

Comparing the Crash Injury Severity Risk Factors at High-Volume and Low-Volume Intersections with Different Traffic Control in Alabama

Sumalatha Kesavareddy*, Kirolos Haleem, Mehrnaz Doustmohammadi, Michael Anderson

Department of Civil and Environmental Engineering, University of Alabama in Huntsville, Huntsville, United States

Abstract Objective: Typically, intersections that carry more traffic will have more crashes; however, these crashes might not be severe. On the other hand, low-volume intersections might have lower number of crashes; however, can be more severe than their high-volume counterparts. Since the geometric and traffic characteristics of signalized and stop-controlled intersections are different, the significant factors affecting crash severity at both intersection types will be also different. This paper identifies and compares those significant factors affecting crash severity at high- and low-volume signalized and stop-controlled intersections in Alabama using five-year crash data from 2010 to 2014. A cut-off value of 1,000 vehicles/day was used to classify intersections as high-volume vs. low-volume. **Method:** A random forest model was used to rank variable importance and a binary logit model was applied to identify the significant factors at both high- and low-volume signalized and stop-controlled intersections. Four discrete models (high-volume signalized, low-volume signalized, high-volume stop-controlled, and low-volume stop-controlled) were developed. Roadway, traffic, vehicle, driver, and environmental characteristics were used as independent variables in the models. **Results:** In all four models, crashes in rural areas showed higher severity compared to urban areas and right-turning maneuver showed relatively lesser severity. Rear-end crashes showed lower severity compared to side impacts at high- and low-volume stop-controlled and high-volume signalized intersections. Head-on crashes, driving under influence (DUI) of alcohol/drugs, and increase in driver age showed higher severity at high- and low-volume signalized intersections. Motorcycles were associated with higher severity at high- and low-volume signalized intersections, as well as high-volume stop-controlled intersections. **Conclusions:** Most of the factors with the highest ranking from the random forest model were found significant in the binary logit models. Strategies to alleviate crash severity at different intersections are suggested. **Practical Applications:** Since the left-turning vehicle maneuver showed higher severity likelihood at high-volume signalized intersections, providing enough sight distance and protected left turn phase (with no permitted phase) in busy intersections is suggested. Also, education programs should be designed disseminating the dangerous effect of DUI on crash severity while crossing signalized intersections.

Keywords Crash Injury Severity, Signalized Intersections, Stop-Controlled Intersections, High-Volume, Low-Volume, Significant Factors

1. Introduction

The purpose of an intersection is to organize vehicles turning maneuvers to allow smooth flow of vehicles from the intersecting approaches. Crashes at intersections may or may not be severe, and this could depend on the traffic volume. Although high-volume intersections typically have more crashes than low-volume ones, it is undetermined whether the crash severity will be higher or not, which mainly

depends on the characteristics of the crash and the geometric characteristics of the intersection. Since the geometric and traffic characteristics of signalized and stop-controlled intersections are different, the significant factors affecting crash severity at both intersection types will be different as well.

Based on the distribution of traffic volumes (either high- or low-volume) at signalized and stop-controlled intersections, the percentage of crash severity varies. For instance, in the state of Alabama, from 2010 to 2014, 34.15% of the total crashes occurred at high-volume signalized intersections, 3.90% occurred at low-volume signalized intersections, 12.37% occurred at high-volume stop-controlled intersections, and 2.16% occurred at low-volume stop-controlled intersections. In this study, a high-volume intersection is that intersection having an

* Corresponding author:

sumalatha.kesavareddy@gmail.com (Sumalatha Kesavareddy)

Published online at <http://journal.sapub.org/statistics>

Copyright © 2018 The Author(s). Published by Scientific & Academic Publishing

This work is licensed under the Creative Commons Attribution International

License (CC BY). <http://creativecommons.org/licenses/by/4.0/>

annual average daily traffic (AADT) greater than 1,000 vehicles/day, whereas a low-volume intersection is that intersection with an AADT less than or equal to 1,000 vehicles/day. This definition of high- and low-volume intersections has been emphasized in previous studies (see for example, Peshkin *et al.* (2009)).

The driver behavior and the inherent geometric and traffic characteristics at signalized and stop-controlled intersections are significantly different; thus, separately analyzing crash severity at each intersection type is needed. Furthermore, the factors that affect crash severity at high- and low-volume signalized and stop-controlled intersections are essential to know, to devise the essential countermeasures to alleviate the crash severity. To the author's knowledge, no study has been conducted to-date to identify and compare the significant factors affecting crash severity at high- and low-volume signalized and stop-controlled intersections. However, few studies have investigated the crash frequencies and travel delays at both signalized and stop-controlled intersections (see for example, Liu (1979); Penmetsa and Pulugurtha (2017)).

The main objective of this study is to identify and compare the significant factors that affect crash severity at high- and low-volume signalized and stop-controlled intersections. The study uses 5-year historical crashes in Alabama from 2010 to 2014. The factors considered in the study include roadway, geometric, traffic, vehicle, environmental, and driver characteristics which are deemed essential to investigate and compare crash severity at these intersection types with high- and low-volumes. The crash severity includes two levels; severe and non-severe. Severe crashes include fatalities and incapacitating (or serious) injuries. On the other hand, non-severe crashes include property damage only (PDO), possible injuries, and non-incapacitating injuries. To identify and compare the significant risk factors of crash severity at intersections types, the binary logit model is applied. In addition, one of the common variable screening methods, the random forest, is used to rank the importance of the independent variables explored in the study.

2. Literature Review

This section reviews the most relevant studies which identified factors that affect crash severity at signalized and stop-controlled intersections. Researchers have adopted many statistical techniques to identify factors that affect crash severity. Among the various models, the binary logit model was widely used for identifying the factors that affect crash severity. For example, Wang and Abdel-Aty (2008) used the binary logit model while analyzing left turn crash severity and concluded that injury severity was higher when left turning vehicles collide with opposing through traffic. They found that driver attributes, vehicular characteristics, geometric design, and environmental factors were the significant variables that affect injury severity for left turn

crashes. Al-Ghamdi (2002) also used the binary logit model to identify the factors that affect crash severity and concluded that location and crash type were the two key significant variables impacting the crash severity.

Different independent variables were considered when analyzing crash severity. According to Dissanayake and Lu (2002), the driver's use of alcohol or drugs, ejection from the vehicle in the crash, point of impact, existence of curve or grade at the crash location, and vehicle speed were determined as factors that increased the severity for young drivers. Savolainen and Mannering (2007) concluded that crashes were less severe under wet pavement conditions and near intersections, because of lower speeds maintained by drivers in those conditions. Kockelman and Kweon (2002) concluded that passenger cars are safer than pickups in single-vehicle crashes, while concluding the opposite for two-vehicle crashes. Duncan *et al.* (2002) identified that high speed, driving at night, female drivers, and driving under the influence increased the risk of severity. Nassar *et al.* (1994) developed severity models for bad weather conditions and concluded that drivers may be more attentive in inclement weather conditions. Zwerling *et al.* (2005) concluded that the total number of crashes, as well as crashes with fatalities or injuries were higher in rural areas compared to urban areas. Kim *et al.* (2013) found that the probability of fatal injury for drivers in single-vehicle crashes increased significantly compared to other crash types, e.g., multi-vehicle crashes.

While focusing on signalized intersections, Huang *et al.* (2008) identified the factors that affected the driver injury severity and vehicle damage. They concluded that severity was lower for crashes that occurred in peak period and in good street lighting condition compared to crashes that occurred at night, T-intersections, vehicles traveling in the right-most lane, and for intersections equipped with red light running cameras. They also found that heavy vehicle drivers were less likely to be injured when compared with two-wheeler riders because of the better protection offer to the driver.

Studies that analyzed severity at unsignalized intersections can also be found in the literature. Wang and Qin (2002) studied driver mistakes that occurred at uncontrolled, stop-controlled, and signalized intersections. Wang and Qin (2002) also identified potential countermeasures and concluded that running stop signs, high speed, driving under the influence (DUI) of alcohol or drugs, and poor visibility can be the causes of increased crash severity. As concluded from Devlin *et al.* (2011), running stop signs was the common error made drivers which increased injury severity. Haleem and Abdel-Aty (2010) studied the factors that relate to crash severity at unsignalized intersections using traffic volume, number of through lanes, geometric factors, shoulder width, number of left turn movements, number of left and right turn lanes on major approach, and driver age.

From the abovementioned literature review, it can be concluded that no single study was conducted to compare the

significant predictors of crash severity at high- and low-volumes at both signalized and stop-controlled intersections.

This study attempts to fill this gap while considering various independent predictors. The detailed description of the data and variables considered is shown in the “Data and Variable Settings” section.

3. Data and Variable Settings

In this study, five years of intersection-related crash data (2010 to 2014) in the state of Alabama were used. The data were extracted from the Critical Analysis Reporting Environment (CARE) database, which is maintained by the Center for Advanced Public Safety (CAPS). After exclusion of other intersections types, such as yield-controlled intersections and unknown intersection types, only crashes that occurred at signalized and stop-controlled intersections were considered. The total numbers of crashes were further categorized into four separate datasets, representing crashes at high-volume signalized, low-volume signalized, high-volume stop-controlled and low-volume stop-controlled intersections. In the analysis, the cut-off value for low traffic volume was less than or equal to a major-approach AADT (or intersection-entering vehicles) of 1,000 vehicles/day, whereas a major-approach AADT greater than 1,000 was considered high traffic volume. Note that due to data unavailability, the minor-approach AADT was not considered in the analysis. The major intersection approach represented the approach with greater number of lanes compared to the minor approach and that approach was usually a state-maintained roadway (for this, the AADT was available). On the other hand, the minor intersection

approach represented the approach with less number of lanes compared to the major approach and that approach was mostly a local roadway (or a non-state-maintained roadway, where the AADT was mostly unavailable).

The use of 1,000 vehicles/day as the cut-off value for separating high- and low-volumes has been also used in previous studies (see for example, the Cornell Local Roads Program (2017)). Close to twenty independent variables were considered. These variables include driver characteristics (driver age, driver gender, and driver condition), roadway characteristics (roadway type, speed limit, roadway condition, and terrain), land use (urban and rural), and environmental characteristics (e.g., weather and lighting conditions).

As previously mentioned, four separate models (high-volume signalized, low-volume signalized, high-volume stop-controlled, and low-volume stop-controlled) were developed for comparing the significant factors contributing to crash injury severity. Twenty independent variables were included in all the four models (as shown in Tables 1 and 2). Descriptive statistics of the variables considered at high- and low-volume signalized intersections, and high- and low-volume stop-controlled intersections are shown in Tables 1 and 2, respectively. In this study, driver age is treated as a continuous independent variable. The AADT at the intersections is calculated by summing up the traffic volumes from both directions on the main road. As shown in Tables 1 and 2, the response (dependent) variable in the analysis was crash severity. The crash severity includes two levels; severe and non-severe. Severe crashes include fatalities and incapacitating (or serious) injuries. On the other hand, non-severe crashes include PDO, possible injuries, and non-incapacitating injuries.

Table 1. Descriptive Statistics of Variables at High- and Low-Volume Signalized Intersections

Variable Description	High-Volume Signalized Intersections	Low-Volume Signalized Intersections
Crash Injury Severity	Non-Severe =21112(95.05%) Severe =1100(4.95%)	Non-Severe =2388(95.71%) Severe =107(4.29%)
Day of Week	Weekday =17665(79.53%) Weekend =4547(20.47%)	Weekday =1937(77.64%) Weekend =558(22.36%)
Hour of Crash	Morning Peak (7am to 9:59 am) =3146(14.16%) Morning Off-Peak (10 am to 3:59 pm) =10011(45.07%) Evening Peak (4 pm to 6:59 pm) =5120(23.05%) Evening Off-Peak (7 pm to 6:59 am) =3935(17.72%)	Morning Peak (7am to 9:59 am) =344(13.79%) Morning Off-Peak (10 am to 3:59 pm) =1070(42.89%) Evening Peak (4 pm to 6:59 pm) =598(23.97%) Evening Off-Peak (7 pm to 6:59 am) =483(19.36%)
Primary Cause	Failed to Yield =4511(20.31%) Followed Too Close =4340(19.54%) Ran Traffic Signal =4307(19.38%) Misjudge Stopping Distance =2883(12.98%) Made Improper Turn =2214(9.97%) Other =1843(8.30%) Distracted by Use/Pass/Fall =1412(6.36%) Driving under Influence (DUI) =453(2.04%) Driving too Fast =249(1.12%)	Failed to Yield =539(21.60%) Followed Too Close =453(18.16%) Ran Traffic Signal =386(15.47%) Misjudge Stopping Distance =334(13.39%) Made Improper Turn =264(10.58%) Distracted by Use/Fall/Pass =174(6.97%) Driving under Influence (DUI) =83(3.33%) Driving too Fast =35(1.44%) Other =226(9.06%)

Variable Description	High-Volume Signalized Intersections	Low-Volume Signalized Intersections
Manner of Crash	Side Impact = 8021(36.11%) Rear-end =10566(47.57%) Angle (Front to Side) =2563(11.54%) Head-on (Front to Front) =362(1.63%) Other =200(0.90%) Casual Vehicle Backing =47(0.21%) Single-Vehicle Crash =421(1.90%) Non-Collision =32(0.14%)	Side Impact =830(33.27%) Rear-end =1163(46.61%) Angle (Front to Side) =364(14.59%) Head-on (Front to Front) =33(1.32%) Other =37(1.48%) Single-Vehicle Crash =68(2.73%)
Lighting Condition	Daylight =17174(77.32%) Dark-Roadway light =2347(10.56%) Dark-Continuous =1992(8.97%) Dusk =546(2.46%) Dawn =153(0.69%)	Daylight =1884(75.51%) Dark-Roadway Light =269(10.78%) Dark-Continuous =233(9.34%) Dusk =87(3.49%) Dawn =22(0.88%)
Weather Condition	Clear =15188(68.38%) Cloudy =4526(20.38%) Rain =1949(8.77%) Other =17(0.08%) Fog =49(0.22%) Mist =435(1.95%) Snow =48(0.22%)	Clear =1703(68.26%) Cloudy =498(19.96%) Rain =233(9.34%) Other =61(2.44%)
Land Use	Rural =1127(5.07%) Urban =21085(94.93%)	Rural =372(14.91%) Urban =2123(85.09)
Driver Gender	Male =11630(52.36%) Female =10582(47.64%)	Male =1334(53.47%) Female =1161(46.53%)
Driver Condition	Apparently Normal =21205(95.46%) Under the Influence =599(2.70%) Asleep/Illness/Emotional =254(1.14%) Physical Impairment =48(0.22%) Other =106(0.48%)	Apparently Normal =2343(93.91%) Under the Influence =102(4.09%) Asleep/Illness/Emotional =37(1.48%) Other =13(0.52%)
Vehicle Maneuver	Slowing/Stopped for Signal =2657(11.96%) Other =122(0.55%) Turning Left = 4725(21.27%) Turning Right =1864(8.39%) Changing Lanes =785(3.53%) Backing =348(1.57%) Entering/Leaving Main Road =123(0.55%) Making U-Turn =89(0.40%) Overtaking =50(0.23%) Movement Essentially =11411(51.37%) Negotiating a Curve =38(0.17%)	Entering/Leaving Main Road =21(0.84%) Slowing/Stopped for Signal =327(13.11%) Turning Left =571(22.89%) Turning Right =222(8.90%) Backing =30(1.20%) Other =23(1.04%) Changing Lanes =99(3.97%) Movement Essentially =1185(47.49%) Making U-Turn =14(0.56%)
Speed Limit	Posted Speed Limit ≤ 45 =16308(73.42%) 45 < Posted Speed Limit ≤ 55 =5904(26.58%)	Posted Speed Limit ≤ 45 =1616(64.77%) 45 < Posted Speed Limit ≤ 55 =879(35.23%)
Roadway Type	Asphalt =22040(99.23%) Concrete =170(0.77%)	Asphalt =2472(99.08%) Concrete =23(0.92%)
Roadway Condition	Dry =18982(85.46%) Wet =3171(14.28%) Other =11(0.055%) Ice =27(0.12%) Snow =21(0.09%)	Dry =2115(84.77%) Wet =374(14.99%) Other =6(0.24%)
Roadway Terrain	Straight and Level =17834(80.29%) Straight at Hillcrest =3615(16.27%) Curve =763(3.44%)	Straight and Level =2111(84.61%) Straight at Hillcrest =313(12.55%) Curve =71(2.85%)

Variable Description	High-Volume Signalized Intersections	Low-Volume Signalized Intersections
Control Function	Yes =21912(98.65%) No =300(1.35%)	Yes =2463(98.72%) No =32(1.28%)
Number of Through Lanes (per Direction)	> Two Lanes = 18117(81.56%) Two Lanes =3997(17.99%) One Lane =98(0.447%)	> Two Lanes =1922(77.03%) Two Lanes =557(22.32%) One Lane =16(0.64%)
Turn Lanes	Both Left & Right =11557(52.03%) Left Turn =6601(29.720%) Right Turn =1188(5.35%) None =2811(12.66%) Other =55(0.25%)	None =295(11.82%) Left Turn =1057(42.36%) Right Turn =118(4.73%) Both Left & Right =1025(41.08%)
One-Way	No =21798(98.14%) Yes =414(1.86%)	No =2474(99.16%) Yes =21(0.84%)
Vehicle Type	Passenger Car =11589(52.17%) Truck/Trailer =801(3.61%) Sport Utility Vehicle =8553(38.51%) Van =1037(4.67%) Motorcycle =69(0.31%) Other =45(0.20%) Bus =39(0.18%) Station Wagon =79(0.36%)	Passenger Car =1261(50.54%) Truck/Trailer =85(3.41%) Sport Utility Vehicle =1013(40.60%) Van =111(4.45%) Other =8(0.32%) Motorcycle =17(0.68%)
Driver Age	Mean=39.176, S.D.=18.538	Mean=38.669, S.D.=18.376

Table 2. Descriptive Statistics of Variables at High- and Low-Volume Stop-Controlled Intersections

Variable Description	High-Volume Stop-Controlled Intersections	Low-Volume Stop-Controlled Intersections
Crash Injury Severity	Non-Severe =7281(89.71%) Severe =835(10.29%)	Non-Severe =1209(86.92%) Severe =182(13.08%)
Day of Week	Weekday =6570(80.95%) Weekend =1546(19.05%)	Weekday =1086(78.07%) Weekend =305(21.93%)
Hour of Crash	Morning Peak (7am to 9:59 am) =1339(16.50%) Morning Off-Peak (10 am to 3:59 pm) =3636(44.80%) Evening Peak (4 pm to 6:59 pm) =1952(24.05%) Evening Off-Peak (7 pm to 6:59 am) =1189(14.65%)	Morning Peak (7am to 9:59 am) =238(17.11%) Morning Off-Peak (10 am to 3:59 pm) =579(41.62%) Evening Peak (4 pm to 6:59 pm) =337(24.23%) Evening Off-Peak (7 pm to 6:59 am) =237(17.04%)
Primary Cause	Failed to Yield =5444(67.08%) Followed Too Close =263(3.24%) Ran Stop Sign =558(6.88%) Misjudge Stopping Distance =279(3.44%) Made Improper Turn =384(4.73%) Distracted by Use/Pass/Fall =211(2.60%) DUI =183(2.25%) Driving too Fast =66(0.81%) Other =728(8.97%)	Failed to Yield =1037(74.55%) Followed Too Close =23(1.65%) Ran Stop Sign =89(6.40%) Misjudge Stopping Distance =39(2.80%) Distracted by Use/Pass/Fall =35(2.52%) Driving too Fast =13(0.93%) DUI =43(3.10%) Other =112(8.05%)
Manner of Crash	Side Impact =5453(67.197%) Rear-end =1002(12.35%) Angle (Front to Side) =1014(12.49%) Head-on (Front to Front) =71(0.87%) Other =88(1.08%) Single-Vehicle Crash =453(5.58%) Non-Collision =17(0.21%) Casual Vehicle Backing =18(0.22%)	Side Impact =912(65.56%) Rear-end =123(8.84%) Angle (Front to Side) =235(16.89%) Head-on (Front to Front) =13(0.95%) Single-Vehicle Crash =92(6.61%) Other =16(1.15%)
Lighting Condition	Daylight = 6578(81.05%) Dark-Roadway light =463(5.70%) Dark-Continuous =821(10.12%) Dusk =190(2.34%) Dawn =64(0.79%)	Daylight =1094(78.65%) Dark-Roadway Light =58(4.17%) Dark-Continuous =175(12.58%) Dusk =40(2.87%) Dawn =24(1.73%)

Variable Description	High-Volume Stop-Controlled Intersections	Low-Volume Stop-Controlled Intersections
Weather Condition	Clear =5680(69.99%) Cloudy =1682(20.72%) Rain =531(6.54%) Fog =76(0.94%) Other =147(1.81%)	Clear =961(69.09%) Cloudy =292(20.99%) Rain =95(6.83%) Fog =20(1.44%) Other =23(1.65%)
Driver Gender	Male =4206(51.82%) Female =3910(48.18%)	Male =740(53.20%) Female =651(46.80%)
Driver Condition	Apparently Normal =7753(95.53%) Under the Influence =231(2.85%) Physical Impairment =14(0.17%) Asleep/Illness/Emotional =81(1.00%) Other =37(0.46%)	Apparently Normal =1326(95.33%) Asleep/Illness/Emotional =13(0.93%) Under the Influence =52(3.74%)
Vehicle Maneuver	Slowing/Stopped for Signal =386(4.76%) Entering/Leaving Main Road =549(6.76%) Making U-turn =15(0.18%) Turning Left =2307(28.43%) Turning Right =612(7.54%) Changing Lanes =14(0.17%) Backing =171(2.11%) Other =135(1.66%) Movement Essentially =3927(48.39%)	Entering/Leaving Main Road =94(6.76%) Slowing/Stopped for Signal =67(4.82%) Other =29(2.08%) Turning Left =449(32.28%) Turning Right =97(6.97%) Backing =17(1.22%) Movement Essentially =638(44.87%)
Speed Limit	Posted Speed Limit ≤ 45 =6770(83.42%) 45 < Posted Speed Limit ≤ 55 =1346(16.58%)	Posted Speed Limit ≤ 45 =1149(82.60%) 45 < Posted Speed Limit ≤ 55 =242(17.40%)
Roadway Condition	Dry =7155(88.16%) Wet =942(11.61%) Other =19(0.23%)	Dry =1218(87.56%) Wet =173(12.44%)
Roadway Terrain	Straight and Level =6328(77.97%) Straight at Hillcrest =1520(18.73%) Curve =268(3.30%)	Straight and Level =1126(80.95%) Straight at Hillcrest =222(15.96%) Curve =43(3.09%)
Road Type	Asphalt =8080(99.56%) Concrete =16(0.20%) Unpaved =20(0.25%)	None
Control Function	Yes =1465(98.79%) No =18(1.21%)	Yes =1260(90.58%) No =131(9.42%)
Number of Through Lanes (per Direction)	> Two Lanes =1054(12.99%) Two Lanes =6931(85.40%) One Lane =131(1.61%)	> Two Lanes =161(11.57%) Two Lanes =1212(87.13%) One Lane =18(1.30%)
Turn Lanes	None =6317(77.83%) Left Turn =451(5.56%) Right Turn =605(7.45%) Other =17(0.21%) Both Left and Right =726(8.95%)	None =1093(78.58%) Left Turn =107(7.69%) Right Turn =67(4.82%) Both Left and Right =124(8.91%)
Vehicle Type	Passenger Car =4266(52.56%) Truck/Trailer =218(2.69%) Van =373(4.60%) Other =40(0.49%) Sport Utility Vehicle =3172(39.08%) Motorcycle =27(0.33%) Bus =20(0.25%)	Passenger Car =697(50.11%) Truck/Trailer =50(3.60%) Sport Utility Vehicle =563(40.47%) Van =65(4.67%) Other =16(1.15%)
One Way	No =7999(98.56%) Yes =117(1.44%)	No =1381(99.28%) Yes =10(0.72%)
Land Use	Rural =2646(32.60%) Urban =5470(67.40%)	Rural =734(52.77%) Urban =657(47.23%)
Driver Age	Mean =41.294, St. Dev. = 20.324	Mean =41.308, St. Dev. = 20.404

As shown in Tables 1 and 2, several categorical variables were considered in this study. Each categorical variable includes two or more levels. One of the levels in each categorical variable was chosen as the reference (or base case) and the remaining levels were estimated in relative to that reference level when fitting the four binary logit models. The effect of each level of the independent variable on crash severity was measured by comparing the estimated (β) value with the base case (Yasmin et al. (2013)). If the estimated value is greater than zero, then this specific level is associated with higher severe injury likelihood when compared to the base case.

In addition to the estimated values, the odds ratio (OR) was used to interpret the implication of each estimate on crash severity. The range of odds ratio is from zero to positive infinity. If OR is less than one (< 1), then it means the variable is generally associated with less likelihood of severe crashes. On the other hand, if OR is greater than one (> 1), then the variable is generally associated with high likelihood of severe crashes. To assess the goodness-of-fit of the fitted binary logit models, the Akaike information criterion (AIC), Pseudo R^2 , and the deviance estimates were used. The lower AIC, residual deviance, and higher Pseudo R^2 values, the better the model, and vice-versa. The stepwise regression technique was used to keep the variables in the models. Stepwise regression is a mix of both the backward deletion and forward addition methods when fitting a regression model. A 90% confidence level (or 10% significance level) was used to judge the inclusion of the variables in the final models.

4. Methodology

4.1. Binary Logit (Logistic Regression) Model

The approach used in this paper to identify the significant predictors of injury severity was the logistic regression model (or logit model). The logistic regression determines the relationship between categorical dependent variables and one or more independent (categorical, continuous, or both categorical and continuous) variables. The response variable in logistic regression can be binary or dichotomous (see for example, Wang and Abdel-Aty (2008); Al-Ghamdi (2002)). Therefore, the response is dichotomous and it takes the values “zero” and “one”, representing “non-occurring” and “occurring” events, respectively (Chang and Yeh (2006)). If the dependent variable is categorized into more than two levels, then the multinomial logistic regression model can be applied (Shankar and Mannering (1996); Islam and Mannering (2006)). In this study, since fatalities were relatively few at the four different model types, fatalities were aggregated with incapacitating injuries to represent severe injuries. The other response level was non-severe injuries. For this the binary logit (logistic regression) model was applied in this study. The binary logistic regression equation takes the following form:

$$g(x) = \text{logit}(x) = \ln[x/(1-x)] = \alpha + \beta x \quad (1)$$

where,

$g(x)$ = crash severity formula as a function of the intercept and vector of independent variables;

x = vector of the independent variables;

α = intercept to be estimated; and

β = vector of coefficients to be estimated.

The likelihood function $L(\alpha, \beta)$ for the pair (x_i, y_i) is given below in Equation (2) (the logistic regression from Carnegie Mellon University (2017)):

$$L(\alpha, \beta) = \prod_{i=1}^n p(x_i)^{y_i} (1 - p(x_i))^{1-y_i} \quad (2)$$

where,

n = number of observations (each observation denotes the i^{th} observed outcome);

x_i = independent/explanatory variable for observation i ;

y_i = dependent/response variable (i.e., the crash severity) for observation i ;

$p(x_i)$ = probability of occurrence of severe injuries (at $y = 1$) for observation i ; and

$1 - p(x_i)$ = probability of occurrence of non-severe injuries (at $y = 0$) for observation i .

It should be noted here that the best estimates of the intercept (α) and β can be calculated by maximizing the likelihood function in Equation (2). In this study, the R package (2016) has been used to estimate the final maximum likelihood estimate of each variable in the four models. The severe crash injury probability (π_1) (or “ $p(x)$ ”) and the non-severe injury probability (π_2) (or “ $1 - p(x)$ ”) are estimated as shown in Equations (3) and (4) below (per Haleem (2016)):

$$\pi_1 = p(x) = \frac{\exp(\alpha + \beta x)}{1 + \exp(\alpha + \beta x)} \quad (3)$$

$$\pi_2 = 1 - p(x) = \frac{1}{1 + \exp(\alpha + \beta x)} \quad (4)$$

4.2. Random Forest

Random forest model was used in this study to rank the importance of variables. This was performed prior to binary logit model to find out the key variables, and then compare the final list of important variables with the final variables in each of the four binary logit models. Ho (1995) created the first algorithm of random decision forest, then it was developed by Breiman (2001) and is considered one of the promising machine learning techniques for screening important variables (see for example, Abdel-Aty and Haleem (2011); Haleem and Gan (2013)). In this technique, a number of trees are grown by selecting some observations randomly from the original data set with substitution, then searching over a randomly selected subset of variables at each split till the variable significant is ranked (Haleem and Gan (2013)). Several studies have used the random forest to analyze binary response variables (see for example, Harb et al. (2009); Sparks (2009)). To rank the variables, the Gini index was used in this study, which tests the homogeneity of the nodes and leaves (or the purity of the variables) in the resulting random forest model. A higher Gini index indicates

a relatively important (or pure) variable, and vice-versa.

5. Results and Discussion

The results of the four binary logit models for analyzing crash severity at both signalized and stop-controlled intersections with high- and low-volumes are shown in Tables 3 and 4, respectively. Table 3 shows the two binary logit models at signalized intersections with high- and low-volumes, while Table 4 shows the two models at stop-controlled intersections with high- and low-volumes. In both tables, only the significant variables are shown. The goodness-of-fit statistics (i.e., the AIC, Pseudo R^2 , and residual and null deviance estimates) are shown in all the four models towards the end of the tables.

5.1. Interpretation of the Binary Logit Models at High- and Low-Volume Signalized Intersections

As shown in Table 3, rural areas in both types of signalized intersections had higher likelihood of severe crashes ($\beta=0.305$, $OR=1.356$ for high-volume and $\beta=0.653$, $OR=1.920$ for low-volume) compared to urban areas. This could be interpreted as drivers tempted to drive faster than the posted speed limit. Moreover, what might encourage speeding in rural areas is the lack of pedestrian activities compared to urban areas. These results are consistent with the findings of Zwerling *et al.* (2005). It is also observed that crash severity is 56.4% higher in low-volume signalized intersections compared to high-volume signalized intersections.

For the low-volume signalized intersections model, PM non-peak hour (7:00 PM to 6:59 AM) ($\beta=0.817$, $OR=2.264$) was associated with a higher severity likelihood compared to the AM non-peak hour (10:00 AM to 3:59 PM). Drivers usually tend to drive faster (and sometimes speed) during night time compared to day time. Rear-end and single-vehicle crash types are significant factors in the high-volume signalized intersection model, while head-on crash type is a significant factor in both the high- and low-volume signalized intersections models. Rear-end crashes ($\beta=-0.452$, $OR=0.637$) showed lesser injury severity than side impact crash types. This might be since rear-end crashes usually occur in congested situations, where the vehicle speed is relatively low. On the other hand, head-on crashes ($\beta=0.568$, $OR=1.765$ for high-volume and $\beta=1.702$, $OR=5.486$ for low-volume) showed higher likelihood of severe injuries than side impact crashes.

These findings are consistent with the results from Yasmin *et al.* (2013) and Al-Ghamdi (2002). Head-on crashes are usually more severe compared to other crash types due to the higher force of impact. Head-on crashes showed similar severity patterns at both signalized intersections; however, head-on crashes at low-volume signalized intersections were associated with much higher severity likelihood compared to high-volume signalized intersections. Interestingly, single-vehicle crash types ($\beta=0.513$, $OR=1.668$) showed

higher likelihood of severe injuries than side impact crashes.

At low-volume signalized intersections, cloudy weather was associated with lesser severe crash likelihood ($\beta=-0.765$, $OR=0.465$) compared to clear weather. This is intuitive since drivers tend to drive more carefully in relatively inclement weather compared to clear weather conditions. Dusk condition was 39.3% less severe compared to daylight condition. This might be since drivers are more cautious and attentive in poor lighting conditions. Concurring with previous studies, female drivers at high-volume signalized intersections were associated with higher severity ($\beta=0.164$, $OR=1.179$) compared to their male counterparts. This is because female drivers have weaker physical conditions compared to males, which make them more vulnerable to higher severity when crashing. This result is consistent with the finding of Yasmin *et al.* (2013).

Crashes involving motorcycles at both types of signalized intersections were associated with (495% and 262%) higher severity likelihood ($OR=5.947$ and $OR=3.617$) compared to crashes involving passenger cars. This finding is consistent with that from Savolainen and Mannering (2007). This can be explained since motorcyclists have no protection when crashing as opposed to drivers inside the vehicle. At high-volume intersections, crashes involving trucks were associated with higher severity ($OR=1.545$) compared to passenger cars.

At both types of signalized intersections, driver condition levels (DUI and depressed/illness) were associated with an increase in severity likelihood compared to normal driver conditions. At high-volume signalized intersections, physical impairment driver condition showed higher severity compared to normal condition. These results are consistent with the finding of Devlin *et al.* (2011). The highest severity increase was for physical impairment ($OR=4.346$) followed by DUI ($OR=3.833$) and finally depressed/ill ($OR=2.866$) at high-volume intersections. At low-volume intersections, the highest severity increase was for DUI ($OR=4.827$) followed by depressed/ill ($OR=3.382$). This shows the negative impact on severity when driving under the influence of alcohol and drugs, being depressed, or ill. It is also noticed that DUI at low-volume intersections was 99.4% higher likelihood of severe injuries compared to high-volume intersections, possibly due to drivers' speeding.

Regarding the vehicle maneuver, the same baseline for each of the high-volume and low-volume signalized intersections model was initially attempted; however, the parameter estimates did not come as expected. For this, different baselines for each model were used. Turning left at high-volume intersections ($\beta=0.509$, $OR=1.665$) showed higher severity likelihood than stopping or slowing (base case). On the other hand, turning right was associated with crash severity reduction at both types of signalized intersections. This is anticipated since lesser conflict points exist when turning right compared to turning left at an intersection. These findings are consistent with those from Wang and Abdel-Aty (2008) and Al-Ghamdi (2002). In both types of signalized intersections, speed limits above 45 mph

and less than or equal to 55 mph (which inherently implies higher vehicle speeds) were associated with higher severity likelihood compared to speed limits less than or equal to 45 mph. Crashes at relatively higher speeds are most often more severe. This result is consistent with the finding of Devlin et

al. (2011). Assessing this finding, low-volume intersections showed 71.5% higher severity compared to high-volume intersections. This result is consistent with the finding of Al-Ghamdi (2002).

Table 3. Binary Logit Models at Signalized Intersection

Variable Description	High-Volume Signalized Intersections			Low-Volume Signalized Intersections		
	Estimate	P-Value	Odds Ratio	Estimate	P-Value	Odds Ratio
Intercept	-3.341	<0.0001	0.035	-3.803	<0.0001	0.022
<i>Land Use (Base: Urban)</i>						
Rural	0.305	0.013	1.356	0.653	0.012	1.920
<i>Time of Day (Base: 10:00 AM – 3:59 PM)</i>						
7:00 PM – 6:59 AM	---	---	---	0.817	0.046	2.264
<i>Manner of Crash (Base: Side Impact)</i>						
Rear-end	-0.452	0.005	0.637	---	---	---
Head-on	0.568	0.001	1.765	1.702	0.001	5.486
Single-vehicle Crash	0.513	0.019	1.668	---	---	---
<i>Weather Condition (Base: Clear)</i>						
Cloudy	---	---	---	-0.765	0.022	0.465
<i>Lighting Condition (Base: Daylight)</i>						
Dusk	-0.499	0.079	0.607	---	---	---
<i>Driver Gender (Base: Male)</i>						
Female	0.164	0.014	1.179	---	---	---
<i>Vehicle Type (Base: Passenger Car)</i>						
Trucks	0.435	0.013	1.545	---	---	---
Motorcycle	1.783	<0.0001	5.947	1.286	0.099	3.617
<i>Driver Condition (Base: Apparently Normal)</i>						
DUI	1.343	<0.0001	3.833	1.574	0.034	4.827
Asleep/Illness/Emotional	1.053	<0.0001	2.866	1.218	0.052	3.382
Physical Impairment	1.469	<0.0001	4.346	---	---	---
<i>Vehicle Maneuver (Base: Slowing/Stopped for Signal) (Base: Turning Left)</i>						
Turning Left	0.509	0.011	1.665	---	---	---
Turning Right	-0.742	0.008	0.476	-2.209	0.033	0.109
<i>Posted Speed Limit (Base: ≤ 45 mph)</i>						
> 45 mph & ≤ 55 mph	0.162	0.030	1.176	0.637	0.007	1.891
<i>Roadway Type (Base: Asphalt)</i>						
Concrete	-2.044	0.042	0.129	---	---	---
<i>Control Function (Base: Yes)</i>						
No	-1.098	0.016	0.333	---	---	---
<i>Number of Through Lanes (per Direction) (Base: Two Lanes)</i>						
> Two Lanes	0.182	0.051	1.199	---	---	---
<i>One-Way (Base: No)</i>						
Yes	-0.471	0.084	0.624	---	---	---
<i>Driver Age</i>						
Driver Age	0.005	0.002	1.005	0.009	0.081	1.009
<i>Goodness-of-Fit Statistics</i>						
AIC	8,000.5			833.08		
Null Deviance	8,756.3			883.28		
Residual Deviance	7,860.5			719.08		
Pseudo R-Squared	0.102			0.186		
<i>Number of Observations</i>						
N	22,213			2,496		

At signalized intersections, concrete roadways (OR=0.129) were associated with 87.1% lesser severity compared to asphalt roadways. This is due to the increased friction and relatively lower speeds when driving on concrete roads. This was also found in the study by Haleem and Abdel-Aty (2010). At high-volume intersections, lesser crash severity likelihood was found when the traffic signal was not functioning (OR=0.333) compared to functioning traffic signals. This can be explained as drivers might be more cautious and attentive when the traffic control stops functioning. In some situations, when the traffic signal is not functioning, police officers organize the traffic at those affected signalized intersections, which could explain the severity reduction.

High-volume signalized intersections located on three lanes or more (per direction) roadways experienced 19.9% more severity (OR=1.199) compared to those intersections located on two lanes (per direction) roadways, which is consistent with the finding of Al-Ghamdi (2002). This can be interpreted as intersections located on relatively wider roadways can increase the driver confusion, especially when undertaking turning maneuvers. Assessing this finding, high-volume signalized intersections located on one-way roadways experienced 37.6% less severity likelihood (OR=0.624) than high-volume intersections located on two-way roadways. This can be due to the reduction in conflict points on one-way roadways as opposed to two-way roadways.

The increase in driver age was associated with an increase in injury severity likelihood at both high- and low-volume signalized intersections. This shows that old and very old drivers had higher severity probability compared to young drivers. This is anticipated and can be attributed to the weaker physical conditions and slower reaction time of old and very old drivers at signalized intersections. It can be noticed that the increase in severity with slightly higher at low-volume intersections compared to high-volume intersections, possibly due to the increase in the probability of speeding at low-volume intersections, which adversely affect older drivers.

When comparing the goodness-of-fit of both models, it can be observed that the low-volume signalized intersections model outperformed the high-volume signalized intersections (i.e., lower AIC and residual deviance and higher Pseudo R^2 estimates).

5.2. Interpretation of the Binary Logit Models at High- and Low-volumes of Stop-Controlled Intersections

As shown in Table 4, at both types of stop-controlled intersections, crashes in rural areas ($\beta=0.798$, OR=2.222 for high-volume; $\beta=1.259$, OR=3.521 for low-volume) were more likely to result in an increased severity likelihood compared to urban areas. Similar to signalized intersections, low-volume stop-controlled intersections were associated with higher severity at rural areas. For the high-volume stop-controlled intersections model, running a stop sign ($\beta=0.406$, OR=1.501) was associated with higher severity

likelihood compared to failing to yield at the intersection. These results are consistent with Devlin *et al.* (2011). Running a stop sign can cause an angular collision with other vehicles, which could possibly increase the likelihood of the injury severity for the drivers and accompanying passengers.

Rear-end and single-vehicle crash types were significant at high-volume signalized intersection model. Rear-end crashes (OR=0.289 for high-volume; OR=0.076 for low-volume) showed lesser injury severity than side impact crash types. Unlike signalized intersections, single-vehicle crashes at high-volume stop-controlled intersections showed lesser likelihood crash severity compared to side impact. This might be because of lower speeds at stop-controlled intersections compared to signalized intersections. Like the previous finding at signalized intersections, cloudy weather was associated with lesser severity likelihood (OR=0.763) at high-volume stop-controlled intersections compared to clear weather.

As found in the signalized intersections model, dusk condition was associated with lesser severity (OR=1.453) at high-volume stop intersections compared to daylight condition. At high-volume stop-controlled intersections, physical impairment driver condition (OR=3.047) showed higher severity compared to normal condition. This is because the drivers with physical impairment have weaker physical condition compared to normal drivers.

Regarding the vehicle maneuver, turning right was associated with severity reduction (OR=0.377) at both types of stop-controlled intersections, which is consistent with the finding from the fitted model at signalized intersections. Surprisingly, the left turning maneuver was associated with injury severity reduction compared to entering/leaving the road at low-volume stop-controlled intersections. The provision of right turning lanes at stop-controlled intersections showed higher likelihood of severity compared to no turn lanes. This might be the reason that drivers may be reckless in turning right, with relatively little attention.

Similar to signalized intersections, crashes involving motorcycles at high-volume stop-controlled intersections were associated with 303.9% higher severity likelihood (OR=4.039) compared to crashes involving passenger cars. SUVs were associated with severity reduction compared to passenger cars, possibly due to the larger inner space in SUVs compared to passenger cars. This result is consistent with the finding of Huang *et al.* (2008). As found at signalized intersections, the age of the driver at high-volume stop-controlled intersections was associated with an increase in injury severity.

When comparing the goodness-of-fit of both models, it can be observed that the low-volume stop-controlled intersections model was better (lower AIC and residual deviance and higher Pseudo R^2 estimates) than the high-volume stop-controlled intersections model. This result is similar to the signalized intersections models, where the low-volume model outperformed the high-volume model.

Table 4. Binary Logit Models at Stop-Controlled Intersections

Variable Description	High-Volume Stop-Controlled Intersections			Low-Volume Stop-Controlled Intersections		
	Estimate	P-Value	Odds Ratio	Estimate	P-Value	Odds Ratio
Intercept	-2.744	<0.0001	0.064	-1.807	<0.0001	0.164
<i>Land Use (Base: Urban)</i>						
Rural	0.798	<0.0001	2.222	1.259	<0.0001	3.521
<i>Primary Cause (Base: Failed to Yield)</i>						
Ran Stop Sign	0.406	0.001	1.501	---	---	---
<i>Manner of Crash (Base: Side Impact)</i>						
Rear-end	-1.240	0.001	0.289	-2.566	0.022	0.076
Single-vehicle Crash	-0.526	0.009	0.591	---	---	---
<i>Weather Condition (Base: Clear)</i>						
Cloudy	-0.270	0.010	0.763	---	---	---
<i>Lighting Condition (Base: Daylight)</i>						
Dusk	-1.826	0.031	0.161	---	---	---
<i>Driver Condition (Base: Apparently Normal)</i>						
Physical Impairment	1.114	0.068	3.047	---	---	---
<i>Vehicle Maneuver (Base: Entering/Leaving Road)</i>						
Turning Right	-0.976	<0.0001	0.377	-1.437	0.002	0.238
Turning Left	---	---	---	-0.631	0.091	0.532
<i>Turn Lanes (Base: None)</i>						
Right Turn Lanes	0.284	0.038	1.328	0.824	0.015	2.280
<i>Vehicle Type (Base: Passenger Car)</i>						
Motorecycle	1.396	0.004	4.039	---	---	---
SUV	---	---	---	-0.555	0.005	0.574
<i>Driver Age</i>						
Driver Age	0.006	<0.0001	1.006	---	---	---
<i>Goodness-of-Fit Statistics</i>						
AIC	4942.70			1007.30		
Null Deviance	5378.80			1079.37		
Residual Deviance	4816.7			905.31		
Pseudo R-Squared	0.105			0.164		
<i>Number of Observations</i>						
N	8117			1392		

5.3. Variable Importance Analysis Using Random Forest

The random forest technique was used in this study for ranking and screening of variables. The finally-screened variables were then compared with the significant variables from the fitted binary logit models (based on the P-value significance rather than the OR).

Using the R package, a random forest model was fitted at each of the four intersection types. It was found that the top variables in the ranking list were mostly matching with the significant variables identified from the four binary logit models. When comparing Figure 1 with the left portion of Table 3, it is observed that all the five variables from the random forest technique were found significant in the binary logit model for high-volume signalized intersections. When

comparing Figure 2 with the right portion of Table 3, it is observed that six out of seven variables from the random forest technique were found significant in the binary logit model for low-volume signalized intersections.

Similarly, when comparing Figure 3 with the left portion of Table 4, it is observed that eight out of ten variables from the random forest list were found significant in the high-volume stop-controlled intersections model. Finally, when comparing Figure 4 with the right portion of Table 4, it is seen that three out of four variables from the random forest list were found significant in the low-volume stop-controlled intersections model.

This evaluates and shows the importance of the resulted variables in each model in affecting the severity at the analyzed intersections. A random forest output for the four

models are shown in Figures 1, 2, 3, and 4. These figures show the top ten variables ranked by the Gini index criterion. Among these top ten variables, many variables were found in

the four binary logit models. The variables were the driver age, primary causes, vehicle type, lane turns, manner of crash, weather, vehicle maneuver, and lighting conditions.

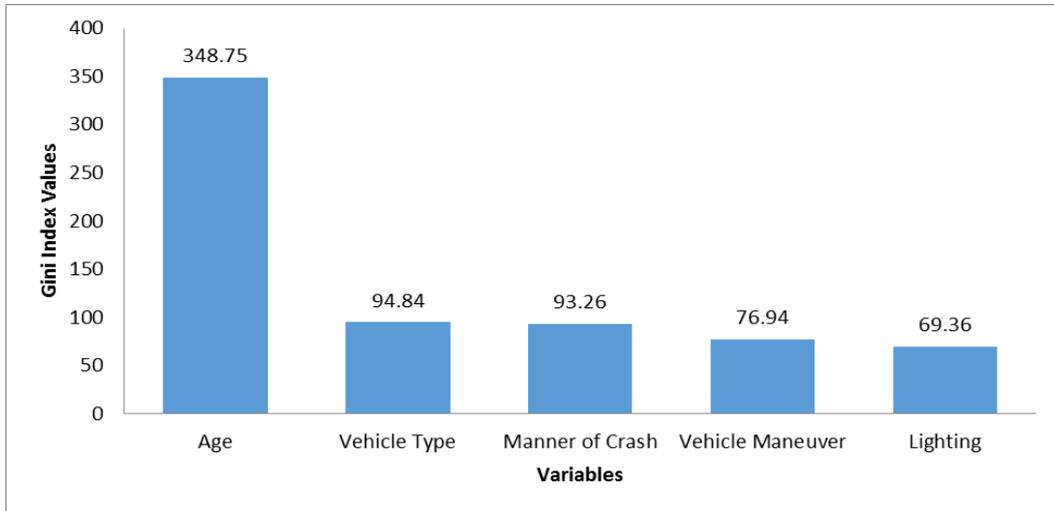


Figure 1. Variable Ranking Using Random Forest Technique at High-Volume Signalized Intersections

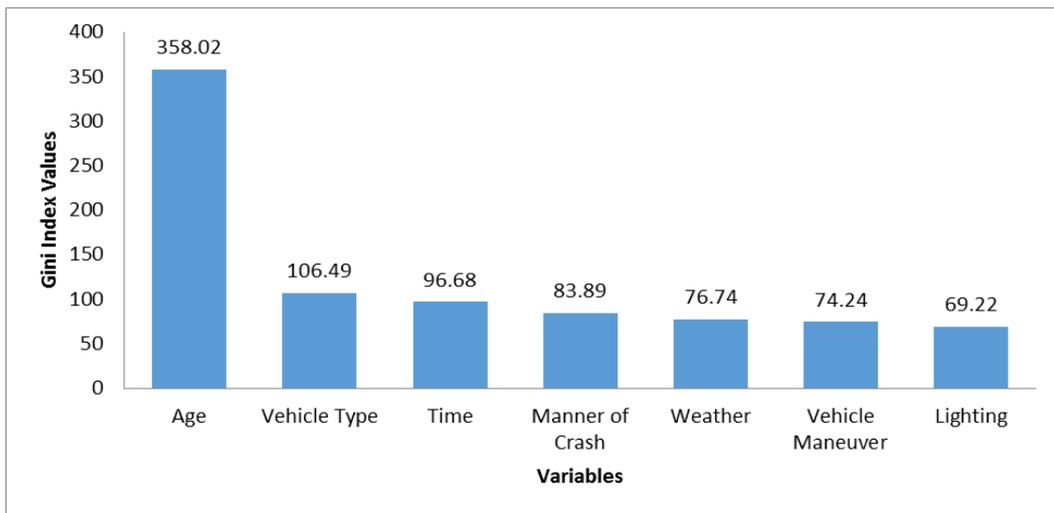


Figure 2. Variable Ranking Using Random Forest Technique at Low-Volume Signalized Intersections

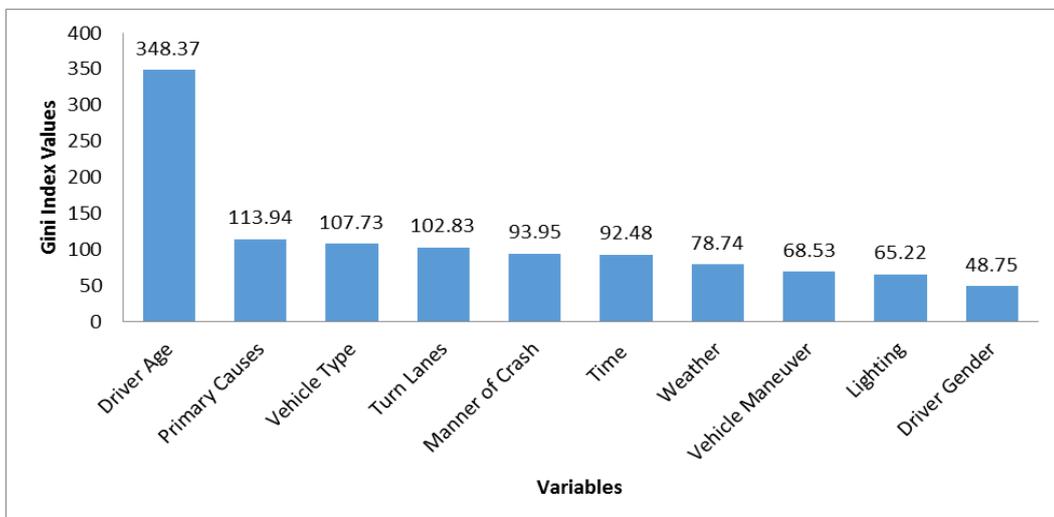


Figure 3. Variable Ranking Using Random Forest Technique at High-Volume Stop-Controlled Intersections

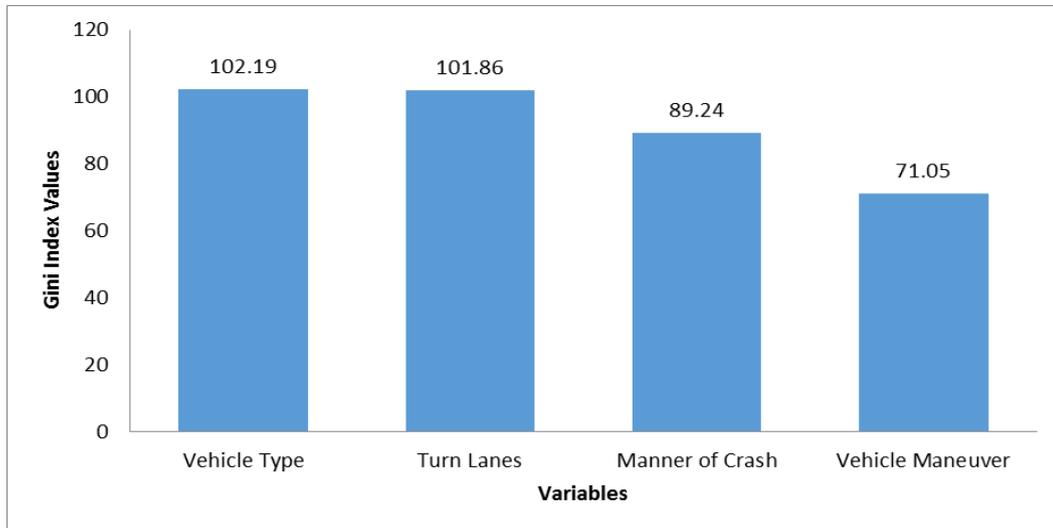


Figure 4. Variable Ranking Using Random Forest Technique at Low-Volume Stop-Controlled Intersections

5.4. Comparison of the Crash Severity Risk Factors at Intersections based on Traffic Volume

This section emphasizes and compares the common significant variables resulted from the binary logit models at signalized and stop-controlled intersections. The common significant variables are shown in Table 5.

There were nine common variables for at least two intersection types, as shown in Table 5. The common variables mentioned in the table were land use, manner of crash, weather condition, lighting condition, driver condition, vehicle maneuver, posted speed limit, and driver age. Rural areas in land use and turning right in vehicle maneuver were the only variables which showed significance in all the four models. Rural areas in all the four models showed higher likelihood of severity compared to their urban counterparts. The highest increase of severity was at low-volume stop-controlled intersections, followed by high-volume stop-controlled intersections, then low-volume signalized intersections, and finally high-volume signalized intersections. In general, drivers tend to drive with higher speeds in rural areas compared to urban areas. Regarding vehicle maneuver, turning right is associated with reduction of severity in all the four models, especially at stop-controlled intersections. Head-on collisions were associated with higher likelihood of severity at both types of signalized intersections. On the other hand, rear-end crashes were showed as reduction of crashes at both types of stop-controlled intersections and at high-volume signalized intersections. Single-vehicle crash type showed contrast results for high-volume signalized and high-volume stop-controlled intersections. This is because of the relatively lower speeds at stop-controlled intersections.

Cloudy weather conditions were associated with lesser severity likelihood at low-volume signalized and high-volume stop-controlled intersections. A higher reduction was found at low-volume signalized intersections (53.5% reduction) compared to high-volume signalized

intersections (23.7% reduction). The increase in traffic volume at these intersections was associated with a reduction in injury severity likelihood. Similarly, dusk condition showed lesser likelihood of severity at signalized and stop-controlled intersections with high traffic volumes. Motorcycles showed higher severity at signalized and high-volume stop-controlled intersections. The highest increase of severity was at high-volume signalized (494.7%), followed by high-volume stop-controlled (303.9%) and then low-volume signalized intersections (261.7%). As expected, all the significant variables in driver condition (DUI and asleep/illness and physical impairment) were associated with higher likelihood of crash severity.

The left-turning maneuver at high-volume signalized intersections (OR=1.665) showed higher severity likelihood, and surprisingly it showed lesser severity likelihood at low-volume stop-controlled intersections. This could be the reason that drivers might be more cautious when performing left turns at stop-controlled intersections. As expected, higher speed limit roadways (> 45 mph & ≤ 55 mph) at signalized intersections experienced higher severity. Higher speed limits at low-volume signalized intersections showed higher severity likelihood (89.1%) compared to high-volume signalized intersections (17.6%).

To show the impact of rural areas on crash severity at both signalized and stop-controlled intersections, Figure 5 is shown. This figure shows the percentage of severity (including both fatalities and severe injuries) at each intersection type. As shown, crashes at stop-controlled intersections in rural areas were associated with higher severity compared to signalized intersections in rural areas. Furthermore, low-volume intersections were associated with higher severity likelihood compared to high-volume intersections. These results concur with the OR interpretation in Table 5, where the OR for stop-controlled intersections were higher than that for signalized intersections.

Table 5. Common Significant Variables at the Different Intersection Types

Variable Description	Signalized Intersections				Stop-Controlled Intersections			
	High-Volume		Low-Volume		High-Volume		Low-Volume	
	Estimate	OR	Estimate	OR	Estimate	OR	Estimate	OR
<i>Land Use</i>								
Rural	0.305	1.356	0.653	1.920	0.798	2.222	1.259	3.521
<i>Manner of Crash</i>								
Head-on	0.568	1.765	1.702	5.486	---	---	---	---
Rear-end	-0.4522	0.637	---	---	-1.240	0.289	-2.566	0.076
Single-vehicle Crash	0.513	1.668	---	---	-0.526	0.591	---	---
<i>Weather Condition</i>								
Cloudy	---	---	-0.765	0.465	-0.270	0.763	---	---
<i>Lighting Condition</i>								
Dusk	-0.499	0.607	---	---	-1.826	0.161	---	---
<i>Vehicle Type</i>								
Motorcycle	1.783	5.947	1.286	3.617	1.396	4.039	---	---
<i>Driver Condition</i>								
DUI	1.343	3.833	1.574	4.827	---	---	---	---
Asleep/Illness/Emotional	1.053	2.866	1.218	3.382	---	---	---	---
Physical Impairment	1.469	4.346	---	---	1.114	3.047	---	---
<i>Vehicle Maneuver</i>								
Turning Right	-0.742	0.476	-2.209	0.109	-0.976	0.377	-1.437	0.238
Turning Left	0.509	1.665	---	---	---	---	-0.631	0.532
<i>Posted Speed Limit</i>								
> 45 mph & ≤ 55 mph	0.162	1.176	0.637	1.891	---	---	---	---
<i>Driver Age</i>								
Driver Age	0.005	1.005	0.009	1.009	0.006	1.006	---	---

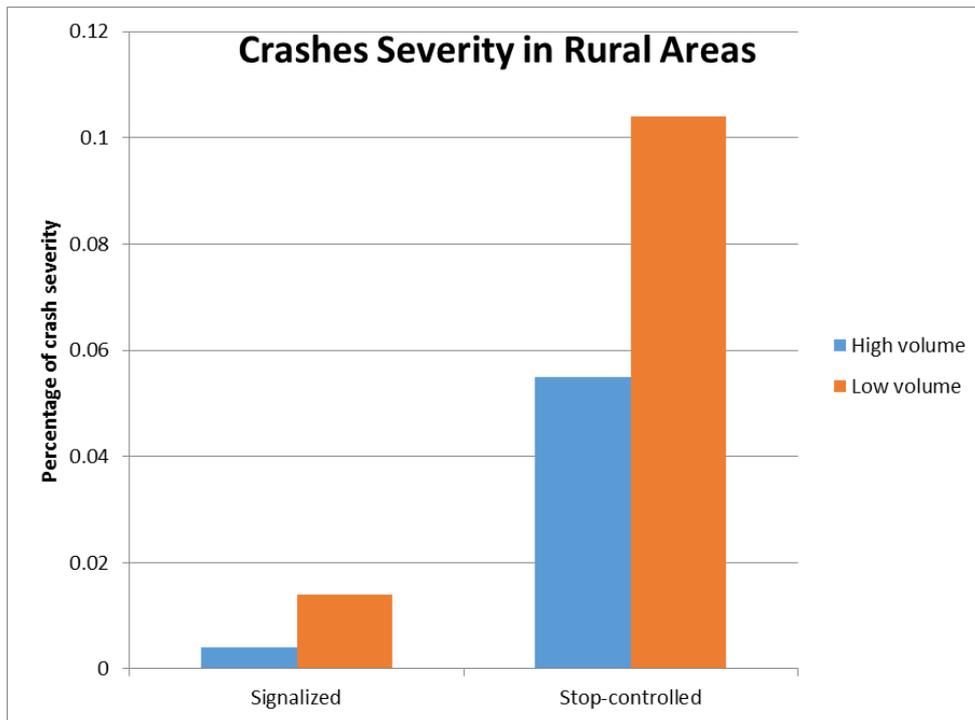


Figure 5. Severity of Rural Areas at High- and Low-Volume Signalized and Stop-Controlled Intersect

6. Conclusions and Study Applications

This study identified and compared the factors that affect crash severity at high- and low-volumes signalized and stop-controlled intersections. The study used a list of variables related to roadway geometric characteristics, traffic characteristics, vehicle characteristics, driver characteristics, and environmental conditions. Four separate binary logits were fitted at high-volume signalized, low-volume signalized, high-volume stop-controlled, and low-volume stop-controlled intersections. Several significant variables were found in the final models. Besides the binary logit model, this study used the random forest technique to rank the importance of the variables. It was concluded that the binary logit model and random forest technique had several significant variables in common.

In all the four models, rural areas experienced higher severity than urban areas. All the significant variables in driver conditions (DUI, asleep/illness, and physical impairment) showed higher crash severity likelihood at signalized and stop-controlled intersections. Rear-end crashes, cloudy weather condition, and dusk lighting condition showed lesser severity at signalized and stop-controlled intersections. Head-on crashes, single-vehicle crashes, turning left maneuver, and higher speed limits showed higher severity at signalized intersections. On the other hand, single-vehicle crash and turning left maneuver showed lesser severity at stop-controlled intersections. There were nine common variables for at least two intersection types. These common variables were land use, manner of crash, weather condition, lighting condition, driver condition, vehicle maneuver, posted speed limit, and driver age. It was concluded that cloudy weather was associated with lesser severity likelihood at low-volume signalized and high-volume stop-controlled intersections. However, a higher reduction was found at low-volume signalized intersections compared to high-volume stop-controlled intersections. Driver age at both types of signalized intersections and high-volume stop-controlled intersections was found to be associated with higher crash injury severity likelihood.

Based on the above-mentioned findings, some recommendations can be proposed to alleviate injury severity at intersections. Since the left-turning maneuver showed higher injury severity likelihood at high-volume signalized intersections, providing enough sight distance and protected left turn phase (with no permitted phase) in busy areas can be suggested. Since crashes in rural areas at signalized and stop-controlled intersections were associated with higher severity, installation of warning signs that discourage speeding is suggested and higher police patrolling activities can be suggested at those locations. In addition, enforcement and education countermeasures can be suggested. For example, enforcing non-speeding by law enforcement officers at signalized and stop-controlled intersections is essential since higher speed limits were associated with increased injury severity likelihood.

Education programs can be performed for drivers to educate them not to operate the vehicle if having unfavorable physical conditions since these conditions were associated with higher crash injury severity likelihood. Similarly, education programs should be designed disseminating the dangerous effect of DUI on crash severity, especially at low-volume signalized intersections. This can also imply more enforcement of DUI drivers at signalized intersections, particularly those with relatively low traffic volume.

Future research can consider the minor-approach traffic volume at both types of intersections when investigating the crash severity. In this study, the minor-approach AADT was unavailable since the minor approaches were mostly located on local (non-state) roads. Considering both major and minor traffic volumes for intersections might better explain the significance of some of the variables.

ACKNOWLEDGMENTS

The authors would like to thank the Alabama Department of Transportation (ALDOT) for providing the traffic volume data at the intersections used in this study.

REFERENCES

- [1] Abdel-Aty, M., Haleem, K. Analyzing Angle Crashes at Unsignalized Intersections Using Machine Learning Techniques. *Accid Anal Prev.* 2011; 43(1): 461–470.
- [2] Al-Ghamdi, S. Using Logistic Regression to Estimate the Influence of Accident Factors on Accident Severity. *Accid Anal Prev.* 2002; 36(3): 447–456.
- [3] Breiman, L. Random Forests. *Machine Learning* 2001; 45(1): 5–32.
- [4] Chang, H., Yeh, T. Risk Factors to Driver Fatalities in Single-Vehicle Crashes: Comparisons between Non-Motorcycle Drivers and Motorcyclists. *ASCE: Journal of Transportation Engineering* 2006; 132(3): 227–236.
- [5] Cornell Local Roads Program. Cornell University, Ithaca, NY. <http://www.clrp.cornell.edu/q-a/151-low-volume.html>. Accessed June 2017.
- [6] Devlin, A., Candappa, N., Corben, B., Logan, D. Designing Safer Roads to Accommodate Driver Error. *Curtin-Monash Accident Research Center* 2011; Project 09-006RSC.
- [7] Dissanayake, S., Lu, J. Analysis of Severity of Young Driver Crashes: Sequential Binary Logistic Regression Modeling. *Transportation Research Record* 2002; 1784: 108–114.
- [8] Duncan, C., Khattak, A., Council, F. Applying the Ordered Probit Model to Injury Severity in Truck-Passenger Car Rear-End Collisions. *Transportation Research Record* 2002; 1635: 63–71.
- [9] Haleem, K. Investigating Risk Factors of Traffic Casualties at Private Highway-Railroad Grade Crossings in the United States. *Accid Anal Prev.* 2016; 95: 274–283.

- [10] Haleem, K., Abdel-Aty, M. Examining Traffic Crash Injury Severity at Unsignalized Intersections. *Journal of Safety Research* 2010; 41(4): 347–357.
- [11] Haleem, K., Gan, A. Effect of Driver's Age and Side of Impact on Crash Severity along Urban Freeways: A Mixed Logit Approach. *Journal of Safety Research* 2013; 43: 67–76.
- [12] Ho, T. Random Decision Forests. Proceedings of the 3rd International Conference on Document Analysis and Recognition, Montreal, Quebec, Canada 1995; 278–282.
- [13] Harb, R., Yan, X., Radwan, E., Su, X. Exploring Precrash Maneuvers Using Classification Trees and Random Forests. *Accid Anal Prev.* 2009; 41(1): 98–107.
- [14] Huang, H., Chin, H., Haque, M. Severity of Driver Injury and Vehicle Damage in Traffic Crashes at Intersections: A Bayesian Hierarchical Analysis. *Accid Anal Prev.* 2008; 40(1): 45–54.
- [15] Islam, S., Mannering, F. Driver Aging and Its Effect on Male and Female Single Vehicle Accident Injuries: Some Additional Evidence. *Journal of Safety Research* 2006; 37(3): 267–276.
- [16] Kim, J., Ulfarsson, G., Kim, S., Shankar, V. Driver-Injury Severity in Single-Vehicle Crashes in California: A Mixed Logit Analysis of Heterogeneity due to Age and Gender. *Accid Anal Prev.* 2013; 50: 1073–1081.
- [17] Kockelman, K., Kweon, Y. Driver Injury Severity: An Application of Ordered Probit Models. *Accident Accid Anal Prev.* 2002; 34(4): 313–321.
- [18] Liu, C. The Effect of Unwarranted Traffic Signals. Joint Highway Research Project. Purdue University. 1979.
- [19] Logistic Regression. Carnegie Mellon University, Pittsburgh, PA. <http://www.stat.cmu.edu/~cshalizi/uADA/12/lectures/ch12.pdf>.
- [20] Nassar, S., Saccomanno, F., Shortreed, J. Road Accident Severity Analysis: A Micro Level Approach. *Canadian Journal of Civil Engineering* 1994; 21: 847–855.
- [21] Penmetsa, P., Pulugurtha, S.S., Modeling crash injury severity by road feature to improve safety. *Traffic Injury Prevention.* 2017; 19: 1,102-109.
- [22] Peshkin, D., Wolters, A., Alvarado, C., Moulthrop, J. Pavement Preservation for High Traffic Volume PCC Roadways: Phase I Findings from SHRP 2 Project R26. National Conference on Preservation, Repair, and Rehabilitation of Concrete Pavements 2009.
- [23] R Package. <http://www.r-project.org/>. Accessed June 2016.
- [24] Savolainen, P., Mannering, F. Probabilistic Models of Motorcyclists' Injury Severities in Single- and Multi-Vehicle Crashes. *Accid Anal Prev.* 2007; 39(5): 955–963.
- [25] Shankar, V., Mannering, F. An Exploratory Multinomial Logit Analysis of Single-Vehicle Motorcycle Accident Severity. *Journal of safety research* 1996; 27(3): 183–194.
- [26] Sparks, J. A Comparison of Data Mining Methods for Binary Response Variables in Direct Marketing. Ph.D. Dissertation, University of Illinois - Chicago, Chicago, IL.2009.
- [27] Wang, K., Qin, X. Exploring Driver Errors at Intersections: Exploring Key Contributors and Solutions. *Transportation Research Record* 2002; 2514: 1–9.
- [28] Wang, X., Abdel-Aty, M. Analysis of Left-Turn Crash Injury Severity by Conflicting Pattern Using Partial Proportional Odds Models. *Accid Anal Prev.* 2008; 40(5): 1674–1682.
- [29] Yasmin, S., Anwar, S., Tay, R. Factors Contributing to School Bus Crashes. *Journal of the Transportation Research Forum* 2013; 53(3): 63–78.
- [30] Zwerling, C., Peek-Asa, C., Whitte, P., Choi, S., Sprince, N., Jones, M. Fatal Motor Vehicle Crashes in Rural and Urban Areas Decomposing Rates into Contributing Factors. *Injury Prevention* 2005; 11(1): 24–28.