Analysis of factors influencing severity of urban expressway accidents in Guangdong province

Chenjun Shi, YuanChang Deng
Guangdong Key Laboratory of Intelligent Transportation Systems,
School of intelligent systems engineering, Sun Yat-sen University
Guangzhou, China
dengych@mail.sysu.edu.cn

Abstract*—In order to investigate the influence factors of traffic accidents on urban expressway, this paper selected the severity of traffic accidents as the dependent variable. 15 factors including gender, vehicle type, road conditions and weather were introduced concerning the 4 aspects of road system, i.e. human, vehicle, road and environment. Then 15 indices were classified and assigned by using categorical variables processing method. Based on logistic model, the analysis model of traffic accident severity on urban expressway was built using the data of 1355 traffic accidents between January 2018 and August 2018 on Guangdong urban expressways. The results showed that the model fit well with the data. Three variables including vehicle type, roadside protection and visibility were significantly correlated with the severity of traffic accidents. Considering the type of the vehicle, the probability of serious accident of truck was the highest. The probability of serious accident with roadside protection was 2.093 times higher than that without roadside protection. The probability of serious accident was highest when visibility was within 50m.

Keywords—Traffic safety; Urban expressway; Logistic model; Severity of accident Introduction (*Heading 1*)

In recent years, with the rapid growth of traffic volume, the traffic accident rate has increased significantly, which has brought an inestimable impact on traffic safety. The People' Republic of China Road Traffic Accidents Annual Statistic Report (2016) (Traffic Management Bureau of the Ministry of public Security, 2017) showed that in 2016, the number of accidents caused by motorcycle are 43196, resulting in 11235 deaths, 52528 injuries, and direct property losses of 99272 million yuan[1].Therefore, we must pay attention to the severity of traffic accidents.

Many scholars have analyzed the relationship between the severity of traffic accidents and human factor, vehicle factor, road factor and environment factor. Sabey and Taylor found that drivers' personal characteristics account for 95% of the factors cause accidents[2]. Some researches generally suggested that male drivers usually exhibit a higher probability of committing traffic violations and consequently have a higher risk of being involved in serious or fatal traffic accidents [3,4,5,6]. Hijar et al. found that road type, street light condition and weather conditions are important factors that affect the severity of vehicle traffic crashes [7]. Besides, weather indicators, such as temperature and hail, are the most important factors that determine traffic safety [8]. Moreover, other factors including time of day and whether the accident occurred on a weekend are found to be significant in determining accident severity [9,10].

To the best of our knowledge, few studies have explored the factors influencing the severity of urban expressway traffic accidents. The road traffic accident database adopted in this study covers traffic accidents that occurred between January 2018 and August 2018 in Guangdong Province, China. This study used data from Guangdong province to assess the impact of various risk factors on the severity of urban expressway accidents, and the results can be helpful in formulating similar measures to reduce traffic casualties and promote road safety in other provinces.

I. MATERIALS AND METHODS

A. Data

We analyzed the traffic accident data for the period January 2018 and August 2018 in Guangdong Province, China. These data are extracted from the Traffic Management Sector Specific Incident Case Data Report. Data were recorded and reported by the traffic police on scene who conducted assessments and provided feed-back immediately to the headquarters of the Traffic Management Department. After deleting the missing value in the data, 1355 valid pieces of data were obtained. Each sample includes demographic information, injury severity, road conditions, accident time and environmental conditions.

The type of accidents in the database are classified into three type: fatal accidents, injuries, and property losses. Considering the role of accident rescue in accident types, this study classifies the severity of accidents into serious accidents and non-severe accidents, with a value of "1", and minor accidents and property losses were defined as non-severe accidents, assigned with "0".

As we all know, the road traffic system is a dynamic system consisting of human, vehicle, roads and environment. Any problem in any link may cause traffic safety problems. The risk factors that were considered include the following: human, vehicle, road and environment, which are considered independent variables.

B. Logistic regression model

Logistic regression model was used to explore the factors affecting the severity of urban expressway traffic accidents. Logistic regression model is a kind of discrete selection model developed abroad in the first half of the 20th century. It has been widely used in biology, economics, psychology, political

978-1-7281-0489-8/19/$31.00 ©2019 IEEE

science and transportation [11]. When a binary categorical variable in a log-linear model is treated as a dependent variable and defined as a function of a series of independent variables. The log-linear model becomes a Logistic regression model, as in equation 1.

$$P(\mathrm{Y}=1\,|X_1, X_2, \cdots, X_n) = \frac{1}{1+e^{(\alpha+\sum_{i=1}^{n}\beta_i X_i)}} \quad (1)$$

In the above formula: X_i is an independent variable; α and β are regression intercept and regression coefficients respectively; P is the conditional probability of event occurrence; $(\alpha + \sum_{i=1}^{n}\beta_i X_i)$ is defined as a series of linear functions that affect the probability of occurrence of an event.

The probability of event occurrence P and the probability of event non-occurrence (1 - P) is called the event occurrence ratio (Odds). Taking Odds as the natural logarithm gives a linear function, as in equation 2:

$$\ln\frac{P}{1-P} = \alpha + \sum_{i=1}^{n}\beta_i X_i \quad (2)$$

C. Risk factors

Drivers' age, driving experience and gender are considered to be potential risk factors. Driver age is divided into five groups: 18-20, 21-30, 31-40, 41-50, ≥51 and driving experience is categorized into five groups: 0-5, 6-10, 11-15, 16-20, ≥21. Although drivers' income, education and marriage status are expected to be potential factors that are associated with accident severity, this information is generally not recorded in the traffic accident database.

In the factors of the vehicle, the license plate issuing place and the vehicle type are considered. The license plate issuing place is divided into two types: provincial and non-provincial. Three major vehicle types, namely passenger vehicles, goods vehicles and motorcycles, are included in the data.

Among the road factors, road line type, road condition, road surface condition, road surface structure and roadside protection facilities are considered to be potential risk factors. Road line types are divided into straight and non-straight. Road conditions are divided into two types: incomplete and intact. Road surface conditions are divided into dry and wet. Road surface structure is divided into asphalt and non-asphalt. Whether or not to set the roadside protection facilities is also considered.

There are five environmental factors: street-light condition, visibility level, weather conditions, whether the accident occurred on a weekend and time of day. Street-light is classified into daylight, good street lighting at night and no street lighting at night. Visibility is divided into four groups: 50m≤, 51-100m, 101-200m, ≥201m. Weather conditions is divided into five groups: sunny, cloudy, rainy, snowy and foggy. Weekends are defined as: 17:00 Friday to 23::59 Sunday, as it is anticipated that the occurrence of Friday traffic accidents after 17:00 is similar to those occurring on Saturday and Sunday [12]. Time of day is divided into two types: day and night.

II. RESULTS

Logistic regression and parameter calculation were preformed using SPSS 22.0. The positive stepwise method was adopted, and the significant level α =0.05. 15 factors including gender, the vehicle type, road line type and weather conditions were substituted into the model.

A. Score test

First, the Score-test was used on the model independent variables to select the variables to select the variables that can meet the modeling requirements. The test results are shown in Table 1.

TABLE I. SCORE TEST

Influencing factor	*Score*	*df*	*P(Sig.)*
Drivers' age	3.806	4	0.433
Driving experience	5.936	4	0.204
Gender	7.542	1	0.006
The license plate issuing place	15.973	1	0.000
Vehicle type	101.449	2	0.000
Road line type	1.265	1	0.261
Road condition	0.607	1	0.436
Road surface condition	0.266	1	0.606
Road surface structure	0.404	1	0.525
Roadside protection facilities	10.389	1	0.001
Street-light condition	4.972	2	0.083
Visibility level	15.397	3	0.002
Weather condition	0.316	2	0.854
Whether the accident occurred on a weekend	0.626	1	0.429
Time of day	0.414	1	0.520

As shown in Table 1, since the significant level α =0.05, according to the Score-test result, the variable with the p value less than 0.05 was selected to enter the model, that is, the four independent variables of the license plate issuing place, vehicle type, roadside protection facilities and visibility level satisfied the modeling requirements.

After further iterative calculation, according to the Wald value test requirements, the final parameter result of the model were shown in Table 2.

TABLE II. ESTIMATION RESULTS OF PARAMETERS

Influencing factor	*B*	*S.E.*	*Wald*	*Sig.*	*Exp(B)*
Vehicle type			82.493	0.000	
Vehicle type(1)	-1.205	0.269	20.132	0.000	0.300
Vehicle type(2)	-1.857	0.206	80.973	0.000	0.156
Roadside protection facilities	0.739	0.250	8.738	0.003	2.093
Visibility level			15.914	0.001	
Visibility level(1)	0.909	0.308	8.704	0.003	2.481
Visibility level(2)	0.291	0.232	1.571	0.210	1.338
Visibility level(3)	-0.362	0.254	2.023	0.155	0.696

978-1-7281-0489-8/19/$31.00 ©2019 IEEE

Influencing factor	*B*	*S.E.*	*Wald*	*Sig.*	*Exp(B)*
Constant	-1.639	0.278	34.765	0.000	0.194

It can be seen from Table 2 that vehicle type, roadside protection facilities and visibility level significantly affect the severity of traffic accidents in urban expressways in Guangdong Province. Among the factors affecting the type of vehicle, the probability of a serious accident of a motorcycle is 0.3 times that of a goods vehicle, and that of a passenger vehicle is 0.156 times that of a goods vehicle. The probability of serious accident with roadside protection was 2.093 times higher that that without roadside protection. Among the influencing factors in the visibility level, the probability of a traffic accident with a visibility level within 50m is 2.481 times the visibility is greater than 200m.

TABLE III. OMNIBUS TEST OF MODEL COEFFICIENTS

	CHI-SQUARE	*DF*	*SIG*
MODEL	180.138	6	0.000

It could be observed from the above table that the value of sig was less than 0.001, which indicating that the model had statistical significance.

TABLE IV. HOSMER AND LEMESHOW TEST

	CHI-SQUARE	*DF*	*SIG*
1	3.600	8	0.891

If the test results of the model have statistical significance, it indicates that the model fitting is poor. As shown in table 4, Hosmer and Lemeshow goodness of fit test has no statistical significance (P=0.891), indicating good fitting of the model.

III. DISCUSSION

The result indicate that some risk factors are found to be significant in the logistic regression analysis for traffic accident severity. As a summary, vehicle type is the most significant factor affecting the severity of traffic accidents. Among them, goods vehicle have the highest probability of serious traffic accidents, probably because the size of the goods vehicle is larger than that of passenger vehicles and motorcycles, and it is more difficult to control. The probability of a serious accident occurring on the roadside protection facility is higher. It may be that the driver will choose to use the roadside protection setting to slow down when a traffic accident occurs, thus increasing the risk of accidents. When the visibility is within 50m, the probability of a serious accident is higher. It is obvious that the lower the visibility, the smaller the drivers' field of vision.

A. Human factor

Some studies have shown that gender, driving experience and age have significant effect with the severity of traffic accidents [9,14, 15,16]. There are also studies showing that education level, income and social status have been shown to be factors associating with road safety [17.18]. But this type of data is not recorded in our database. In our study, the effect of drivers' gender, age and driving experience are not found to be significant on both traffic accident severity. This is similar to research results of Wang et al. [18], which can be as a reference for subsequent research.

B. Vehicle, Road and enverionment factors

It should be noted that goods vehicles exhibit significantly higher risk of accident severity. In fact, this finding is consistent with Hijar et al. [7] and Ghamdi et al. [19]. The license plate issuing place have no effect with accidents severity. Strengthening vehicle safety and overload status checks, especially for goods vehicle, is also critical to reducing the severity of accidents. Our research has found that setting roadside protection facilities increases the probability of serious accidents, which may serve as a reference for other studies. Moreover, bad visibility exerts pressure on the risk on traffic accident severity, which is the same as Zhang et al [16].

C. Limitations and future research

Since the data comes from the Traffic Management Sector Specific Incident Case Data from the Guangdong Provincial Security Department, the report is the only official source of cross-section traffic accident data available in the province, so data quality can be guaranteed. However, there are many missing values in the data, so there may be possible deviations caused by insufficient reporting. Furthermore, issues such as omitted variable bias, endogeneity and within-crash correlation should be aware of when looking into the results [20]. The data is from January 2018 to August 2018, less than one year. There may be deviations in statistical results, and future studies try to use data for one year or integer years. Besides, traffic violations can also result in different levels of traffic accident severity. For example, speeding, drunk driving of fatigue driving each has its own significant risk factors. Future research work will look into type specific traffic violations events and assess their corresponding risk factors.

ACKNOWLEDGMENT

Special thanks to the data support department of the Traffic Safety Department of the Guangdong Provincial Department of Safety.

REFERENCES

[1] Traffic Management Bureau of the Ministry of Public Security. The People's Republic of China road traffic accident annual statistic report (2017). Wuxi, China, 2015.

[2] B.E. Sabey, H. Taylor,The known risks we run: the high way. In: Report SR567. Transport Research Laboratory, Crowthorn, Berks, 1980.

[3] F.H. Norris, B.A. Matthews , J.K. Riad, Characterological, situational, and behavioral risk factors for motor vehicle accidents: a prospective examination. Accident Analysis and Prevention, vol. 32, pp. 505–515, 2000.

[4] J. Zhang , J. Lindsay, K. Clarke, G. Robbins, Y. Mao, Factors affecting the severity of motor vehicle traffic crashes involving elderly drivers in Ontario. Accident Analysis and Prevention, vol. 32, pp. 117–125, 2000.

[5] R.A. Derrig, M. Segui-Gomez, A. Abtahi, L.L. Liu, The effect of population safety belt usage rates on motor vehicle-related fatalities. Accident Analysis and Prevention, vol. 34, pp. 101–110, 2002.

[6] F. Valent, F. Schiava, C. Savonitto, T. Gallo, S. Brusaferro, F. Barbone, Risk factors for fatal road traffic accidents in Udine, Italy. Accident Analysis and Prevention, vol. 34, pp. 71–84, 2002.

978-1-7281-0489-8/19/$31.00 ©2019 IEEE

[7] M. Hijar, C. Carrillo, M. Flores, R. Anaya, V. Lopez, Risk factors in highway traffic accidents: a case control study. Accident Analysis and Prevention, vol. 32, pp. 703–709, 2000.

[8] P.A. Yakovlev, M. Inden, Mind the weather: a panel data analy- sis of time-invariant factors and traffic fatalities. Economics Bulletin, vol. 30, pp. 2685–2696, 2010.

[9] D.L. Massie, K.L. Campbell, A.F. Williams, Traffic accident involvement rates by driver age and gender. Accident Analysis and Prevention, vol. 27, pp. 73–87, 1995.

[10] L.T. Lam, R. Norton, M. Woodward, J. Connor, S. Ameratunga, Passenger carriage and car crash injury: a comparison between younger and older drivers. Accident Analysis and Prevention, vol. 35, pp. 861–867, 2003.

[11] S. Kaplan, C.G. Prato, Risk factors associated with bus accident severity in the United States:A generalized ordered logit model .Journal of Safety Research, vol. 43, pp. 171-180, 2012.

[12] K.E. MacLeod, J.B. Griswold, L.S. Arnold, D.R. Ragland, Factors associated with hit-and-run pedestrian fatalities and driver identification. Accident Analysis and Prevention, vol. 45, pp. 366–372, 2011.

[13] H. Hayakawa, P.S. Fischbeck, B. Fischhoff, Traffic accident statistics and risk perceptions in Japan and the United States. Accident Analysis and Prevention, vol. 32, pp. 827–835, 2000.

[14] M. Bedard, G.H. Guyatt, M.J. Stones, J.P. Hirdes, The independent contribution of driver, crash, and vehicle characteristics to driver fatalities. Accident Analysis and Prevention, vol. 34, pp. 717–727, 2002.

[15] G. Zhang, K.K.W. Yau, G. Chen, Risk factors associated with traffic violations and accident severity in China. Accident Analysis and Prevention, vol. 59, pp. 18-25, 2013.

[16] M. Hasselberg, M. Vaeza, L. Laflamme, Socioeconomic aspects of the circumstances and consequences of car crashes among young adults. Social Science and Medicine, vol. 60, pp. 287–295, 2005.

[17] R. Factor, D. Mahalel, G. Yair, Inter-group differences in road-traffic crash involvement. Accident Analysis and Prevention, vol. 40, pp. 2000–2007, 2008.

[18] Lei Wang, Pu Lv, Yongjie Lin. Traffic accidents on fressways: influencing factors analysis and injury severity eveluation. China Safety Science Journal, vol. 26, pp. 86-90, 2016.

[19] A.S. Al-Ghamdi, Using logistic regression to estimate the influence of accident factors on accident severity. Accident Analysis and Prevention, vol. 34, pp. 729–741, 2002.

[20] P.T. Savolainen, F.L. Mannering, D. Lord, M.A. Quddus, The statistical analysis of highway crash-injury severities: a review and assessment of methodological alternatives. Accident Analysis and Prevention, vol. 43, pp. 1666–1676, 2011

978-1-7281-0489-8/19/$31.00 ©2019 IEEE