

Research Article

Exploring Factors Contributing to Crash Injury Severity in the Rural-Urban Fringe of the Central City

Biao Wu ¹, Xingyu Wang,¹ Tuo Liu,¹ Naibao Dong,² and Yun Li ³

¹School of Automobile and Traffic Engineering, Heilongjiang Institute of Technology, Harbin 150050, China

²Research Center of Transportation Information and Science of Heilongjiang Province, Harbin 150080, China

³School of Automobile and Transportation Engineering, Guangdong Polytechnic Normal University, Guangzhou 510665, China

Correspondence should be addressed to Yun Li; lyhyit@163.com

Received 12 May 2021; Revised 31 July 2021; Accepted 14 August 2021; Published 30 August 2021

Academic Editor: Xinqiang Chen

Copyright © 2021 Biao Wu et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

To analyze the risk factors influencing the crash injury severity in rural-urban fringes, crash data in rural-urban fringes were collected from Harbin, China. Four risk factors, namely, time of day, vehicle type, road feature, and crash type, were investigated associated with the severity of rural-urban fringe crashes. The crash injury severity was divided into two categories, including fatal and nonfatal crash. The logistic regression was applied to explore the relationships between the severity outcomes and time of day, vehicle type, road feature, and crash type. The test methods of goodness-of-fit and badness-of-fit are conducted to examine the validity of estimation results. The results show considerable matching of the number of different crash types between calculated results and actual data. Compared with the other influencing factors, the time of day is significant factor for crash injury severity based on the study. As such, the proposed calibration procedure and the factors of choice are recommended as a validated approach to analyze and identify the main factors influencing crash injury severity in rural-urban fringes.

1. Introduction

With the rapid growth of China's economy, the development level of urbanization is accelerating. By the end of 2020, the urbanization rate is more than 60% in China, which is about 16% higher than that in 2006. The acceleration of urbanization leads to the gradual expansion of economic and social activities to the periphery of the city and the expansion of the scope of urban built-up areas, thus forming the transitional zone of urban-rural fringe. Rural-urban fringes refer to the transitional zones of urban and rural areas with the nature of urban and rural land use. As an important part of the road network, rural-urban fringes have become inclined to be involved in more traffic crashes with multiple interference points and serious mixed traffic problems, etc. [1, 2]. Rural-urban fringes crash not only comprise a substantial portion of traffic crashes but also comprise a considerable portion of traffic crash casualties. To reduce traffic conflicts in rural-urban fringes and prevent them from becoming crash-prone locations, why and how traffic crashes occur in rural-urban fringes should be addressed.

For the rural-urban fringe roads, they are obviously different from urban roads and highways in road traffic characteristics. Rural-urban fringe roads carry the urban transit traffic, the internal traffic in the urban-rural fringe, and the arrival and departure traffic between urban and rural-urban fringe. Due to the imperfect infrastructure, weak traffic management and serious mixed traffic flows on rural-urban fringe roads impose significant safety issues. More than 50% of the total crashes are related to the rural-urban fringe roads in China [3]. However, much attention and effort have been paid to exploring factors contributing to crash injury severity from various roadway entities, focusing on urban roads [4] and highways [5, 6]. Few studies have been conducted to investigate the factors influencing crash injury severity in rural-urban fringes.

Because the traffic characteristics in rural-urban fringes are different from inside the city, the time of day, vehicle type, road feature, and crash type are mainly considered as influencing factors of crash injury severity in this study. Methods for identifying major factors are discussed and validated using a logistic regression model based on field

data. The contribution of this study is to identify the obvious risk factors to put forward targeted improvement measures.

2. Literature Review

There must be good knowledge of the risk factors contributing to crash injury severity to improve rural-urban fringe safety. To determine an appropriate model on exploring factors contributing to crash injury severity in rural-urban fringes, it is carried out by reviewing the findings of previous studies, focusing on the safety of rural-urban fringes, the risk factors influencing rural-urban fringe safety, and the modelling approach for crash injury severity in rural-urban fringes used by the researchers. The following section describes the findings from previous studies in this regard.

2.1. Relevant Studies on Safety of Rural-Urban Fringes. Many studies have focused on the safety of the rural-urban fringe. Liu et al. [7] put forward some countermeasures to improve traffic safety based on developing the law of rural-urban fringe crashes. Wei et al. [8] explored the causes of high-risk traffic safety areas in rural-urban fringe from the perspective of highway traffic safety service level, taking Tongzhou district in Beijing City as an example. Wei and Mou [9] established the traffic safety evaluation system of urban-rural fringe based on the attribute recognition model. Zhang et al. [10] discussed the conflict and safety between pedestrian and vehicle in rural-urban fringe using the indicators of the time-to-collisions (TTC) and post-encroachment time (PET). Zhang et al. [11] proposed a speed control section division scheme based on the sequential cluster method according to the spot speed distribution characteristics of low-grade roads in urban-rural fringe. Zhang and Shao [12] established a multiobjective optimal model for signal timing under the comprehensive consideration of several factors, such as the delay, capacity, number of stops, and vehicle emissions.

2.2. Relevant Studies on Influencing Factors to Rural-Urban Fringe Safety. Existing research results have conducted preliminary studies on identifying the risk factors that affect the crash injury severity in rural-urban fringe. For instance, Li [13] revealed the factors influencing the probability of traffic crashes in rural-urban fringe in China by questionnaire survey and established multivariate logistic regression models for multiple unsafe driving behaviours. He found that the weather, road features, crash types, and time of day contribute to severe rural-urban fringe crashes. Niu et al. [14] used a polytomous ordered logit model to analyze the factors contributing to urban-rural fringe crashes in China. They found that the drivers' age, gender, crash type, weather, and road feature had a great impact on the severity of rural-urban fringe crashes.

2.3. Relevant Studies on the Modelling Approach for Crash Injury Severity. So far, many research methods have been developed to investigate risk analysis and prediction in

previous studies [15–17], and statistical regression approaches and methods based on machine learning (ML) have been the primary method on investigating the relationship between crash injury severity and risk factors [18, 19]. For example, Xie et al. developed a random-parameter ordered probit model to explore risk factors with crash severity on two-lane rural roads in China [20]. Pervez et al. established a random-parameter logit model to examine the factors associated with the motorcycle injury severity [21]. Zeng and Huang applied Artificial Neural Networks (ANN) to predict crash injury severity [22]. Tang et al. proposed a stacking framework combining Random Forests (RF), Adaptive Boosting (Adaboost), Gradient Boosting Decision Tree (GBDT), and logistic regression model to predict the crash injury severity at freeway diverge areas [23].

Logistic regression has proven to be an effective and reliable method to explore the relationship between the response and explanatory variables in the field of traffic crash. Zhang et al. [24] conducted a population-based cross-sectional study to examine the factors affecting the severity of motor vehicle traffic crashes (MVTCS) involving elderly drivers in Ontario by logistic regression. Al-Ghamdi [25] applied the logistic regression to crash-related data collected from traffic police records to examine the contribution of several variables to crash severity in Riyadh. Yau [26] used the stepwise logistic regression models to identify the factors affecting the severity of single-vehicle traffic crashes in Hong Kong. Yan et al. [27] developed a multiple logistic regression model to analyze the characteristics of rear-end crashes at signalized intersections in Florida. Sze and Wong [28] developed a binary logistic regression to reveal the associations between the probability of fatality and severe injury and all contributory factors. Harb et al. [29] created a multiple and conditional logistic regression model to identify factors contributing to freeway work-zone crashes in Florida. Ma et al. [30] used the logistic regression model to analyze the impact of crash time, collision type, weather, and daily standard passenger car traffic volume to AADT ratio on the serious situation of highway tunnel traffic crashes. Chen et al. [31] applied logistic regression to reveal the significant risk factors affecting the severity of intersection crashes in Victoria. Feng et al. [32] used the binary logistic model to study the impact of 10 factors such as crash period on the severity of traffic crash of ring expressway. Zhang et al. [33] analyzed the influence of weather conditions, vehicle types, plane line type, and other factors on the crash injury severity in continuous downhill sections in mountainous areas. Ji et al. [34] used the severity of traffic crashes as the dependent variable and time, traffic operating environment, and traffic crash participants as independent variables and established a model on traffic crashes severity based on the logistic model. Rudisill et al. [35] used logistic regression model to explore the effect of driver age, collision location, and other factors on the cause of traffic crashes. Wang et al. [36] applied a classification tree-based logistic regression model to identify the risk factors affecting crash injury severity for different types of e-bike riders.

In summary, the existing literature focused on exploring factors on the severity of rural-urban fringe crashes in China is limited. Using a categorical data technique to develop the factors contributing to the crash injury severity on rural-urban fringe roads is relatively rare. Considering the influencing factors to the severity of rural-urban fringe crashes in China would be valuable to better understand the issues on traffic safety. Therefore, this paper applies the logistic regression model to reveal the potential risk factors affecting the severity of rural-urban fringe crashes in Harbin, China. The risk factors included temporal characteristics, vehicle features, road factors, and crash characteristics. The results will provide valuable information to assist road safety stakeholders in developing appropriate countermeasures for severe and fatal crashes in rural-urban fringes.

3. Study Area and Data Collection

3.1. Study Area. The characteristics of nature, society, and ecology are special and the situation of land use is complex in rural-urban fringe areas [37]. According to previous research efforts [38], this paper determines the scope of the rural-urban fringes as follows: taking the outermost circular highway of the city as the inner boundary and the closed curve formed by the administrative outer boundary of towns connected with urban built-up areas as the outer boundary, this specific area is divided into the rural-urban fringe, as shown in Figure 1.

This study uses data collected in Harbin City, China, which involves 14 towns and 138 villages as depicted in Figure 2. As can be seen, the rural-urban fringes in Harbin City include Qulin Rural Area, Songbei Town, Songpu Town, Xingfu Town, Tuanjie Town, Wanggang Town, Liming Town, Hulan street, Minzhu Town, Wanbao Town, Xinfa Town, Yushu Town, Chenggaozi Town, and Chaoyang Town.

3.2. Data Collection. To analyze the crash injury severity in rural-urban fringe, we collected crash data from Harbin Public Security Bureau Traffic Police Station, and the crash file contains the crash time, location, vehicle types, road features, driver information, passenger information, crash injury severity, and other environmental variables.

Firstly, the crashes sample data is preprocessed, treating crashes samples with large amounts of missing data as invalid samples. Traffic crashes for rural-urban fringe roads in 2014–2017 were obtained, including a total of 722 crash data points that occurred. To ensure the correctness of the analysis results, the duplicate and unknown data points were removed. Finally, 661 crash data points with complete information were used to analyze the characteristics of traffic crashes in this paper. The valid crash sample size is generally recommended above 10 times the number of explanatory variables [39], and the potential explanatory variables of the proposed model include 4 factors of vehicle, road, environment, and crash, so the samples of 667 valid crashes meet the modelling requirements. The frequency of the two categories in 2014–2017 is given in Table 1, respectively.

4. Variable Selection

4.1. Explanatory Variables. In terms of comprehensive and practical principles, the explanatory variables in establishing the cash risk measurement model were mainly selected. These variables should be independent of each other and comprehensive enough to accommodate all aspects related to the occurrence of crashes. Besides, the selected variables should be suitable for collected data. Therefore, four potentially influencing factors were selected from four aspects, vehicle, road, environment, and crash, according to the roadway characteristics and crash information of the rural-urban fringe roads. Among them, vehicle characteristics include the composition of the crash vehicle; road attributes include straight road section, curved road section, and intersection; driving environment attributes include the time of day; and the crash attribute is in the form of the crash including collision, vehicle-pedestrian crash, rear-end crash, rollover, and collision fixtures. Table 2 shows the descriptive statistics of explanatory variables in 2014–2017.

Referring to the existing research results and considering the characteristics of a traffic crash in urban-rural fringes, four risk factors are selected as the specific variables, including time of day, vehicle types, road features, and crash types, as shown in Table 3.

The four influencing factors in Table 3 are classified variables after being assigned, so dummy variables should be assigned. If the dummy variable has k categories, the dummy variable is converted to $k-1$ variables. The factors including time of day, vehicle types, road features, and crash types are treated as dummy variables. The specific results are shown in Table 4.

4.2. Definition of Crash Injury Severity. Based on the “Law of the People’s Republic of China on Road Traffic Safety” in 2004 (modified in 2011), the crash injury severity can be divided into four levels, namely, minor crashes, general crashes, serious crashes, and malignant crashes in China. The criteria for the classification of serious crashes and malignant crashes are divided according to the number of personal deaths; that is, the division of crashes is based on the severity. Therefore, in this paper, the crash injury severity was divided into two levels, fatal and nonfatal crash concerning the standards, as seen in Table 5.

5. Methodology

5.1. Principle of the Logistic Model. Logistic regression is a probabilistic nonlinear regression model, which is a multivariate analysis method to explore the relationship between explanatory variables and a discrete response variable. Considering the conditional probability $p(y = 1/x) = p$ of the vector $x = (x_1, x_2, \dots, x_n)$ with n independent variables as the probability of an event occurrence according to the observed measurement, the logistic regression model can be expressed as

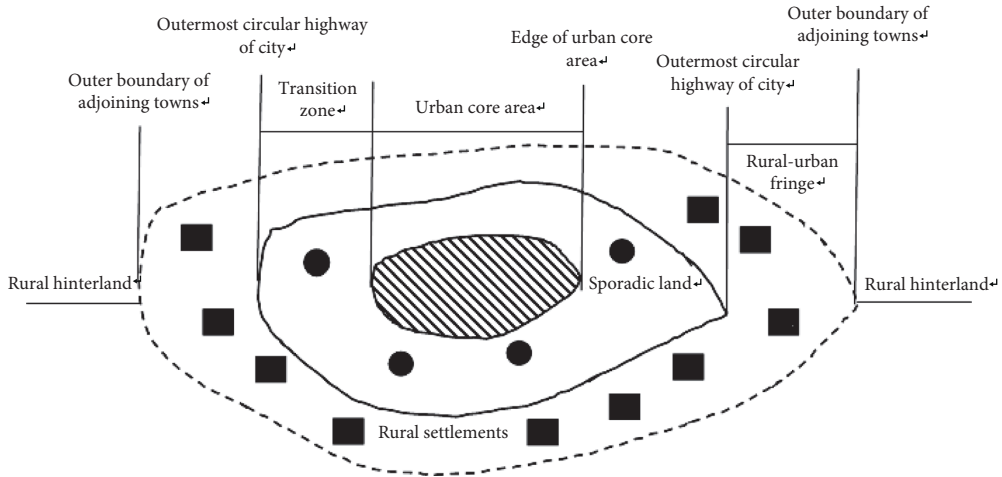


FIGURE 1: Schematic diagram of rural-urban fringes.

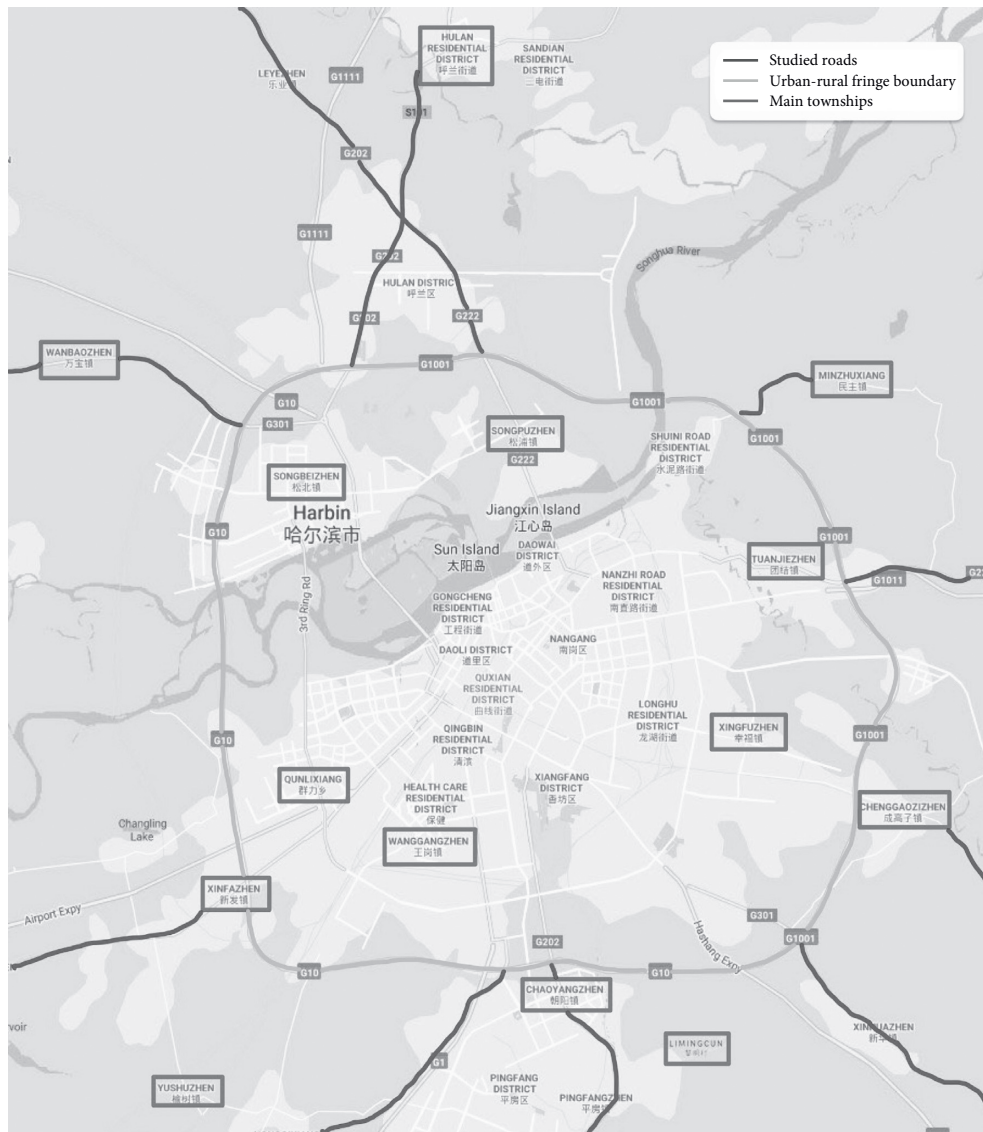


FIGURE 2: Geographical location and surrounding environment of research subjects.

$$p(y = 1/x) = \pi(x) = \frac{1}{1 + e^{-g(x)}} \quad (1)$$

5.2. Binary Logistic Regression Model Building. In this paper, the crash injury severity is taken as the response variable, which is a binary variable. Therefore, a binary logistic regression is used to test the relationship between the response variable and the related potential factors and to rank the relative importance of the explanatory variables. Binary logistic regression is used since the response variable Y (crash injury severity) can only take on two values: $Y=1$ for the fatal crash and $Y=0$ for the nonfatal crash. The probability that a fatal crash in rural-urban fringe will occur or not is modelled as logistic distribution in

$$p = p(y_i = 1|x_1, x_2, \dots, x_m) = \frac{\exp(\beta_0 + \beta_1 x_1 + \beta_2 x_2 + \dots + \beta_m x_m)}{1 + \exp(\beta_0 + \beta_1 x_1 + \beta_2 x_2 + \dots + \beta_m x_m)} \quad (2)$$

where p is the probability of fatal crashes in rural-urban fringe, x_i is the explanatory variable, β_0 is a constant term, and $\beta_1 \sim \beta_m$ is the model coefficient.

5.3. Model Checking. To analyze the fitting effect of the logistic regression model, it is necessary to test the goodness-of-fit and badness-of-fit of this model. The goodness-of-fit test is to convert the natural logarithm of the likelihood ratio function into Chi-square value and then conduct a significance test by observing Chi-square distribution [40]. It is shown that the model works when the Chi-square value is greater than the given significance level. The method of Hosmer–Lemeshow test is used for badness-of-fit, which takes Chi-square distribution as the standard, and the calculated Chi-square value needs to be lower than the critical value and greater than the significant level.

6. Results

In this paper, 661 traffic crashes are used as the modelling data, which occurred on the urban-rural fringe roads in Harbin City from 2014 to 2017.

$$p = \frac{\exp(0.159 - 1.376x_3(3) - 0.9x_3(4) - 1.497x_3(5) - 0.716x_3(6) - 0.74x_3(7))}{1 + \exp(0.159 - 1.376x_3(3) - 0.9x_3(4) - 1.497x_3(5) - 0.716x_3(6) - 0.74x_3(7))} \quad (3)$$

where $x_3(3)$ is 06:00–08:59, $x_3(4)$ is 09:00–11:59, $x_3(5)$ is 12:00–14:59, $x_3(6)$ is 15:00–17:59, and $x_3(7)$ is 18:00–20:59.

According to formula (3), the model coefficients are negative, which indicates that the probability of fatal crashes is small in the time interval of 06:00–20:59. From Table 6, the probability of fatal crashes is 25.3% in the time interval of 06:00–08:59, 40.7% in the time interval of 09:00–11:59, 22.4% in the time interval of 12:00–14:59, 48.9% in the time

TABLE 1: Frequency distribution of crash injury severity category.

Year	Fatal		Nonfatal		Total
	Frequency	Percentage	Frequency	Percentage	
2014	36	29.27	87	70.73	123
2015	44	28.76	109	71.24	153
2016	65	33.51	129	66.49	194
2017	53	27.75	138	72.25	191

6.1. Model Explanatory Variable Selection. The statistical analysis software of SPSS 22.0 is used to estimate the parameter of binary logistic regression. The input method is applied for iterative analysis and the iterative condition is selected as significantly less than 0.05. A total of 4 steps of iteration are carried out to obtain the corresponding logistic regression model. Finally, the factor of time of day is selected as the explanatory variable, as shown in Table 6.

6.2. Model Fitting Test. Based on the methods of goodness-of-fit and badness-of-fit test, the specific values of the fitting degree test are obtained, as shown in Table 7.

According to Table 7, the calculated value of the likelihood ratio Chi-square is 20.559, which is greater than the critical value of Chi-square 9.488, and the Sig. value is 0.001, less than the significance level of 0.05, indicating that the model has high goodness of fit. However, the test value of Hosmer–Lemeshow is 7.025, less than the critical value of Chi-square 9.488, and the Sig. value is 0.534, which is greater than the significance level of 0.05, indicating that the model has low badness of fit.

6.3. Analysis of Model Results. According to the above results, a model on the factor of crash time contributing to crash injury severity in urban-rural fringe is established. According to formula (2), the probability of a fatal crash in urban-rural fringes is

interval of 15:00–17:59, and 47.7% in the time interval of 18:00–20:59. The analysis results are consistent with the temporal characteristics of rural-urban fringe crashes.

6.4. Model Validation. To validate the universality of the calibrated model for crash injury severity, 120 rural-urban fringe crashes in Harbin City from 2014 to 2017 were selected to verify (3). The results are shown in Table 8.

TABLE 2: Descriptive statistics of explanatory variables.

Variables	Fatal		Nonfatal		
	Frequency	Percentage	Frequency	Percentage	
Time of day	00:00–02:59	5	2.53	17	3.67
	03:00–05:59	25	12.63	41	8.86
	06:00–08:59	25	12.63	71	15.33
	09:00–11:59	33	16.67	70	15.12
	12:00–14:59	31	15.66	59	12.74
	15:00–17:59	35	17.68	95	20.52
	18:00–20:59	30	15.15	77	16.63
	21:00–23:59	14	7.07	33	7.13
Vehicle type	Large passenger vehicle	7	3.54	11	2.38
	Small passenger vehicle	24	12.12	58	12.53
	Car	71	35.86	192	41.47
	Heavy vehicle	40	20.20	77	16.63
	Motorcycle	27	13.64	53	11.45
	Other	29	14.65	72	15.55
Road features	Straight road section	156	78.79	346	74.73
	Curved road section	8	4.04	28	6.05
	Intersection	34	17.17	89	19.22
Crash type	Collision	103	52.02	297	64.15
	Vehicle-pedestrian crash	58	29.29	113	24.41
	Rear-end crash	26	13.13	32	6.91
	Rollover	3	1.52	6	1.30
	Collision fixtures	8	4.04	15	3.24

TABLE 3: Variable assignment.

No.	Influence factor	Variables	Assignment and meaning
1	Time of day	x_1	(00:00–02:59) = 1, (03:00–05:59) = 2, (06:00–08:59) = 3, (09:00–11:59) = 4, (12:00–14:59) = 5, (15:00–17:59) = 6, (18:00–20:59) = 7, (21:00–23:59) = 8
2	Vehicle type	x_2	Large passenger vehicle = 1, small passenger vehicle = 2, car = 3, heavy vehicle = 4, motorcycle = 5, other = 6
3	Road feature	x_4	Straight road section = 1, curved road section = 2, intersection = 3
4	Crash type	x_3	Collision = 1, vehicle-pedestrian crash = 2, rear-end crash = 3, rollover = 4, collision fixture = 5

The estimated results shown in Table 8 show close values as actual crash data. The absolute number difference is only 2, and the absolute error is 1.7%, indicating that the model has high universality.

7. Discussion

Based on the analysis results of the logistic regression model, we found that the “time of day” was the significant factor contributing to crash injury severity in the rural-urban fringe. It should be pointed out that the result is consistent with an existing study (e.g., Li, 2018) according to which the time of day exhibits a positive effect on severe crashes on rural-urban fringe roads. Hence, the temporal characteristic can be analyzed in this section.

The 24-hour variation of traffic crashes and the time distribution of casualties in rural-urban fringe from 2014 to 2017 are shown in Figures 3 and 4, respectively.

It can be clearly seen from Figure 3 that the number of rural-urban fringe crashes accords with “bimodal distribution” in temporal characteristics mainly focusing on 7:

00–11:59 and 17:00–19:59. These two peak periods have a total of 8 hours, accounting for 33.33% of the whole day. The number of crashes accounts for 47.14% of the total number of traffic crashes in the whole day. A large traffic volume in the morning and evening peak hours is the primary cause of the high incidence of rural-urban fringe crashes. During the time interval of 17:00–19:59, the light changes from bright to dark, the sight distance of drivers becomes smaller, and the judgment ability decreases, which is easy to induce traffic crashes.

Figure 4 shows that the time distribution of traffic crashes casualties in rural-urban fringes is consistent with the time distribution of the number of crashes. The number of traffic crashes injuries was the highest at 7:00–7:59, 9:00–9:59, 18:00–18:59, and 19:00–19:59, which were 62, 57, 55, and 56, respectively. The injury rate of traffic crashes was 1.77 (the highest) at 7:00–7:59, and that was 1.08 persons/per time at 19:00–19:59. The death toll of traffic crashes was 23 persons, the highest number at 19:00–19:59, and 0.78 persons/per time at 4–5 (the highest). In addition, 55.23% of the fatal crashes occurred between 17:00 and 05:00.

TABLE 4: Assign transform of dummy variables.

Influence factor	Category	Variable 1	Variable 2	Variable 3	Variable 4	Variable 5	Variable 6	Variable 7	Variable 8
Time of day	00:00–02:59	1	0	0	0	0	0	0	0
	03:00–05:59	0	1	0	0	0	0	0	0
	06:00–08:59	0	0	1	0	0	0	0	0
	09:00–11:59	0	0	0	1	0	0	0	0
	12:00–14:59	0	0	0	0	1	0	0	0
	15:00–17:59	0	0	0	0	0	1	0	0
	18:00–20:59	0	0	0	0	0	0	1	0
	21:00–23:59	0	0	0	0	0	0	0	1
Vehicle type	Large passenger vehicle	1	0	0	0	0	—	—	—
	Small passenger vehicle	0	1	0	0	0	—	—	—
	Car	0	0	1	0	0	—	—	—
	Heavy vehicle	0	0	0	1	0	—	—	—
	Motorcycle	0	0	0	0	1	—	—	—
	Other	0	0	0	0	0	—	—	—
Road features	Straight road section	1	0	—	—	—	—	—	—
	Curved road section	0	1	—	—	—	—	—	—
	Intersection	0	0	—	—	—	—	—	—
Crash type	Collision	1	0	0	0	—	—	—	—
	Vehicle-pedestrian crash	0	1	0	0	—	—	—	—
	Rear-end crash	0	0	1	0	—	—	—	—
	Rollover	0	0	0	1	—	—	—	—
	Collision fixtures	0	0	0	0	—	—	—	—

TABLE 5: Classification of crash injury severity in urban-rural fringes.

Crash severity level	Crash injury severity	Criteria for the classification
I	Fatal	Personnel deaths during the period of validity, indicating that the crash is more serious
II	Nonfatal	Personnel injuries and property losses, indicating that the crash is relatively light

TABLE 6: Estimation result of parameters.

Time of day	Parameter value	Standard error	Wald value	Sig. value	Exp (B)	Confidence interval at 95% exp (B)	
						Lower limit	Upper limit
06:00–08:59	-1.376	0.392	12.323	0.000	0.253	0.117	0.545
09:00–11:59	-0.900	0.368	5.975	0.015	0.407	0.198	0.837
12:00–14:59	-1.497	0.405	13.687	0.000	0.224	0.101	0.495
15:00–17:59	-0.716	0.350	4.179	0.041	0.489	0.246	0.971
18:00–20:59	-0.740	0.362	4.178	0.041	0.477	0.235	0.970
Constant	0.159	0.553	0.089	0.766	1.172	—	—

TABLE 7: Fitting degree test.

Goodness-of-fit test (likelihood ratio Chi-square)		Badness-of-fit test (Hosmer–Lemeshow test value)	
Calculated value	Sig. value	Calculated value	Sig. value
20.559	0.001	7.025	0.534

TABLE 8: Accuracy of the model.

Crash type	Fatal	Nonfatal
Actual data	31	89
Proportion (%)	25.8	74.2
Model calculation	33	87
Proportion (%)	27.5	72.5
Absolute error value (%)	1.7	—

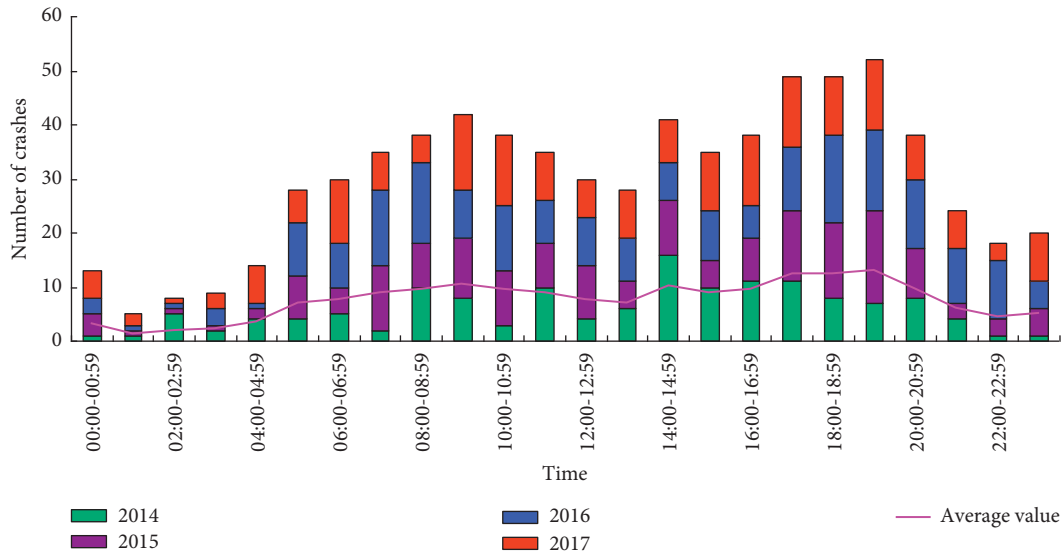


FIGURE 3: Temporal characteristic of traffic crashes in the rural-urban fringe.

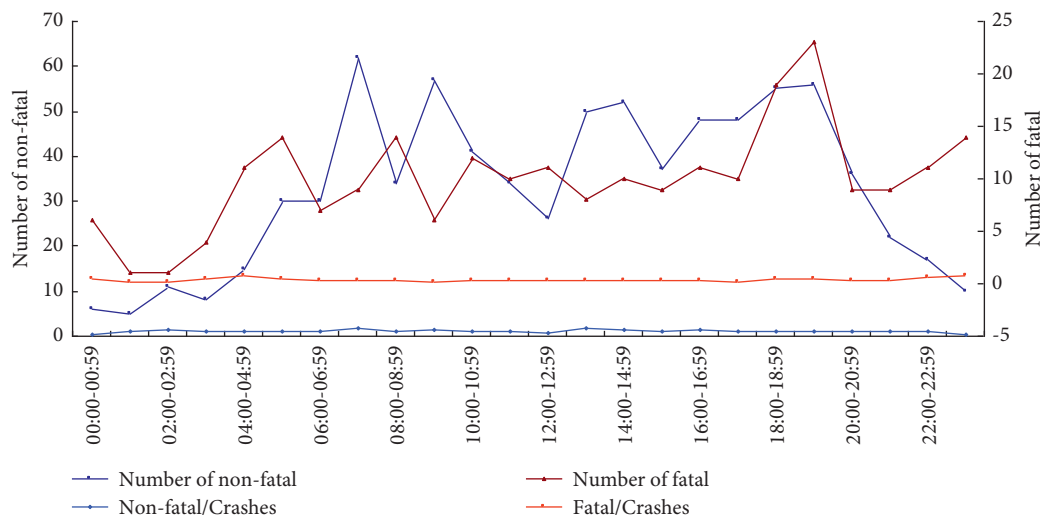


FIGURE 4: 24-hour time distribution of the number of casualties in traffic crashes.

8. Conclusions

This paper attempted to investigate the crash injury severity in rural-urban fringes using the crash data obtained from Harbin Public Security Bureau Traffic Police Station. A logistic regression model was developed to analyze the crash injury severity based on 661 crashes on the rural-urban junction roads selected from January 1st, 2014, to December 31st, 2017. Based on previous research findings on applied

variables, four explanatory variables were selected including time of day, vehicle type, road feature, and crash type, and two crash injury severity levels, i.e., fatal and nonfatal, were considered in this paper. The model of logistic regression was adopted to reveal the relationship between the specific variable and crash injury severity. The results showed that the factor of time of day had a significant influence on the crash injury severity and the number of crashes accords with “bimodal distribution” in temporal characteristics in the

rural-urban fringe. Its feasibility and practicality were verified by the methods of goodness-of-fit and badness-of-fit test, which showed a better fitting effect used in this model. This study aimed to propose and validate the method to identify the main factors influencing crash injury severity by using the logistic model to supply the basis of traffic safety measures.

There are some limitations of this study. It should be noted that the present study was based on a sample of 661 crashes collected on rural-urban fringe roads in Harbin City, which has limitations such as unavailability of some relevant variables and underreporting of injury crashes. Furthermore, the logistic regression model is a conventional approach in exploring factors contributing to crash injury severity. Therefore, more reliable statistical methods and machine learning methods should be utilized for empirical researches with large sample data to analyze the crash injury severity in the rural-urban fringe in the future.

Data Availability

The research data in this paper are mainly from Harbin Public Security Bureau Traffic Police Station.

Conflicts of Interest

The authors declare that they have no conflicts of interest regarding the publication of this paper.

Authors' Contributions

Conceptualization and methodology were developed by B. W. and Y. L.; software was provided by X. W.; validation and data curation were carried out by N. D. and X. W.; formal analysis was done by T. L. and X. W.; the original draft was prepared by X. W. and B. W.; reviewing and editing were performed by T. L., B. W., and Y. L. All authors have read and agreed to the published version of the manuscript.

Acknowledgments

The authors would like to thank Harbin Public Security Bureau Traffic Police Station for providing related data for the case study. The authors would like to thank Transportation Science and Technology Project of Heilongjiang Province (No. HJK2016A004), the Fundamental Research Funds for the Provincial-Level Colleges and Universities in Heilongjiang Province (No. 2018CX09), the Program for Provincial-Level Leading Talents Team Training of Heilongjiang Institute of Technology (No. 2020LJ04), and the Open Foundation for Jiangsu Key Laboratory of Traffic and Transportation Security Funded Project (No. TTS2017-06) for their financial support to this research project.

References

- [1] J. F. Tang, *Study on Planning of Small and Medium-Sized City Road Network and Transportation*, Tongji University, Shanghai, China, 2010.
- [2] N. Zhong, Y. F. Ma, W. J. Zhang, and J. Lu, "Study on selection methods of speed control measures for low grade roads in rural-urban fringe," *Proceedings of the 7th International Conference on Green Intelligent Transportation System and Safety*, vol. 419, pp. 305–316, 2016.
- [3] X. M. Fu, *Study on Traffic Safety Management under the Condition of Snow and Ice in Urban-Rural Binding Region*, Dalian Maritime University, Dalian, China, 2012.
- [4] P. Kopelias, F. Papadimitriou, and K. Papandreou, "Urban freeway crash analysis: geometric, operational, and weather effects on crash number and severity," *Transportation Research Record Journal of the Transportation Research Board*, vol. 2015, pp. 123–131, 2007.
- [5] P. T. Savolainen, F. L. Mannering, D. Lord, and M. A. Quddus, "The statistical analysis of highway crash-injury severities: a review and assessment of methodological alternatives," *Accident Analysis & Prevention*, vol. 43, no. 5, pp. 1666–1676, 2011.
- [6] J. Xu, W. Lin, X. Wang, and Y. M. Shao, "Acceleration and deceleration calibration of operating speed prediction models for two-lane mountain highways," *Journal of Transportation Engineering, Part A: Systems*, vol. 143, pp. 1–13, Article ID 4017024, 2017.
- [7] T. Z. Liu, B. M. Tang, J. M. Ling, and L. Liu, "Analysis on traffic crashes law and its countermeasures of rural-urban roads," *Journal of Chongqing Jianzhu University*, vol. 27, no. 5, pp. 767–771, 2008.
- [8] Z. H. Wei, J. Cao, and Y. S. Chen, "Cause analysis of high risk district in the urban-rural conjunction areas in Beijing," *Journal of Beijing University of Technology*, vol. 36, no. 2, pp. 223–228, 2010.
- [9] Q. Wei and R. F. Mou, "Road traffic safety in urban-rural conjunction region based on attribute recognition mathematic model," *Journal of Shijiazhuang Tiedao University (Natural Science)*, vol. 23, no. 3, pp. 106–110, 2010.
- [10] H. L. Zhang, L. Y. Yang, and G. C. Ao, "Analysis of the conflict and safety between pedestrian and vehicle at rural-urban fringe," *Highway*, vol. 62, no. 1, pp. 134–140, 2017.
- [11] W. J. Zhang, Y. Zhao, Y. Y. Zhang, and J. Lu, "Research on speed control section division of low-grade roads in urban-rural fringe," *CICTP*, vol. 419, pp. 2683–2693, 2019.
- [12] X. Y. Zhang and C. F. Shao, "Multi-objective signal timing optimal model for rural-urban fringe area intersection," *Journal of System Simulation*, vol. 32, no. 4, pp. 709–717, 2020.
- [13] Y. G. Li, *Analysis on the Distribution and Influencing Factors of Traffic Crashes in Urban and Rural Areas*, Chang'an University, Xi'an, China, 2018.
- [14] S. F. Niu, G. Q. Li, and E. Q. Liu, "Safety analysis of different vehicle types in urban-rural binding regions based on logit model," *Journal of Chongqing Jianzhu University*, vol. 39, no. 1, pp. 28–33, 2020.
- [15] X. Q. Chen, L. Qi, Y. S. Yang et al., "Video-based detection infrastructure enhancement for automated ship recognition and behavior analysis," *Journal of Advanced Transportation*, vol. 2020, Article ID 7194342, 12 pages, 2020.
- [16] B. Wu, Y. H. Tang, X. P. Yan, and C. G. Soaresc, "Bayesian Network modeling for safety management of electric vehicles transported in RoPax ships," *Reliability Engineering & System Safety*, vol. 209, pp. 1–15, 2021.
- [17] J. J. Tang, W. Q. Yin, C. Y. Han, X. Y. Liu, and H. L. Huang, "A random parameters regional quantile analysis for the varying effect of road-level risk factors on crash rates," *Analytic Methods in Accident Research*, vol. 29, pp. 1–20, 2021.

- [18] J. J. Tang, L. L. Zheng, C. Y. Han et al., "Statistical and machine learning methods for clearance time prediction of road incidents: a methodology review," *Analytic Methods in Accident Research*, vol. 27, 2020.
- [19] A. Jamal and W. Umer, "Exploring the injury severity risk factors in fatal crashes with neural network," *International Journal of Environmental Research and Public Health*, vol. 17, no. 20, p. 7466, 2020.
- [20] S. K. Xie, X. F. Ji, W. C. Yang, R. Fang, and J. J. Hao, "Exploring risk factors with crash severity on China two-lane rural roads using a random-parameter ordered probit model," *Journal of Advanced Transportation*, vol. 2020, Article ID 8870497, 14 pages, 2020.
- [21] A. Pervez, J. Lee, and H. L. Huang, "Identifying factors contributing to the motorcycle crash severity in Pakistan," *Journal of Advanced Transportation*, vol. 2021, Article ID 6636130, 10 pages, 2021.
- [22] Q. Zeng and H. Huang, "A stable and optimized neural network model for crash injury severity prediction," *Accident Analysis & Prevention*, vol. 73, pp. 351–358, 2014.
- [23] J. Tang, J. Liang, C. Han, Z. Li, and H. Huang, "Crash injury severity analysis using a two-layer Stacking framework," *Accident Analysis & Prevention*, vol. 122, pp. 226–238, 2019.
- [24] J. Zhang, J. Lindsay, K. Clarke, G. Robbins, and Y. Mao, "Factors affecting the severity of motor vehicle traffic crashes involving elderly drivers in Ontario," *Accident Analysis & Prevention*, vol. 32, no. 1, pp. 117–125, 2000.
- [25] A. S. Al-Ghamdi, "Using logistic regression to estimate the influence of accident factors on accident severity," *Accident Analysis & Prevention*, vol. 34, no. 6, pp. 729–741, 2002.
- [26] K. K. W. Yau, "Risk factors affecting the severity of single vehicle traffic accidents in Hong Kong," *Accident Analysis & Prevention*, vol. 36, no. 3, pp. 333–340, 2004.
- [27] X. Yan, E. Radwan, and M. Abdel-Aty, "Characteristics of rear-end accidents at signalized intersections using multiple logistic regression model," *Accident Analysis & Prevention*, vol. 37, no. 6, pp. 983–995, 2005.
- [28] N. N. Sze and S. C. Wong, "Diagnostic analysis of the logistic model for pedestrian injury severity in traffic crashes," *Accident Analysis & Prevention*, vol. 39, no. 6, pp. 1267–1278, 2007.
- [29] R. Harb, E. Radwan, X. Yan, A. Pande, and M. Abdel-Aty, "Freeway work-zone crash analysis and risk identification using multiple and conditional logistic regression," *Journal of Transportation Engineering*, vol. 134, no. 5, pp. 203–214, 2008.
- [30] Z. L. Ma, C. F. Shao, and X. Li, "Analysis of factors affecting crash severity in highway tunnels based on logistic model," *Journal of Jilin University (Engineering and Technology Edition)*, vol. 40, no. 2, pp. 423–426, 2010.
- [31] H. Chen, L. Cao, and D. B. Logan, "Analysis of risk factors affecting the severity of intersection crashes by logistic regression," *Traffic Injury Prevention*, vol. 13, no. 3, pp. 300–307, 2012.
- [32] Z. X. Feng, Y. W. Lei, H. Zhang, K. Wang, and S. Han, "Analysis on traffic crash severity influenced by road environment on circular highway," *China Journal of Highway and Transport*, vol. 29, no. 5, pp. 116–123, 2016.
- [33] J. Zhang, W. Q. Zhu, K. Wang, and K. Tao, "Analysis on traffic crash severity of continuous long downgrade section the mountainous area based on logistic," *Xi'an University of Architecture & Technology (Natural Science Edition)*, vol. 49, no. 6, pp. 841–846, 2017.
- [34] X. F. Ji, D. L. Li, and W. C. Yang, "Temporal evolution analysis on causes of traffic crashes occurring on secondary highways in mountain areas," *China Safety Science Journal*, vol. 29, no. 4, pp. 31–36, 2019.
- [35] T. M. Rudisill, S. Menon, B. Hendricks, M. Zhu, and G. S. Smith, "Differences between occupational and non-occupational-related motor vehicle collisions in West Virginia: a cross-sectional and spatial analysis," *PLoS One*, vol. 14, no. 12, p. e0227388, 2019.
- [36] Z. Wang, S. Huang, J. Wang, D. Hao, and A. Kuang, "Risk factors affecting crash injury severity for different groups of e-bike riders: a classification tree-based logistic regression model," *Journal of Safety Research*, vol. 76, pp. 176–183, 2021.
- [37] M. Li, S. Y. Zang, C. S. Wu, and Y. Tian, "Spatial and temporal variation and its driving forces of urban impervious surface in urban-rural continuum of Harbin," *Acta Geographica Sinica*, vol. 72, no. 1, pp. 105–115, 2017.
- [38] W. L. Lou, *Issues on the Land Use in Urban-Rural-Integration-Area in China*, Wuhan University, Wuhan, China, 2005.
- [39] G. Shieh, "Sample size calculations for logistic and Poisson regression models," *Biometrika*, vol. 88, no. 4, pp. 1193–1199, 2001.
- [40] X. Chen, Z. Li, Y. Yang, L. Qi, and R. Ke, "High-resolution vehicle trajectory extraction and denoising from aerial videos," *IEEE Transactions on Intelligent Transportation Systems*, vol. 22, no. 5, pp. 3190–3202, 2021.