Research Article

Distinguishing between Rural and Urban Road Segment Traffic Safety Based on Zero-Inflated Negative Binomial Regression Models

Xuedong Yan,[1] Bin Wang,[1] Meiwu An,[2] and Cuiping Zhang[1]

[1] *MOE Key Laboratory for Transportation Complex Systems Theory and Technology, School of Traffic and Transportation, Beijing Jiaotong University, Beijing 100044, China*

[2] *Saint Louis County Highways and Traffic and Public Works, 121 South Meramec Avenue, Saint Louis, MO 63105, China*

Correspondence should be addressed to Xuedong Yan, xdyan@bjtu.edu.cn

In this study, the traffic crash rate, total crash frequency, and injury and fatal crash frequency were taken into consideration for distinguishing between rural and urban road segment safety. The GIS-based crash data during four and half years in Pikes Peak Area, US were applied for the analyses. The comparative statistical results show that the crash rates in rural segments are consistently lower than urban segments. Further, the regression results based on Zero-Inflated Negative Binomial (ZINB) regression models indicate that the urban areas have a higher crash risk in terms of both total crash frequency and injury and fatal crash frequency, compared to rural areas. Additionally, it is found that crash frequencies increase as traffic volume and segment length increase, though the higher traffic volume lower the likelihood of severe crash occurrence; compared to 2-lane roads, the 4-lane roads have lower crash frequencies but have a higher probability of severe crash occurrence; and better road facilities with higher free flow speed can benefit from high standard design feature thus resulting in a lower total crash frequency, but they cannot mitigate the severe crash risk.

1. Introduction

Previous studies have been focused on distinguishing between rural and urban traffic safety using traffic crash data, but the influence of rural or urban settings on segment safety is controversial. The fatal traffic crash research indicated that fatality rates in rural areas are higher than in urban areas [1–3]. The higher fatality and injury rates in rural road facilities

have been attributed to various reasons, such as longer emergency response time and further distance to crash locations [3]. The other explanations include higher speed limit and worse driving habits (e.g., alcohol, a lower rate of seat belt use, and safety precaution use), road conditions, and accessibility to trauma care [4–6]. On the other hand, some studies reported that crash frequencies in urban areas were higher than those in rural areas [7, 8], because urban regions involve more complex traffic conditions, high traffic volume, congestion, poor pavement conditions, and so forth [9]. While the risk of severe crashes appears higher in rural segments, no differences were identified in the cause of injury and place of injury between urban and rural drivers [10].

Neither crash frequencies nor fatality rate can entirely represent the influence of rural or urban settings on the segments. Many other factors lead to the occurrence of traffic crashed, such as traffic characteristics, road design characteristics, demographic features, and pavement maintenance conditions [11–17]. Therefore, numerous cross-sectional studies have been conducted to characterize the relationships between factors and road segment-involved crashes. In the prior studies, the Poisson models are the most common ones which have been widely used [18, 19]. It is known that a Poisson model would be appropriate only when the mean and the variance of the crash frequencies are approximately equal. However, this assumption has been proved invalid for modeling traffic crash frequencies [20, 21], because the variances of crash frequencies were generally greater than means. Therefore, the negative binomial (NB) regression models were introduced to overcome this overdispersion problem, which had a more flexible mean-variance equality constraint [16, 22]. Nevertheless, both Poisson and NB models cannot deal with the property of crash frequency data with a large density of zeros (no crashes occur on roads during the observation period). Correspondingly, the zero-inflated count regress models were developed and applied for analyzing and predicting crash frequencies. The zero-inflated count regression models are capable of handling the apparent "excess" zeros crash data and generally have a more statistical suitability for modeling crash data than Poisson and NB regression models [23].

In order to understand the role of rural or urban settings in segment safety, the crash rate, crash frequencies, and the injury and fatality frequencies are taken into consideration in this study to distinguish between rural and urban traffic safety. The GIS-based crash data during four and half years in Pikes Peak Area, USA were applied for analyses. The GIS techniques for traffic data process have been proved effective to analyze and visualize crash data [24] and have advantages in data display, clear presentation of spatial relationship, and convenient query of relevant data [25, 26]. Since previous studies have discussed the suitability of various models in the prediction of crash frequencies, we adopt zero-inflated negative binomial (ZINB) regression models for crash frequency analysis and prediction, because zero-crash segments account for more than 40% of the total data in this study.

2. Methodology

2.1. Data Preparation

Accident data were obtained from the department of revenue (DOR) and were calculated by the total accidents recorded from 2006 July to 2010 December. It contains useful traffic information, such as crash location, severity, weather condition, and segment type, and the data were geocoded into GIS databases by the PPACG (Pikes Peak Area Council of Governments).

Based on the GIS process of spatial join between whole road network and urban boundary, the road segments were classified into two categories: rural segments and urban

Table 1: ZINB models for segment crash frequency analyses and predictions.

Role	Variables	Types	Description
Dependent variable	Num_crsh	Continuous	Number of crashes (crash frequencies)
Independent variable	ADT_1000	Continuous	Average thousand daily traffic (thousand vehicles per day)
	Length	Continuous	The length of a segment (mile)
	FFS	Continuous	Free flow speed in road segment (mph)
	Daily_VMT	Continuous	Average daily Vehicle Mile Traveled (VMT)
	RorU	Discrete	Rural or urban: 0 = urban; 1 = rural
	FC	Discrete	Segment functional classification: 0 = interstate; 1 = expressway; 2 = arterial; 3 = M_Arterial
	Numberofla	Discrete	Number of lanes: 0 = 2 lanes; 1 = 4 lanes

segments. Before analyzing segment crashes, the crashes at intersections were separated from the databases. Thus, the 200-ft intersection buffers were first created, and the crashes within these intersection buffers were deleted from the segment crash analyses. Then, with a road-segment layer separated from the road network geodatabase, the crashes associated with segments needed to be further separated from all other crashes. Because these segments may have wide cross-sections, a 150-foot buffer on both sides of an arterial centerline was adopted to capture most crashes associated with the segments only. After the 150 foot buffers were created, the crashes within these buffers were selected and aggregated in their corresponding segments.

Because different categories of road facilities vary by characteristics of highway design, traffic operation, and environments, the crash data associated with a specific type of highways needed to be separated from the other types of highways. In this study, the crash risk was calculated and analyzed not only for the overall segment network, but also for interstate, expressway, principal arterial, and minor arterial, respectively. The segments belonging to other road types were excluded from these segments. The combined data set was further organized according to the following criteria.

(i) These accidents were divided into three categories: fatal, injury, and property-damage only (PDO) accounting for the accident severity.

(ii) Road segments with 2 and 4 lanes were selected, because 6 lanes segments exist in urban areas only.

(iii) ADT was calculated by 1000, because the change in crash frequency with increment of one vehicle is meaningless.

The cleaned accident data were overlaid with the GIS-based network and distributed into each segment in rural and urban areas. The segments were first analyzed and compared in terms of crash rate based on the comparative statistics of the four types of road segments. Then, ZINB models for segment crash frequency analyses and predictions were developed, in which variables are described in Table 1.

2.1.1. Zero-Inflated Negative Binomial Regression

For a Poisson crash frequency model, it assumes that the observed crash count data y_i, given the vector of covariate x_i, follows a Poisson distribution. The density function of y_i can be expressed as follows:

$$P(Y_i = y_i \mid x_i) = \frac{e^{-\mu_i}\mu_i^{y_i}}{y_i!}, \quad y_i = 0, 1, 2, \ldots, \tag{2.1}$$

where the parameter u_i, conditional mean number of events for each covariate x_i, is given by

$$\mu_i = \mathrm{Exp}(x_i'\beta), \tag{2.2}$$

where β is a $(k+1) \times 1$ parameter vector (β_0 is the coefficient for intercept, and β_1, β_1, β_2, \ldots, β_k are for k regressors).

In the Poisson regression, the conditional variance of the count variable is equal to the conditional mean as follows:

$$V(y_i \mid x_i) = E(y_i \mid x_i) = \mu_i, \tag{2.3}$$

where x_i is the covariate of road segment geometric and traffic features in each record including the intercept; u_i is the conditional mean of the crash frequency y_i. Since this assumption is contradict to the fact that the vehicle accident data are always significantly overdispersed relative to its mean, the NB regression model was developed with a heterogeneity component accounting for unobserved heterogeneity in the crash count data as follows:

$$u_i = \mathrm{Exp}(x_i\beta + \varepsilon_i) = \mathrm{Exp}(x_i\beta)\mathrm{Exp}(\varepsilon_i), \tag{2.4}$$

where β is the parameter coefficients vector to be estimated for independent variables including intercept; $\mathrm{Exp}(\varepsilon_i)$ is a heterogeneity component accounting for unobserved heterogeneity in the crash count data, which is independent of x_i. However, there is always a large density of zeros in crash count data, which cannot accurately be predicted by traditional NB models. For this situation, the zero-inflated regression models were developed in the crash frequency-related research area.

Zero-inflated count models provide a way of modeling the excess zeros in addition to allowing for overdispersion. For each road segment, there are two possible data generation processes. Process 1 is chosen with probability ω_i and process 2 with probability $1-\omega_i$. Process 1 generates only zero counts, whereas process 2 generates counts from either a poisson or a negative binomial model. In this paper, the probability ω_i depends on the geometric and traffic features of segment i, can be obtained from the logistic function F, as follows:

$$\omega_i = F(z_i'\gamma) = \Lambda(z_i'\gamma) = \frac{\exp(z_i'\gamma)}{1 + \exp(z_i'\gamma)}, \tag{2.5}$$

Table 2: Original statistics for the length and mileage.

Area	Number	Mean length	Minimum	Maximum	Std. deviation	Mileage
Rural	374	0.968	0.063	5.998	0.989	362.100
Urban	2387	0.301	0.051	3.992	0.382	719.239
Network	2761	0.392	0.051	5.998	0.525	1081.339

where z_i' is the vector of independent variables specified in the logistic regression model (road facility and traffic features) and intercept; γ is the vector of zero-inflated coefficients to be estimated.

The probability of crash frequency for segment i can be expressed as follows:

$$p(y_i \mid x_i, z_i) = \begin{cases} \omega_i + (1 - \omega_i)g(y_i \mid x_i), & y_i = 0, \\ (1 - \omega_i)g(y_i \mid x_i), & y_i > 0, \end{cases} \tag{2.6}$$

where $g(y_i \mid x_i)$ follows either Poisson distribution or NB distribution; x_i is the vector of covariates of observation i specified in the model.

In this study, ZINB models were used for regression efforts because zero-crash segments account for more than 40% of the total data.

3. Results

3.1. Comparative Statistical Analyses of Rural and Urban Traffic Safety

During the observation period of four and a half years, there were 9651 crashes occurring in the study areas, consisting of 1057 records in rural segments and 8594 records in urban segments. Among the crashes in the rural segments, there were 15 fatal and 176 injured accidents. On the other hand, 46 fatal and 1038 injury crashes happened in urban areas. Table 2 shows the descriptive statistics for rural and urban segment lengths, which indicate that average mileage of rural segments (0.968 mile) is longer than urban segments (0.293 mile) because of a lower density of intersections in rural networks. Figure 1 displays the road segment crash rate distribution, calculated as the number of crashes per 100 million VMT, where the double line is the boundary between rural and urban areas. It shows that the percentage of segments with higher crash rates within the urban region is more than rural areas.

Table 3 displays the t-test statistics of rural and urban segment comparison for different types of facilities. It shows that there is a significant difference between rural and urban in terms of crash rates using both crash per lane*miles*year and crash per 100 million VMT in 2-lane segments. The crash rates in rural segments are consistently lower than urban segments. The 2-lane expressway is exceptional mainly because of the small sample size of 2-lane rural expressway. However, there is no statistical difference between rural and urban 4-lane arterial segments.
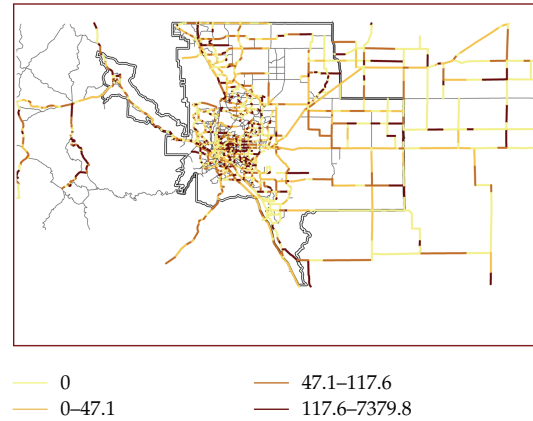
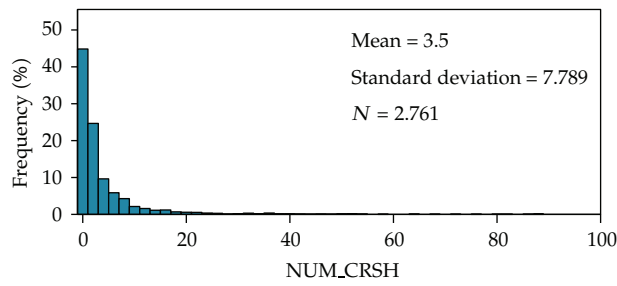**Figure 1:** Road segment crash rate distribution in terms of the number of crashes per 100 million VMT.

Figure 2: Crash frequencies distribution.

3.2. ZINB Regression Analyses

The crash frequencies distribution histogram (Figure 2) clearly illustrates that there are excessive zeros (over 40%) in the crash data. The P values in Kolmogorov-Smirnov, Cramer-von Mises, and Anderson-Darling normality tests are all less than 0.05. Therefore, it strongly supports the null hypothesis that the crash data do not follow the normal distribution. Therefore, the ZINB models are suitable to the crash count data regression analyses.

ZINB models were developed using the software SAS 9.2. We chose the crash frequency in segment (Num_crsh) as the dependent variable, and the regressors included segment length (length), number of lanes (Numberofla), thousand average annual day traffic (ADT_1000), free flow speed (FFS), and RoU (rural or urban). The segment type was not considered in this model since it was highly correlated with FFS and RoU.

Table 4 shows the parameter estimates of ZINB model for total crash frequency in segment, and only significant variables ($P < 0.05$) were included in the model. The ZINB model parameter estimates include 2 parts: NB regression and logistic regression. In the NB regression process, it can be found that the number of lanes, rural or urban, ADT, length, and FFS are all significantly correlated with the number of crashes. Further, the measure of Alpha in Table 4 is 1.435, with a P value less than 0.001, displaying a very strong overdispersion effect and indicating the superiority of the ZINB model over the zero-inflated Poisson (ZIP) model. ADT_1000 and LENGTH are positive associated with the crash frequency, suggesting that crash frequencies increase with increments of traffic volume and segment length. The

Table 3: *t*-test statistics for rural and urban segment comparison.

Road function	Lane	Variable	District	N	Mean	Standard deviation	Mean difference	P value
Interstate	2	Per_lane_ Mile	Rural	30	2.06	1.12	−2.456	0.000
			Urban	86	4.52	4.63		
		Per_100 MVMT	Rural	30	45.31	25.07	−31.83	0.001
			Urban	86	77.14	79.69		
	4	Per_lane_ Mile	Rural	1	0.00	—	−1.167	—
			Urban	4	1.16	2.21		
		Per_100 MVMT	Rural	1	0.00	—	−30.184	—
			Urban	4	30.18	56.91		
Expressway	2	Per_lane_ Mile	Rural	2	0.44	0.62	−1.873	0.074
			Urban	151	2.31	3.37		
		Per_100 MVMT	Rural	2	19.83	28.05	−55.233	0.171
			Urban	151	75.06	112.07		
	4	Per_lane_ Mile	Rural	0	—	—	—	—
			Urban	77	1.26	2.34		
		Per_100 MVMT	Rural	0	—	—	—	—
			Urban	77	43.08	74.91		
P_Arterial	2	Per_lane_ Mile	Rural	91	0.57	0.76	−0.621	0.000
			Urban	162	1.19	1.97		
		Per_100 MVMT	Rural	91	34.01	44.34	−25.837	0.006
			Urban	162	59.85	102.08		
	4	Per_lane_ Mile	Rural	9	0.54	0.44	−0.363	0.546
			Urban	592	0.90	1.79		
		Per_100 MVMT	Rural	9	23.99	19.57	−67.630	0.571
			Urban	592	91.62	357.28		
M_Arterial	2	Per_lane_ Mile	Rural	241	0.29	0.78	−0.307	0.000
			Urban	783	0.60	1.20		
		Per_100 MVMT	Rural	241	108.70	424.56	−24.454	0.444
			Urban	783	133.16	436.68		
	4	Per_lane_ Mile	Rural	0	—	—	—	—
			Urban	532	0.59	1.41		
		Per_100 MVMT	Rural	0	—	—	—	—
			Urban	532	105.56	246.18		
Total comparison								
		Per_lane_ Mile	Rural	374	0.51	0.93	−0.480	0.000
			Urban	2387	0.99	2.07		
		Per_100 MVMT	Rural	374	82.64	343.16	−20.32	0.133
			Urban	2387	102.96	331.91		

results are consistent with many previous research conclusions [7, 9, 27]. FFS is negatively associated with the crash frequency, indicating that crash frequencies are decreasing with increment of roadway free flow speed. Since FFS is correlated with the design standard of road facilities, it would be more appropriate to be explained that a better road facility with higher FFS has a lower crash rate compared to the facilities with lower FFS. In this study, FFS can be treated as a surrogate of speed limit but it can more accurately reflect the actual traffic operation status in road segments than speed limit. Previous research finding is less conclusive about the impact of speed limit on crash frequency [28]. In addition, four-lane roadways were found to be associated with a lower number of crashes than 2-lane roadways

Table 4: Parameter estimates of ZINB model for total crash frequency.

Parameter	DF	Estimate	Standard error	t value	P value
Negative binomial regression part					
Intercept	1	0.850	0.161	5.29	<0.0001
Numberofla (2 lane)	0	—	—	—	—
Numberofla (4 lane)	1	−0.438	0.107	−4.10	<0.0001
RoU (rural)	1	−0.270	0.095	−2.84	0.0045
RoU (urban)	0	—	—	—	—
ADT_1000	1	0.075	0.004	18.20	<0.0001
LENGTH	1	0.395	0.070	5.61	<0.0001
FFS	1	−0.011	0.004	−2.99	0.0028
Logistic regression part					
Inf_Intercept	1	8.214	0.913	8.99	<0.0001
Inf_ADT	1	−0.028	0.013	−2.20	0.0277
Inf_LENGTH	1	−71.713	8.739	−8.21	<0.0001
_Alpha	1	1.435	0.065	22.20	<0.0001

in this model. This is reasonable because this comparison was based on the assumption of same traffic exposure so that the segments with 4 lanes should have lower traffic volume per lane. More importantly, the urban regions appear to have a higher crash frequency than rural areas, which is consistent with the crash rate analyses results. The logistic regression part of the model predicts the likelihood of zero crash occurrences. The modeling results reveal that the variables of ADT_1000 and LENGTH are significant in estimating the probability of segments belonging to the zero crash occurrence group. According to the parameter coefficients estimated, the higher the traffic exposure (thousand of AADT and segment length), the lower the possibility of zero crash occurrences, which is consistent with all the previous study conclusions.

Furthermore, Table 5 shows the parameter estimates of ZINB model for injury and fatal crash frequency in a segment (Alpha is 1.074, with a P value less than 0.001). The NB regression indicates that Numberofla, RoU, ADT_1000, and LENGTH are significant variables to predict injury and fatal crash frequency, which displays a very similar result to that for total crash frequency except for FFS. It implies that although the better road facilities with higher FFS benefit from high standard design features resulting in a lower total crash frequency (as shown in Table 4), they would not mitigate the severe crash risk. A previous study reported that by controlling the other factors, purely increasing operation speed in road segments by 1% would approximately result in 2% increment in injury crash rate and 4% increment in fatal crash rate [29]. On the other hand, compared to the total crash frequency model, the logistic regression results for injury and fatal crash frequency model are quite different though the effect of LENGTH keeps similarity. First, the number of lanes is a significant variable for estimating the probability of zero injury and fatal crash occurrence in segment. Compared to 2-lane roads, the 4-lane roads have a lower severe crash frequency but have a lower probability of zero crashes. A possible explanation is that changing lane maneuver in 4-lane segments would increase the severe crash risk. Second, the effect of ADT_1000 in the Logistic regression of injury and fatal crash model is reverse from the total crash model. It shows that as traffic volume increases, the likelihood of zero severe crashes decreases. This interesting

Table 5: Parameter estimates of ZINB model for injury and fatal crash frequency.

Parameter	DF	Estimate	Standard error	t value	P value
		Negative binomial regression part			
Intercept	1	−1.272	0.093	−13.73	<0.0001
Numberofla (2 lane)	0	—	—	—	—
Numberofla (4 lane)	1	−0.398	0.098	−4.08	<0.0001
RoU (rural)	1	−0.232	0.118	−1.96	0.0496
RoU (urban)	0	—	—	—	—
ADT_1000	1	0.060	0.004	15.58	<0.0001
LENGTH	1	0.338	0.068	5.00	<0.0001
		Logistic regression part			
Inf_Intercept	1	5.404	0.850	6.35	<0.0001
Numberofla (2 lane)	0	—	—	—	—
Numberofla (4 lane)	1	−0.967	0.352	−2.75	0.0060
Inf_ADT_1000	1	0.035	0.016	2.21	0.0268
Inf_LENGTH	1	−38.325	6.436	−5.95	<0.0001
_Alpha	1	1.073	0.124	8.64	<0.0001

finding is consistent with the previous conclusion in a crash severity study, which explains that lower ADT could mean higher speeds that more often lead to severe/fatal crashes [30].

4. Conclusion and Discussions

There have been numerous studies to clarify the role of rural or urban settings in segment safety, but it was still controversial to make a conclusion. Before reaching the common agreement on the difference between rural and urban traffic safety, it is important to clarify the definition of "rural." Generally, to distinguish from urban environments, rural areas have the attributes associated with demographic features (e.g., low population size and density, outside boundary of urban area), economic statues (low economic indicators, farming, and agriculture), social structure (e.g., intimate, informal, and homogeneous forms of social interaction, limited social resources), cultural characteristics (e.g., traditional, conservative, provincial, slow to change), and so forth. The above features are often used to explain the statistical fact that the death rate from many common causes in US is significantly higher in rural compared to urban areas [1, 6], as well as in different countries [31–33].

However, these thresholds should not be universally applied to make local transportation safety analyses. For many developed regions, although districts are clearly separated into rural and urban regions according to their demographic, economic, or social attributes, the transportation facilities are well connected to each other and formed more standardized road networks. Thus, it was reported that there are relatively high numbers of crashes in urban regions because the heavy traffic volume and complex driving environments in urban lead to more conflicts between vehicles [34]. Therefore, for a specific safety evaluation project, this study supports the argument that more detailed crash risk comparisons between rural and urban transportation road segments should be performed at a comparable level. In this paper, the crash rate comparison and ZINB regression for both total crash frequency and injury and fatal crash frequency in road segment were conducted to discriminate between rural and urban traffic safety. It was found that compared to urban areas, the measures for

traffic safety in rural areas show lower crash rates, total crash frequencies, and injury and fatal crash frequencies. The results based on the ZINB regression models also showed the following.

(i) Segment crash frequencies increase as traffic volume and segment length increase. However, higher traffic volume will lower the likelihood of severe crash occurrence.

(ii) Compared to 2-lane roads, the 4-lane roads have a lower crash frequency but have a higher probability of severe crash occurrence.

(iii) Better road facilities with higher free flow speed benefit from high standard design feature resulting in a lower total crash frequency but would not mitigate the severe crash risk.

Finally, it can be concluded that in the research area traffic safety of rural segments is better than urban segments, which implies that a priority for traffic safety improvement should be put on the urban highway segments.

Acknowledgments

References

[1] M. S. Eberhardt, D. D. Ingram, and D. M. Makuc, *Urban and Rural Health Chart-Book. Health, United States, 2001*, National Center for Health Statistics, Hyattsville, Md, USA, 2001.

[2] L. Kmet, P. Brasher, and C. Macarthur, "A small area study of motor vehicle crash fatalities in Alberta, Canada," *Accident Analysis and Prevention*, vol. 35, no. 2, pp. 177–182, 2003.

[3] R. P. Gonzalez, G. R. Cummings, H. A. Phelan, S. Harlin, M. Mulekar, and C. B. Rodning, "Increased rural vehicular mortality rates: roadways with higher speed limits or excessive vehicular speed?" *The Journal of Trauma*, vol. 63, no. 6, pp. 1360–1363, 2007.

[4] L. H. Brown, A. Khanna, and R. C. Hunt, "Rural versus urban motor vehicle crash death rates: 20 Years of FARS data," *Prehospital Emergency Care*, vol. 4, no. 1, pp. 7–13, 2000.

[5] D. A. Borgialli, E. M. Hill, R. F. Maio, C. P. Compton, and M. A. Gregor, "Effects of alcohol on the geographic variation of driver fatalities in motor vehicle crashes," *Academic Emergency Medicine*, vol. 7, no. 1, pp. 7–13, 2000.

[6] M. E. Rakauskas, N. J. Ward, and S. G. Gerberich, "Identification of differences between rural and urban safety cultures," *Accident Analysis and Prevention*, vol. 41, no. 5, pp. 931–937, 2009.

[7] M. A. Abdel-Aty and A. E. Radwan, "Modeling traffic accident occurrence and involvement," *Accident Analysis and Prevention*, vol. 32, no. 5, pp. 633–642, 2000.

[8] C. Zwerling, C. Peek-Asa, P. S. Whitten, S. W. Choi, N. L. Sprince, and M. P. Jones, "Fatal motor vehicle crashes in rural and urban areas: decomposing rates into contributing factors," *Injury Prevention*, vol. 11, no. 1, pp. 24–28, 2005.

[9] X. Jiang, X. Yan, B. Huang, and S. H. Richards, "Influence of curbs on traffic crash frequency on high-speed roadways," *Traffic Injury Prevention*, vol. 12, no. 4, pp. 412–421, 2011.

[10] M. Leff, L. Stallones, T. J. Keefe, R. Rosenblatt, and M. Reeds, "Comparison of urban and rural non-fatal injury: the results of a statewide survey," *Injury Prevention*, vol. 9, no. 4, pp. 332–337, 2003.

[11] L. Aarts and I. van Schagen, "Driving speed and the risk of road crashes: a review," *Accident Analysis and Prevention*, vol. 38, no. 2, pp. 215–224, 2006.

[12] C. Y. Chan, B. Huang, X. Yan, and S. Richards, "Effects of asphalt pavement conditions on traffic accidents in tennessee utilizing pavement management system," in *Proceedings of the 88th TRB Annual Meeting*, Washington, DC, USA, 2009.

[13] P. J. Cooper, "The relationship between speeding behaviour (as measured by violation convictions) and crash involvement," *Journal of Safety Research*, vol. 28, no. 2, pp. 83–95, 1997.

[14] N. J. Garber and A. A. Ehrhart, "Effect of speed, flow, and geometric characteristics on crash frequency for two-lane highways," *Transportation Research Record*, vol. 1717, pp. 76–83, 2000.

[15] C. N. Kloeden, A. J. McLean, and G. Glonek, "Reanalysis of travelling speed and the rate of crash involvement in Adelaide South Australia," Tech. Rep. CR 207, Australian Transport Safety Bureau ATSB, Canberra, Australia, 2002.

[16] J. Milton and F. Mannering, "The relationship among highway geometrics, traffic-related elements and motor-vehicle accident frequencies," *Transportation*, vol. 25, no. 2–4, pp. 395–413, 1998.

[17] H. Renski, A. J. Khattak, and F. M. Council, "Effect of speed limit increases on crash injury severity: analysis of single-vehicle crashes on North Carolina Interstate highways," *Transportation Research Record*, no. 1665, pp. 100–108, 1999.

[18] S. P. Miaou, "The relationship between truck accidents and geometric design of road sections: poisson versus negative binomial regressions," *Accident Analysis and Prevention*, vol. 26, no. 4, pp. 471–482, 1994.

[19] H. Huang and H. C. Chin, "Modeling road traffic crashes with zero-inflation and site-specific random effects," *Statistical Methods & Applications*, vol. 19, no. 3, pp. 445–462, 2010.

[20] V. N. Shankar, F. L. Mannering, and W. Barfield, "Effect of roadway geometrics and environmental factors on rural freeway accident frequencies," *Accident Analysis and Prevention*, vol. 27, no. 3, pp. 371–389, 1995.

[21] A. Vogt and J. Bared, "Accident models for two-lane rural segments and intersections," *Transportation Research Record*, vol. 1635, pp. 18–29, 1998.

[22] J. Carson and F. Mannering, "The effect of ice warning signs on ice-accident frequencies and severities," *Accident Analysis and Prevention*, vol. 33, no. 1, pp. 99–109, 2001.

[23] D. Lord, S. P. Washington, and J. N. Ivan, "Poisson, poisson-gamma and zero-inflated regression models of motor vehicle crashes: balancing statistical fit and theory," *Accident Analysis and Prevention*, vol. 37, no. 1, pp. 35–46, 2005.

[24] D. Fletcher, "Geographic information systems for transportation: a look forward," in *Transportation in the New Millenium: State of the Art and Future Directions, Transportation Research Board*, CDROM, Washington, DC, USA, 2000.

[25] X. Qiaojun, P. Fuquan, L. Jian, Z. Guoqiang, and W. C. Eidson, "GIS-based traffic safety management for highway intersections," in *Proceedings of the 1st International Symposium on Transportation and Development Innovative Best Practices (TDIBP '08)*, pp. 198–203, American Society of Civil Engineers, Beijing, China, April 2008.

[26] S. Fang, Z. Yang, and Y. Chen, "Design of GIS-based road safety information management system," *Journal of Tongji University*, vol. 34, no. 5, pp. 629–633, 2006.

[27] B. Persaud and L. Dzbik, "Accident prediction models for freeways," *Transportation Research Record*, vol. 1401, pp. 55–60, 1993.

[28] B. N. Fildes and S. J. Lee, "The speed review: road environment, behavior, speed limits, enforcement and crashes," Tech. Rep. CR 127 (FORS) CR 3/93 (RSB), Federal Office of Road Safety and the Road Safety Bureau, Roads and Traffic Authority of New South Wales, New South Wales, Australia, 1993.

[29] G. Nilsson, *Traffic Safety Dimensions and the Power Model to Describe the Effect of Speed on Safety*, vol. 221 of *Bulletin*, Lund Institute of Technology, Lund, Sweden, 2004.

[30] A. Das, M. Abdel-Aty, and A. Pande, "Using conditional inference forests to identify the factors affecting crash severity on arterial corridors," *Journal of Safety Research*, vol. 40, no. 4, pp. 317–327, 2009.

[31] B. P. Y. Loo, W. S. Cheung, and S. Yao, "The rural-urban divide in road safety: the case of China," *The Open Transportation Journal*, vol. 5, pp. 9–20, 2011.

[32] M. Y. Renwick, G. G. Olsen, and M. S. Tyrrell, "Road fatalities in rural New South Wales. Weighing the causes," *Medical Journal of Australia*, vol. 1, no. 7, pp. 291–294, 1982.

[33] J. P. Thouez, M. F. Joly, A. Rannou, Y. Bussiere, and R. Bourbeau, "Geographical variations of motor-vehicle injuries in Quebec, 1983–1988," *Social Science and Medicine*, vol. 33, no. 4, pp. 415–421, 1991.

[34] S. Eiksund, "A geographical perspective on driving attitudes and behaviour among young adults in urban and rural Norway," *Safety Science*, vol. 47, no. 4, pp. 529–536, 2009.