Effects of Pavement Surface Conditions on Traffic Crash Severity

Jaeyoung Lee, M.ASCE[1]; BooHyun Nam, A.M.ASCE[2]; and Mohamed Abdel-Aty, M.ASCE[3]

Abstract: Improving road safety through proper pavement maintenance is one of the goals of pavement management. Many studies have found that pavement conditions significantly influence traffic safety. Although several studies have explored the relationship between pavement conditions and crash occurrence, the effect of poor pavement conditions on crash severity levels has not been investigated, especially by using a discrete model that can handle ordered data. This paper focuses on the development of the relationship between poor pavement conditions and crash severity levels using a series of Bayesian ordered logistic models for low/medium/high speed roads and single/multiple collision cases. The Bayesian ordered logistic regression models indicated that the poor pavement condition decreases the severity of single-vehicle collisions on low-speed roads whereas it increases their severity on high-speed roads. On the other hand, the poor pavement condition increases the severity of multiple-vehicle crashes on all roads. Findings of this study can assist transportation agencies at the federal, state, and local levels to select appropriate pavement maintenance and rehabilitation strategies to reduce traffic crash severity levels. **DOI: 10.1061/(ASCE)TE.1943-5436.0000785.** © *2015 American Society of Civil Engineers.*

Author keywords: Pavement management; Traffic safety; Bayesian modeling; Ordered logistic model; Severity analysis; GIS analysis; Random forest; Variable importance.

Introduction

With consistent efforts of transportation engineers, federal, state, and local government officials, both fatalities and fatality rates from road traffic crashes in the United States have steadily declined from 2006 to 2011. Nevertheless, fatalities from traffic crashes slightly increased in 2012 [National Highway Safety Traffic Administration (NHTSA) 2013]. About 33,500 lives were lost from road traffic crashes in the year 2012, and road traffic crashes are still one of the leading causes of deaths, according to the Centers for Disease Control and Prevention (Hoyert and Xu 2012). According to NHTSA (2008), human factors contribute to 93% of traffic crashes, whereas 16 and 7% of traffic crashes are due to roadway environmental factors and vehicle factors, respectively. From this perspective, pavement condition is considered as an important factor for traffic safety because the pavement condition itself is one of the road environmental factors and it also affects human and vehicle factors simultaneously. Thus, improving traffic safety through well-maintained pavements is one of the key objectives of pavement management systems. Currently, traffic safety is considered as a separate area, but it should be incorporated into the pavement management system. When pavements are evaluated along with traffic safety (e.g., crash count and severity level), many factors related to pavement engineering properties are to be considered; for example, pavement condition, materials and mix design, geometric design, road color or visibility. Previous studies have discussed limited analyses of traffic safety associated with pavement condition, focusing on individual pavement properties and utilizing simple trend analyses or simple statistical models. Tighe and Haas (1998) addressed that partially paved shoulders can cause significant safety and economic benefits. Zegeer and Council (1995) studied road safety mainly with respect to pavement geometric elements (e.g., width of lane and shoulder, shoulder type, and design of median) and concluded that lane widening can decrease traffic crashes by 40%. Karlaftis and Golias (2002) explored the effect of roadway geometric characteristics on rural roadway crash rates using hierarchical tree-based regression models. They found that both geometric design and pavement condition variables are the two most key factors affecting traffic crash rates. Henry (1996), Wambold (1988), and Heaton et al. (1990) have conducted active research on pavement friction and surface texture measurement, and developed the international friction index (IFI). Karan et al. (1976) investigated the relationships between roughness and vehicle speed. They revealed that the average speed and riding comfort index has a positive and high correlation ($r = 0.77$). This finding implies that drivers increase their speeds as pavement condition is better. Al-Masaeid (1997) reported that pavement roughness, international roughness index (IRI), or present serviceability rating (PRS), had a large impact on crash rates.

Miller and Johnson (1973) conducted a before-and-after study for pavement resurfacing in the United Kingdom. The authors analyzed the number of crashes on the M4 Motorway 2 years before and 2 years after resurfacing for over 500 crashes. The results indicated that the number of crashes was reduced by 28 and 63% in dry and wet conditions, respectively. The average number of crashes was reduced by 45%. Öberg (1981) explored the relationship between pavement friction coefficients and stopping distances for passenger vehicles on rural roads in Sweden. The author showed that required stopping distance exponentially increases at friction coefficient below 0.3. Gothie (1996) carried out an analysis of the cause and effect relationship between road surface

[1]Safety Program Director, Center for Advanced Transportation Systems Simulation, Dept. of Civil, Environmental and Construction Engineering, Univ. of Central Florida, Orlando, FL 32816-2450 (corresponding author). E-mail: jaeyoung@knights.ucf.edu

[2]Assistant Professor, Dept. of Civil, Environmental and Construction Engineering, Univ. of Central Florida, Orlando, FL 32816-2450.

[3]Professor and Chair, Dept. of Civil, Environmental and Construction Engineering, Univ. of Central Florida, Orlando, FL 32816-2450.

Note. This manuscript was submitted on October 6, 2014; approved on March 19, 2015; published online on May 22, 2015. Discussion period open until October 22, 2015; separate discussions must be submitted for individual papers. This paper is part of the ***Journal of Transportation Engineering***, © ASCE, ISSN 0733-947X/04015020(11)/$25.00.

properties and crash risks in France. The result from this study showed that crash rates increase by at least 50% on wet pavement when moving from a segment with a sideway force coefficient more than 0.6 to that with a side force coefficient less than 0.5.

Several researchers have looked into the relationship between pavement surface condition and traffic safety using statistical models. Anastasopoulos and Mannering (2009) used a random-parameters negative binomial (NB) model and they found that several pavement factors such as IRI, pavement rutting, and pavement condition rating (PCR) significantly influence the frequency of crashes. Anastasopoulos et al. (2012) applied a multivariate Tobit model to examine crash rates by three injury severity levels (i.e., no injury, possible injury, and injury crashes) with pavement conditions. Labi (2011) adopted NB models and analyzed three levels of crash severity, property damage only (PDO), injury, and fatal/injury crashes. The study revealed that both pavement surface friction and pavement condition play an important role in crash occurrence. It was shown that the higher friction number decreases crash counts regardless of severity levels on rural major/minor arterials, and the better pavement condition lowers the crash frequency of all severity levels on the rural principal arterials. Abdel-Aty et al. (2009) focused on the effect of pavement resurfacing projects on multilane arterials with partially limited access. They applied the empirical Bayes (EB) method and found that the pavement resurfacing projects resulted in a reduction in the number of total, severe, and rear-end crashes. Chan et al. (2010) conducted a comprehensive study on the effects of asphalt pavement conditions on traffic crashes based on the pavement management system (PMS). That study focused on urban interstate highways only with asphalt pavements, divided median types, and 55 mph speed limits. They also developed a series of NB models, including various pavement condition variables such as IRI, RD, and PSI.

The previous studies summarized above explored the relationship between pavement conditions and crash occurrence; however, the effect of poor pavement conditions on individual crash severity levels has not been addressed, particularly using a model such as the ordered logistic regression that can handle an ordered response variable. As a prerequisite for the development of a safety-incorporated pavement management system, an investigation on how pavement condition affects crash severity level was carried out in this study, with an emphasis on the evaluation of the relationship between poor pavement conditions and crash severity levels using a series of Bayesian ordered logistic models. It is expected that the findings of this study can assist transportation agencies at the federal, state, and local levels to select appropriate pavement maintenance and rehabilitation strategies to reduce traffic crash severity levels.

Data Preparation

The overall data preparation process is shown in Fig. 1. Data were obtained from Florida Department of Transportation (FDOT) and the variable importance was analyzed using random forests. A nested structure was developed for more specific statistical investigation. Subsequently, the data were processed based on the nested structure. Basically, the severity analysis was conducted using all candidate variables and full samples. Nevertheless, roadway segments with poor pavement conditions are rare. Some may argue that the modeling result, particularly for poor pavement, is biased since the roadway with poor pavement condition is very rare. In order to validate the result, crash cases with poor pavement condition was matched to the comparison group, which was a set of randomly selected crash cases with normal pavement conditions in an approximately 1:4 ratio. After that, the severity model was estimated using the poor pavement variable only with the sampled data.

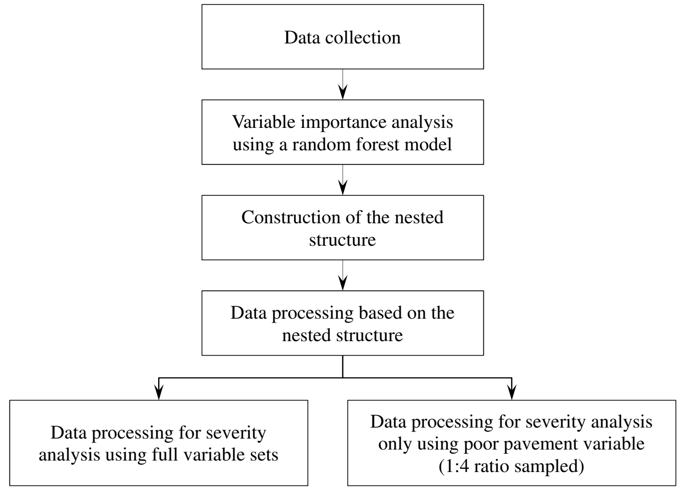


Fig. 1. Overall data preparation process

Data Collection and Variable Importance Analysis

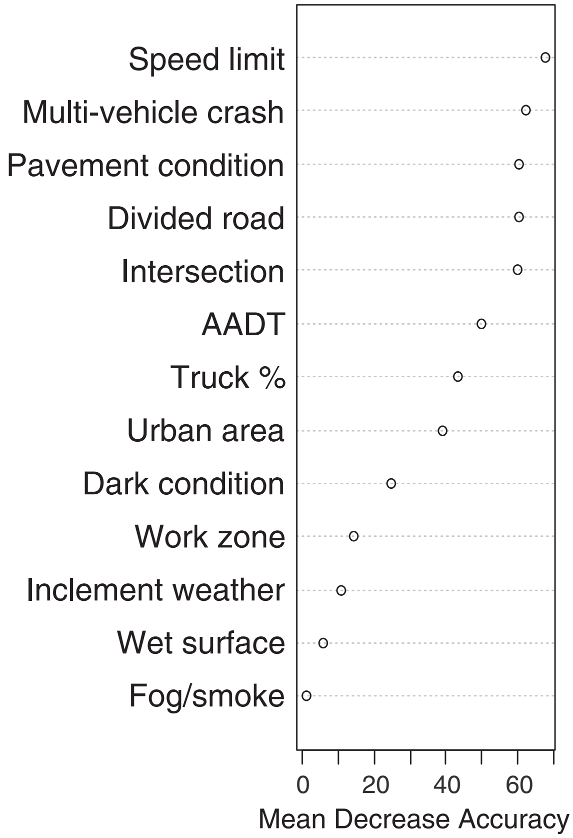
Roadway, traffic, and crash data for the specific information of individual crashes on state-maintained roads for 2012 were collected from the FDOT. These individual crash data were obtained for the severity analysis. The crash data provides five severity levels (i.e., PDO, possible injury, nonincapacitating injury, incapacitating injury, and fatal crashes); however, the five severity levels were reduced to three levels for several reasons. First, property damage alone and possible injury crashes were combined because these two levels are not clearly distinguished in the investigation. Second, incapacitating and fatal crashes were merged because the number of fatal crashes is too small to analyze separately. Thus, all crashes were divided into three levels of injury severity, i.e., minor, intermediate, and severe crashes (Table 1).

The random forest model consists of multiple unpruned decision trees (Breiman 2001). The random forest model has been widely used to identify important variables for the categorical target variable. The mean decrease accuracy, which is a scaled average of the prediction accuracy of each variable, was used for identifying the variable importance. The calculation is based on a process of randomly permuting the values of available variables across the observations and measuring the impact on the predictive accuracy of the resulting tree. If a variable has a larger impact, it is considered a more important variable for explaining the target variable (i.e., crash severity level).

As shown in Fig. 2, *speed limit* and *multi-vehicle crash* have the largest importance for the crash severity. In order to determine the

Table 1. Severity Levels before and after the Combination

Before the combination			After the combination		
Severity	Count	Proportion (%)	Severity	Count	Proportion (%)
PDO	114,181	56.09	Minor	161,272	79.22
Possible injury	47,091	23.13			
Non-incapacitating injury	30,557	15.01	Intermediate	30,557	15.01
Incapacitating injury	10,347	5.08	Severe	11,749	5.77
Fatal	1,402	0.69			



Variable	Mean Decrease Accuracy
Speed limit	67.66
Multi-vehicle crash	62.21
Pavement condition	60.46
Divided road	60.26
Intersection	60.13
AADT*	49.73
Truck %	43.40
Urban area	39.09
Dark condition	24.61
Work zone	14.34
Inclement weather	10.82
Wet surface	5.92
Fog/smoke	1.22

***Annual Average Daily Traffic**

Fig. 2. Result of variable importance analysis

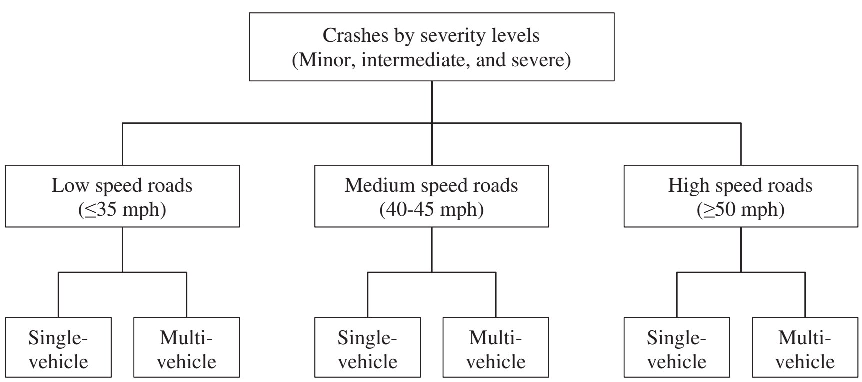


Fig. 3. Nested structure

more specific effects of pavement conditions on traffic safety, a nested structure was developed based on the variable importance from the random forest model. Each crash case was divided into three speed limit levels. The low, medium, and high speed roads were defined as those with speed limits of 35 mph or lower, between 40 and 45 mph, and 50 mph or higher, respectively. Subsequently, the crashes were again divided into single-vehicle and multi-vehicle crashes. The nested structure is presented in Fig. 3.

Pavement Conditions

The pavement condition data were obtained from the Roadway Characteristics Inventory (RCI) of the FDOT Transportation Statistics Office. The pavement condition is scaled from 0 to 5. The definition of each pavement condition is presented in Table 2 (FDOT 2013). The FDOT annually sends out trained raters for the field condition survey (visual interpretation of the roadway surface) and evaluates the pavement condition over the state's road systems. The pavement condition data are used for highway performance monitoring system (HPMS), work programs, and pavement design and the condition assessment is conducted in all paved principal arterial system roadways, National highway system (NHS) routes, all paved (HPMS) standard samples, and all strategic intermodal system (SIS)-related roadways. Thus, the pavement condition index is determined in the range from 0 (poorest) to 5 (best), which represents a comprehensive pavement condition, accounting for

Table 2. Definition, Length, and Proportion of Each Pavement Condition (Data from FDOT 2013)

Pavement condition index	Definition	Length (mi)	Percentage	Category
0.0–1.0	Very poor: virtually impassable. 75% or more deteriorated	5.61	0.1	Poor pavement
1.0–2.0	Poor: large potholes and deep cracks exist. Discomfort at slow speeds	36.54	0.5	
2.0–3.0	Fair: rutting, map cracking, and extensive patching	489.88	7.0	Normal pavement
3.0–4.0	Good: first class ride with only slight surface deterioration	3,957.39	56.4	
4.0–5.0	Very good: only new or nearly new pavement	2,526.25	36.0	

Table 3. CMH Tests for Selecting Poor Pavement Threshold

Poor pavement threshold	Pavement condition	Minor (%)	Intermediate (%)	Severe (%)	CMH (p)
<1.0	Normal	79.22	15.01	5.77	2.652 (0.103)
	Poor	68.42	15.79	15.79	
<1.5	Normal	79.22	15.01	5.77	1.004 (0.316)
	Poor	71.74	21.74	6.52	
<2.0[a]	Normal	79.23	15.01	5.77	10.097 (0.002)
	Poor	73.02	16.27	10.71	
<2.5	Normal	79.23	15.00	5.77	0.659 (0.417)
	Poor	78.21	16.01	5.78	
<3.0	Normal	79.23	15.09	5.68	1.526 (0.217)
	Poor	79.20	14.83	5.98	

[a]Significant at 95% confidence level.

pavement roughness and different types of surface distresses (e.g., cracking, rutting, pothole, patches).

The pavement condition index was converted to a binary variable due to the following two reasons: First, because the pavement condition index is ordinal, differences between indexes of 1 and 2 are not equal to the different between indexes of 3 and 4. Thus, it was necessary to generalize the index for the modeling purpose. A series of models were developed to analyze whether the binary dummy variable showing poor and normal pavement conditions has a significantly meaningful difference in crash severity. Second, one of the objectives of this study was to suggest a threshold for pavement condition to effectively reduce crash severity. Therefore, Cochran-Mantel-Haenszel (CMH) tests were conducted to find a reasonable threshold for the poor pavement. If a CMH value is statistically significantly large, it implies that there is a meaningful difference in the proportion of each severity level between poor and normal pavement conditions for a given criterion. Initially, 1.0, 1.5, 2.0, 2.5, and 3.0 were attempted as poor pavement criteria as shown in Table 3. Among the five candidate poor pavement criteria, only 2.0 showed a significant difference in severity levels between normal and poor pavement conditions. Thus, a variable was defined, i.e., *poor pavement*, which indicates poor or very poor pavement conditions with a pavement condition index of less than 2.0. Approximately 0.6% of roads in Florida are categorized under poor pavement conditions (Table 2).

Data Processing

Data were processed based on the nested structure shown in Fig. 3 and Tables 4–6 summarize the processed data on low-speed, medium-speed, and high-speed roads, respectively. As mentioned earlier in this section, it is possible that the modeling results may be affected by small proportion of poor pavement conditions (0.6%). Therefore, crash cases with poor pavement condition were matched to the comparison group, which was a set of randomly selected crash cases with normal pavement conditions in about a 1:4 ratio. These data are summarized in Table 7.

Moreover, Fig. 4(a) shows pavement conditions of roadway segments. It was observed that the roads under poor condition are distributed mostly in the rural area whereas pavement conditions in the urban area are generally good. This trend is also confirmed in Fig. 4(b), which presents the average pavement condition indices by counties.

Statistical Modeling

A series of Bayesian ordered logistic regression models were estimated using full data and sampled data, respectively, based on the nested structure. Different from the classical models, Bayesian models do not depend on the assumption of asymptotic normality. Sampling-based methods of Bayesian estimation focus on estimating the entire density of parameters as compared with the traditional classical estimation methods, which are intended for finding a

Table 4. Description of the Low-Speed Roads Full Data

	Single-vehicle crashes				Multi-vehicle crashes			
Variable	Mean	SD	Minimum	Maximum	Mean	SD	Minimum	Maximum
Sample size	$N = 5{,}765$				$N = 47{,}618$			
Crash severity[a]	1.375	0.646	1	3	1.183	0.459	1	3
Poor pavement	0.004	0.064	0	1	0.003	0.058	0	1
Truck %	1.421	2.960	0	41	2.358	3.258	0	31
Speed limit	31.621	4.528	10	35	32.383	3.611	10	35
Intersection	0.241	0.428	0	1	0.479	0.500	0	1
Dark condition	0.495	0.500	0	1	0.215	0.411	0	1
Divided	0.282	0.450	0	1	0.305	0.461	0	1
Fog/smoke	0.003	0.051	0	1	0.0004	0.020	0	1
Wet surface	0.183	0.387	0	1	0.124	0.330	0	1
Work zone	0.015	0.122	0	1	0.011	0.106	0	1
Urban	0.927	0.260	0	1	0.986	0.116	0	1

[a]1 = minor; 2 = intermediate; and 3 = severe crashes.

J. Transp. Eng., 2015, 141(10): 04015020

Table 5. Description of the Medium-Speed Roads Full Data

Variable	Single-vehicle crashes				Multi-vehicle crashes			
	Mean	SD	Minimum	Maximum	Mean	SD	Minimum	Maximum
Sample size		$N = 8{,}482$				$N = 90{,}636$		
Crash severity[a]	1.434	0.691	1	3	1.243	0.531	1	3
Poor pavement	0.001	0.034	0	1	0.0003	0.017	0	1
Truck %	2.689	3.596	0	41	3.258	3.163	0	44
Speed limit	43.622	2.234	40	45	43.373	2.343	40	45
Intersection	0.264	0.441	0	1	0.500	0.500	0	1
Dark condition	0.510	0.500	0	1	0.213	0.410	0	1
Divided	0.545	0.498	0	1	0.660	0.474	0	1
Fog/smoke	0.003	0.051	0	1	0.000	0.022	0	1
Wet surface	0.217	0.412	0	1	0.141	0.348	0	1
Work zone	0.010	0.098	0	1	0.010	0.098	0	1
Urban	0.870	0.336	0	1	0.980	0.140	0	1

[a]1 = minor; 2 = intermediate; and 3 = severe crashes.

Table 6. Description of the High-Speed Roads Full Data

Variable	Single-vehicle crashes				Multi-vehicle crashes			
	Mean	SD	Minimum	Maximum	Mean	SD	Minimum	Maximum
Sample size		$N = 14{,}505$				$N = 36{,}572$		
Crash severity[a]	1.428	0.687	1	3	1.308	0.601	1	3
Poor pavement	0.001	0.031	0	1	0.001	0.024	0	1
Truck %	9.171	6.931	0	52	6.665	4.928	0	43
Speed limit	61.794	6.807	50	70	57.623	6.607	50	70
Intersection	0.077	0.267	0	1	0.276	0.447	0	1
Dark condition	0.407	0.491	0	1	0.225	0.417	0	1
Divided	0.723	0.448	0	1	0.803	0.398	0	1
Fog/smoke	0.003	0.054	0	1	0.002	0.045	0	1
Wet surface	0.330	0.470	0	1	0.191	0.393	0	1
Work zone	0.043	0.202	0	1	0.043	0.202	0	1
Urban	0.605	0.489	0	1	0.854	0.353	0	1

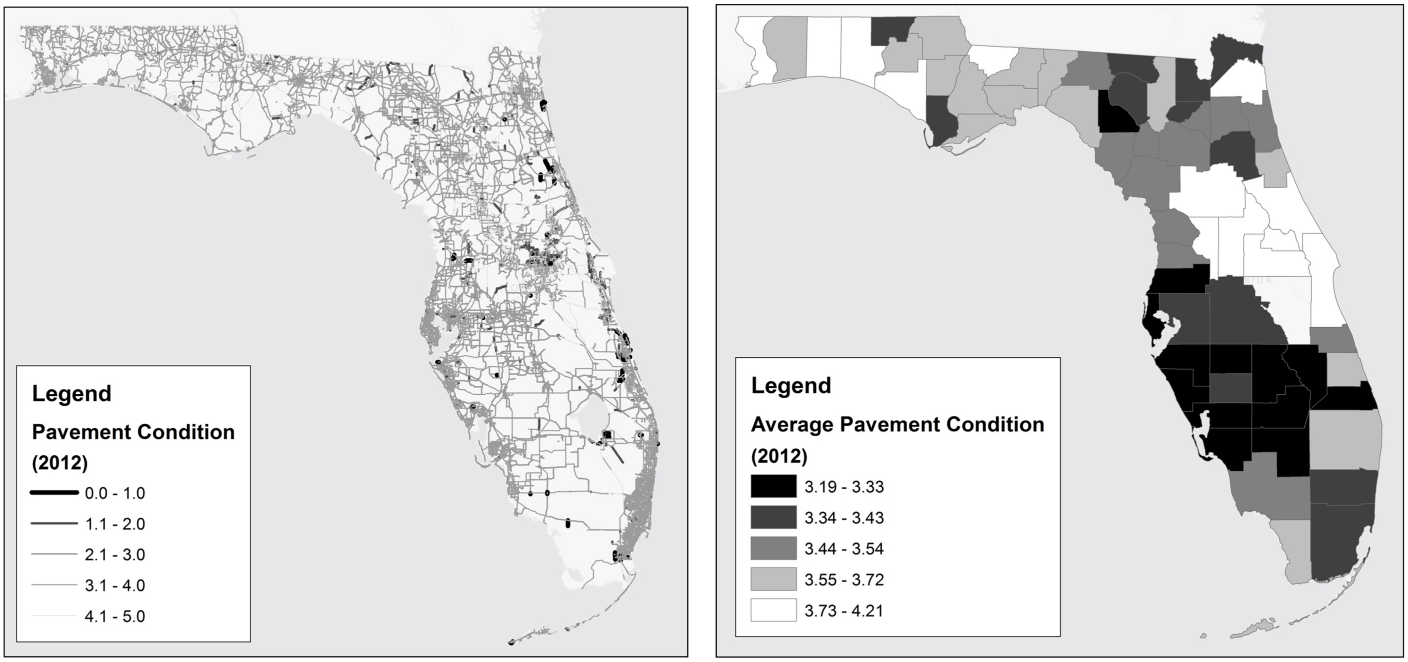
[a]1 = minor; 2 = intermediate; and 3 = severe crashes.

Table 7. Description of the Sampled Data for the Validation

Category	Variable	Single-vehicle crashes				Multi-vehicle crashes			
		Mean	SD	Minimum	Maximum	Mean	SD	Minimum	Maximum
	Sample size			$N = 120$				$N = 790$	
Low-speed roads	Crash severity[a]	1.325	0.610	1	3	1.187	0.462	1	3
Poor pavement	0.200	0.402	0	1	0.200	0.400	0	1	
Medium-speed roads	Sample size			$N = 50$				$N = 125$	
	Crash severity[a]	1.380	0.602	1	3	1.288	0.565	1	3
	Poor pavement	0.200	0.404	0	1	0.200	0.402	0	1
High-speed roads	Sample size			$N = 70$				$N = 105$	
	Crash severity[a]	1.400	0.710	1	3	1.429	0.705	1	3
	Poor pavement	0.200	0.403	0	1	0.200	0.402	0	1

[a]1 = minor; 2 = intermediate; and 3 = severe crashes.

single point estimate using the maximum likelihood approach (Congdon 2003). Admittedly, sometimes point estimates may be more convenient for the practical application since it clearly suggest a single point. However, the Bayesian approach has a significant advantage over the maximum likelihood estimation. The Bayesian estimation determines posterior density for each parameter under consideration. This density estimation is the outcome of a process where a long run or a series of long runs of samples are taken from the posterior density based on the prior information about the parameter and data. Accordingly a Bayesian approach provides a considerable interpretive advantage since posterior estimates reflect the probabilities that the analyst is primarily interested in, the probability of the null hypothesis being true called a Bayesian credible interval (BCI). In contrast, classical confidence intervals on parameter estimates provide the probability of observing data given that a parameter takes on a specific value. This distinction of the Bayesian approach provides a substantial philosophical and practical advantage (Mitra and Washington 2007).



(a) (b)

Fig. 4. Spatial distributions of pavement conditions: (a) pavement conditions by roadway segments; (b) average pavement condition indices by county

Therefore, the Bayesian approach is thought to be more suitable compared with the classical likelihood-based inference methods and thus have been popular in recent traffic safety research. This study also adopted a Bayesian approach because of the reasons above.

Regarding the ordered logistic model, some researchers have adopted the ordered logistic models for severity analysis (Breiman 2001; Abdel-Aty et al. 2011). In ordinal response models, a series of latent thresholds are formulated. A one-dimension space of real number is divided into three intervals according to two thresholds (γ_1, γ_2) that link to the three ordered severity levels: minor ($k = 1$), intermediate ($k = 2$), and severe crashes ($k = 3$). For any crash i, the observed ordinal severity levels (y_i) are related to a latent response variable (y_i^*) as follows:

$$y_i = \begin{cases} 1 & \text{if } -\infty < y_i^* \le \gamma_1 \\ 2 & \text{if } \gamma_1 < y_i^* \le \gamma_2 \\ 3 & \text{if } \gamma_2 < y_i^* \le +\infty \end{cases} \tag{1}$$

The latent response variable (y_i^*) is formulated with the following linear specification:

$$y_i^* = \theta_i + \varepsilon_i \tag{2}$$

$$\theta_i = \beta_0 + \beta X_i \tag{3}$$

where β_0 = intercept; β = estimated coefficient vector; X_i = covariates; and ε_i = random error term, which is assumed to have a logistic distribution, with F-distribution cumulative density function. The cumulative response probabilities for the three levels of the ordinal outcome could be expressed as follows:

$$P_{i(k)} = Pr(y_i \le k) = F(\gamma_{(k)} - \theta_i) = \frac{\exp[\gamma_{(k)} - \theta_i]}{1 + \exp[\gamma_{(k)} - \theta_i]} \tag{4}$$

Concerning the variable selection, the candidate variables with 90% BCI were chosen in the final model. Deviance information criterion (DIC), which is a Bayesian generalization of AIC (Akaike information criterion), was calculated to measure the model complexity and fit (Spiegelhalter et al. 2002)

$$\text{DIC} = 2 \times \bar{D} - \hat{D} \tag{5}$$

where $\bar{D}$ = posterior mean of deviance, D, $\hat{D} = 2 \times [p(y|\theta)]$, and $\bar{\theta}$ = posterior mean of θ.

Each submodel with the smallest DIC in the nested structure was selected as a final model.

As shown in the variable importance analysis using random forest, annual average daily traffic (AADT) was found to be one of the important variables; however, AADT was not included in the ordered logistic regression models since AADT was highly correlated with speed limit ($r = 0.514$). Also, the severity models (i.e., ordered logistic regression models) do not typically include AADT, which is different from the crash frequency modeling.

The modeling results of low-speed, medium-speed, and high-speed roads are presented in Tables 8–10, respectively. The results indicated that the poor pavement condition has statistically significant effects on the crash severity, at the 90% confidence level, except for single-vehicle crashes on medium speed roads. In case of multi-vehicle crashes on low- and medium-speed roads, and both single- and multi-vehicle crashes on high speed roads, the poor pavement condition contributes to the likelihood of more severe crashes. On the contrary, the poor pavement condition decreases the likelihood of crash severity in single-vehicle crashes on low-speed roads.

Other than the poor pavement, there are several factors affecting the crash severity. It was revealed that the percentage of trucks reduces the severity for multi-vehicle crashes on low- and medium-speed roads, and single-vehicle on medium-speed roads. The speed limit has commonly positive effects on all crashes except for single-vehicle crashes on high-speed roads. Furthermore, if the crash location is at the intersection, the crash severity is likely to decrease for single-vehicle crashes on low- and medium-speed

Table 8. Bayesian Ordered Logistic Modeling Results for Crash Severity for Low-Speed (≤35 mph) Roads (Full Models)

	Single-vehicle crash ($N = 5{,}765$)						Multi-vehicle crash ($N = 47{,}618$)					
			Bayesian credible interval						Bayesian credible interval			
Variable	Mean	SD	2.5%	5.0%	95.0%	97.5%	Mean	SD	2.5%	5.0%	95.0%	97.5%
γ_1	−2.321	0.217	−2.753	−2.669	−1.964	−1.905	−4.106	0.164	−4.408	−4.385	−3.846	−3.822
γ_2	−0.918	0.212	−1.352	−1.250	−0.565	−0.526	−2.361	0.163	−2.663	−2.634	−2.106	−2.072
Poor pavement	−1.316	0.671	−2.829	−2.574	−0.323	−0.157	0.378	0.190	−0.004	0.049	0.676	0.760
Truck %	—	—	—	—	—	—	−0.064	0.005	−0.073	−0.072	−0.056	−0.055
Speed limit	0.029	0.006	0.018	0.019	0.039	0.042	0.041	0.004	0.035	0.035	0.048	0.049
Intersection	−0.280	0.072	−0.428	−0.405	−0.167	−0.154	0.270	0.028	0.219	0.224	0.313	0.324
Divided road	0.135	0.065	−0.001	0.024	0.234	0.247	—	—	—	—	—	—
Urban area	−0.905	0.098	−1.112	−1.081	−0.746	−0.721	−0.684	0.084	−0.848	−0.822	−0.540	−0.519
Work zone	—	—	—	—	—	—	—	—	—	—	—	—
Dark condition	—	—	—	—	—	—	—	—	—	—	—	—
Wet surface	−0.324	0.077	−0.473	−0.454	−0.192	−0.170	−0.128	0.040	−0.209	−0.199	−0.056	−0.051
DIC	8,820.890						47,498.100					

Table 9. Bayesian Ordered Logistic Modeling Results for Crash Severity for Medium-Speed (40–45 mph) Roads (Full Models)

	Single-vehicle crash ($N = 8{,}482$)						Multi-vehicle crash ($N = 90{,}636$)					
			Bayesian credible interval						Bayesian credible interval			
Variable	Mean	SD	2.5%	5.0%	95.0%	97.5%	Mean	SD	2.5%	5.0%	95.0%	97.5%
γ_1	−2.425	0.486	−3.294	−3.236	−1.715	−1.647	−4.598	0.3245	−5.067	−5.030	−3.845	−3.793
γ_2	−1.142	0.486	−2.002	−1.945	−0.442	−0.381	−3.045	0.3241	−3.510	−3.475	−2.288	−2.226
Poor pavement	0.574	0.610	−0.640	−0.393	1.578	1.685	0.901	0.423	−0.019	0.160	1.589	1.699
Truck %	−0.014	0.007	−0.027	−0.025	−0.002	0.000	−0.043	0.003	−0.049	−0.048	−0.039	−0.038
Speed limit	0.020	0.011	0.002	0.004	0.039	0.040	0.050	0.006	0.033	0.034	0.058	0.059
Intersection	−0.164	0.057	−0.279	−0.260	−0.071	−0.058	0.255	0.017	0.221	0.227	0.282	0.285
Divided road	0.169	0.050	0.070	0.083	0.248	0.264	—	—	—	—	—	—
Urban area	−0.499	0.073	−0.647	−0.619	−0.386	−0.365	−0.567	0.071	−0.726	−0.708	−0.462	−0.439
Work zone	−0.504	0.274	−1.034	−0.958	−0.066	0.014	0.193	0.081	0.036	0.061	0.324	0.344
Dark condition	—	—	—	—	—	—	0.197	0.020	0.158	0.165	0.230	0.235
Wet surface	−0.356	0.058	−0.464	−0.452	−0.257	−0.239	−0.186	0.026	−0.238	−0.231	−0.144	−0.135
DIC	14,041.800						108,371.000					

Table 10. Bayesian Ordered Logistic Modeling Results for Crash Severity for High-Speed (≥50 mph) Roads (Full Models)

	Single-vehicle crash ($N = 14{,}505$)						Multi-vehicle crash ($N = 36{,}572$)					
			Bayesian credible interval						Bayesian credible interval			
Variable	Mean	SD	2.5%	5.0%	95.0%	97.5%	Mean	SD	2.5%	5.0%	95.0%	97.5%
γ_1	−1.665	0.036	−1.733	−1.726	−1.604	−1.594	−2.481	0.132	−2.234	−2.261	−2.696	−2.754
γ_2	−0.368	0.033	−0.428	−0.418	−0.312	−0.302	−1.112	0.131	−0.859	−0.899	−1.321	−1.382
Poor pavement	1.279	0.520	0.203	0.360	2.081	2.275	0.670	0.403	−0.133	0.002	1.270	1.408
Truck %	—	—	—	—	—	—	—	—	—	—	—	—
Speed limit	—	—	—	—	—	—	0.008	0.002	0.004	0.004	0.011	0.012
Intersection	—	—	—	—	—	—	0.355	0.028	0.294	0.307	0.405	0.413
Divided road	—	—	—	—	—	—	−0.246	0.034	−0.303	−0.297	−0.188	−0.176
Urban area	−0.305	0.036	−0.374	−0.364	−0.244	−0.234	−0.608	0.035	−0.676	−0.666	−0.552	−0.536
Work zone	−0.319	0.092	−0.498	−0.474	−0.174	−0.143	−0.155	0.064	−0.280	−0.251	−0.047	−0.043
Dark condition	−0.132	0.036	−0.206	−0.192	−0.073	−0.061	0.346	0.028	0.283	0.297	0.384	0.399
Wet surface	−0.536	0.041	−0.614	−0.609	−0.473	−0.455	−0.087	0.029	−0.143	−0.133	−0.042	−0.030
DIC	23,729.400						49,560.000					

roads whereas the severity tends to increase for multi-vehicle crashes at the intersection, regardless of the speed limits. This is because single-vehicle crashes more frequently occur on roadway segments whereas multi-vehicle crashes happen more frequently at intersections.

It was shown that crashes are more severe in single-vehicle crashes on low- and medium-speed roads if the road is physically divided. However, crashes are less severe for multi-vehicle crashes on high-speed divided roads. It was also found that the severity is lower in the urban area all of the time, regardless of

Table 11. Bayesian Ordered Logistic Modeling Results for Crash Severity for Low-Speed (≤35 mph) Roads (Simple Models)

	Single-vehicle crash ($N = 120$)						Multi-vehicle crash ($N = 790$)					
			Bayesian credible interval						Bayesian credible interval			
Variable	Mean	SD	2.5%	5.0%	95.0%	97.5%	Mean	SD	2.5%	5.0%	95.0%	97.5%
γ_1	−2.434	0.387	−3.233	−3.134	−1.838	−1.740	−3.606	0.211	−4.002	−3.947	−3.258	−3.222
γ_2	−0.943	0.231	−1.420	−1.330	−0.549	−0.479	−1.782	0.115	−2.007	−1.966	−1.606	−1.563
Poor pavement	−1.136	0.726	−2.739	−2.389	−0.023	0.119	0.420	0.233	−0.069	0.025	0.797	0.845
DIC	175.174						810.923					

Table 12. Bayesian Ordered Logistic Modeling Results for Crash Severity for Medium-Speed (40–45 mph) Roads (Simple Models)

	Single-vehicle crash ($N = 50$)						Multi-vehicle crash ($N = 125$)					
			Bayesian credible interval						Bayesian credible interval			
Variable	Mean	SD	2.5%	5.0%	95.0%	97.5%	Mean	SD	2.5%	5.0%	95.0%	97.5%
γ_1	−3.359	0.725	−4.950	−4.676	−2.275	−2.063	−3.247	0.441	−4.215	−4.019	−2.527	−2.390
γ_2	−1.050	0.370	−1.775	−1.656	−0.474	−0.367	−1.471	0.259	−1.976	−1.916	−1.035	−0.977
Poor pavement	1.184	0.786	−0.409	−0.227	2.514	2.693	1.071	0.481	0.146	0.275	1.901	2.070
DIC	81.718						168.610					

Table 13. Bayesian Ordered Logistic Modeling Results for Crash Severity for High-Speed (≥50 mph) Roads (Simple Models)

	Single-vehicle crash ($N = 70$)						Multi-vehicle crash ($N = 105$)					
			Bayesian credible interval						Bayesian credible interval			
Variable	Mean	SD	2.5%	5.0%	95.0%	97.5%	Mean	SD	2.5%	5.0%	95.0%	97.5%
γ_1	−2.638	0.467	−3.599	−3.452	−1.930	−1.777	−2.252	0.339	−2.901	−2.832	−1.720	−1.661
γ_2	−1.481	0.332	−2.144	−2.044	−0.944	−0.851	−1.027	0.242	−1.494	−1.440	−0.641	−0.577
Poor pavement	2.039	0.649	0.767	0.979	3.084	3.276	0.855	0.473	−0.073	0.132	1.669	1.804
DIC	104.489						175.138					

single/multi-vehicle crashes and speeds. It can be explained in two ways. First, drivers may increase their speeds more in the rural area because of the low traffic volume. Second, monotonous driving conditions in the rural area may result in drowsiness and/or inattention. In addition, crash severity is lower in the work zone for single-vehicle on medium-speed roads and all crashes on high-speed roads but it is higher in the work zone for multi-vehicle crashes on medium speed roads.

Furthermore, some environmental conditions such as the dark condition and wet surface have meaningful effects on the crash severity. It was uncovered that crashes are more severe in dark conditions in multi-vehicle crashes on medium-speed roads whereas they are less severe in dark conditions in multi-vehicle crashes on high-speed roads. Lastly, it is worth noting that the severity is lower with the wet surface condition in all crash types. This finding may not be intuitive; however, Yamamoto and Shankar (2004) and Qudddus et al. (2002, 2009) also found the same results. According to Yamamoto and Shankar (2004) and Quddus et al. (2009), drivers tend to reduce their speeds and be more careful on wet surface, and crash severity levels are more likely to decrease when crashes occur.

Meanwhile, a series of simple Bayesian ordered logistic regression models developed using the sampled data for validation. Also, only the poor pavement variable was included to confirm that the effect of poor pavement conditions is consistent in the sampled data. Tables 11–13 exhibit the results of the simple models and it was shown that the signs and significance levels of poor pavement in six models are coherent with full models (Tables 8–10). Thus, it is concluded that the results of the poor pavement in the full models are reliable despite the large difference in proportions of poor and normal pavement conditions.

Discussion of the Results

Fig. 5(a) shows the odds ratio of higher severity levels for poor pavement. The odds ratio is a measure of association that approximates relative risk, or in other words, how much more likely it is for the outcome (severity level) to be present among those with $x = 1$ (poor pavement) than among those with $x = 0$ (normal pavement) (Hosmer and Lemeshow 2004). The results show that the multi-vehicle crashes under poor pavement conditions are more likely to cause higher injury-severity levels for all three speed-level roads. On the other hand, when it comes to the single-vehicle crashes, the poor pavement decreases the severity level on low-speed roads whereas it results in more severe crashes on high-speed roads [Fig. 5(b)].

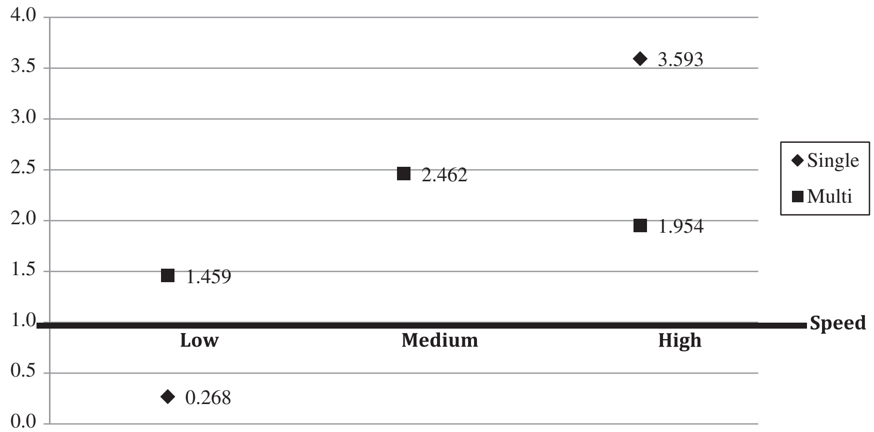
Although it was shown that the poor pavement condition has a significant effect on severity levels, explaining how and why the poor pavement condition affects the severity of crashes is still required. Fig. 6 compares the crash types of single-vehicle crash [Fig. 6(a)] and multi-vehicle crash [Fig. 6(b)] under normal and poor pavement conditions. In the case of single-vehicle crashes, the proportion of *hit fixed objects* and *ran into ditch/water* increased

Category	Low-speed		Medium-speed		High-speed	
	Estimate	Odds ratio	Estimate	Odds ratio	Estimate	Odds ratio
Single-vehicle	-1.316 (-2.574, -0.323)	0.268 (0.076, 0.724)	#0.574 (-0.393, 1.578)	#1.775 (0.675, 4.845)	1.279 (0.360, 2.081)	3.593 (1.433, 8.012)
Multi-vehicle	0.378 (0.049, 0.676)	1.459 (1.050, 1.966)	0.901 (0.160, 1.589)	2.462 (1.174, 4.899)	0.670 (0.002, 1.270)	1.954 (1.002, 3.561)

The numbers in parentheses indicate 90% credible interval for the estimates

Not significant at 90% credible interval

(a)



(b)

Fig. 5. Summary of odds ratio of severity levels for the poor pavement condition: (a) odds ratio and confidence level of severity levels; (b) graphical representation of odds ratio of severity levels

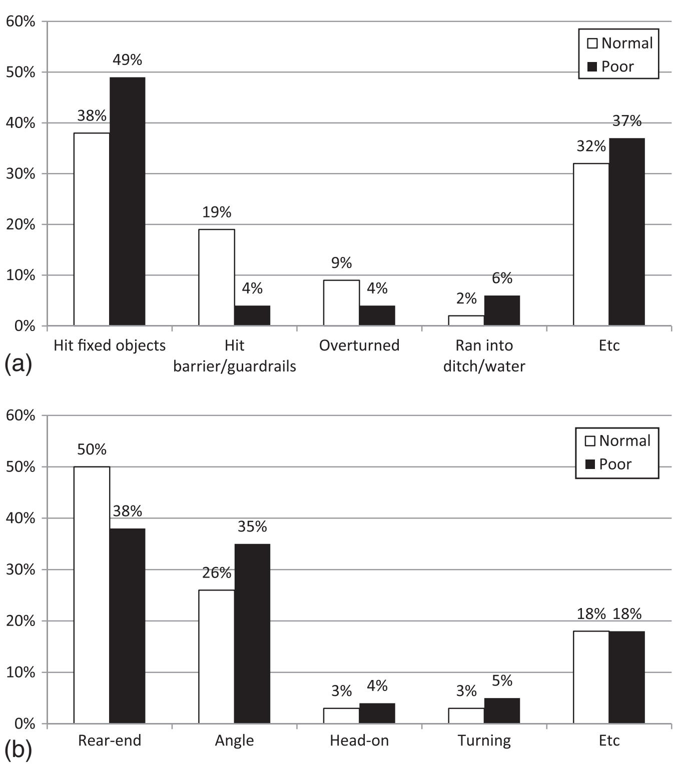


Fig. 6. Comparison of crash types: (a) single-vehicle crashes; (b) multi-vehicle crashes

from 40 to 55% in poor pavement conditions. In contrast, the proportion of *hit barrier/guardrails* and *overturned* cases decrease in poor pavement conditions. It may imply that drivers may have enough time to reduce their speeds to minimize the impact of the collision with fixed objects if they observed the poor pavement condition ahead on the low-speed roads. However, when it comes to the high-speed roads, drivers may not have enough time to react and also the collision with fixed object at high-speed results in the higher severity levels. Concerning the multi-vehicle crashes, Rear-end crashes are less likely while head-on, angle, turning crashes increase in the poor pavement condition. In general, the rear-end crashes cause lower severity levels whereas the head-on crashes result in quite severe damages (or injuries), and both angle and turning crashes have relatively higher crash severity levels. The difference in crash type between normal and poor pavement conditions may explain why and how the poor pavement condition affects the severity levels of crashes.

As seen in the results of this study, road safety is apparently affected by the pavement surface condition, particularly for crash severity. Poor pavement conditions at low-speed roads result in less severe crashes for single-vehicle collisions but more severe crashes for multi-vehicle collisions. In the case of single-vehicle collisions at low-speed and multi-vehicle collisions at medium- and high-speed have higher severity levels when pavement conditions are poor. Nevertheless, almost all states (including Florida) do not use the safety analysis in PMS or vice versa. Incorporating safety analysis into the pavement management (or vice versa) is one of the urgent needs for more effective and safer management of roadway systems. Therefore, one of meaningful alternatives suggested by the authors is to develop a pavement condition-based safety index that can quantify the effect of pavement surface condition on road safety. This pavement-related safety index can be developed by evaluating and quantifying the effects of each pavement distress (e.g., crack, rut, and roughness) on both crash count and severity level. Subsequently, threshold values of this index can be constructed so that road sections below the threshold values may be considered as different categories of urgent project. This pavement condition-based safety index can be constructed for each segment matched with PMS database; thus, a pavement management engineer can take into account during his decision procedure.

Conclusions and Recommendations

This study explored the effect of pavement conditions on crash severity levels. The importance of the variables was analyzed using the random forest model and revealed that the speed limits and single/multi-vehicle collision have the largest influence on the severity of traffic crashes. Based on this result, the nested structure that divides crashes by speed limits and single/multi-vehicle crashes was developed. A series of six Bayesian ordered logistic regression models in the nested structure were estimated and revealed that there are various factors such as roadway, traffic, and environmental factors affecting crash severity. Among these factors, the poor pavement condition increases the severity of multi-vehicle crashes on all three speed-level roads. In addition, the poor pavement increases the severity of single-vehicle crash on high-speed roads. However, in the low-speed roads, the results indicate that the severity of single-vehicle crashes decrease as pavement condition gets poorer.

The results from this study suggest that the severity levels of most of crash types can be reduced when the pavement condition is well maintained. Thus, the recommendation is to maintain the pavement condition above an acceptable level: minimum 2.0 of pavement condition index or higher. Although this study revealed several interesting findings, it has two limitations. First, the severity analysis did not take into account the different types of pavement distresses such as roughness, cracking, and rutting separately. A specific distress type can have higher impacts on the crash severity. In addition, the safety analysis can be applied to different pavement types (e.g., flexible, rigid, and composite pavements) because different pavement types involve different surface distresses. These two main issues could be addressed in follow-up studies collecting more objective and reliable pavement condition indices for the analysis. The authors plan to collect a PMS data set [e.g., rut, roughness (IRI), crack] and will evaluate the effect of each parameter on both crash count and severity level. This quantification analysis will be conducted throughout the state of Florida as well as each district because the effect of those parameters on road safety may vary depending on geographical location. As discussed above, a pavement condition-based safety index can be developed; thus, road safety can be quantitatively incorporated into existing pavement management systems.

Acknowledgments

The authors wish to thank the Florida Department of Transportation for providing data.

References

Abdel-Aty, M., Devarasetty, P. C., and Pande, A. (2009). "Safety evaluation of multilane arterials in Florida." *Accid. Anal. Prev.*, 41(4), 777–788.

Abdel-Aty, M. A., Ekram, A., Huang, H., and Choi, K. (2011). "A study on crashes related to visibility obstruction due to fog and smoke." *Accid. Anal. Prev.*, 43(5), 1730–1737.

Al-Masaeid, H. R. (1997). "Impact of pavement condition on rural road accidents." *Can. J. Civ. Eng.*, 24(4), 523–531.

Anastasopoulos, P. C., and Mannering, F. L. (2009). "A note on modeling vehicle accident frequencies with random-parameters count models." *Accid. Anal. Prev.*, 41(1), 153–159.

Anastasopoulos, P. C., Shankar, V. N., Haddock, J. E., and Mannering, F. L. (2012). "A multivariate Tobit analysis of highway accident-injury-severity rates." *Accid. Anal. Prev.*, 45(1), 110–119.

Breiman, L. (2001). "Random forests." *Mach. Learn.*, 45(1), 5–32.

Chan, C. Y., Huang, B., Yan, X., and Richards, S. (2010). "Investigating effects of asphalt pavement conditions on traffic accidents in Tennessee based on the pavement management system (PMS)." *J. Adv. Transp.*, 44(3), 150–161.

Congdon, P. (2003). *Bayesian statistical modeling*, Wiley, New York.

FDOT (Florida Department of Transportation). (2013). "Roadway characteristics inventory (RCI) features & characteristics handbook." ⟨http://www.dot.state.fl.us/planning/statistics/rci/⟩ (Jan. 2015).

Gothie, M. (1996). "Relationship between surface characteristics and accidents." *Proc., 3rd Int. Symp. on Pavement Surface Characteristics*, ARRB Group Limited, Vermont South, Australia, 271–281.

Heaton, B., Henry, J. J., and Wambold, J. C. (1990). "Texture measuring equipment vs skid testing equipment." *Proc., Conf. of the Australian Road Research Board Pavements and Materials*, ARRB Group Limited, Vermont South, Australia, 53–64.

Henry, J. J. (1996). "Overview of the international PIARC experiment to compare and harmonise texture and skid resistance measurements: The international friction index." *Proc., 3rd Int. Symp. on Pavement Surface Characteristics*, ARRB Group Limited, Vermont South, Australia.

Hosmer, D. W., and Lemeshow, S. (2004). *Applied logistic regression*, Wiley, New York.

Hoyert, D. L., and Xu, J. (2012). "National vital statistics report 61(6)." U.S. Dept. of Health and Human Services, Centers for Disease Control and Prevention, Hyattsville, MD.

Karan, M. A., Haas, R., and Kher, R. (1976). "Effects of pavement roughness on vehicle speeds." *Transportation Research Record, 602*, 122–127.

Karlaftis, M. G., and Golias, I. (2002). "Effects of road geometry and traffic volumes on rural roadway accident rates." *Accid. Anal. Prev.*, 34(3), 357–365.

Labi, S. (2011). "Efficacies of roadway safety improvements across functional subclasses of rural two-lane highways." *J. Saf. Res.*, 42(4), 231–239.

Miller, M. M., and Johnson, H. D. (1973). "Effects of resistance to skidding on accidents: Surface dressing on an elevated section of the M4 motorway." *Laboratory Rep. LR 542*, Transport and Road Research Laboratory, Crowthorne, U.K.

Mitra, S., and Washington, S. (2007). "On the nature of over-dispersion in motor vehicle crash prediction models." *Accid. Anal. Prev.*, 39(3), 459–468.

NHTSA (National Highway Traffic Safety Administration). (2008). "National motor vehicle crash causation survey report to congress." U.S. Dept. of Transportation, Washington, DC.

NHTSA (National Highway Traffic Safety Administration). (2013). "2012 motor vehicle crashes: Overview." U.S. Dept. of Transportation, Washington, DC.

Öberg, G. (1981). "Friction and journey speed on roads with various winter road maintenance." *VTI Rapport 218*, National Swedish Road and Traffic Research Institute, Linköping, Sweden.

Quddus, M. A., Noland, R. B., and Chin, H. C. (2002). "An analysis of motorcycle injury and vehicle damage severity using ordered probit models." *J. Saf. Res.*, 33(4), 445–462.

Quddus, M. A., Wang, C., and Ison, S. G. (2009). "Road traffic congestion and crash severity: Econometric analysis using ordered response models." *J. Transp. Eng.*, 10.1061/(ASCE)TE.1943-5436.0000044, 424–435.

Spiegelhalter, D. J., Best, N. G., Carlin, B. P., and Van Der Linde, A. (2002). "Bayesian measures of model complexity and fit." *J. R. Stat. Soc. Ser. B (Stat. Method.)*, 64(4), 583–639.

Tighe, S. L., and Haas, R. (1998). "Safety and economic benefits of partially paved shoulders." *Proc., Annual Conf. of Transportation Association of Canada*, Transportation Association of Canada, Ottawa.

Wambold, J. C. (1988). "Road characteristics and skid testing." *Transportation Research Record, 1196*, 294–299.

Yamamoto, T., and Shankar, V. N. (2004). "Bivariate ordered-response probit model of driver's and passenger's injury severities in collisions with fixed objects." *Accid. Anal. Prev.*, 36(5), 869–876.

Zegeer, C. V., and Council, F. M. (1995). "Safety relationships associated with cross-sectional roadway elements." *Transportation Research Record, 1512*, 29–36.