M., 2011. Pedestrian injury analysis with consideration of the selectivity bias in linked police–hospital data. Accid. Anal. Prev. 43, 1689–1695.
Train, K., 1999. Halton Sequences for Mixed Logit. University of California Berkeley, Department of Economics, Berkeley, CA.
Train, K., 2009. Discrete Choice Methods with Simulation, 2nd ed. Cambridge University Press Publication, Cambridge, UK.
Transportation for America (T4A), 2011. http://t4america.org/ (accessed December 2011).
Ulfarsson, G., Kim, S., Booth, K., 2010. Analyzing fault in pedestrian–motor vehicle crashes in North Carolina. Accid. Anal. Prev. 42, 1805–1813.
Ulfarsson, G., Mannering, F., 2004. Differences in male and female injury severities in sport-utility vehicle, minivan, pickup and passenger car accidents. Accid. Anal. Prev. 36, 135–147.
Washington, S., Karlaftis, M., Mannering, F., 2011. Statistical and Econometric Methods for Transportation Data Analysis, second ed. Chapman and Hall/CRC, Boca Raton, FL.
Zajac, S., Ivan, J., 2003. Factors influencing injury severity of motor vehicle–crossing pedestrian crashes in rural Connecticut. Accid. Anal. Prev. 35, 369–379.
Zhang, H., 2010. Identifying and Quantifying Factors Affecting Traffic Crash Severity in Louisiana. PhD Dissertation. Louisiana State University, Baton Rouge, LA.
Zhao, H., Yang, G., Zhu, F., Jin, X., Begeman, P., Yin, Z., Yang, K., Wang, Z., 2013. An investigation on the head injuries of adult pedestrians by passenger cars in China. Traffic Inj. Prev. 14, 712–717.