JOURNAL OF TRANSPORTATION SAFETY & SECURITY
https://doi.org/10.1080/19439962.2020.1796865

Developing context-specific safety performance functions for Florida intersections to more accurately predict intersection crashes

Ghalia Gamaleldin[a], Haitham Al-Deek[a], Adrian Sandt[a], John McCombs[a], Alan El-Urfali[b], and Nizam Uddin[c]

[a]Department of Civil, Environmental, and Construction Engineering, University of Central Florida, Orlando, Florida, USA; [b]Florida Department of Transportation, Tallahassee, Florida, USA; [c]Department of Statistics and Data Science, University of Central Florida, Orlando, Florida, USA

ABSTRACT

Safety performance functions (SPFs) are vital tools used to predict and reduce intersection crashes. Because SPFs developed in the Highway Safety Manual (HSM) only use data from certain states, several states have developed region-specific SPFs. However, these SPFs typically only utilize the three roadway categories in the HSM. This research developed SPFs based on a new context classification system used by the Florida Department of Transportation (FDOT) which categorizes intersections into eight different categories. Zero-inflated negative binomial (ZINB), zero-inflated Poisson, and hurdle models were developed and compared to the commonly used negative binomial (NB) and Poisson models for four context classification groups. To develop these context-specific SPFs, data for 29 variables were collected based on the Model Inventory of Roadway Elements 2.0, allowing for standard data collection across agencies. A statistically significant linear regression model (adjusted $R^2 = 0.684$) was built to predict missing minor AADT volumes. ZINB models outperformed the other models for the two unsignalized intersection groups, whereas NB models performed the best for the two signalized intersection groups. The influential variables differed for each group, showing how FDOT's context classification system can identify specific crash-influencing factors for different classifications, helping agencies better reduce intersection crashes.

KEYWORDS

Safety performance functions; context classification system; MIRE 2.0; intersection crashes; regression modeling

1. Introduction

Improving safety by reducing traffic crashes is a major goal of traffic agencies. In 2016, traffic crashes caused ~2.5% of all deaths and were the leading cause of death for people between the ages of 5 and 29 years (World Health Organization, 2018). About 40% of all vehicle-related crashes in the

CONTACT Haitham Al-Deek haitham.al-deek@ucf.edu Department of Civil, Environmental, and Construction Engineering, University of Central Florida, 12800 Pegasus Drive, Suite 211, P.O. Box 162450, Orlando, Florida, USA.

United States (US) occur at intersections, with ~2.5 million intersection accidents reported every year by the Federal Highway Administration (FHWA) (Eun-Ha, 2010). The main reason for intersections having a high number of crashes is conflicting traffic movements. Previous research has shown that the annual average daily traffic (AADT) is the major significant factor in crash prediction models, but the relationship between crash frequency and AADT can be nonlinear due to the presence of other factors, including area type, roadway functions, and intersection and roadway design features (Lord & Mannering, 2010; Lu, 2013; Srinivasan & Carter, 2011). By classifying intersections into detailed categories based on shared geometric and traffic factors, it can be easier to determine how these factors impact intersection crashes.

State and local agencies currently use crash prediction models known as safety performance functions (SPFs) to determine and quantify the various factors that affect the number of crashes at intersections. These SPFs can also be used as network screening tools to identify high crash-risk locations and evaluate the overall safety of intersections (Srinivasan & Carter, 2011). The Highway Safety Manual (HSM) has developed SPFs for three categories: rural two-lane roads, rural multilane highways, and urban and suburban arterial highways (American Association of State Highway and Transportation Officials [AASHTO], 2010). However, using a more detailed classification system could result in more accurate SPFs which provide agencies with a better understanding of the relationships between crash frequency and intersection attributes compared to SPFs developed using the HSM.

The Florida Department of Transportation (FDOT) has developed a new context classification system which classifies roadways and intersections into eight different categories based on land use and other parameters (FDOT, 2017). This system allows for SPFs to be developed for up to 32 different intersection types (signalized and unsignalized 3-leg and 4-leg intersections for each of the eight categories), compared to the 10 intersection types in the HSM. With these context-specific SPFs, local agencies can better understand the geometric and traffic factors which affect crash frequency for different intersection types and identify intersections with high potential for crash reductions. These crash reductions are especially important for Florida, which has an estimated total cost of $3.02 billion in crash-related deaths per year (Centers for Disease Control & Prevention, 2015).

To showcase the potential of FDOT's context classification system, this article develops SPFs for 4 of the 32 context classification groups for which data were collected at the time this article was written. Further research will collect data and develop SPFs for the remaining groups. No previous research has developed SPFs using this type of context classification system.

The results show how the significant factors differ between groups; these differences might not have been identified if only the HSM roadway categories were used. Comparisons are also made with previously developed SPFs for Florida as applicable to show the improved performance of these context-specific SPFs.

For each of the four studied context classification groups, multiple modeling methodologies are considered and compared. The standard models used for developing SPFs in the HSM are negative binomial (NB) models (Young & Park, 2013). Alternative methods have been studied including hurdle models, zero-inflated Poisson (ZIP) models, and zero-inflated negative binomial (ZINB) models to deal with issues associated with NB and Poisson models (Basu & Saha, 2017; Chen & Xie, 2016; Shiyuka, 2018; Zhang, Xie, & Li, 2012). These different modeling techniques are used as appropriate for each group based on their individual data characteristics and compared to determine the model which best predicts the crash frequency for each studied intersection group.

Although the context classification system and SPFs discussed in this article are for Florida intersections, this classification system and methodologies can be transferred to other states. To assist with this transferability, data were collected using the Model Inventory of Roadway Elements (MIRE) 2.0, which was developed by FHWA as a data collection standard (Lefler et al., 2017). Using MIRE 2.0 allows the data collection procedures from this study to be easily transferred between agencies, allowing them to more quickly develop SPFs, identify significant factors and high-risk intersections, and make appropriate modifications to mitigate crashes.

2. Literature review

Because the SPFs provided in the HSM were fitted and validated with data collected from a limited number of states (Shirazi, Lord, & Geedipally, 2017), they might not be accurate for all states. A study conducted on Florida and Ohio intersections showed that using crash modification factors (CMFs) and SPFs from the HSM gave significantly different results when applied to Florida intersections than Ohio intersections (Wang, Abdel-Aty, & Lee, 2015). This suggests that it is not suitable to apply HSM SPFs to all states (Wang et al., 2015). SPFs developed using characteristics specific to the study region can be more accurate because the relationships between crashes and roadway characteristics can differ compared to the HSM (Kaaf & Abdel-Aty, 2015). A study conducted in Regina, Saskatchewan, Canada illustrated these differences by comparing calibrated HSM SPF models developed for three intersection types with NB models developed specifically for Regina (Young & Park, 2012). The calibration factors for the studied signalized intersections showed that the HSM SPFs underpredicted the total collisions by 56% and underpredicted property damage only collisions by 64% (Young & Park, 2012).

Several states have developed their own region-specific SPFs rather than using the HSM SPFs. These region-specific SPFs generally outperform the HSM SPFs by providing more accurate crash estimates (Donnell, Gayah, & Jovanis, 2014; Garber & Rivera, 2010; Srinivasan & Carter, 2011; Wang & Abdel-Aty, 2007; Xie & Chen, 2016). Some studies have also tried to address the problems associated with the significant amounts of zeros that can be present in crash data. In these situations, zero-inflated models and hurdle models typically perform better than conventional generalized linear models (GLMs) because they can handle data with excessive zeros (Basu & Saha, 2017). A study conducted in Malaysia on five years of road accidents showed that ZINB model performed better than Poisson and NB models, indicated by a lower Akaike information criterion (AIC) value (Prasetijo et al., 2019).

Sufficient data quantity and quality are important to develop accurate SPFs. As previously mentioned, AADT volumes are essential factors in crash prediction models. These volumes are usually available for major roadways but are not always available for minor roads. Therefore, it could be necessary to predict these minor AADT values using roadway and intersection characteristics. Very little previous research utilized minor AADT prediction models. A study in Oregon developed a linear regression model to predict minor AADT volumes for use in SPFs (Dixon et al., 2015). Another study in Florida developed AADT prediction models for three categories of road areas: rural, small–medium urban, and large metropolitan areas (Pan, 2008). The results of these previous studies are compared with a minor AADT model developed in this article (Section 5.1).

This literature review shows that many studies have developed region-specific SPFs using various modeling techniques and compared these SPFs to calibrated HSM SPFs, but most have used the same roadway classifications described in the HSM. Additionally, no studies have collected data using the MIRE 2.0 standard to develop SPFs. The research discussed in this article addresses these shortcomings by using a more detailed context classification system and multiple modeling methodologies to develop SPFs tailored specifically to certain intersections types and utilizing the MIRE 2.0 data standard so agencies can easily collect data and develop their own SPFs.

3. Research goal and objectives

The goal of this article is to show how FDOT's context classification system can improve traffic safety through the development of context-specific

Figure 1. FDOT context classification system (FDOT, 2017).

SPFs. These SPFs are tailored to specific groups of intersections with similar characteristics, showing the various factors that impact crash frequency for different intersection classifications. This information can improve network screening and help agencies best identify where to implement countermeasures to most effectively reduce crashes. To illustrate how agencies can develop context-specific SPFs, four intersection groups are studied. Data are collected for each group using the MIRE 2.0 data standard. Because minor AADT values are important for SPFs, a linear regression model is developed to predict these values for intersections where they are missing. Multiple crash prediction models (NB, Poisson, hurdle, ZIP, and ZINB) are developed and compared to identify the best SPF for each group based on mean absolute percentage error (MAPE), mean squared prediction error (MSPE), AIC, and Bayesian information criterion (BIC). These SPFs could be used to identify intersections with high crash risk and determine modifications to best reduce this crash risk.

4. Data and methodology

The development of context-specific SPFs required the categorization of Florida intersections into their appropriate context classification group, collection of intersection and roadway data (mainly following the MIRE 2.0 data standard), and development and comparison of crash prediction models to determine the most accurate models for each studied intersection group. These steps are detailed in sections 4.1–4.3.

4.1. Classifying intersections using FDOT's context classification system

The context classification system used by FDOT classifies intersections/roadways into one of eight categories (shown in Figure 1): C1-Natural, C2-Rural, C2T-Rural Town, C3R-Suburban Residential, C3C-Suburban Commercial, C4-Urban General, C5-Urban Center, and C6-Urban Core (FDOT, 2017). Land area characteristics are first identified for each intersection, then primary measures (such as building height) are used to

Table 1. Number of intersections for each context classification group per FDOT district.

Context classification category	District 1 (D1)	District 2 (D2)	District 3 (D3)	District 4 (D4)	District 5 (D5)	District 6 (D6)	District 7 (D7)
C1-Natural	13	35	83	33	52	3	11
C2-Rural	122	108	98	2	48	11	26
C2T-Rural Town	69	182	157	0	39	18	46
C3R-Suburban Residential	132	83	74	113	81	30	107
C3C-Suburban Commercial	125	87	112	37	163	19	67
C4-Urban General	42	81	52	144	88	162	49
C5-Urban Center	72	17	14	21	27	80	89
C6-Urban Core	0	93	0	9	1	22	4

correctly classify each intersection into a classification group (FDOT, 2017). Using these context categories can provide insights into how the impacts of roadway characteristics and traffic factors on crashes differ across classification groups. For example, C5 and C6 classifications have relatively higher population and intersection densities compared to C1 and C2 classifications. Therefore, C5 and C6 classifications are expected to have more pedestrians, bicyclists, and transit users compared to other classifications (FDOT, 2018). Understanding these types of differences can help agencies determine the features that should be present at intersections in certain classifications (such as bike lanes and pedestrian crossings) to reduce crash risks.

For this research, context classification information was provided by FDOT for 3453 randomly selected intersections (excluding roundabouts) throughout Florida. Table 1 shows the distribution of these intersections by context category and FDOT district and Figure 2 shows the locations and classifications of these intersections. This table and figure show the regional variability among classifications, with more rural intersections (C1 and C2) in D2 and D3 and more suburban and urban classifications (C3C and C4) in D5 and D7. The number of intersections per context group will increase as FDOT classifies more intersections. All classifications were finalized by FDOT, so no intersections could be reclassified. However, intersections with significantly different geometric features, traffic volumes, or area characteristics compared to other intersections in the same group were excluded from modeling.

The 3453 intersections were classified into 32 different groups (3-leg and 4-leg unsignalized and signalized intersections for each of the eight context categories). Three years of crash data (2013–2015) for all crash types were collected for these intersections. Based on the HSM recommended minimum sample sizes of 50 intersections and 100 crashes per year (AASHTO, 2010), SPFs can be developed for 21 intersection groups using the currently

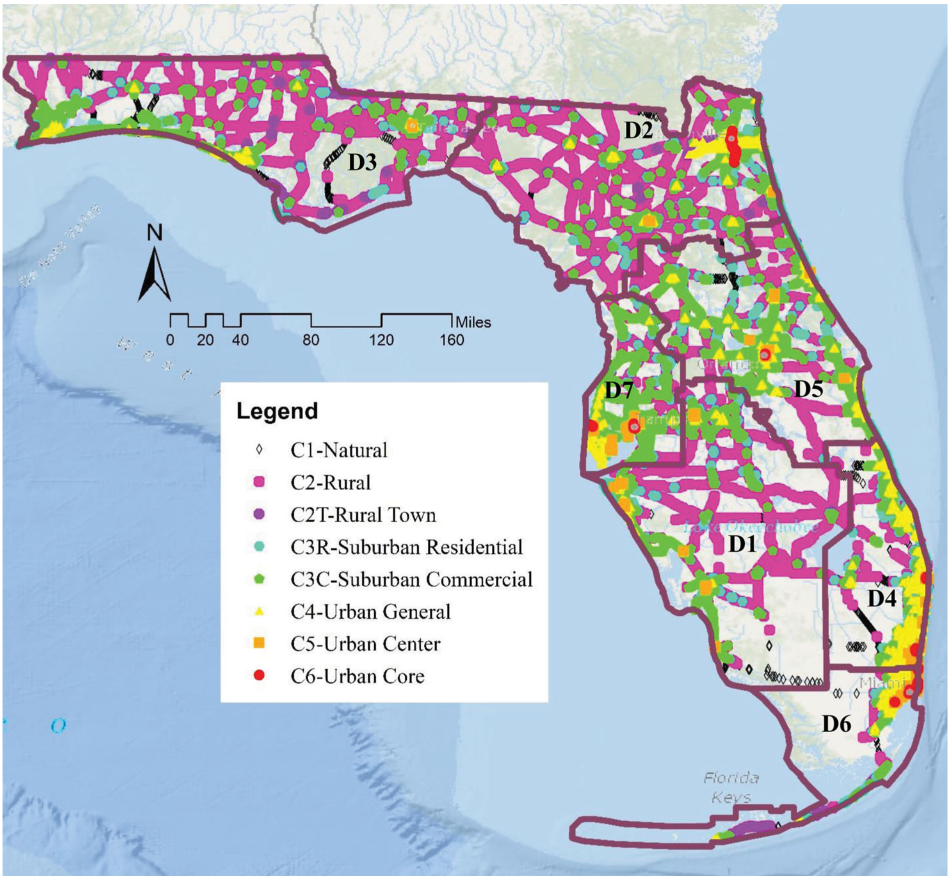

Figure 2. Map of Florida intersections categorized by context classification (developed by UCF research team). *Colors are needed for this figure.

available data. As FDOT classifies more intersections, SPFs could be developed for the 11 remaining groups, assuming they meet both sample size recommendations.

4.2. Data collection and preparation

Quality data are essential for agencies to make accurate and informed decisions regarding the safety and design of intersections. Roadway and intersection variables from the MIRE 2.0 data standard were collected for four intersection groups that met the HSM sample size guidelines: C2T unsignalized 3-leg, C3R signalized 4-leg, C3C unsignalized 4-leg, and C3C signalized 4-leg. MIRE 2.0 standardizes the data collection procedures for several roadway and traffic elements, which allows for easier collaboration between agencies (Lefler et al., 2017). By using MIRE 2.0, the data collection procedures in this article are transferable to other agencies.

Data were collected for 29 variables using Geographic Information System (GIS) files from FDOT (2019) and aerial and street view imagery from Google Maps (Google, 2019). Brief descriptions of all variables are provided in Table 2. The collected data included some variables not listed in the MIRE 2.0 data elements for intersections (functional class, speed limit, road width, and FDOT district). The first three variables were included as potential independent variables for the minor AADT prediction model, whereas the district variable was included to identify any differences in crash frequencies within an intersection group across FDOT districts. Before modeling, the major and minor AADT values were logarithmically (log) transformed.

4.3. Crash prediction model development

To develop SPFs for each of the four studied groups, multiple modeling methodologies were considered and compared using various performance measures. The best performing model was then selected as the SPF for that group. Descriptions of all modeling methodologies considered in this article are discussed in the following sections, with their applications to the four studied groups discussed in Section 5.

4.3.1. Minor AADT model

Before developing SPFs for each group, a multiple linear regression model of log(minor AADT) using statistically significant (5% significance level) intersection characteristics was developed to predict minor AADT at intersections without this information. This model used data from all four studied intersection groups. A log transformation of minor AADT was used to ensure the validity of the regression model's statistical assumptions. The model has the form shown in Equation 1 (Washington, Karlaftis, & Mannering, 2011). Stepwise selection was used to determine the significant variables in the model, with the final model chosen based on performance diagnostics and best fit statistics.

$$\mathrm{Log}(\text{minor AADT}) = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \ldots + \beta_n x_n, \tag{1}$$

where n is the number of significant input variables in the model, β_0 is the intercept, β_i is the variable coefficient, and x_i is the significant variable.

4.3.2. Poisson regression model

In a Poisson regression model, the probability of a specific site i having y crashes per year is given by Equation 2 (Basu & Saha, 2017). The Poisson distribution restricts the mean and variance to be equal and therefore might not always give accurate results. This model was considered for all four studied intersection groups.

Table 2. Description of variables considered in SPF development.

Variable name	Description	Type	Categories
Total crashes	Number of crashes at an intersection from 2013 to 2015	Continuous	None
Log (major AADT)	AADT on major roadway	Continuous	None
Log (minor AADT)	AADT on minor roadway	Continuous	None
District	District Number	Categorical	1–7
School zone	Indication of whether the intersection is in a school zone	Binary	Yes (1), No (0)
RR zone	Indication of whether the intersection is in a railroad zone	Binary	Yes (1), No (0)
Intersect angle	The smallest angle between any two legs of the intersection	Binary	=90 (1), <90 (0)
Lighting	Indication of whether the intersection contains a source of light	Binary	Yes (1), No (0)
Major exclusive left turn number	Number of exclusive left-turn lanes on the major approach	Continuous	None
Major exclusive left turn length	Storage length of exclusive left-turn lanes on the major approach	Continuous	None
Major exclusive right turn number	Number of exclusive right-turn lanes on the major approach	Continuous	None
Major exclusive right turn length	Storage length of exclusive right-turn lanes on the major approach	Continuous	None
Minor exclusive left turn number	Number of exclusive left-turn lanes on the minor approach	Continuous	None
Minor exclusive left turn length	Storage length of exclusive left-turn lanes on the minor approach	Continuous	None
Minor exclusive right turn number	Number of exclusive right-turn lanes on the minor approach	Binary	Yes (1), No (0)
Minor exclusive right turn length	Storage length of exclusive right-turn lanes on the minor approach	Continuous	None
Major median	Median type separating opposing traffic lanes on the major approach	Binary	Undivided (1), Divided (0)
Minor median	Median type separating opposing traffic lanes on the minor approach	Binary	Divided, Undivided
Major bike lane presence	Presence of a bike lane on the major approach	Binary	Yes (1), No (0)
Minor bike lane presence	Presence of a bike lane on the minor approach	Binary	Yes (1), No (0)
Major crosswalk	Type of crosswalk crossing the major approach	Binary	Unmarked (1), Marked (0)
Minor crosswalk	Type of crosswalk crossing the minor approach	Binary	Unmarked (1), Marked (0)
Major through lanes	Total number of through lanes on major approach	Continuous	None
Minor through lanes	Total number of through lanes on minor approach	Continuous	None
Functional class major	Functional classification of the major approach	Binary	Collector (1), Arterial (0)
Functional class minor	Functional classification of the minor approach	Categorical	Local road (2), Collector (1), Arterial (0)
Speed limit major	Posted speed limit on the major approach	Continuous	None

(*continued*)

Table 2. Continued.

Variable name	Description	Type	Categories
Speed limit minor	Posted speed limit on the minor approach	Continuous	None
Road width major	Average paved width of the major approach	Continuous	None
Road width minor	Average paved width of the minor approach	Continuous	None

$$P(y_i) = \frac{\exp(-\lambda_i)\ \lambda_i^y}{y!},\ y = 0, 1,\ \ldots \quad (2)$$

where $E(Y|i) = \lambda_i$ (the expected number of crashes per year for the given site i) and $\mathrm{var}(Y|i) = \lambda_i$ (the variance of the number of crashes per year for the given site i).

λi is modeled as:

$$\lambda_i = \exp\ (\beta_0 + \beta_1 x_1 + \beta_2 x_2 + \ldots + \beta_n x_n). \quad (3)$$

The above Poisson model is a basic "count data" model that can be used for modeling count data as it is easy to interpret (Srinivasan & Bauer, 2013). However, Poisson models cannot handle over-dispersed data (variance greater than the mean) well (which is typical for crash data) and are negatively influenced by the low sample mean and sample size biases (Basu & Saha, 2017).

4.3.3. NB model

One way to account for over-dispersion is to model crash counts using NB regression. NB models are used in the HSM to develop SPFs. The NB model has the form as shown in Equation 4 (Donnell et al., 2014). Like Poisson models, NB models are easy to interpret, but they can better model data with significant over-dispersion (Donnell et al., 2014). This model was considered for all four studied intersection groups

$$\lambda_i = \exp(\beta_i x_i + \varepsilon_i) \quad (4)$$

where β_i is the variable coefficient, x_i is the significant variable, and $\exp(\varepsilon_i)$ is a Gamma-distributed disturbance term with mean 1 and variance α.

4.3.4. ZIP and ZINB models

The disadvantage of using GLM models like Poisson and NB is their inability to account for excessive zero counts that can be present in crash data (especially for rural areas with lower traffic volumes). To address this issue, ZIP and ZINB models can be used. These models allow for different sets of variables to model the zero state and the count state (Prasetijo et al., 2019).

The equations for these models are the same as the equations for the Poisson model (Equation 2) and NB model (Equation 4), respectively, with the zero-inflated models having an additional model of the same form for the zero counts. These models were considered for the studied groups with excess zero counts, as they are not applicable for groups with few zero counts. Groups with a percentage of zero crash data larger than 10% were considered to have excess zero counts.

4.3.5. Hurdle model

The hurdle model can also be used to model data with excess zeros. Unlike the zero-inflated models, which assume that a zero value indicates a safe intersection with no crashes, hurdle models assume that all intersections have crash potential and the zero state does not remain permanently at any intersection (Shiyuka, 2018). The hurdle model combines a count data model ($f_{\text{count}}(y; x; \beta)$) and a zero-hurdle model ($f_{\text{zero}}(0; z; \gamma)$) as shown in Equation 5 (Shiyuka, 2018)

$$f_{\text{hurdle}} = \begin{cases} f_{\text{zero}}(0; z; \gamma) & \text{if } y = 0 \\ 1 - f_{\text{zero}}(0; z; \gamma) . \dfrac{f_{\text{count}}(y; x; \beta)}{f_{\text{count}}(0; x; \beta)} & \text{if } y > 0 \end{cases} \tag{5}$$

where β and γ are model parameters estimated by maximum likelihood.

A logit model is used to distinguish counts of zeros from large counts (effectively collapsing the count distribution into two categories) and then a truncated Poisson model is used for the positive counts (Shiyuka, 2018). The hurdle model was only considered for the studied groups with excess zero counts.

5. Modeling results and discussion

Using the collected data and discussed modeling methodologies, SPFs were developed for each of the four studied intersection groups (C2T unsignalized 3-leg, C3C unsignalized 4-leg, C3R signalized 4-leg, and C3C signalized 4-leg) in Florida. This application of FDOT's context classification shows how agencies can develop context-specific SPFs using this system or a similar system. It also shows how the models' results compare to each other and previous studies and identifies similarities and differences between the significant variables for the different intersection groups.

5.1. Minor AADT multiple linear regression model

Before developing SPF models, a minor AADT model was developed to estimate minor AADT for intersections without these data. The four studied groups contained 818 intersections, with 300 having available minor AADT. These 300 intersections were used to build the model. SAS 9.4 (2013) was used to fit a multiple linear regression model and determine the statistically significant variables to predict minor AADT. Using the stepwise variable selection method, significant variables at 5% significance level were identified and are displayed in Table 3. Equation 6 shows the minor AADT linear regression model equation

$$\begin{aligned}\mathrm{Log}(\text{minor AADT}) = {} & 4.32163 + 0.43804\log(\text{Major AADT}) \\ & - 0.08533\left(\mathrm{Through}_{\mathrm{Major}}\right) + 0.22562\left(\mathrm{Through}_{\mathrm{Minor}}\right) \\ & + 0.40356\ \left(\text{Exclusive }\mathrm{LT}_{\mathrm{Minor}}\right) \\ & + 0.35157\left(\text{Exclusive }\mathrm{RT}_{\mathrm{Minor}}\right) \\ & - 0.63957\left(\mathrm{FC}_{\mathrm{Minor}} = 2\right) - 0.65650\left(\mathrm{FC}_{\mathrm{Minor}} = 1\right) \\ & + 0.01671\left(\mathrm{SL}_{\mathrm{Minor}}\right) - 0.0040\left(\mathrm{RW}_{\mathrm{Major}}\right) \\ & + 0.00691\left(\mathrm{RW}_{\mathrm{Minor}}\right).\end{aligned} \tag{6}$$

This model shows that an increase in major AADT, minor through lanes, minor road speed limit, minor road width, and number of minor exclusive left and right turn lanes increase the minor AADT (keeping all other variables constant), whereas an increase in major road width and major through lanes decrease the minor AADT (keeping all other variables constant). Additionally, the minor road functional class has a negative coefficient for both local and collector roads, indicating that these roads have lower minor AADT volumes compared to arterial roads. Variance inflation factors (VIFs) were calculated for all of the significant variables to ensure there is no collinearity; these values are shown in Table 3. VIF values from 5 to 10 indicate collinearity (JMP, n.d.). All the calculated VIF values are less than 3, so there is no significant collinearity between the independent variables.

An analysis of variance was conducted to determine the significance of the developed model; this showed that the model was significant (F-ratio = 67.02, p value < 0.0001). In addition, the adjusted R^2 was 0.684. The data was divided into a training set (80% of the data) and a test set (20% of the data) to determine the prediction performance of the developed model. The test MAPE was 5.6% and the MSPE was 0.37, indicating that the model accurately predicts minor AADT volumes. A random forest model was also developed and compared to the linear regression model shown in Equation 6. The performance measures' results for the random forest model are MAPE = 6.24% and MSPE was 0.84. Comparing these values

Table 3. Minor AADT model estimates and statistical significance.

Variable	Variance inflation factor	Estimate	Standard error	F value	p value
Intercept		4.32163	0.78930	29.98	$<$0.0001
Log (major AADT)	2.62	0.43804	0.08663	25.57	$<$0.0001
Major through lanes ($\text{Through}_{\text{Major}}$)	2.66	−0.08533	0.03510	5.91	0.0159
Minor through lanes ($\text{Through}_{\text{Minor}}$)	1.49	0.22562	0.03768	35.86	$<$0.0001
Minor exclusive left turn number ($\text{Exclusive LT}_{\text{Minor}}$)	1.71	0.40356	0.06920	34.01	$<$0.0001
Minor exclusive right turn number ($\text{Exclusive RT}_{\text{Minor}}$)	1.31	0.35157	0.08403	17.51	$<$0.0001
Functional class minor = 2 ($\text{FC}_{\text{Minor}} = 2$)	2.59	−0.63957	0.14193	20.31	$<$0.0001
Functional class minor = 1 ($\text{FC}_{\text{Minor}} = 1$)	2.05	−0.65650	0.11097	35.00	$<$0.0001
Speed limit minor (SL_{Minor})	1.56	0.01671	0.00651	6.58	0.0111
Road width major (RW_{Major})	1.92	−0.00400	0.00183	4.77	0.0301
Road width minor (RW_{Minor})	1.99	0.00691	0.00260	7.09	0.0084

between both models shows that the linear regression model performs better.

The developed minor AADT model also performed better than similar models from previous studies. For the six minor AADT models developed by Pan (2008) in Florida, the highest adjusted R^2 was 0.418 (higher is better) and the lowest MAPE was 31.99% (lower is better). The minor regression model developed by Dixon et al. (2015) for Oregon intersections had the same coefficient signs for the three variables common between that model and this paper's model, with a positive coefficient for log(major AADT) and negative coefficients for major through lane and minor functional class. However, the Oregon model had a lower adjusted R^2 (0.5658) and higher MAPE (52.4%) (Dixon et al., 2015). The Oregon model had a smaller sample size (66 intersections) than this article's model (300 intersections), which could have caused the Oregon model's lower accuracy. Because the developed model had good performance compared to all other existing models nationwide, it was used to determine the missing minor AADT values for all four studied intersection groups.

5.2. SPF model for C2T – rural town unsignalized 3-leg intersections

The first SPF model was developed for the C2T unsignalized 3-leg intersection group, which contained 218 intersections. Figure 3 shows the distribution of the number of crashes per intersection from 2013 through 2015.

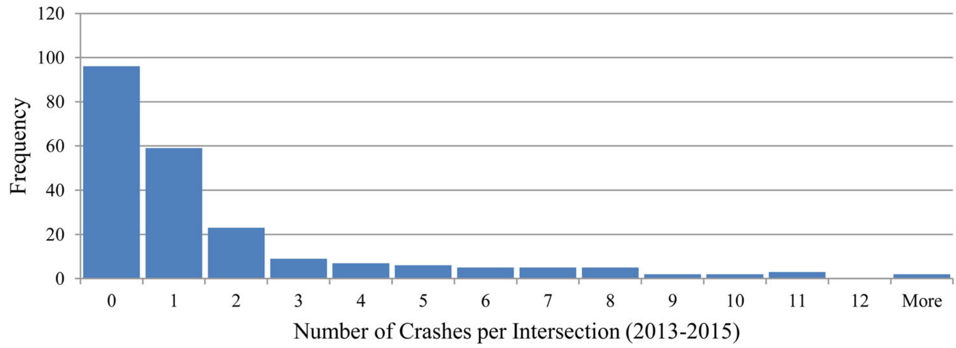

Figure 3. Crash distribution for C2T unsignalized 3-leg intersections from 2013 to 2015.

The percentage of zero crash data for this group is 42%, which could be due to the rural town setting of this group. Because this group has excess zeros (greater than 10%), hurdle, ZIP, and ZINB models were considered. Additionally, an over-dispersion test was performed in SAS 9.4 (2013) and the results were significant (p value < 0.0001), verifying the presence of over-dispersion. Therefore, a NB model will likely perform better than a Poisson model, but both models were considered.

For each considered model, the MAPE, MSPE, AIC, and BIC were compared. Table 4 shows that the ZINB model had the lowest values for all four measures, indicating that it was the best fitting model and more accurately predicted crashes than the other four models. Therefore, the ZINB was chosen as the SPF for C2T unsignalized 3-leg intersections. Table 5 shows the coefficient estimates and p values for the ZINB model, with the SPF shown in Equation 7. All variables were significant at 5% significance level. Major AADT and minor AADT had a positive relationship with the number of crashes, indicating that an increase in the major or minor AADT would increase crashes at the intersection (with all other variables fixed). Additionally, the functional class of the minor road had a negative coefficient for local roads ($FC_{minor} = 2$). Equation 8 shows the zero-state model, in which major AADT was the only significant variable. The negative coefficient indicates that intersections with higher major AADT volumes have a lower probability of having zero crashes

$$\text{Log}(\text{Total Crashes}) = -10.547 + 0.9083\ \log(\text{Minor AADT}) + 0.3894\log(\text{Major AADT}) - 1.0602(FC_{\text{Minor}} = 2) \tag{7}$$

$$\text{Log}(\text{Total Crashes}) = 51.2901 - 6.5991\log(\text{Major AADT}) \tag{8}$$

To show the improved accuracy of this context-specific SPF, it was compared to a previously developed Florida-specific SPF for rural 3-leg

Table 4. Performance measures for C2T unsignalized 3-leg intersection SPF models.

Performance measure	Poisson	Negative binomial	Hurdle	ZIP	ZINB
MAPE	0.69	0.64	0.61	0.67	0.53
MSPE	9.55	8.96	8.23	8.88	7.83
AIC	522	472	471	514	468
BIC	540	493	479	532	470

Table 5. Results of zero-inflated negative binomial model for unsignalized 3-leg intersections in C2T-rural town classification.

Variable	Estimate	Standard error	Chi-square	*p* value
Intercept	−10.547	3.1299	11.36	0.0008
Log(Minor AADT)	0.9083	0.5520	7.71	0.0099
Log(Major AADT)	0.3894	0.3207	4.47	0.0246
Functional class minor = 2 ($FC_{Minor} = 2$)	−1.0602	0.2770	14.65	0.0001

intersections (Lu, 2013). The MSPE for the C2T 3-leg intersection SPF (Equation 7) is 7.83 and the mean squared error (MSE) is 7.34. For the previously developed rural 3-leg intersection SPF (which used different data), MSPE = 187.14 and MSE = 192.47 (Lu, 2013). Based on these performance measures, the context-specific SPF developed in this article performs better than the previously developed SPF for Florida rural 3-leg intersections. This could be because the context-specific SPF accounts for excess zeros in the crash data, includes additional variables, and is developed specifically for C2T intersections. Similar comparisons to previous research could not be conducted for the remaining three context-specific SPFs developed in this article as comparable Florida-specific SPFs were not available.

5.3. SPF model for C3C – suburban commercial unsignalized 4-leg intersections

The second SPF model was developed for the C3C unsignalized 4-leg intersection group, which contained 120 intersections. The distribution of the response variable, shown in Figure 4, and the significant results of the over-dispersion test (p value < 0.0001) indicate that over-dispersion is present for this group. Like the previous group, this group's crash data also has excessive zeros (11% of the intersections in this group have zero crashes), so hurdle, ZIP, and ZINB models were developed and compared to NB and Poisson models.

Comparing the performance measures of the five models in Table 6 shows that the ZINB model had the lowest values for all measures, so it was chosen as the SPF for C3C unsignalized 4-leg intersections. Table 7 shows the coefficient estimates and p values for the ZINB model, with Equations 9 and 10 showing the SPF and zero-state model, respectively. All variables were significant at 5% significance level. Like the previous group,

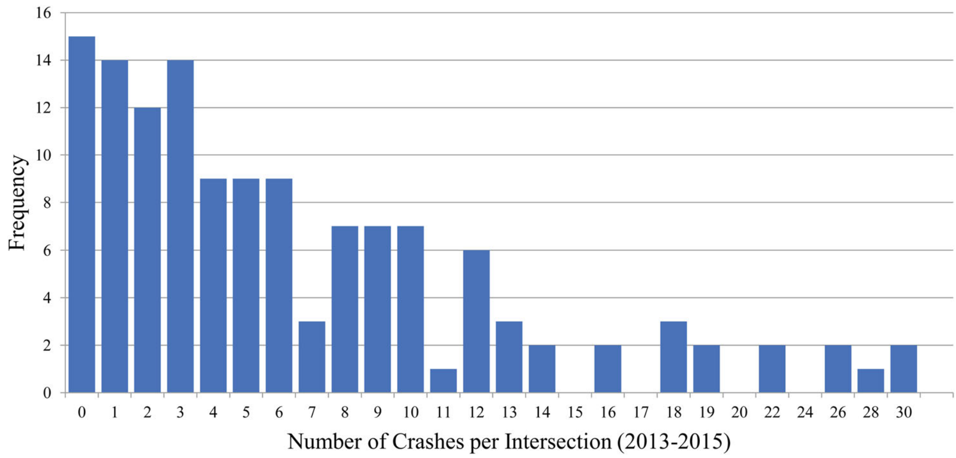

Figure 4. Crash distribution for C3C unsignalized 4-leg intersections from 2013 to 2015.

Table 6. Performance measures for C3C unsignalized 4-leg intersection SPF models.

Performance measure	Poisson	Negative binomial	Hurdle	ZIP	ZINB
MAPE	0.65	0.62	0.73	0.523	0.36
MSPE	16.5	15.68	23.3	23	14.23
AIC	622	477	639	589	468
BIC	641	492	661	620	482

Table 7. Results of zero-inflated negative binomial model for unsignalized 4-leg intersections in C3C-suburban commercial classification.

Variable	Estimate	Standard error	Wald chi-square	*p* value
Intercept	−6.6257	2.5185	6.92	0.0085
District = 7	−0.6855	0.2480	7.64	0.0057
Log(Major AADT)	0.9794	0.2578	14.44	0.0001
Intersect angle = 90	0.5386	0.2114	6.49	0.0109
Major median ($\text{Median}_{\text{Major}}$)	0.6770	0.2770	7.61	0.0074

the zero-state model had major AADT as the only significant variable with a negative coefficient. This group's SPF also had major AADT with a positive coefficient, which is expected and consistent with the previous group

$$\begin{aligned}\text{Log}(\text{Total Crashes}) = {} & -6.6257 - 0.6855(\text{District} = 7) \\ & + 0.9794\log(\text{Major AADT}) \\ & + 0.5386(\text{Intersect Angle}) + 0.6770\left(\text{Median}_{\text{Major}}\right),\end{aligned} \tag{9}$$

$$\text{Log}(\text{Total Crashes}) = 11.3943 - 1.5655\log(\text{Major AADT}). \tag{10}$$

The remaining variables in Equation 9 were not in the previous group's SPF. Both intersection angle and major median type had positive coefficients, indicating that intersections with angles less than 90 degrees had fewer crashes than 90-degree angled intersections and intersections with medians on the major roadway had fewer crashes than intersections with undivided major roadways. Drivers might be more cautious and travel at lower speeds while turning at intersections with smaller angles (due to reduced sight distances), thereby reducing the chances of a crash at these intersections, while medians can prevent crashes between vehicles going opposite directions. The other significant variable was the district variable for FDOT D7. Because this was the only district with a statistically significant coefficient, it was compared to all other districts. The negative estimate for this variable indicates that intersections from this group in D7 are expected to have fewer crashes than intersections from this group in other districts. This variation among districts could be due to differences in driver populations, signalization, or other factors that vary across districts. Future research will look at potential contributing factors to identify reasons for this variation. Including this variable in the models is important as it shows FDOT and local agencies areas with increased or reduced crash risk that can be further studied to improve safety.

Comparing both studied unsignalized intersection groups shows that even though they both used the same type of model (ZINB) for their SPFs, each group had a different set of significant variables. Some variables were significant in both SPFs, whereas some were only significant in one. This demonstrates the importance of using a detailed classification system, like FDOTs, to understand the various factors that contribute to intersection crashes for different classifications.

5.4. SPF model for C3R – suburban residential signalized 4-leg intersections

The third SPF model was developed for the C3R signalized 4-leg intersection group, which contained 192 intersections. Because these intersections are signalized, there are additional factors related to signalization (such as signal coordination and phasing) that could be considered when modeling this group which were not applicable to the previous unsignalized intersection groups. However, data on these factors were not available at the time of this research. In the future, data for these additional signalization-related variables will be collected and used when developing SPFs for signalized intersection groups. Like the previous groups, the over-dispersion test was significant (p value < 0.0001), verifying the presence of over-dispersion. Only two intersections in this group had zero crashes (less than 1%), so the hurdle, ZIP, and ZINB models were not considered for this group. The NB and Poisson models had the following performance measure values: AIC of 1064 (NB) and 1878 (Poisson), BIC of 1101 (NB) and 1921 (Poisson), MAPE of 0.57 (NB) and 0.83 (Poisson), and MSPE of 413.6 (NB) and 428.0 (Poisson). Because the NB model had lower values for all four

Table 8. Results of negative binomial model for signalized 4-leg intersections in C3R-suburban residential classification.

Variable	Estimate	Standard error	Wald chi-square	p value
Intercept	−9.1623	1.3853	43.74	<0.0001
Log (Minor AADT)	0.5869	0.1007	33.97	<0.0001
Log (Major AADT)	0.7580	0.1620	21.91	<0.0001
Major exclusive left turn length (Exclusive LT $\text{Length}_{\text{Major}}$)	0.0008	0.0004	3.44	0.0638

measures, it was chosen as the SPF model for the C3R signalized 4-leg intersection group.

Table 8 shows the coefficient estimates and p values for the NB model, with the SPF shown in Equation 11. All variables were significant at 5% significance level except for the major exclusive left turn length variable, which was significant at the 10% significance level. Major AADT and minor AADT had a positive relationship with total crashes, which is consistent with the previous groups. The major exclusive left turn length also had a positive coefficient, indicating that intersections with longer left turn lanes tend to have more crashes. This positive coefficient could be an indication of the increased crash risk due to higher left turn volumes (because roads with higher turn volumes typically have longer turn lanes) and the corresponding increase in conflicting traffic movements. However, left turn volume data were not used in this study due to their unavailability, so the exact nature of this relationship is unclear

$$\begin{aligned}\text{Log}(\text{Total Crashes}) = & -9.1623 + 0.5869\log(\text{Minor AADT}) \\ & + 0.7580\ \log(\text{Major AADT}) \\ & + 0.0008\left(\text{Exclusive LT Length}_{\text{Major}}\right)\end{aligned} \tag{11}$$

5.5. SPF model for C3C – suburban commercial signalized 4-leg intersections

The fourth SPF model was developed for the C3C signalized 4-leg intersection group, which contained 288 intersections. The over-dispersion test for this group was significant (p value < 0.0001), indicating that over-dispersion is present. This group only had three intersections (1%) with zero crashes, so the hurdle, ZIP, and ZINB models were not considered for this group. The NB and Poisson models had the following performance measure values: AIC of 1086 (NB) and 1921 (Poisson), BIC of 1103 (NB) and 1958 (Poisson), MAPE of 0.57 (NB) and 0.60 (Poisson), and MSPE of 364.4 (NB) and 465.9 (Poisson). Because the NB model had lower values for all four measures, it was chosen as the SPF model for the C3C signalized 4-leg intersection group.

Table 9 shows the coefficient estimates and p values for the NB model, with the SPF shown in Equation 12. All variables were significant at 5% significance level. Like the previous groups, major and minor AADT had

Table 9. Results of negative binomial model for signalized 4-leg intersections in C3C-suburban commercial classification.

Variable	Estimate	Standard error	Wald chi-square	*p* value
Intercept	−7.7505	1.1790	43.21	<0.0001
Log(Minor AADT)	0.5946	0.1016	34.24	<0.0001
Log(Major AADT)	0.5354	0.1204	19.78	<0.0001
Major exclusive left turn length	0.0009	0.0005	3.84	0.0501
Major median	0.2460	0.1191	4.27	0.0389

positive coefficients. The major exclusive left turn length and major median variables also had positive coefficients, which agree with these variable's coefficients in the C3R signalized 4-leg SPF and the C3C unsignalized 4-leg SPF, respectively. These similarities make sense because this group is signalized like the C3R signalized 4-leg group and is the same context classification as the C3C unsignalized 4-leg group. However, the combination of variables and coefficient values for this group differ from the other groups, showing the importance of developing SPFs for different intersection groups

$$\begin{aligned}\mathrm{Log}(\mathrm{Total\ Crashes}) = &-7.705 + 0.5946\log(\mathrm{Minor\ AADT}) \\ &+ 0.5354\log(\mathrm{Major\ AADT}) \\ &+ 0.0009\left(\mathrm{Exclusive\ LT\ Length}_{\mathrm{Major}}\right) \\ &+ 0.2460\left(\mathrm{Median}_{\mathrm{Major}}\right) \qquad (12)\end{aligned}$$

For all four developed SPFs, tolerance and VIF values were calculated to ensure there was no collinearity between the significant variables. Tolerance values below 0.1 indicate collinearity, whereas VIF values of 5 to 10 indicate collinearity. The lowest tolerance value for any of the SPFs was 0.406, whereas the highest VIF value was 2.464. These values indicate that there were no multicollinearity problems in the developed SPFs.

6. Summary and conclusions

Properly understanding the factors that influence crashes at different intersection types is important for agencies to effectively reduce crashes and improve safety. The HSM provides default SPFs and calibration factors that agencies can use, but these SPFs might not be accurate for some areas and are only developed for three different intersection/roadway types. Using a more detailed classification system can help local and state agencies better understand how influential crash factors differ across classifications and what modifications will be most effective in reducing crashes. This article developed SPFs using FDOT's innovative context classification system, which contains eight context categories based on area type, land use, and other parameters. By using this system, more detailed SPFs can be developed rather than using the HSM's three categories.

To illustrate the usefulness and benefits of FDOT's context classification system, context-specific SPFs were developed for four intersection groups (C2T unsignalized 3-leg, C3R signalized 4-leg, and C3C signalized and unsignalized 4-leg). Data were collected for 29 roadway and traffic variables using the national MIRE 2.0 standard as a data collection template. This is the first use of MIRE 2.0 to develop SPFs. By following this national standard, it is easier for other state and local agencies to accurately collect the same data and utilize the methodologies discussed in this article. During data collection, it was discovered that several intersections had missing minor AADT values. Because minor AADT is an important variable in SPFs, a linear regression model was developed to predict minor AADT at intersections where these data were unavailable. The developed minor AADT model was significant at $\alpha = 0.05$ with nine significant factors, adjusted R^2 of 0.684, and MAPE of 5.6% when applied to a test data set. These performance measures are better than similar models developed in previous research.

SPFs were then developed for the four studied groups of intersections by comparing Poisson, NB, ZIP, ZINB, and hurdle models as appropriate. Multiple performance measures (MAPE, MSPE, AIC, and BIC) were used to determine the best fitting and performing model for each classification group. The developed minor AADT model was used to estimate missing minor AADT values for all four groups. All variables included in the developed SPFs were significant at 5% significance level (except for the major exclusive left turn length variable in the C3R signalized 4-leg intersection group). For C2T unsignalized 3-leg and C3C unsignalized 4-leg intersections, the ZINB models had lower performance measure values than the other models and were therefore used as the SPFs for these groups. One reason for the ZINB models' better performance was the high number of zero crash counts in these two groups. The C2T unsignalized 3-leg SPF was compared to a Florida-specific rural SPF developed in previous research using different data; this comparison showed that the context-specific SPF developed in this study performed better than the previously developed SPF. Similar Florida SPFs from previous research were not available for comparison with the other groups. For the C3R signalized 4-leg and C3C signalized 4-leg intersections, the NB models had better prediction performance than Poisson models and were used as the SPFs for these groups. The other three models were not considered for these groups because there were very few intersections with zero crashes. The developed SPFs differed between the groups and included variables not present in the HSM. One of these variables was a regional variable which was significant for the C3C unsignalized 4-leg intersection group and showed that D7 had fewer crashes than other intersections in this group. Including regional factors provides agencies with a better understanding of crash variations due to regional aspects and can identify areas that merit further study. There were also common variables across the different context-specific SPFs, including major median and major exclusive left turn length. The coefficient signs for these common variables were consistent whenever these variables were significant, indicating they had similar effects for different context classification categories.

Overall, this article illustrates that the significant and influential variables which affect crash frequency differ across different classification groups, even between classifications which are similar or between intersections with the same context classification, but different signalization. By utilizing FDOT's context classification system or a similar system, agencies can better understand the safety differences between intersections in different classification groups and tailor countermeasures to address these factors. Future expansions to this research will develop context-specific SPFs for the remaining intersection groups with sufficient sample sizes. Additional variables will be included in the minor AADT model (such as signalized/unsignalized and intersection context classification) and the SPF modeling process (such as signalization variables for signalized intersection groups) to improve the accuracy of the developed SPFs. Comparisons will be made with SPFs calibrated for Florida to determine any benefits provided by the context-specific SPFs. Once developed, the context-specific SPFs will also be used to identify the intersections in different classification groups that would benefit the most from modifying geometric and traffic features. These improvements will help reduce crashes and increase intersection safety throughout Florida. It is expected that other agencies which implement a similar context classification system will obtain similar benefits.

Declaration of interest

None.

Funding

The authors wish to acknowledge the financial support of the Florida Department of Transportation. The opinions, findings, and conclusions expressed in this paper are those of the authors and not necessarily those of the Florida Department of Transportation or the U.S. Department of Transportation.

ORCID

Ghalia Gamaleldin http://orcid.org/0000-0002-3815-7164
Haitham Al-Deek http://orcid.org/0000-0002-0508-3250
Adrian Sandt http://orcid.org/0000-0002-7708-3591
John McCombs http://orcid.org/0000-0001-6098-0942

References

American Association of State Highway and Transportation Officials. (2010). *Highway safety manual* (1st ed.). Washington, DC: American Association of State Highway and Transportation Officials.

Basu, S., & Saha, P. (2017). Regression models of highway traffic crashes: A review of recent research and future research needs. *Procedia Engineering*, *187*, 59–66. doi:10.1016/j.proeng.2017.04.350

Centers for Disease Control & Prevention. (2015, December 14). *Motor vehicle crash deaths: Costly but preventable [Fact sheet]*. Retrieved from https://www.cdc.gov/motorvehiclesafety/pdf/statecosts/FL-2015CostofCrashDeaths-a.pdf.

Chen, C., & Xie, Y. (2016). Modeling the effects of AADT on predicting multiple-vehicle crashes at urban and suburban signalized intersections. *Accident Analysis & Prevention*, *91*, 72–83. doi:10.1016/j.aap.2016.02.016

Dixon, K., Monsere, C., Avelar, R., Barnett, J., Escobar, P., Kothuri, S., & Wang, Y. (2015). *Improved safety performance functions for signalized intersections* (Report No. FHWA-OR-RD-16-03). Salem, OR: Oregon Department of Transportation Research Section.

Donnell, E. T., Gayah, V. V., & Jovanis, P. (2014). *Safety performance functions* (Report No. FHWA-PA-2014-007-PSU WO 1). University Park, PA: Thomas D. Larson Pennsylvania Transportation Institute.

Eun-Ha, C. (2010). *Crash factors in intersection-related crashes: An on-scene perspective* (Report No. DOT HS 811 366). Washington, DC: National Highway Traffic Safety Administration.

Florida Department of Transportation. (2017). FDOT context classification. Retrieved from https://fdotwww.blob.core.windows.net/sitefinity/docs/default-source/roadway/completestreets/files/fdot-context-classification.pdf?sfvrsn=12be90da_2.

Florida Department of Transportation. (2018, January). FDOT district five complete streets & context classification [Presentation]. Retrieved from https://www.r2ctpo.org/wp-content/uploads/Context-Classification-Presentation.pdf.

Florida Department of Transportation. (2019). Geographic Information System (GIS) [GIS shapefiles]. Retrieved from https://www.fdot.gov/statistics/gis/default.shtm.

Garber, N. J., & Rivera, G. (2010). Safety performance functions for intersections on highways maintained by the Virginia Department of Transportation (Report No. FHWA/VTRC 11-CR1). Charlottesville, VA: University of Virginia, Department of Civil & Environmental Engineering.

Google. (2019). Google Maps. Retrieved from https://www.google.com/maps/.

JMP. (n.d.). *Statistics Knowledge Portal: Multicollinearity*. Retrieved from https://www.jmp.com/en_us/statistics-knowledge-portal/what-is-multiple-regression/multicollinearity.html.

Kaaf, K. A., & Abdel-Aty, M. (2015). Transferability and calibration of Highway Safety Manual performance functions and development of new models for urban four-lane divided roads in Riyadh, Saudi Arabia. *Transportation Research Record: Journal of the Transportation Research Board*, *2515*(1), 70–77. doi:10.3141/2515-10

Lefler, N., Zhou, Y., Carter, D., McGee, H., Harkey, D., & Council, F. (2017). *Model inventory of roadway elements – MIRE 2.0* (Report No. FHWA-OR-RD-16-03). Vienna, VA: Vanasse Hangen Brustlin, Inc. (VHB); Chapel Hill, NC: University of North Carolina Highway Safety Research Center.

Lord, D., & Mannering, F. (2010). The statistical analysis of crash-frequency data: A review and assessment of methodological alternatives. *Transportation Research Part A: Policy and Practice*, *44*(5), 291–305. doi:10.1016/j.tra.2010.02.001

Lu, J. (2013). *Development of safety performance functions for SafetyAnalyst applications in Florida* (Doctoral dissertation). Retrieved from https://digitalcommons.fiu.edu/etd/880/.

Pan, T. (2008). *Assignment of estimated average annual daily traffic volumes on all roads in Florida* (Doctoral dissertation). Retrieved from https://scholarcommons.usf.edu/cgi/viewcontent.cgi?referer=https://search.yahoo.com/&httpsredir=1&article=1441&context=etd

Prasetijo, J., Musa, W. Z., Jawi, Z. M., Zainal, Z. F., Anuar, K., Ridho, M.A., & Aron, M. A. M. (2019). Vehicle road accident prediction model along federal road FT050 Kluang-A/Hitam-B/Pahat Route using excess zero data. *Journal of the Society of Automotive Engineers Malaysia*, *3*(1), 38–49.

SAS 9.4. (2013). *Computer software.* Cary, NC: SAS Institute Inc.

Shirazi, M., Lord, D., & Geedipally, S. R. (2017). A procedure to determine when safety performance functions should be recalibrated. *Journal of Transportation Safety & Security*, *9*, 457–469.

Shiyuka, N. (2018). *Hurdle negative binomial model for motor vehicle crash injuries in Namibia* (Master's thesis). Retrieved from http://repository.unam.edu.na/bitstream/handle/11070/2232/shiyuka2018.pdf?sequence=1&isAllowed=y.

Srinivasan, R., & Bauer, K. (2013). *Safety performance function development guide: Developing jurisdiction-specific SPFs* (Report No. FHWA-SA-14-005). Chapel Hill, NC: University of North Carolina Highway Safety Research Center.

Srinivasan, R., & Carter, D. (2011). *Development of safety performance functions for North Carolina* (Report No. FHWA/NC/2010-09). Chapel Hill, NC: University of North Carolina Highway Safety Research Center.

Wang, X., & Abdel-Aty, M. (2007). Right-angle crash occurrence at signalized intersections. *Transportation Research Record: Journal of the Transportation Research Board*, *2019*(1), 156–168. doi:10.3141/2019-19

Wang, J., Abdel-Aty, M., & Lee, J. (2015). Examination of the transferability of safety performance functions for developing crash modification factors using empirical Bayes method. *Transportation Research Record: Journal of the Transportation Research Board*, *2583*(1), 73–80. doi:10.3141/2583-10

Washington, S. P., Karlaftis, M. G., & Mannering, F. L. (2011). *Statistical and econometric methods for transportation data analysis* (2nd ed.). Boca Raton, FL: Taylor and Francis Group, LLC.

World Health Organization. (2018). Global status report on road safety 2018. *Geneva: World Health Organization*. Retrieved from World Health Organization website https://www.who.int/violence_injury_prevention/road_safety_status/2018/en/.

Xie, Y., & Chen, C. J. (2016). *Calibration of safety performance functions for Massachusetts urban and suburban intersections* (Report No. UMTC 16.01). Lowell, MA: University of Massachusetts Lowell.

Young, J., & Park, P. Y. (2012). Comparing the Highway Safety Manual's safety performance functions with jurisdiction-specific functions for intersections in Regina. Paper presented at the 2012 Annual Conference of the Transportation Association of Canada, Fredricton, New Brunswick.

Young, J., & Park, P. Y. (2013). Benefits of small municipalities using jurisdiction-specific safety performance functions rather than the Highway Safety Manual's calibrated or uncalibrated safety performance functions. *Canadian Journal of Civil Engineering*, *40*(6), 517–527. doi:10.1139/cjce-2012-0501

Zhang, Y., Xie, Y., & Li, L. (2012). Crash frequency analysis of different types of urban roadway segments using generalized additive model. *Journal of Safety Research*, *43*(2), 107–114. doi:10.1016/j.jsr.2012.01.003