

Risk Factors of Secondary Accidents Based on the Bayesian Hierarchical Model

Zhang Z.Y.¹, and Zhang Z.^{2,*} and Yang J.F.³

¹School of Naval Architecture, Ocean and Civil Engineering, Shanghai Jiao Tong University, Shanghai 200240, China

²School of Design, Shanghai Jiao Tong University, Shanghai 200240, China

³School of Information Systems, Singapore Management University, 188065, Singapore

Received 1 August 2018; Accepted 22 November 2018

Abstract

Prevention of secondary accidents is an important component of safety management in the transport sector, which is determined by their serious consequence. Recognizing the risk factors of secondary accidents and determining the influences of different risk factors are the research frontiers at present. However, existing studies mainly focus on risk factors at the accident level but hardly consider the influencing factors of roads. Moreover, the hierarchical structure and interaction of risk factors are overlooked. To analyze the influencing factors of secondary accidents and reduce the interaction effects of influencing factors in risk analysis, important influencing factors of secondary accidents were determined using the Bayesian hierarchical model in this study based on 18,231 data of secondary accidents in Florida from 2012 to 2017. These factors covered 10 aspects, including weather conditions, place of accidents, time of accidents, vehicle speed, traffic signal control, road alignment, road conditions, light conditions, speed limit and number of straight lanes. Then, the fitness of the Bayesian hierarchical model was calculated through the deviance information criterion (DIC). The fitting degree of the Bayesian hierarchical model was compared with that of the Bayesian logistic model. Finally, effects of the recognized influencing factors on the occurrence probability of secondary accidents were analyzed. Results show that (1) vehicle speed, speed limit of road sections, place of accidents, light conditions, time of accidents, and road alignment remarkably influence the occurrence of secondary accidents. Moreover, major influencing factors at special locations (e.g., intersections and bridges) are different from those in other places. (2) The DIC value of the Bayesian hierarchical model is lower than that of the Bayesian logistic model, thereby indicating that the former can decrease the nesting effects of influencing factors on the results and offset shortages of the traditional risk model. Research conclusions provide not only a new idea to study the risks of traffic accidents but also a new decision-making idea and method to control and prevent secondary accidents.

Keywords: Secondary accidents, Bayesian hierarchical model, Risk analysis

1. Introduction

Secondary accidents often cause greater losses than common traffic accidents. These accidents have been reported [1,2] to account for 20% of all accidents on highways and cause a fatality of 18%. Secondary accidents often involve several vehicles and thereby cause further serious consequences. The fatality and losses involving secondary accidents are remarkably higher than those of common traffic accidents. The occurrence of secondary accidents not only influences the traffic operation and causes further delay but also destroys road facilities and decreases the traffic capacity of roads, thereby bringing serious damages to the traffic system. In fact, improving accident management has been proven to reduce the occurrence of secondary accidents effectively [3]. Therefore, studying the influencing factors of secondary accidents and providing a transportation sector guidance for preventing secondary accidents are key problems that must be solved at present.

However, existing studies on the influencing factors of secondary accidents have mainly focused on traditional

statistical methods [4,5]. Nevertheless, the lumpability and interactions of the different influencing factors have been ignored. Existing studies mainly use logit and probit regression models [6,7] to define the relationships between secondary accidents and influencing factors. A conventional regression analysis of secondary accidents cannot meet the hypothesis of independent observation and obtains small values of standard deviation due to the interaction of different influencing factors, thus intensifying the model error. Additionally, no comprehensive study on risk factors has been reported yet due to the incomplete data on accidents. Road factors are hardly considered in existing studies. Generally, existing studies of secondary accidents mainly analyze the factors related to accidents rather than road factors. Furthermore, the interaction of risk factors is neglected. Data in the traffic accident database in Florida were classified in the present study with considerations of the connections between individuals of secondary accidents and the subordination group. The 11 factors, including weather conditions, place of accidents, and time of accidents, were divided into two levels, namely, road and accident. Then, influences of different factors on secondary accidents were explored, which were vital to the occurrence of secondary accidents.

*E-mail address: zhanzhang@sjtu.edu.cn

ISSN: 1791-2377 © 2018 Eastern Macedonia and Thrace Institute of Technology. All rights reserved.

doi:10.25103/jestr.115.07

2. State of the art

Existing studies on the influencing factors of secondary accidents mainly determine the correlation between secondary accidents and influencing factors by using fitting methods of accident data. However, key attentions are paid to accident factors, such as vehicle conditions and speed, but influences of road factors are neglected. Yang et al. [8] developed a logit model and found that collision mode is an important influencing factor of secondary accidents. However, they neglected the effects of factors related to road facilities (e.g., road condition and alignment) on secondary accidents. Zhan et al. [9] studied the relationships of vehicle speed, type of collision, number of vehicles, time, vehicle type, and number of lanes on secondary accidents based on the probit regression model, but the influences of speed limit were neglected. Xu et al. [10] proposed a random effect logit model to explore the relationship between real-time traffic conditions (e.g., traffic flow) and the occurrence probability of secondary accidents, but the influences of traffic facilities (e.g., light and road conditions) on secondary accidents were overlooked.

Moreover, a risk analysis of secondary accidents mainly uses the traditional regression model. Vlahogianni et al. [11] viewed the number of involved vehicles and lanes as two independent variables in the logistic model, but they neglected the hidden relationship between these factors. Yang [12] studied the influences of the number of lanes, regions, and trucks on secondary accidents by using the logistic model. However, the relationship between the number of trucks and regions was omitted. The influencing factors of secondary accidents were not completely independent, and secondary accidents that occur under the same road conditions are related to some extent. Hence, the traditional regression analysis may cause remarkably ambiguous influencing factors and increase the error term in the model, which determines the limitations of the traditional regression analysis. While the Bayesian hierarchical model has been applied to other fields to process the group effect of influencing factors, thus enabling the offset of the abovementioned limitations [13]. Yu et al. [14] divided the influencing factors of accidents into road and season levels by using the Bayesian hierarchical model to discuss the occurrence characteristics of accidents on highways in mountainous areas with considerations of random error and variable information at different levels. Kim et al. [15] constructed a Bayesian hierarchical model and applied it to study the influencing factors of the severity of accidents based on the city and accident levels. The macroscopic and microscopic influencing factors were combined. Zeng [16] analyzed the factors that influence the severity of accidents by using the Bayesian hierarchical model and divided them into the driver and vehicle levels with considerations of the group effect of influencing factors. On the basis of the full considerations of data nesting, Xie [17] analyzed hit-and-run accidents from running away and hurt degrees by using the Bayesian hierarchical model. Cheng [18] explored the relationship between the severity of accidents and influencing factors based on the Bayesian hierarchical model. He divided the influencing factors into accident and vehicle levels, which avoided the interaction of relevant factors.

In conclusion, existing studies on secondary accidents mainly have focused on factors at the accident level (e.g., speed, number of vehicles, and weather conditions) but paid less attention to factors at the road level (e.g., speed limit,

road alignment, and road conditions). Meanwhile, only few studies on the group effects of influencing factors have been reported. To address the existing research limitations, the influencing factors of secondary accidents were divided into road and accident levels in the present study. A Bayesian hierarchical model was constructed to discuss the relationship between factors at different levels and in secondary accidents. The model was verified by using deviance information criterion (DIC).

The remainder of this study is organized as follows. Section 3 describes the accident database and Bayesian hierarchical model. Section 4 analyzes the influences of recognized factors on the occurrence probability of secondary accidents. Differences in the risk factors of secondary accidents in different places are also discussed. Section 5 summarizes the conclusions.

3. Methodology

3.1 Acquisition of basic data

Secondary accident data were collected from the database of the Florida Department of Highway Safety and Motor Vehicles (DHSMV). This database covers four parts, namely, the characteristics of accidents, characteristics of vehicles, features of people, and law enforcement. A total of 205,021 accident data were in the DHSMV database, and each datum contained 312 variables, including time, place, vehicle information, and driver information. In this study, 18,231 data of secondary accidents from 2012 to 2017 were extracted to analyze the relevant influencing factors. Here, 11 relevant variables, including weather conditions, place of accidents, time of accidents, vehicle speed, traffic signal control, road alignment, road conditions, light conditions, number of straight lanes, and speed limit, were divided into road and accident levels. The former covered speed limit, road conditions, light conditions, traffic signal control, road alignment, and number of straight lanes. The latter contained days of the week, weather conditions, vehicle speed, time of accidents, and place of accidents. Table 1 lists the mean and standard deviation of the 11 variables. A modeling analysis based on 10,772 data of secondary accidents at different places was conducted to evaluate the differences of secondary accidents at special places (e.g., intersection and bridge) with those at common places. Table 2 lists the statistical information of data.

3.2 Bayesian framework

The theoretical framework of Bayesian inference is expressed as follows [19]:

$$\pi(y|\theta) = \frac{L(y|\theta)\pi(\theta)}{\int L(y|\theta)\pi(\theta)d\theta} \quad (1)$$

where y is the vector of observed data, θ denotes the vector of parameters required for the likelihood function, $L(y|\theta)$ represents a likelihood function, $\pi(\theta)$ indicates the prior distribution of θ , $\int L(y|\theta)\pi(\theta)d\theta$ refers to the marginal distribution of the observed data, and $\pi(y|\theta)$ signifies the posterior distribution of θ when y is fixed.

Table 1. Statistical information of variables

Variables	Descriptions	Mean	Standard Deviation(S.D.)
Accident level			
Days of the week	If the accident occurs on a weekend, then it is 0; otherwise, it is 1.	0.25	0.43
Weather conditions	If the accident occurs on a sunny day, then it is 0; otherwise, it is 1.	0.27	0.44
Occurrence place	If the accident occurs at intersections and bridges, then it is 1; otherwise, it is 0.	0.65	0.48
Occurrence time	If the accident occurs during the day, then it is 0; otherwise, it is 1.	0.29	0.45
Vehicle speed	If the vehicle speed is ≤ 30 mile/h, then it is 0; otherwise, it is 1.	0.33	0.48
Road level			
Traffic signal control	If a traffic signal control exists, then it is 0; otherwise, it is 1.	0.51	0.50
Road alignment	If accidents occur on straight roads, then it is 0; otherwise, it is 1.	0.10	0.30
Road conditions	If the road is dry, then it is 0; otherwise, it is 1.	0.13	0.33
Light conditions	If the light conditions are good, then it is 0; otherwise, it is 1.	0.33	0.47
Number of straight lanes	If at least six straight lanes exist, then it is 0; otherwise, it is 1.	0.99	0.08
Speed limit			
	If the speed limit is ≤ 30 mile/h, then it is 0.		
	If $30 < \text{speed limit} \leq 60$ mile/h, then it is 1.	0.75	0.82
	If the speed limit is higher than 60 mile/h, then it is 2.		

Table 2. Statistical information of secondary accident data at different places

Variables	Special places		Common places	
	Mean	S.D.	Mean	S.D.
Accident level				
Days of the week	0.25	0.43	0.25	0.44
Weather conditions	0.27	0.44	0.27	0.45
Occurrence time	0.28	0.45	0.32	0.47
Vehicle speed	0.27	0.45	0.42	0.51
Road level				
Traffic signal control	0.45	0.50	0.62	0.49
Road alignment	0.08	0.27	0.13	0.34
Road conditions	0.12	0.33	0.14	0.35
Light conditions	0.31	0.46	0.37	0.48

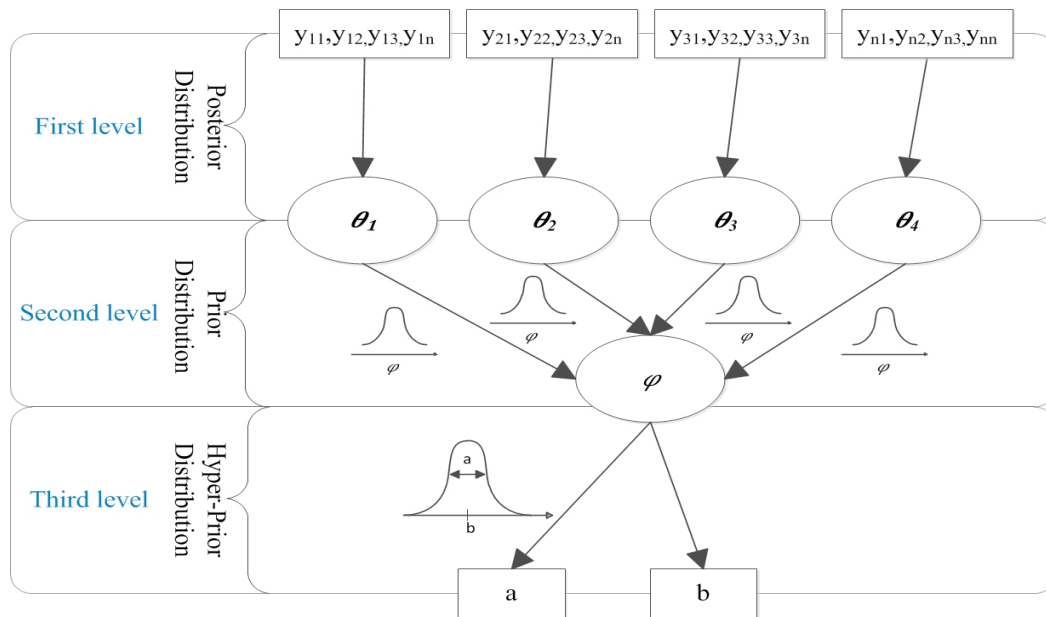


Fig. 1. Generic structure of a two-level hierarchical model

3.3 Bayesian random effects binary logistic model

$y = 1$ represents the occurrence of secondary accidents and $y = 0$ reflects no accident. The occurrence probabilities are p and $1 - p$ respectively.

$$Y_i : \text{Bernoulli}(\theta_i) \tag{2}$$

$$\text{logit}(\theta_i) = \log\left(\frac{\theta_i}{1 - \theta_i}\right) = \mathbf{X}\boldsymbol{\beta} + \mu_i \tag{3}$$

where \mathbf{X} denotes the vector of explanatory variables. $\boldsymbol{\beta}$ is the coefficient vector of explanatory variables and it

constructs the non-informative priors $\text{Normal}(0,1000)$ in accordance with the normal distribution. μ_i represents the random effects at the accident level and i refers to the sample size. The variance of random effects obeys the inverse gamma distribution $(10^{-3}, 10^{-3})$.

3.4 Bayesian hierarchical model

Fig. 1 shows the structure of the Bayesian hierarchical model. Parameters in the rectangle frames and ovals are known and unknown. The Bayesian hierarchical model is expressed as follows:

$$Y_{ik} : \text{Bernoulli}(\theta_{ik}) \tag{4}$$

The first-level model:

$$\text{logit}(\theta_{ik}) = \text{log}\left(\frac{\theta_{ik}}{1-\theta_{ik}}\right) = \beta_{0k} + \sum_{b=1}^B \beta_{bk} X_{bik} \quad (5)$$

The second-level model:

$$\beta_{0k} = \gamma_{00} + \sum_{a=1}^A \gamma_{0a} W_{ak} + \varepsilon_{0k} \quad (6)$$

$$\beta_{1k} = \gamma_{10} \quad (7)$$

$$\beta_{2k} = \gamma_{20} \quad (8)$$

⋮

$$\beta_{bk} = \gamma_{b0} \quad (9)$$

On the basis of these models, the overall model is as follows:

$$\text{log}\left(\frac{\theta_{ik}}{1-\theta_{ik}}\right) = \gamma_{00} + \sum_{a=1}^A \gamma_{0a} W_{ak} + \varepsilon_{0k} + \sum_{b=1}^B \beta_{bk} X_{bik} \quad (10)$$

where Y_{ik} is a binary observation value, θ_{ik} represents the occurrence probability of accidents ($Y_{ik} = 1$), γ_{00} denotes the intercept, W_{ak} refers to a variable at the road level, and X_{bik} indicates a variable at the accident level. γ_{0a} and γ_{b0} are estimation regression coefficients. ε_{0k} signifies the random effect at the road level.

3.5 Model estimation

A Bayesian model generally estimates parameters by using the Markov Chain Monte Carlo (MCMC) method [20]. The open source software WinBUGS was applied for calculating the models. Without reliable prior information, all variable coefficients were hypothesized to conform to the normal distribution $(0,10^3)$ and the variable of stochastic effect to

obey to the inverse gamma distribution $(10^{-3},10^{-3})$. A total of 100,000 MCMC chain iterations were performed, and the first 4,000 iterations were not used in the results. Gelman–Rubin criteria were used to judge whether the MCMC was converged. The variable was deemed substantial when a 95% Bayesian confidence interval (BCI) did not cover 0 [21].

3.6 Model evaluation

The DIC are standards used to measure the model. A low DIC value indicates the good fitting data of the model. The DIC is calculated as follows:

$$DIC(m) = 2\overline{D(\theta_m, m)} - D(\overline{\theta_m}, m) = D(\overline{\theta_m}, m) + 2p_m \quad (11)$$

where $D(\overline{\theta_m}, m)$ is a usual deviance measure, $\overline{D(\theta_m, m)}$ is the posteriori mean.

$$D(\overline{\theta_m}, m) = -2\text{log} f(y|\overline{\theta_m}, m) \quad (12)$$

$\overline{\theta_m}$ denotes the posterior mean of parameters.

$$p_m = D(\theta_m, m) - D(\overline{\theta_m}, m) \quad (13)$$

p_m refers to the number of valid parameters in model m .

4. Result analysis and discussions

In this study, the influencing factors of secondary accidents were analyzed using the Bayesian hierarchical model based on secondary accident data in Florida. Model verification results (Table 3) indicated that the DIC value of the Bayesian hierarchical model was 6103.6, which was smaller than that of the Bayesian logistic model (7652.34). This finding showed that the Bayesian hierarchical model was superior to the Bayesian logistic model in studying secondary accidents, and it could efficiently reflect the connections of influencing factors at different levels.

Table 3. Results of Bayesian models

Variables	Bayesian logistic model				Bayesian hierarchical model			
	Effect estimate		95% BCI odds ratio		Effect estimate		95% BCI odds ratio	
	Mean (S.D.)	Odds ratio	2.50%	97.50%	Mean (S.D.)	Odds ratio	2.50%	97.50%
Intercept	-0.85 (0.023)	0.43	0.41	0.46	-0.01 (0.032)	0.99	0.93	0.99
Accident level	—	—	—	—	0.10 (0.028)	1.11	1.05	1.17
Days of the week	—	—	—	—	0.32 (0.027)	1.37	1.30	1.45
Vehicle speed	0.16 (0.027)	1.17	1.11	1.24	-0.13 (0.027)	0.87	0.83	0.92
Occurrence place	-0.40 (0.026)	0.67	0.63	0.70	0.11 (0.028)	1.11	1.05	1.18
Occurrence time	—	—	—	—	—	—	—	—
Road level								
Speed limit	-0.28 (0.022)	0.76	0.72	0.79	-0.06 (0.023)	0.94	0.90	0.99
Light conditions	—	—	—	—	0.10 (0.028)	1.11	1.05	1.17
Road alignment	0.06 (0.029)	1.06	1.01	1.13	0.12 (0.029)	1.13	1.07	1.20
Random effect	0.16 (0.09)	—	—	—	0.10 (0.09)	—	—	—
DIC value	7652.34				6103.6			

4.1 Model comparison results

Table 3 lists the parameter estimation results of variables and DIC values. On the basis of the DIC value, the Bayesian hierarchical model shows better degree of fitting than the Bayesian logistic model. The trace diagram (Fig. 2) reveals that the Bayesian hierarchical model converges finally.

Moreover, factors in both models present the same positive or negative effects on secondary accidents, which reflect the consistency of both models. However, the Bayesian hierarchical model covers additional substantial variables. This case reveals that the Bayesian hierarchical model can reduce the influences of variable nesting on the results and

discloses the relationships between variables and secondary accidents. Thus, this model is applicable to the characteristics of influencing factors of secondary accidents. The Bayesian hierarchical model recognizes important influencing factors of secondary accidents at the road (e.g., speed limit, light conditions, and road alignment) and

accident (e.g., days of the week, vehicle speed, place, and time) levels. Key attention shall be paid to these factors when formulating the prevention and control measures of secondary accidents.

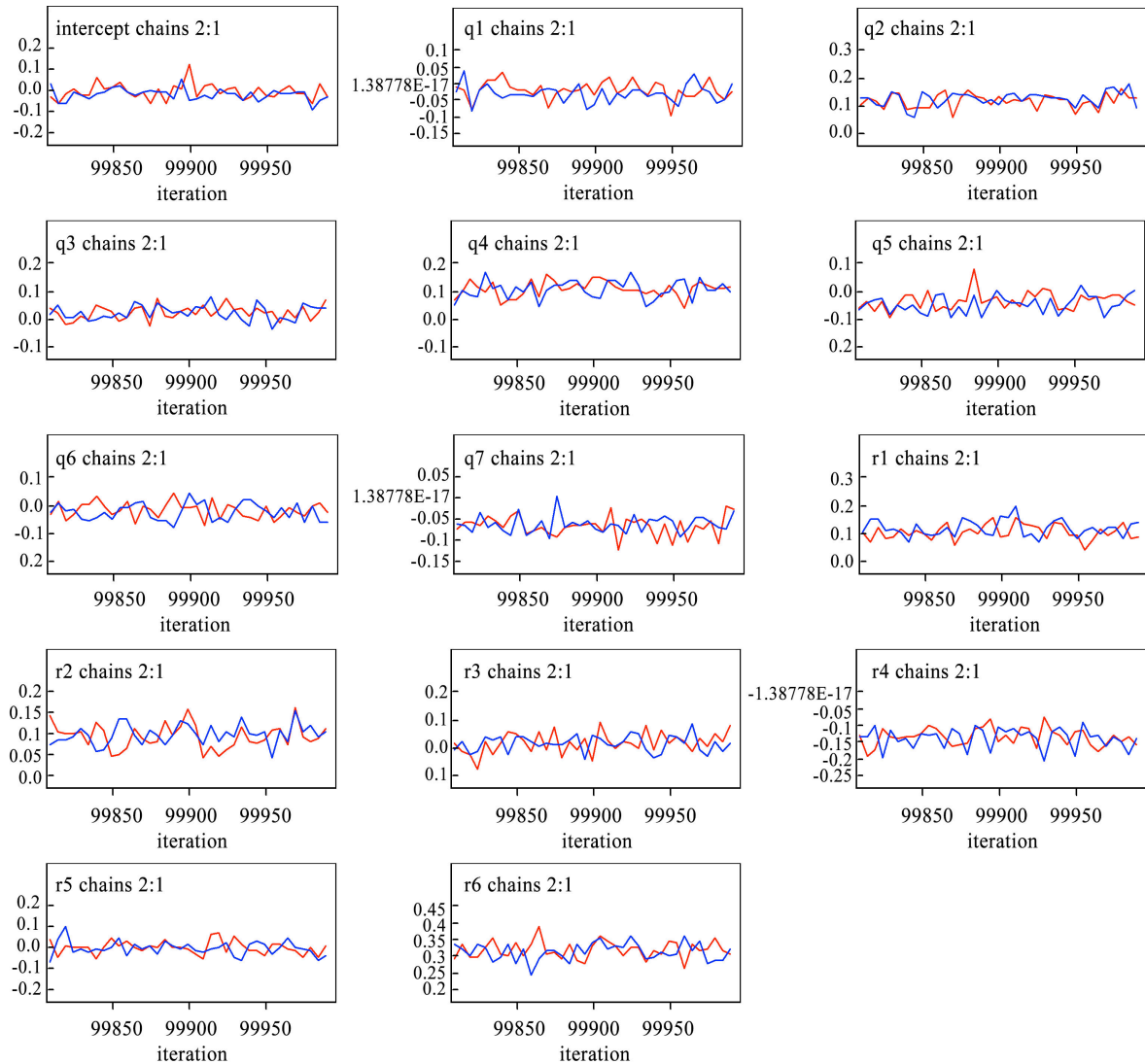


Fig. 2. Trace diagram of the influencing factors in the Bayesian hierarchical model

4.2 Explanation of variables

4.2.1 Speed

Results of the Bayesian hierarchical model (Table 3) indicate that vehicle speed remarkably influences the occurrence of secondary accidents. The probability of occurrence of secondary accidents increases by 37% ($e^{0.32} - 1$) for each increase in vehicle speed by 20 mile/h. Therefore, a high vehicle speed causes further secondary accidents. Table 4 presents the relationships between the occurrence probability of secondary accidents and other factors under a fixed speed. When the vehicle speed is higher than 30 mile/h and it crosses turns, intersections, or bridges, the occurrence probability of secondary accidents is increased by 50%, which agrees with the research results of Kopitch [22]. On the basis of the statistical information diagram on vehicle speed in secondary accidents (Fig. 3), the number of secondary accidents obeys the normal distribution

with changes in speed and reaches the peak in the interval of 30–40 mile/h. Although results of the Bayesian hierarchical model reveal that a high speed causes a high risk of secondary accidents, the Department of Accident Management should not focus on road sections with high speed due to different traffic flows in different speed intervals. The speed interval of 30–40 mile/h accounts for the highest occurrence frequency of secondary accident and deserves further attention from the department. Speed distribution generally presents a normal distribution [23]. In Fig. 4, the accident rates reach the lowest when the mean speed is in the range of 15%–20%. The speed deviation is slightly higher or lower than the mean, whereas the accident rates are increased. Therefore, controlling the vehicle speed in the appropriate mean range is an effective method for preventing accidents. This study recommends that the Transportation Safety Department control the mean speed in the range of 30–40 mile/h to reduce the occurrences of secondary accidents.

Table 4. Effects of speeds

Variables	High speed (>30 mile/h)	
	Value	Increase of risks
Accident level		
Days of the week	Weekend	+48%
Occurrence place	Intersection/bridge	+50%
Occurrence time	Night	+48%
Road level		
Speed limit	Low	+31%
Light conditions	Dusk	+48%
Road alignment	Curve upgrade/downgrade	+50%

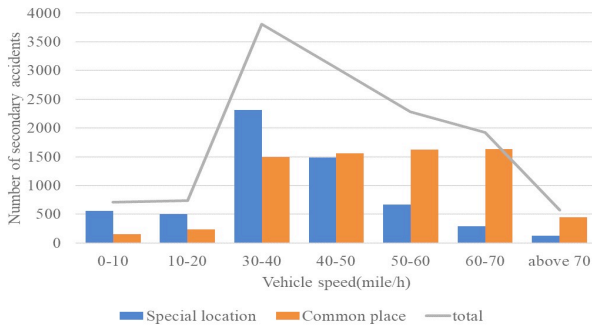


Fig. 3. Statistical information on vehicle speed in secondary accidents

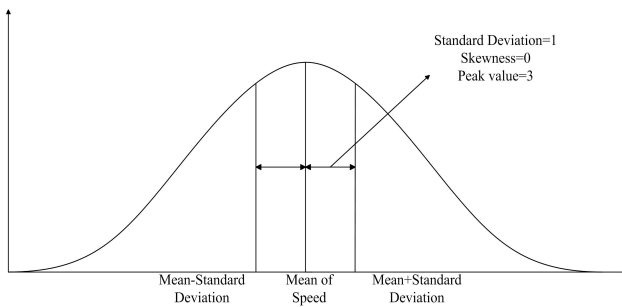


Fig. 4. Distribution of vehicle speed samples

Table 3 shows that the speed limit is negatively correlated with secondary accidents. The occurrence probability of secondary accidents decreases by 6% ($1 - e^{-0.06}$) when the speed limit increases from 30 mile/h to 60 mile/h. Generally, speed limit is related road level. The reason might be that the traffic environment in low-level roads is complicated and it easily causes secondary accidents. Hence, further attention should be paid to low-level roads.

4.2.2 Occurrence places

The Bayesian hierarchical model reflects that the occurrence probability of secondary accidents at intersections and bridges is 13% ($1 - e^{-0.13}$) higher than that in other places. This result conforms to the common sense. Drivers easily neglect the surrounding environment and traffic conditions at intersections and bridges due to road designs and geological conditions, thus resulting in the high occurrence of secondary accidents. In Fig. 3, the occurrence probability of secondary accidents at special places is higher than that at common places when the speed is lower than 40 mile/h. On the contrary, the occurrence probability of secondary accidents at common places is higher than that at special places when the speed is higher than 40 mile/h. The reason might be that the vehicle speed at special places is low due to limited geological conditions; thus, most secondary accidents mainly occur at common places under high vehicle speeds.

4.2.3 Occurrence time

Fig. 5 illustrates that secondary accidents mainly occur on weekends. Statistical data show that the number of secondary accidents during the day is slightly higher than that during night, which may be attributed to the small traffic flow early in the morning. The occurrence frequency of secondary accidents in the evening of weekends is higher than that during the day. Therefore, weekend evening is a time interval of the frequent occurrence of secondary accidents. The Bayesian hierarchical model results (Table 3) also demonstrate that the occurrence probabilities of secondary accidents on weekends and in the evening are 11% ($e^{0.10} - 1$) higher than those in common times. On the one hand, drivers cannot observe the external environment well due to poor visual conditions. On the other hand, traffic conditions are further complicated at weekends when the traffic flow is large, which influences the judgment of drivers.

4.2.4 Road characteristics

Results demonstrate that light conditions and road alignment can increase the occurrence probability of secondary accidents. The occurrence probability of secondary accidents on road sections with good light conditions is 11% ($e^{0.10} - 1$) lower than that on road sections without lights. Secondary accidents are mainly caused by ignoring the changes in traffic conditions in front or delayed response. Light improves the visual conditions of drivers, such that they have additional time to respond, thus decreasing the occurrence probability of secondary accidents. Similarly, the occurrence probability of secondary accidents on straight road sections is lower than those on upslope, downhill, and turns, which also might be the consequence of a poor field of view.

4.3 Differences at various places

Given the unique traffic conditions and environment in special places, such as intersections and bridges, the risk factors of secondary accidents are different from those in conventional road sections. Table 5 shows the model results at special and common places. In common places, the occurrence time, vehicle speed, road alignment, and light conditions play an important role on the occurrence of secondary accidents. However, days of the week, vehicle speed, and traffic signal become the major influencing factors in special places. Obviously, speed is a key factor in all places; it is positively related to the number of secondary accidents.

The field of view of drivers is a key factor in common places and is closely related with the occurrence time, vehicle speed, road alignment, and light conditions. Secondary accidents easily occur in the evenings because the field of view during the day is better than that in the evening. Similar to light conditions, the poor road alignment can also affect the field of view of drivers and cause secondary accidents.

In special places, the occurrence time (weekend or working days) and traffic signal control become prominent factors, which might be related to the large traffic flow and complicated traffic conditions during weekends. Vehicles are in chaos on road sections without traffic signal control, which can increase the occurrence probability of secondary accidents.

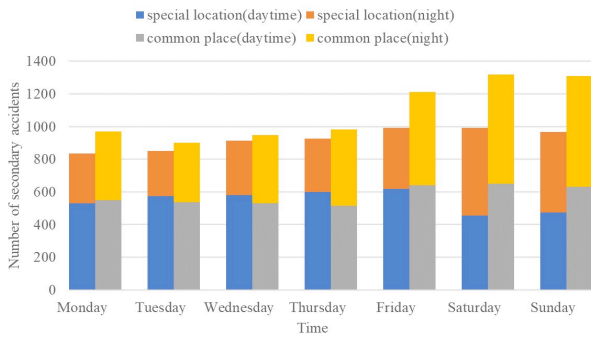


Fig. 5. Statistical information on the occurrence time of secondary accidents

5. Conclusions

A Bayesian hierarchical model was constructed in this study to solve the nesting and interaction problems of influencing factors of secondary accidents. Here, 11 influencing factors were divided into road and accident levels to analyze their effects on secondary accidents. Then, the Bayesian hierarchical model was compared with the Bayesian logistic model, which demonstrated the group effect of influencing factors. Finally, differences of influencing factors of secondary accidents at special and common places were discussed. Conclusions could be drawn as follows:

Table 5. Bayesian hierarchical model results on occurrence places

Variable	Special places				Common places			
	Effect estimate		95% BCI odds ratio		Effect estimate		95% BCI odds ratio	
	Mean (S.D.)	Odds ratio	2.50%	97.50%	Mean (S.D.)	Odds ratio	2.50%	97.50%
Intercept	-0.01 (0.031)	0.99	0.93	1.05	-0.01 (0.031)	0.99	0.93	1.05
Accident level								
Days of the week	0.05 (0.029)	1.05	1.00	1.12	0.05 (0.030)	1.06	0.995	1.12
Weather conditions	-0.01 (0.029)	0.99	0.93	1.05	0.03 (0.030)	1.03	0.97	1.09
Occurrence time	0.05 (0.029)	1.05	0.99	1.11	0.06 (0.030)	1.06	1.004	1.13
Vehicle speed	0.16 (0.029)	1.17	1.10	1.24	0.15 (0.029)	1.17	1.10	1.24
Road level								
Traffic signal control	-0.03 (0.029)	0.97	0.92	1.03	-0.02 (0.030)	0.98	0.92	1.04
Road alignment	0.04 (0.031)	1.04	0.98	1.11	0.06 (0.031)	1.07	1.003	1.13
Road condition	0.004 (0.030)	0.996	0.94	1.06	0.02 (0.030)	1.02	0.96	1.08
Light conditions	0.05 (0.030)	1.05	0.99	1.12	0.06 (0.030)	1.06	1.003	1.13
Number of straight lanes	-0.02 (0.031)	0.98	0.92	1.04	-0.02 (0.031)	0.98	0.92	1.04
Speed limit	-0.04 (0.026)	0.96	0.92	1.01	-0.05 (0.027)	0.05	0.90	1.002
Random utility	0.12 (0.10)	—	—	—	0.13 (0.11)	—	—	—
DIC value	4002.8				4018.9			

*Numbers in bold reflects that parameter estimation is substantial at the 95% confidence level.

(1) The Bayesian hierarchical model can accurately recognize the influencing factors of secondary accidents. Days of the week, occurrence time, and light conditions are three unsubstantial factors in the Bayesian logistic model, but they considerably influence the occurrence of secondary accidents in the Bayesian hierarchical model. The DIC value also proves the superiority of the Bayesian hierarchical model to the Bayesian logistic model. Therefore, the proposed Bayesian hierarchical model can reduce the error caused by the interaction of influencing factors.

(2) Vehicle speed, days of the week, occurrence time, light conditions, and road alignment are positively correlated with the occurrence of secondary accidents, whereas speed limit and occurrence place are negatively correlated with the occurrence of secondary accidents. Traffic signal control and time (weekend or working days) are important influencing factors of secondary accidents at special places, such as intersections and bridges. On the contrary, the occurrence

time, vehicle speed, road alignment, and light conditions are further important at common places.

These research conclusions provide new techniques and opinions for analyzing secondary accidents. However, this study has limitations. The relationship between the secondary and corresponding primary accidents is neglected due to inadequate data. Hence, future studies must explore a hierarchical model with 3–4 layers for evaluating such relationship.

Acknowledgements

This study was supported by the National Natural Science Foundation of China (Grant Number: E080701/51508325).

This is an Open Access article distributed under the terms of the Creative Commons Attribution Licence



References

- Owens, N., Armstrong, A., Sullivan, P., Mitchell, C., Newton, D., "Traffic Incident Management Handbook" [EB/OL]. Retrieved from https://ops.fhwa.dot.gov/eto_tim_pse/publications/timhandbook/tim_handbook.pdf, 2010-01-01/2018-06-25.
- Tedesco, S. A., Alexiadis, V., Loudon, W. R., Margiotta, R., Skinner, D., "Development of a model to assess the safety impacts of implementing IVHS user services", In: *Proceedings of the IVHS America Annual Meeting*, Atlanta, United States: Transportation Research Board, 1994, pp.343-352.
- Sun, C. C., Chilukuri, V., "Dynamic incident progression curve for classifying secondary traffic crashes". *Journal of Transportation Engineering*, 136(12), 2010, pp.1153-1158.
- Chilukuri, V., Sun, C., "Use of dynamic incident progression curve for classifying secondary accidents". In: *85th Annual Meeting of the Transportation Research Board*, Washington D.C., United States: Transportation Research Board, 2006, pp:1-12.
- Junhua, W., Boya, L., Lanfang, Z., Ragland, D. R., "Modeling secondary accidents identified by traffic shock waves". *Accident analysis and prevention*, 87, 2015, pp.141-147.
- Karlaftis, M. G., Latoski, S. P., Richards, N. J., Sinha, K. C., "Its impacts on safety and traffic management: an investigation of secondary crash causes". *Journal of Intelligent Transportation Systems*, 5(1), 1999, pp.39-52.

7. Yang, H., Ozbay, K., Xie, K., "Assessing the risk of secondary crashes on highways". *Journal of Safety Research*, 49(6), 2014, pp.143.e1-149.
8. H. Yang., B. Bartin., K. Ozbay., "Mining the characteristics of secondary crashes on highways". *Journal of Transportation Engineer.* 140(4), 2014, pp.04013024.
9. Zhan, C., Gan, A., Hadi, M., "Identifying secondary crashes and their contributing factors". *Transportation Research Record: Journal of the Transportation Research Board*, 2102(1), 2009, pp.68-75.
10. Xu, C., Liu, P., Yang, B., Wang, W., "Real-time estimation of secondary crash likelihood on freeways using high-resolution loop detector data". *Transportation Research Part C: Emerging Technologies*, 71, 2016, 406-418.
11. Eleni I. V., Matthew, G. K., Foteini P. O., "Modeling the effects of weather and traffic on the risk of secondary incidents". *Journal of Intelligent Transportation Systems*, 16(3), 2012, pp.109-117.
12. Yang, H., Bartin, B., Ozbay, K., "Use of Sensor Data to Identify Secondary Crashes on Freeways". In: *Transportation Research Board 92nd Annual Meeting*, Washington D.C., United States: Transportation Research Board, 2013, pp.82-92.
13. Qiang Z., Jiaren S., Huiying W., "Bayesian Hierarchical Modeling Monthly Crash Counts on Freeway Segments with Temporal Correlation," *Journal of Advanced Transportation*, 2017, 2017, pp.1-8.
14. Yu, R., Abdel-Aty, M., Ahmed, M., "Bayesian random effect models incorporating real-time weather and traffic data to investigate mountainous freeway hazardous factors". *Accident Analysis and Prevention*, 50(2), 2013, pp.371-376.
15. Kim, M., Kho, S. Y., Kim, D. K., "Hierarchical ordered model for injury severity of pedestrian crashes in south korea". *Journal of Safety Research*, 61, 2017, pp.33-40.
16. Zeng, Q., Wen, H., Huang, H., "The interactive effect on injury severity of driver-vehicle units in two-vehicle crashes". *Journal of Safety Research*, 59, 2016. pp.105-111.
17. Meiquan, X., Wen, Cheng., Gill, G.S., Zhou, J., Xudong, J., Choi, S., "Investigation of hit-and-run crash occurrence and severity using real-time loop detector data and hierarchical Bayesian binary logit model with random effects". *Traffic injury prevention*, 19(2), 2018, pp.207-213.
18. Chen, C., Zhang, G., Liu, X. C., Ci, Y., Huang, H., Ma, J., Chen, Y., Guan, H., "Driver injury severity outcome analysis in rural interstate highway crashes: a two-level bayesian logistic regression interpretation". *Accident analysis and prevention*, 97, 2016, pp.69-78.
19. Spiegelhalter, D. J., Best, N. G., Carlin, B. P., Linde, A. V. D., "Bayesian measures of model complexity and fit". *Journal of the Royal Statistical Society*, 64(4), 2003, pp.583-616.
20. Gelman, A., Hill, J., "Data Analysis Using Regression and Multilevel/Hierarchical Models". London: Cambridge University Press, United Kingdom, 2007, pp.408-410
21. Shaheed, M. S., Gkritza, K., Carriquiry, A. L., Hallmark, S. L., "Analysis of occupant injury severity in winter weather crashes: a fully bayesian multivariate approach". *Analytic Methods in Accident Research*, 11, 2016, pp.33-47.
22. Kopitch, L., Saphores, J., "Assessing the effectiveness of changeable message signs on secondary crashes". In: *90th Annual Meeting of the Transportation Research Board*, Washington D.C., United States: Transportation Research Board, 2011, pp.16.
23. D.Solomon., "Accident on Main Rural Highways Related to Speed, Driver, and Vehicle" [EB/OL]. Retrieved from https://safety.fhwa.dot.gov/speedmgmt/ref_mats/fhwasa1304/Resourc es3/40 - Accidents on Main Rural Highways Related to Speed, Driver, and Vehicle.pdf, 1964-07-01/2018-06-25.