UDC: 656.13.084
725.39
DOI: http://dx.doi.org/10.7708/ijtte.2018.8(3).04

DETERMINANTS OF SEVERE INJURY AND FATAL TRAFFIC ACCIDENTS ON URBAN AND RURAL HIGHWAYS

Aschalew Kassu[1], Michael Anderson[2]

[1] *Department of Mechanical, Civil Engineering and Construction Management, Alabama A&M University, Normal, AL 35762, USA*

[2] *Department of Civil & Environmental Engineering, University of Alabama in Huntsville, Huntsville, AL 35899, USA*

Received 14 April 2018; accepted 15 May 2018

Abstract: Based on the recent Safety Fact Sheet, highway traffic accidents are becoming one of the leading causes of death. The main goal of this work is to study the correlation, impacts and the association of several highway pavements and geometric design elements, the prevailing traffic characteristics and environmental conditions on severe traffic crashes occurred over the five years period from 2010 to 2014 on selected two and four-lane rural and urban highways in the state of Alabama. Several state urban and rural highways with homogeneous variances, equal mean values and similar distributions of the crashes are identified and combined to form crash datasets. The significance of the initial categorical variables on the likelihood of generating severe crashes on dry and wet pavement surfaces are evaluated. The negative binomial regression model is used to estimate the observed crash data and identify the principal variables associated with the crashes.

Keywords: fatal crashes, severe crashes, severe injury, traffic accidents, negative binomial model, urban and rural highways, two and four-lane highways.

1. Introduction

Based on the recent report released by the National Safety Council (NSC), the total motor vehicle-related fatalities in the United States in 2014, 2015 and 2016 were 35,398, 37,757 and 40,200 respectively (NSC, 2017). The three years annual increase in the number of death due to traffic accidents from 2014 to 2016 were about 7 and 6 percent respectively. In 2006 alone, the estimated cost of traffic crash-related property damages, injuries, and fatalities were estimated to be $432.5 billion, which again was 12 percent higher than the 2015 estimate. The report also indicated that in the state of Alabama, the motor-vehicle related fatalities for the three years period from 2014 to 2016 were 821, 846, and 1,044 respectively, with 23 percent increase from the year 2015 to 2016 making Alabama one of the six states with a record of over 20 percent increase in motor-vehicle deaths from the year 2015 to 2016. The highest increase in fatalities from 2015 to 2016 was reported to be 34 percent in New Mexico, and the highest decrease in death was 23 percent in Wyoming. If we see the rate of fatalities across the ten regions of the United States categorized by the National Highway Traffic Safety Administration (NHTSA), from 2014 to 2015, an increase in fatalities was reported in nine out of the ten NHTSA

[1] Corresponding author: aschalew.kassu@aamu.edu

Regions. The only region whereby a decrease in fatalities was reported was Region 6 (NH, OK, TX, LA, and MS). The highest increase in fatalities (about 20 percent) was reported in Region 10 (WA, OR, ID and MT) followed by 14 percent increase in a fatality in Region 4 (AL, TN, GA, SC, FL) (NHTSA, 2016). The 2015 traffic fatalities in Alabama was reported to be 849, which was four percent higher than 2014. The fatalities rates per 100,000 population; 100,000 licensed drivers and 100 million VMT in Alabama and the national statistics for the United States were 17.47, 21.73, 1.26 and 10.92, 16.09, 1.13 respectively (NHTSA, 2017). In 2014, the national fatalities and injury per 100 million VMT were 1.08 and 7.7 respectively (NHTSA, 2014).

2. Literature Review

Several methodological approaches were adopted to study the effects of factors associated with crash severity on highways. Using negative binomial regression, (Hadi *et al.,* 1995) derived a mathematical model to estimate crashes on Florida roadways. The selected predictive variables were mainly highway cross-section design elements including lane, shoulder and median width, median type, speed limit, type of friction course (open or dense graded) and AADT. The impacts of weather and environmental factors such as rain and lighting were not included in the analysis. Injury, fatal and total crash rates were used as dependent variables. The crash data of these three dependent variables were divided into nine categories, and hence multiple models were derived for each dependent variable. The study concluded that: (1) for all highway sections, higher crash frequencies were associated with higher AADT values, (2) the pavement friction course was not significant to the occurrence of a crash, and (3) the two major predictors of fatal crashes in their study area were AADT and section length. Ordered probit models were used by Abdel-Aty (2003) to study drivers injury severity levels for three facility types including intersection, toll plazas and roadway sections for central Florida area over the two years period ranging from 1996-1997. The variables included were drivers' age and gender, speed, weather condition, type of vehicle (passenger car, van, light truck), seatbelt use, alcohol use, lighting condition, rural versus urban section, and peak period. The results on the roadway sections indicated alcohol use, weather condition, and peak period were not significant and female drivers and older drivers (over 65 years of age) had a higher probability of severe injuries. The study also found that, as compared with urban sections of the roadways, there was a higher probability of injury severity level during daylight in rural areas as compared with urban sections of the roadways. Likewise, (Krull *et al.,* 2000) reported the higher probability of rollover crashes in rural areas than urban sections of the highways. They applied a logistic regression model to examine the effects of rollover crashes on driver injury severity levels (fatal, incapacitating injury and non-injury) for Michigan and Illinois. The study concluded that in both of these states, injury severity level increased with alcohol use, failure to use a seatbelt, and with an increase in posted speed limits. The findings also suggested that higher probabilities of severe injuries were observed during daylight, dry pavement and rural section as opposed to a dark condition, wet pavement surface condition and urban sections of the roadway.

(Caliendo *et al.,* 2007) applied Poisson, negative binomial and negative multinomial regression techniques to model crash frequency on four-lane rural roads in Italy, using five years crash data collected from 1999-2003 on 46.6 km (29 miles) of highway segment both during dry and wet pavement conditions, with AADT values ranging from 17,600 to 47,400. The authors proposed separate models for curve and tangent sections of wet and dry pavement surface conditions of the highways. The terms included in the study were AADT, segment length, sight distance, curvature, side friction coefficient, longitudinal slope, and rainfall. The results identified the presence of junctions, segment length, and AADT were the main factors contributed to severe crashes. When pavement surface condition was included as a variable, their model showed that wet pavement condition was found to be a statistically significant variable, and the number of crashes occurred during wet pavement condition increased by a factor of 2.32 for tangent sections as compared with the crashes on dry pavement surfaces. This result was different from the findings of (Krull *et al.,* 2000), which suggested more crashes were likely to occur on dry pavement surface conditions.

The 2008 and 2009 traffic crash data on Texas highways, were used by (Li *et al.,* 2013), to analyze the impacts of pavement conditions rating including pavement distress, ride quality (smoothness), skid resistance, and International Roughness Index (IRI), on highway crash severity. Based on Texas Department of Transportation, the IRI scores between 1-95, 96-170 and 171-950 in/mi are considered very good, fair and poor respectively. The data analysis approach mainly used in the crash analysis was Pearson Chi-square of the crash severity outcomes and the predictor variables associated with pavement surface conditions followed by multiple comparison tests using Tukey-Kramer and Fisher's Least Significant Difference tests. The results of the study indicated that the impacts of pavement conditions were insignificant on freeways and were more significant on crash severity mainly when driving at high-speed limits on non-freeway multilane highways, during daylight on dry pavement surface conditions. (Oňa *et al.,* 2011), and (Wu *et al.,* 2014); however, reported the influence of lighting condition on severe crashes differently. On their work on the analysis of injury and fatal crashes on rural highways, (Oňa *et al.,* 2011) showed that driving with no street lighting was associated with severe injury and fatality as compared with driving during daylight. (Wu *et al.,* 2014) also reported that driving during dark roadway conditions increased multivehicle crash driver fatality by about 113 percent as compared with daylight condition.

Using ordered-probit models and a paired-comparison of the crash on similar highway segments, one with elevated speed limits with another stretch where the speed limits were kept the same, (Renski *et al.,* 1999) examined the effects of posted speed limit increases on single-vehicle crash severity on Interstate highways in North Carolina. The results of the study reported that increasing posted speed limits from 65 to 70 mph (~105 to 123 km/hr) on Interstate highways in North Carolina did not show any significant impact on crash severity. (Roh *et al.,* 2017) assessed the impacts of the percentage of heavy vehicles, flow rate, and variations in the average speed of traffic on four, six and eight-lane highways in Seoul, Korea over a period of two months. The results of the study indicated that the three highways considered show characteristics differences. For four-lane highways, as the flow rate increases and the percentage of trucks increases from zero to 35 percent, the average traffic speed reduction increases. Applying logistic regression models on three years crash data collected from 31 different highways in Canada, (Usman *et al.,* 2016) studied the factors contributing to injury severity crashes. Based on the report, use of alcohol increased the likelihood of major injury and fatality by about 0.8 percent.

3. Methodology

The five years severe injury and fatal traffic accident data ranging from 2010-2014 were extracted from Critical Analysis Reporting Environment (CARE) managed by the Center for Advanced Public Safety (CAPS) at the University of Alabama. The data include severe crashes occurred on two and four-lane urban and rural state highways on dry and wet pavement surfaces. Since fatal crashes were relatively few, severe injury and fatal crash records were combined to form severe crashes. The association of thirteen explanatory variables on the likelihood of occurrences of severe crashes on the selected highways was examined. The variables included in the analysis were, segment length, AADT, TADT, IRI, rut depth, cross-slope, grade, macrotexture, posted speed limit, a number of lanes (two or four), rural or urban designation of the segments, lighting, and weather. Among these, the variables that were not statistically significant at 5% level were excluded from the final models. The descriptive statistics of the crash data and some of the variables used are presented in Table 1. The severe crash rate (CR) expressed as crashes per 100 million vehicle- miles of travel for each segment is calculated using the corresponding AADT and vehicles-miles-traveled (VMT) as Eq. (1) and Eq. (2) (FHWA, 1990):

$$CR = \frac{C\ x\ 10^8}{AADT\ x\ 365\ x\ N\ x\ L} \quad (1)$$

and

$$VMT = AADT\ x\ 365\ x\ L \quad (2)$$

where, C = crash count for the segment; N = number of years of data; and L = length of roadway segment.

Several earlier studies confirmed that multiple linear regression (MLR) approaches for modeling traffic safety is unsuitable. The MLR models are formulated based on the assumption of normality and equality of variances. To determine the statistical approached to be adopted in this work, the normality of the severe crash data was tested. Both graphical and numerical techniques were used to check the normality of the severe crashes for individual routes and the combined dataset. For numerical analysis, Kolmogorov-Smirnov and Shapiro-Wilk tests were used. After testing the normality of the individual crash data recorded on each highway segments, multiple non-parametric statistical comparison tests including homogeneity of variances, equality of means, and similarity in the distributions of the crash data of the selected routes in Alabama Department of Transportation (ALDoT) highway systems were examined using Levene's test, Welch's test and Kruskal-Wallis (K-W) tests respectively. If the results of these test are not statistically significant ($\alpha > 0.05$), we have no evidence to reject the null hypothesis, which suggest that the mean (Welch's test), the variances (Levene's test) and distribution of crash rates (K-W test) are the same across the categories of the highways selected. However, if the tests result in p-values < 0.05, it implies that there is strong evidence to suggest that the crash rates for at least one of the highways are statistically significantly different from the others. To identify which highway or group of highways are different from the others, a grouping of the individual severe crash data and multiple comparisons were performed using Games-Howell *post hoc* test. The routes with homogeneous variances, equal means, and similar distributions were combined to form severe crash datasets occurred on dry and wet pavement surfaces. Next, the association between the crash dataset and the candidate explanatory variables were analyzed using negative binomial model, given by Eq. (3):

$$\ln(\mu_i) = \beta_0 + \sum_j x_{ij}\,\beta_j + \varepsilon, \qquad (3)$$

where μ, is the mean of the distribution (Wood, 2002; Lord and Mannering, 2010; Oh *et al.,* 2006).

To determine the significance of the categorical variables including the urban-rural designation of the segments, pavement surface condition (dry or wet), lighting (light or dark), number of lanes (two or four), on the probability of occurrence of severe crashes and decide whether these variables need to be included in the regression models or not, statistical pairwise comparison of the crash rates across the categories were performed. These include homogeneity of variances, equality of means and similarity in distribution of the severe crashes across the categories. If the crash rates across the categorical variables are found to have a significant difference from each other, the variables will be dummy coded and considered as potential explanatory variables. To identify the key predictors of the initial model, both step-wise and best subset regression models were performed using crash rates on dry, wet and the combined data set as independent variables. In this study, step-wise and best subset regression approaches were also used for preliminary analysis and evaluation of the candidate explanatory variables. The statistically significant predictors identified by the MLR approach were used as input variables for Poisson and negative binomial regression models (not shown). Poisson regression model assumes (Wood, 2002) that the expected number of crash rate (the mean) is equal to the variances (equi-dispersed). Test for over-dispersion performed on all the severe crash dataset across the various state highways (dry, wet, aggregate) was found to be statistically significant, indicating that the dataset was all over-dispersed. This indicates Poisson regression model is not suitable to fit the data. However, to compare improvements of the negative binomial model, as compared with Poisson models evaluate the goodness-of-fit tests Akaike's Information Criterion (AIC), Bayesian Information Criterion (BIC), Log Likelihood, and Deviance values are used.

Table 1
Descriptive Statistics of Severe Crashes and the Variables

Variable	Dry pavement crashes, N=2097				Wet pavement crashes, N=2478				Total severe, N=3344			
	Mean	St.Dv.	Min.	Max.	Mean	St.Dv.	Min.	Max.	Mean	St.Dv.	Min.	Max.
Crash rate	21.53	35.97	0	172	2.996	0.213	0	75	6.32	15.22	0	76
Crash count	0.89	1.66	0	19	0.13	0.01	0	5	0.36	1.01	0	12
Seg. Length	1.11	1.24	0.001	9.54	1.13	1.28	0.001	10.49	1.09	1.25	0.001	9.54
AADT	11975	11274	360	113460	11240	12447	360	113460	12058	12326	360	113460
TADT	11.48	6.84	1	40	10.95	6.52	1	43	11.41	6.64	1	40
Cross Slope			-7	7			-7	7			-7	7
Grade			-6.8	8			-7.1	10			-7.1	8
IRI	87.25	77.52	25	700	93.47	83.26	25	700	89.17	78.51	25	700
Rut depth	0.18	0.084	0	0.67	0.18	0.09	0	0.67	0.18	0.08	0	0.66
Macrotexture	0.54	0.501	0	11.05	0.55	0.54	0	11.05	0.55	0.56	0	11.05
Speed Limit	52.48	9.51	15	70	54.56	9.77	25	70	53.95	9.79	15	70

**N=Number of segments*

4. Results and Discussions

Among the sixteen state highways selected for the study, the majority of them appeared to have mean values which were not significantly different. Tests of homogeneity of variances and similarity in distributions of the severe crash data across most of the highways were not statistically significantly different as well. These were determined from the results of Welch's, Levene's and Kruskal-Wallis tests followed by Games-Howell post-hoc pair-wise comparison tests. The routes combined (shown in Tables 2-4) were all found to be statistically non-significant at a level of $\alpha=0.05$. Table 2 and 3 show the list of eleven and thirteen homogenous routes combined to form severe crash data set occurred on dry and wet pavement surfaces respectively. To combine the dry and wet pavement crashes, the common nine routes from each group were selected and dummy coded to categorize the dry and wet pavement severe crashes forming a single crash data for the aggregate dataset (Table 4). The dummy coded variables are used as the categorical variables in the analysis. Based on both tests of normality (Shapiro-Wilk and Kolmogorov-Smirnov), all the crash rates at the individual routes and the aggregate data were not normally distributed at a significance level of $\alpha = .001$. Similarly, tests of normality and the descriptive statistics of the data across the categorical variables were performed for severe crashes occurred on dry, wet pavement surfaces and aggregate data sets and presented in Tables 5, 6 and 7 respectively.

Table 2

Tests of Normality and Descriptive Statistics of the Severe Crash Rates occurred on Dry Pavement Surfaces for the Selected Eleven State Highways

Route ID	ǂN	Mean	Sd. Dv.	Normality Test		
				Skewness	Kurtosis	*Sig.
AL 001	385	27.01	40.07	1.611	1.898	0.001
AL 002	158	21.08	35.32	2.161	4.676	0.001
AL 003	387	18.93	33.94	2.289	5.480	0.001
AL 005	153	18.90	36.21	2.192	4.310	0.001
AL 006	164	18.08	29.65	1.938	4.024	0.001
AL 008	130	19.68	31.69	1.677	2.065	0.001
AL 010	127	19.34	36.23	2.035	3.763	0.001
AL 012	171	25.26	38.92	1.921	3.439	0.001
AL 013	191	20.08	30.91	2.068	4.814	0.001
AL 014	125	22.70	39.01	1.789	2.192	0.001
AL 015	106	20.98	40.02	1.812	2.114	0.001
¥Combined	2097	21.53	35.97	1.958	3.480	0.001

**Significance level for both Shapiro-Wilk and Kolmogorov-Smirnov tests.*
¥All routes combined
Number of segments

Table 3

Tests of Normality and Descriptive Statistics of the Severe Crash Rates occurred on Wet Pavement Surfaces for the Selected Thirteen State Highways

Route ID	N	Mean	Std. Dev.	Normality Test		
				Skewness	Kurtosis	Sig.
AL 001	423	3.56	10.94	3.955	17.049	0.001
AL 003	363	1.94	8.40	5.426	32.945	0.001
AL 004	98	2.09	10.20	5.570	32.867	0.001
AL 005	153	3.41	10.74	3.494	11.893	0.001
AL 006	178	3.06	10.16	3.984	16.916	0.001
AL 007	176	4.27	13.04	3.217	10.025	0.001
AL 008	141	3.16	11.05	4.660	24.564	0.001
AL 009	205	2.38	10.55	5.073	26.826	0.001
AL 010	130	3.26	11.84	3.975	16.067	0.001
AL 013	188	3.09	8.49	3.096	9.640	0.001
AL 014	125	1.85	7.73	5.266	31.660	0.001
AL 015	104	2.79	11.47	4.015	14.868	0.001
AL 017	194	3.78	13.07	3.746	13.685	0.001
Combined	2478	3.00	10.60	4.213	18.777	0.001

Table 4

Tests of Normality and Descriptive Statistics of the Aggregate Severe Crash Rates occurred on Both Dry and Wet Pavement Surfaces for the Selected Nine State Highways

Route ID	N	Mean	St. Dv.	Normality Test		
				Skewness	Kurtosis	Sig.
AL 001	739	7.07	15.71	2.468	5.479	0.001
AL 003	714	6.17	15.12	2.641	6.368	0.001
AL 005	282	4.61	12.63	3.085	9.230	0.001
AL 006	327	6.97	15.94	2.410	5.087	0.001
AL 008	258	7.14	16.69	2.550	5.827	0.001
AL 010	240	5.54	15.79	2.934	7.619	0.001
AL 013	362	8.02	15.72	2.048	3.450	0.001
AL 014	231	4.84	13.08	3.139	10.244	0.001
AL 015	191	3.87	13.97	3.587	11.655	0.001
Combined	3344	6.32	15.22	2.616	6.213	0.001

Table 5

Tests of Normality and Descriptive Statistics of the Severe Crash Rates occurred on Dry Pavement Surfaces across the Categorical Variables for the Selected Eleven State Highways

Variable		N	Mean	St. Dv.	Normality Test		
					Skewness	Kurtosis	Sig.
Rural/Urban	Rural	1112	22.18	36.65	1.906	3.175	0.001
	Urban	985	20.79	35.19	2.021	3.880	0.001
No. of Lanes	Two	1090	21.82	38.59	1.892	2.861	0.001
	Four	1007	21.21	32.91	2.016	4.264	0.001
Lighting	Dark	732	20.75	35.51	1.975	3.586	0.001
	Light	1365	21.95	36.22	1.951	3.439	0.001
Weather	Cloudy	422	19.78	34.97	2.197	4.661	0.001
	Clear	1675	21.97	36.21	1.904	3.239	0.001

Table 6

Tests of Normality and Descriptive Statistics of the Severe Crash Rates occurred on Wet Pavement Surfaces across the Categorical Variables for the Selected Thirteen State Highways

Variable		N	Mean	St. Dv.	Normality Test		
					Skewness	Kurtosis	Sig.
Rural/Urban	Rural	1513	3.10	10.79	4.078	17.430	0.001
	Urban	965	2.83	10.29	4.453	21.319	0.001
No. of Lanes	Two	1307	3.16	11.38	3.987	15.992	0.001
	Four	1171	2.81	9.66	4.500	23.037	0.001
Lighting	Dark	884	3.30	10.62	3.795	15.219	0.001
	Light	1594	2.83	10.59	4.455	20.901	0.001
Weather	Rainfall	1872	2.99	10.57	4.214	18.853	0.001
	Normal	606	3.02	10.69	4.220	18.709	0.001

Table 7

Tests of Normality and Descriptive Statistics of the Aggregate Severe Crash Rates occurred on Both Dry and Wet Pavement Surfaces across the Categorical Variables for the Selected Nine State Highways

Variables		N	Mean	St. Dv.	Normality Test		
					Skewness	Kurtosis	Sig.
Surface Cond.	Dry	1600	11.79	20.02	1.564	1.230	0.001
	Wet	1744	1.31	4.88	4.104	17.102	0.001
Rural/Urban	Rural	1911	6.08	15.06	2.710	6.755	0.001
	Urban	1433	6.65	15.43	2.501	5.580	0.001
No. of Lanes	Two	1679	5.64	15.23	2.814	7.059	0.001
	Four	1665	7.01	15.18	2.433	5.496	0.001
Lighting	Dark	1163	6.34	15.50	2.661	6.416	0.001
	Light	2181	6.31	15.07	2.590	6.098	0.001
Weather	Rainfall	1370	1.49	5.67	4.837	30.137	0.001
	Normal	1974	9.67	18.51	1.881	2.442	0.001

To determine whether the severe crash rates across the categorical variables were significantly different or not, tests of equality of means, homogeneity of variances, and similarity in distributions across the categories were performed (Table 8-10). For the crash rates which were not statistically significantly different with respect to Welch's, Levene's and Kruskal-Wallis tests, the results suggested that the particular categorical variable did not have any influence on the likelihood of the occurrences of severe crashes on these highways. For those significantly different, the results suggested that the variables need to be retained, dummy coded and used as categorical variables in the statistical models developed to fit the severe crash data. From Table 8, we can see that except for the distribution of the crash across the number of lanes (p-value=.001), the mean, variances, and distributions of severe crashes on dry pavement surfaces across the categories of rural-urban designation of the segment, number of lanes, lighting and the weather conditions were not significantly different.

For wet-weather related severe crashes (Table 9), the distributions of the data across the number of lanes and lighting conditions were found to be statistically significant (similar in distribution) at a level of p-value=0.049 and p-value=0.037 respectively. However, the means and variances of the severe crashes across both categories were not found to be significantly different. These were also confirmed by Games-Howell post-hoc tests at p-value=0.417 and 0.288 for the number of lanes and lighting conditions. Across urban-rural designations of the segments, as well as the weather conditions (rainfall versus normal weather), all the tests confirmed that there were no significant differences in means, variances, and distributions of the severe crash data across these categories. The results in Table 8 and 9 suggest that the use of these four categorical variables in formulating a mathematical relationship between the severe crashes occurred on dry and wet pavements will turn out to be statistically insignificant. For the sake of completeness, all the variables listed in the methodology section of this paper were used in formulating negative binomial models.

The results of a similar analysis performed on the aggregate data consisting of the severe crashes occurred both on dry and wet pavement surfaces is shown in Table 10. Here the means, variances and the distributions of the severe crashes across the categories of the pavement surfaces (dry versus wet), number of lanes and weather conditions (rainfall versus normal weather) were all different at a significance level of 0.05. These warrants dummy coding and inclusion of these categorical variables in the statistical model. Again for the sake of completeness, and to further confirm the level of significances, all the variables listed in the methodology section were used in formulating negative binomial models for the aggregate data. Interestingly, for the aggregate dataset, the number of lanes and pavement surface condition were found to be key variables associated with the likelihood of occurrences of severe crashes at a significance level of 0.0001 (Table 11).

Table 8

Test for Equality of Mean, Homogeneity of Variances, And Similarity in Distribution of the Dry Pavement Severe Crashes across the Categorical Variables

Variables		Levene's Test	Welch's Test	Kruskal-Wallis Test	Games-Howell Test
Rural/Urban	Test Stat.	0.780	0.790	0.340	0.890
	P-Value	0.376	0.375	0.559	0.375
No. of Lanes	Test Stat.	0.150	0.160	12.55	0.400
	P-Value	0.694	0.692	0.001	0.692
Lighting	Test Stat.	0.530	0.540	1.120	0.730
	P-Value	0.466	0.464	0.291	0.464
Weather	Test Stat.	1.240	1.290	1.100	1.140
	P-Value	0.266	0.256	0.294	0.256

Table 9

Test for Equality of Mean, Homogeneity of Variances, and Similarity in Distribution of the Wet-Weather Severe Crashes across the Categorical Variables

Variables		Levene's Test	Welch's Test	Kruskal-Wallis Test	Games-Howell Test
Rural/Urban	Test Stat.	0.370	0.380	0.040	0.620
	P-Value	0.543	0.538	0.847	0.538
No. of Lanes	Test Stat.	0.650	0.660	3.860	0.810
	P-Value	0.421	0.417	0.049	0.417
Lighting	Test Stat.	1.130	1.130	4.340	1.060
	P-Value	0.287	0.288	0.037	0.288
Weather	Test Stat.	0.000	0.000	0.010	0.070
	P-Value	0.945	0.945	0.926	0.945

Table 10

Test for Equality of Mean, Homogeneity of Variances, and Similarity in Distribution of the Full Crash Dataset Containing Both Severe Crashes observed on Dry and Wet Road Surfaces across the Categorical Variables

Variables		Levene's Test	Welch's Test	Kruskal-Wallis Test	Games-Howell Test
Surface Condition (Dry/Wet)	Test Stat.	448.710	415.650	348.030	20.390
	P-Value	0.001	0.001	0.001	0.001
Rural/Urban	Test Stat.	1.170	1.170	2.390	1.080
	P-Value	0.279	0.280	0.122	0.280
No. of Lanes (2 lane/4 Lane)	Test Stat.	6.810	6.810	36.560	2.610
	P-Value	0.009	0.009	0.001	0.009
Lighting (Dark/Light)	Test Stat.	0.000	0.000	0.050	0.040
	P-Value	0.964	0.965	0.820	0.965
Weather (Rain/ No Rain)	Test Stat.	251.140	339.560	206.460	18.430
	P-Value	0.001	0.001	0.001	0.001

Table 11

Parameter Estimates for Severe Crashes observed Dry, Wet Pavement and the Full Dataset Using the Statistically Significant Variables

Parameter	Dry Pavement			Wet Pavement			Total		
	Estimate	Sig.	IRR	Estimate	Sig.	IRR	Estimate	Sig.	IRR
Constant	3.188	0.001	24.24	-1.238	0.001	0.29	-0.016	0.954	0.984
Seg. Length	0.491	0.001	1.634	1.008	0.001	2.739	1.317	0.001	3.731
AADT	0.000015	0.06	1.000	0.000074	0.001	1.000	0.000094	0.001	1
TADT	-0.042	0.001	0.959				-0.037	0.004	0.963
No of Lanes									
4 Lane=1	_	_	_	_	_	_	0.538	0.001	1.712
2 Lane=0									
Lighting									
Light=1	_	_	_	_	_	_	-0.282	0.05	0.755
Dark =0									
Surface Cond.									
Wet=1	_	_	_	_	_	_	-2.858	0.001	0.057
Dry=0									
Disp. Coeff.	8.88			38.42			14.23		
Goodness-of-Fit Statistics									
AIC	11970			4167			9531		
BIC	12004			4190			9580		
LL	-5980			-2080			-4758		
Deviance	1609			589			1440		

The final mathematical model developed to fit the observed severe crashes on dry, wet and aggregate datasets on the homogenous state highways is shown in Table 11. The negative binomial model developed for the three datasets used the factors identified at a statistical significance level of 0.05. Segment length and AADT (except for dry pavement condition) were found reliably associated with severe crashes. For dry pavement conditions and the aggregate datasets, the higher the percentage of heavy trucks along the segments is the less likely the occurrences of severe crashes. A unit increase in the percentage of trucks in the segments will likely reduce the severe crashes on dry pavements (B=-0.042, IRR=0.959) and the aggregate data (B=-0.037, IRR 0.963) by 4.1 and 3.7 percent respectively, which corresponds with the results of (Usman *et al.,* 2016). They observed that the higher the traffic volume, the lower the likelihood of occurrences of severe injury and fatal crashes. The reasons suggested were the likelihood of congestion due to increase in traffic volume, which in turn decreases operating speed. (Roh *et al.,* 2017) also reported that for four-lane highways, as the percentage of truck increases on the segments, there was a decrease in the average speed of the traffic. As a result of the operating speed reduction posted speed limit was not found to be an influencing factor leading to severe injury and fatality on dry pavement, wet-weather and aggregate dataset. Similarly, regardless of the pavement surface condition (dry or wet), weather condition (rainfall versus normal), and highway design elements including rut depth, macrotextures, IRI, grade and pavement cross-slope were not found to be key determinants (Ihs *et al.,* 2011; Gunaratne *et al.,* 2012; Li *et al.,* 2013) of severe injury and fatal crashes on two and four-lane urban and rural highways selected.

For the aggregate severe crash dataset, lighting and pavement surface conditions were also the principal determinants of severe injury and fatal crashes. The results indicate that the likelihood of severe injury and fatality on the selected two and four-lane rural and urban highways during day and street lighting condition (B=-0.282, IRR=0.755) is 24.5 percent lower as compared with driving during dark roadway conditions. This result corresponds with the results of (Edwards, 1998; Oňa *et al.,* 2011; Wu *et al.,* 2014). As compared with daylight condition, a 113 percent increase in multivehicle crash driver fatality was reported during dark roadway conditions (Wu *et al.,* 2014). Pavement surface condition was also a determinant factor for the occurrences of severe injury and fatality on the selected highways. The results of this study suggested that the likelihood of occurrences of severe injury and fatal crashes on wet pavement surface (B=-2.858, IRR=0.057) on the selected highways was 94.3 percent less than the crashes observed on dry pavement surfaces. This was possibly attributed to the short duration of wet-weather condition per year, reduction in operating speed and traffic volume due to rainfall and bad weather conditions (Keay and Simmonds, 2005; Gunaratne *et al.,* 2012) and, drivers taking extra preventive actions to avoid accidents due to rainfall and slippery road surface conditions (Edwards, 1998; Nassar *et al.,* 1994; Jung *et al.,* 2011). Using the principal factors identified in Table 11, for the aggregate severe injury and fatal crashes per 100 million VMT (CR) can be represented by the following negative binomial regression, Eq. (4):

$$\begin{aligned} CR = \exp(&-0.016 + 1.317\ SL + 9.5\times10^{-5}\ AADT - 0.037\ TADT + \\ &+ 0.0\ TLn + 0.538\ FLn + 0.0\ DK - 0.282\ LT + 0.0\ DRY - 2.858\ WET) \end{aligned} \tag{4}$$

where, SL=segment length, TLn=Two Lane, FLn=Four Lane, DK= Dark, LT = Light.

The negative binomial regression models developed for the severe crashes occurred on dry and wet pavements individually, and combined were evaluated by using goodness-of-fit tests including, Akaike's Information Criterion (AIC), Bayesian Information Criterion (BIC), Deviance, and Log Likelihood. To evaluate the relative improvements of the negative binomial models (NBM), the AIC and the deviance values were compared with Poisson regression model (PRG) models formulated with the same principal variables. For the aggregate severe crash data observed on both dry and wet pavement surfaces the AIC and deviance values of PRM (not shown), were 55,268 and 51,892 respectively. From the results of the PRM, it is observed that the AIC value is equivalent to the adjusted R-square value of about 32 percent. The corresponding AIC and deviance values of the NBM shown in Table 11 were 9,531 and 1440 respectively. This substantial reduction in both AIC and deviance values indicate that the NBM developed is powerful model than the PRG (Fridstrøm *et al.,* 1995; Abdel-Aty and Radwan, 2000) in approximating the statistical fit to the severe injury and fatal crashes observed in the selected two and four-lane urban and rural highways in Alabama.

5. Conclusions

This study evaluated the significance of thirteen factors related to pavement and geometric design elements of two and four-lane urban and rural highways, environmental factors and traffic characteristics on severe injury and fatal crashes reported for the five years period ranging from 2010 to 2014 on the selected state routes in Alabama. Using these factors, negative binomial regression models were estimated for severe injury and fatal crashes observed on dry pavement, wet-weather condition and for the aggregate dataset. The results indicate that the principal factors influencing severe crashes for the aggregate crashes on dry and wet pavement surfaces were segment length, AADT, TADT, number of lanes, lighting and pavement surface condition. Interestingly, irrespective of the pavement surface and the weather conditions (rainfall versus normal), traffic characteristics of the segments appear to influence the likelihood of severe injury and fatal crashes as compared with the highway design elements. Overall, the roadway pavement and geometric features, such as rut depth, macrotextures, IRI, grade and pavement cross-slopes of the study segments were insignificant factors in determining the likelihood of severe crashes. This result agrees well with prior studies by (Ihs *et al.,* 2011; Gunaratne *et al.,* 2012; and Li *et al.,* 2013). As shown in Tables 8-10, the severe crashes observed in rural sections of the segments for the three different datasets were not significantly different from urban sections. This indicates there was no clear association between the urban-rural designations of the highways and the likelihood of occurrences of severe crashes. The results of the negative binomial regression model also confirmed that urban-rural settings of the segments were not significantly associated with the severe crashes observed. As can be seen in Table 11, depending on the pavement surface condition where the observed severe crashes were observed, the contributing factors, the model variables, the estimates and the goodness-of-fit test statistics values were different.

Acknowledgements

The authors would like to thank the Center for Advanced Public Safety (CAPS) at the University of Alabama for providing access to the CARE data, and the Alabama Department of Transportation (ALDOT) for providing the data on traffic volume, geometric and pavement conditions of the highways.

References

Abdel-Aty, M.; Radwan, A.E. 2000. Modeling traffic accident occurrence and involvement, *Accident Analysis and Prevention* 32(5): 633-642.

Abdel-Aty, M. 2003. Analysis of driver injury severity levels at multiple locations using ordered probit models, *Journal of Safety Research* 34(5): 597-603.

Caliendo, C.; Guida, M.; Parisi, A. 2007. A crash-prediction model for multilane roads, *Accident Analysis and Prevention* 39(4): 657-670.

Edwards, J.B. 1998. The relationship between road accident severity and recorded weather, *Journal of Safety Research* 29(4): 249-262.

Fridstrøm, L.; Ifver, J.; Ingebrigtsen, S.; Kulmala, R.; Thomsen, L. K. 1995. Measuring the contribution of randomness, exposure, weather, and daylight to the variation in road accident counts, *Accident Analysis & Prevention* 27(1): 1-20.

Gunaratne, M.; Lu, Q.; Yang, J.; Metz, J.; Jayasooriya, W.; Yassin, M.; Amarasiri, S. 2012. *Hydroplaning on multi lane facilities. Final Report.* Florida Department of Transportation. 130 p.

Hadi, M.A.; Aruldhas, J.; Chow, L.; Wattleworth, J.A. 1995. Estimating safety effects of cross-section design for various highway types using negative binomial regression, *Transportation Research Record. Journal of the Transportation Research Board* 1500: 169-177.

FHWA. 1990. *Highway Performance Monitoring System, Field Manual*. U.S. Department of Transportation, FHWA Publication 5600.1A, Washington, D.C.

Ihs, A.; Gustafsson, M.; Eriksson, O.; Wiklund, M.; Sjögren, L. 2011. *Road user effect models-the influence of rut depth on traffic safety*. The Finnish Transport Agency, VTI rapport 731A.

Jung, S.; Qin, X.; Noyce, A.D. 2011. Modeling highway safety and simulation in rainy weather, *Transportation Research Record. Journal of the Transportation Research Board* 2237: 134-143.

Keay, K.; Simmonds, I. 2005. The association of rainfall and other weather variables with road traffic volume in Melbourne, Australia, *Accident Analysis and Prevention* 37(1): 109-124.

Krull, K.; Khattak, A.; Council, F. 2000. Injury effects of rollovers and events sequence in single-vehicle crashes, *Transportation Research Record. Journal of the Transportation Research Board* 1717: 46-54.

Li, Y.; Liu, C.; Ding, L. 2013. Impact of pavement conditions on crash severity, *Accident Analysis and Prevention* 59: 399-406.

Lord, D.; Mannering, F. 2010. The Statistical Analysis of Crash Frequency Data: A Review and Assessment of Methodological Alternatives, *Transportation Research Part A* 44(5): 291-305.

Nassar, S.; Saccomanno, F.F.; Shortreed, J.H. 1994. Road accident severity analysis: A micro-level approach, *Canadian Journal of Civil Engineering* 21(5): 847-855.

NHTSA. 2016. *National Center for Statistics and Analysis. Early estimate of motor vehicle traffic fatalities for 2015. (Brief Statistical Summary. Report No. DOT HS 812 269)*. Washington, DC: National Highway Traffic Safety Administration (NHTSA).

NHTSA. 2017. *National Center for Statistics and Analysis. State Traffic Data: 2015 Data (Traffic Safety Facts. Report No. DOT HS 812 412)*. Washington, DC: National Highway Traffic Safety Administration (NHTSA).

NHTSA. 2014. *National Center for Statistics and Analysis. A compilation of motor vehicle crash data from the fatality analysis reporting system and the general estimates system. (Traffic Safety Facts. Report No. DOT HS 812 261)*. Washington, DC: National Highway Traffic Safety Administration (NHTSA).

NSC. 2017. NSC Motor Vehicle Fatality Estimate. National Safety Council (NSC). Available from internet: <http://www.nsc.org/NewsDocuments/2017/12-month-estimates.pdf>.

Oh, J.; Washington, S.; Nam, D. 2006. Accident prediction model for railway-highway interfaces, *Accident Analysis and Prevention* 38(2): 346-356.

Oña, J.; Mujalli, R.O.; Calvo, F.J. 2011. Analysis of traffic accident injury severity on Spanish rural highways using Bayesian networks, *Accident Analysis and Prevention* 43(1): 402-411.

Renski, H.; Khattak, A. J.; Council, F.M. 1999. Effect of speed limit increases on crash injury severity: analysis of single-vehicle crashes on North Carolina interstate highways, *Transportation Research Record. Journal of the Transportation Research Board* 1665: 100-108.

Roh, C.G.; Park, B.J.; Kim, J. 2017. Impact of heavy vehicles on highway traffic flows: case study in the Seoul metropolitan area, *Journal Transportation Engineering, Part A: Systems* 143(9): 05017008.

Usman, T.; Fu, L.; Miranda-Moreno, L.F. 2016. Injury severity analysis: comparison of multilevel logistic regression models and effects of collision data aggregation, *Journal of Modern Transportation* 24(1): 73-87.

Wood, G.R. 2002. Generalized linear accident models and goodness of fit testing, *Accident Analysis and Prevention* 34(4): 417-427.

Wu, Q.; Chen, F.; Zhang, G.; Liu, X.C.; Wang, H.; Bogus, S.M. 2014. Mixed logit model-based driver injury severity investigations in single- and multi-vehicle crashes on rural two-lane highways, *Accident Analysis and Prevention* 72: 105-115.