



A crash risk identification method for freeway segments with horizontal curvature based on real-time vehicle kinetic response

Changjian Zhang^a, Jie He^{a,*}, Mark King^b, Ziyang Liu^a, Yikai Chen^c, Xintong Yan^a, Lu Xing^d, Hao Zhang^a

^a School of Transportation, Southeast University, Nanjing, 210018, China

^b Centre for Accident Research and Road Safety, Queensland University of Technology, Brisbane, 4059, Australia

^c School of Automobile and Traffic Engineering, Hefei University of Technology, Hefei, 230009, China

^d School of Traffic and Transportation Engineering, Changsha University of Science and Technology, Changsha, China



ARTICLE INFO

Keywords:

Freeway safety
Segments with horizontal curvature
Vehicle kinetic response
Negative binomial regression model
Random effects negative binomial regression model

ABSTRACT

With the development and maturation of vehicle-based data acquisition technology, in-vehicle data is increasingly being used to explore road safety. This paper reports on research that analyzed the real-time tire force data (kinetic response) obtained from vehicle kinetic experiments, and constructed a new approach for identifying the high-risk of crashes on freeway segments with horizontal curvature. First, the road was divided into 1km units. Then, taking into account the characteristics of freeway alignment, each segment with horizontal curve was selected as the object of subsequent analysis. Automotive instrumentation was used to obtain a measure of tire force in the course of normal driving. The entire data set was preprocessed according to rate of change and the density of the data was reduced. By defining the outliers of the kinetic data and conducting factor analysis, two representative crash risk indicators of longitudinal and lateral stability were obtained. Negative binomial regression model (NBR model) and random effects negative binomial regression model (RENBR model) were constructed and jointly applied based on the new indicators to predict the risk value of horizontal curve segments. The method showed good prediction performance (71.8 %) for high-risk road segments with design flaws, but the predicted effect for low-risk road segments was not ideal. This study not only illustrated the effectiveness of in-vehicle data in assessing road crash risk by coupling multiple kinetic parameters, but also provided support for freeway safety research using surrogate measures of risk when there is a lack of crash statistics.

1. Introduction

Freeway traffic safety is a topic of widespread research interest. Although many methods, technologies and indicators have been applied to the prediction and identification of traffic safety conditions (Dixit et al., 2014; Gettman et al., 2008; Li et al., 2017), crashes on freeways remain high and severe in many countries. China is the country with the most road traffic crashes in the world. In 2016, there were 8934 traffic crashes on China's freeways, causing 5947 deaths and 11,956 injuries. Although freeways only account for 2.8 % of the total length of public roads in China, the number of traffic crashes, injuries and deaths accounted for 7.7 %, 9.4 %, and 13.7 % of the total traffic crashes, injuries and deaths respectively. Compared with other grades of road, the crash rate, injury rate, and mortality rate per 100 km of freeway are 3.0 times, 3.8 times, and 5.1 times higher, respectively.

Traditional traffic safety analysis methods are mainly focused on traffic crash data. Road segments that may have potential safety hazards can be identified by mining historical crash data (Geurts et al., 2004; Geurts and Wets, 2003). However, this method cannot provide a deep understanding of the causes of traffic crashes, is retrospective (so that a problem must become evident through crashes before it is addressed) and it is therefore difficult to improve the safety of freeways in a targeted manner. Alternative approaches are being increasingly explored, relying on a knowledge of the factors known to contribute to traffic crashes (such as driver behavior, road linear characteristics, road environments, weather, traffic flow, etc) and then utilizing theory, statistical analysis and model construction to predict safety outcomes and develop targeted suggestions for improving the road environment (Xing et al., 2020; Dadashova et al., 2016; Ellison et al., 2015; Urie et al., 2016; Wang et al., 2015).

* Corresponding author.

E-mail address: hejie@seu.edu.cn (J. He).



Fig. 1. Schematic map of study road segment scope.

Only a few studies have focused on vehicle factors. The reason may be that it is difficult to obtain the motion parameters of a vehicle in a crash. For this reason, some scholars have shifted the research focus to simulation experiments (P. Li and He, 2016; Lv et al., 2018; Yang et al., 2019). Unfortunately, because of the assumptions and simplifications of the various parameters involved in the simulation, it is impossible to describe the true condition of the road completely and comprehensively, which limits the accuracy of the research results.

With the development and popularization of on-board GIS and GPS, it has become possible to use in-vehicle data to investigate road safety. It is worth mentioning that "naturalistic driving data" as a branch of "in-vehicle data" is widely used in the field of traffic safety analysis. Some of this research has recognized the link between driving events and the potential risk of crashes (Pande et al., 2017). These events are called "crash surrogate events", and have proved useful in studying the prediction and identification of road locations with potential risk (Guo et al., 2010). An even broader range of data can be obtained through naturalistic driving studies (NDS) that have been conducted over the past decade to collect objective data. Examples of such studies are the "100-car naturalistic driving study", which used naturalistic driving data from 100 cars to explore the characteristics of rear-end crashes, then the US Strategic Highway Research Program Phase 2 deployed 2800 cars to investigate and analyze various road safety issues (Antin et al., 2011; Ghasemzadeh and Ahmed, 2019; Hallmark et al., 2013).

Historical crash data can be combined with in-vehicle data to enable predictions of safety on roadways based on crash surrogate events and other driving data. Wu (2012) used the logit model to verify the causal relationship between crash surrogate events and crashes by studying the data of 100 vehicles. It was found that when crash surrogate events exceed a certain threshold, it may lead to a crash, so that this threshold raised the potential risk of the crash.

This study focuses on vehicle kinetics as crash surrogate events. It has been found that when a vehicle's lateral acceleration exceeded 0.7 g, the potential risk of a crash increased 24 times (Wu and Jovanis, 2013). Additionally, Dingus (2006) proposed the use of vehicle longitudinal acceleration ≥ 0.7 g and lateral acceleration ≥ 0.6 g as crash surrogate events. Based on this, Kluger et al. (2016) used longitudinal acceleration as an indicator to find "abnormal events "from a large amount of

in-vehicle data, and predicted crash-prone locations on road segments, with a prediction accuracy rate of 78 %. Gitelman et al. (2018) explored the relationship between in-vehicle data (braking and speed alert), road characteristics, and traffic crashes, and built a negative binomial regression model to predict crashes and determine high-risk locations on road segments. Moreover, kinematic parameters such as the vehicle's deceleration rate, jerks, and yaw rate have been included in the study of "crash surrogate events" with significant effects (Dingus et al., 2006; Hydén, 1987; Pande et al., 2017).

These studies have confirmed a strong correlation between single kinematic parameters collected by in-vehicle recorder and traffic crashes. Nevertheless, in reality, crashes are often the result of the transformation of multiple factors. A single indicator can easily make the research results have certain one-sidedness. Therefore, this paper attempts to combine the effects of various forces and moments experienced by the vehicle during normal driving, and discover the mathematical relationship between vehicle kinetic parameters and crash risk.

In general, crash-prone segments on freeways frequently exhibit adverse factors, such as cross winds, poor alignment, severe road damage, and transition to transportation facilities such as tunnels. These factors tend to directly affect the vehicle, causing its kinetic parameters to produce different outputs compared to safe driving. This paper describes a real-car experiment using a vehicle equipped with the automotive instrumentation to examine multiple vehicle kinetic parameters on freeway segments with horizontal curvature under normal driving conditions. Based on analyses that reduce the data to two factors - a longitudinal stability indicator and a lateral stability indicator - the road crash risk value was predicted. This development enables the strong correlation between vehicle kinetic parameters and the high-risk of freeway segments with horizontal curvature, thereby reducing dependence on historical crash data.

Therefore, the results are also applicable to new freeways with kinetic simulation data and other roads that lack crash statistics. For newly-built roads or roads lacking historical crash data statistics (such roads are widespread in China), similar experiment can be used to collect tire kinetic data on these roads. The road segments that may have a high-risk of crashes will be initially identified and handed over. Based on the recognition results, the department can formulate measures to

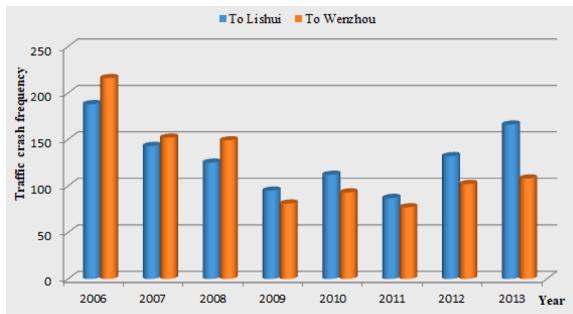


Fig. 2. Wenli freeway crash distribution 2006-13.

improve road safety in a targeted manner. For the roads that are still in the design stage, people only need to build a road model in the simulation software that is consistent with the design drawing, and then build a vehicle model and a driving control model, the tire kinetic data can be obtained of the road and the safety analysis can be implemented.

2. Data preparation and preprocessing

2.1. Objective segments and road crash analysis

The experimental segments are K117 ~ K189 of Wenli Freeway in Zhejiang Province, with a maximum speed limit of 100 km/h, as shown in Fig. 1. Considering that most segments of Wenli Freeway (K117 ~ K189) are separated lanes, Liu (2020) divided them into the lanes traveling to Wenzhou (to the southeast) and to Lishui (to the northwest), giving 144 segments of 1 km length. Then according to the horizontal curve radius, each segment was divided into straight and curved segments, of which 36 were straight segments and 108 were segments with horizontal curvature. This paper will analyze the distribution of road crashes based on the results of the above studies.

A total of 2026 traffic crashes occurred on the experimental segments from 2006 to 2013, including 1048 traveling to Lishui and 978 traveling to Wenzhou. In this study, in order to match the kinetic data of the road segments, the crash data used in this study has been divided according to the direction. Fig. 2 shows the distribution of crashes.

For the purpose of facilitating the analysis of road safety, the crashes on each road segment were converted to a number of "equivalent crashes" by weighting historical crashes according to their severity, to account for the varying levels of social and economic impact associated with different levels of severity. The calculation formula is:

$$N_e = N + \alpha N_1 + \beta N_2 + \gamma N_3 \quad (1)$$

Where N_e , N , N_1 , N_2 and N_3 are equivalent crashes, total crashes, fatal crashes, severe crashes, and minor crashes, respectively; α , β and γ are equivalence coefficients, which are generally take values $\alpha = 2.0$, $\beta = 1.5$, $\gamma = 1.2$ (in China) (Geng and Peng, 2018; Y. Wang and Wang, 2019; Yuan et al., 2010; Zhang and Ma, 2010). These coefficients are generally determined according to the characteristics of different countries and regions, such as the degree of social development, economic conditions, transportation policies and laws. If the equivalence coefficients were not used, every fatal crash, severe crash, minor crash, and no casualty crash would be treated in the same amount, which could confuse the identification of high crash risk locations.

Compared with other straight segments, segments with horizontal curvature have a higher crash rate. Therefore, when identifying potential road crash risks, segments with horizontal curvature are usually considered (Dahir and Hassan, 2019). By analyzing the distribution of equivalent crashes on the experimental road, it is found that there are more segments of horizontal curvature and more crashes on them, compared with straight segments (3 times and 1.52 times respectively). This paper focuses on segments with horizontal curvature, and explores

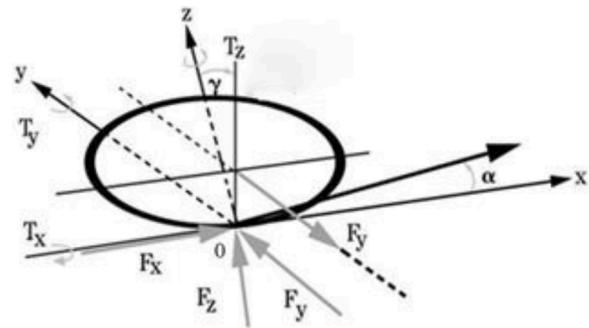


Fig. 3. Wheel coordinate system and force analysis diagram.

whether there are mathematical relationships between the kinetic response of a vehicle negotiating the segment and the risk of a road crash under normal driving conditions.

2.2. Acquisition of road kinetic data

When the car is driving, the tire is subjected to forces in three directions, and to moments about three axes, which together constitute the six forces acting on the tire. The International Association of Automotive Engineers (SAE) specifies a standard tire motion coordinate system, and defines the tire force and torque and related motion variables (Zhou et al., 2007), as shown in Fig. 3. The force and torque acting on the wheels are the ground tangential reaction force "Fx", the ground lateral reaction force "Fy", and the ground normal reaction force "Fz", while the moments are the overturning torque "Tx", the rolling resistance torque "Ty", and the returning torque "Tz" generated by the ground reaction force.

An experimental HONDA CR-V car and automotive instrumentation were used for the experimental study. Automotive instrumentation consisted of a 6 component force axle load sensor SLW-ND, a wheel alignment measuring system, a wireless signal receiver DT-24R, a 6 component force measuring analyzer MFT-306 T, control software WAM-701A, a multifunction car recorder TMR-200 and a laptop PC. The wheel alignment measuring system consisted of wheel alignment sensor WAD-1A and special measuring analyzer WAM-1A, were installed together with the SLW-ND in the right front wheel of the HONDA CR-V car, shown in Fig. 4.

Lateral tire forces were measured by the SLW-ND and then were transferred to the MFT-306 T by the DT-24R. The WAD-1A ensured the signal acquisition precision of the SLW-ND, while the SLW-ND ensured the signal processing precision for the WAM-1A.

The kinetic data were collected by the instrumented vehicle. The specific test condition is shown in Table 1. In this study, the driver's driving experience was 15 years. He was familiar with the purpose of this experiment and had sufficient rest before driving. There was no behavior such as drinking and drug use that had a detrimental influence on driving control. The driver remained sober and focused throughout the driving. Such an experimental design not only can minimize the interference of driver factors on the relationship between kinetic parameters and crash risk, but also ensures that this method can be more convenient and widely applied. However, it does some existing problems, for the reason that it is impossible to accurately control and describe the driver variable. In the future, when the traffic safety management department uses this method to identify high crash risk road segments, the prediction results may produce a certain deviation (due to different drivers). Considering these issues, we applied cruise mode in the experiment in an attempt to further weaken the impact of driver characteristics.

In addition, for the purpose of expanding the data samples and improving the accuracy of the study, the experiment was conducted on the experimental road for 2 cycles under the same conditions, and data



Fig. 4. Experimental setup and testing: (a)–(d) Experimental HONDA CR-V car and automotive instrumentation; (e) field test.

Table 1
Experimental conditions.

Conditions	Content
Weather	Sunny, breeze
Road segment	Wenli Freeway towards Lishui and towards Wenzhou, K117-K189
Road surface	Class A highway pavement
Experimental vehicle	Dongfeng Honda CR-V
Driving speed	100 km / h uniform speed
Collection data type	Longitudinal force / moment, lateral force / moment, vertical force / moment, speed
Data collection interval	0.01 s
Test wheel	Right front wheel

was therefore obtained for a total of $108 \times 2 = 216$ segments with horizontal curvature (recorded $asi = 1, 2, \dots, 108, 109, \dots, 216$). Note that the method of collecting the vehicle data means that, even though the same segment may have had two sets of measurements taken, they were not at the same locations and are effectively independent for the purposes of analysis.

2.3. Vehicle kinetic data preprocessing

Preprocessing was undertaken since the original data was too dense (data collection interval 0.01 s), and given that the absolute size of the kinetic response was closely related to the body weight and the vehicle load, the absolute values were transformed into rates of change. Based on this, this paper first calculated the average value of various types of kinetic data for every 0.1 s interval and recorded it as $data_{t,n}^{0.1}$, where t represents the t -th 0.1 s, and n represents a certain type of kinetic parameter. Then, the change rates of various types of kinetic data within each 0.1 s interval compared with its predecessor was calculated according to Eq. (2).

$$x_{t,n}^{0.1} = |data_{t+1,n}^{0.1} - data_{t,n}^{0.1}| \quad (2)$$

Where $x_{t,n}^{0.1}$ is the change rate of a certain type of kinetic data in the t -th 0.1 s.

In order to simplify the data set, the average value of the 10 rates of change of various types of kinetics within each second was calculated and recorded as $x_{t,n}$, the average rate per 0.1 s for the t -th 1 s. The seven types of parameters used for this study were: the change rate of “Fx” (named CRFx), the change rate of “Fy” (named CRFy), the change rate of

Table 2
Descriptive statistics of the original data.

Variable	Obs	Mean	Std. Dev.	Min	Max
CRFx	6871	339.678	109.421	65.2	1316.4
CRFy	6871	263.968	97.023	55.6	1108
CRFz	6871	1209.561	674.303	244	6785.6
CRTx	6871	.08	.034	.015	.393
CRTy	6871	.033	.026	.007	.454
CRTz	6871	.029	.009	.007	.088
CRV	6871	.785	.365	.15	7.741

Table 3
Descriptive statistics of anomaly floating data.

Variable	Obs	Mean	Std. Dev.	Min	Max
CRFx	694	558.149	93.274	476.8	1316.4
CRFy	688	472.718	100.9	380.4	1108
CRFz	688	2749.087	728.945	2044	6785.6
CRTx	689	.156	.036	.122	.393
CRTy	689	.094	.042	.056	.454
CRTz	696	.046	.005	.041	.088
CRV	687	1.583	.466	1.207	7.741

"Fz" (named CRFz), the change rate of "Tx" (named CRTx), the change rate of "Ty" (named CRTy), the change rate of "Tz" (named CRTz), and the change rate of "V" (named CRV).

On this basis, 216 segments with horizontal curvature were screened based on road geometric alignment and video data from the driving recorder. Thirty segments were discarded, including segments with missing data, rainwater, and construction site segments. Only the data for segments with horizontal curvature that had dry and good road conditions during the drive were retained. After screening, 186 fin. 1 valid segments were determined (recorded as segment = 1,2,3...i...186).

2.4. Data set description

After filtering, a total of 6871 observations across the 186 road segments were generated, which was called the original data set. Table 2 shows the statistics for original data sets.

Based on the mean and standard deviation information, it can be inferred that a part of the data has a large degree of deviation. The occurrence of outliers may be caused by the following 4 reasons: 1) The geometric design of some road segments is flawed, causing abnormal fluctuations in the force of the vehicle; 2) When entering or exiting a special transportation facility such as a tunnel, the data may fluctuate abnormally due to sudden changes in the traffic environment; 3) Due to the discontinuous transition between road alignment and transportation facilities, traffic flow is disrupted. For example, the tunnel exit is connected to a long downhill section, or the tunnel exit is connected to a sharp bend; 4) Other factors such as road surface damage and improper application of traffic signs cause drivers to make discontinuous input to the vehicle.

In order to mine the characteristics of abnormal floating data for each road segment, the largest 10 % of the various types of kinetic data sets were defined as the anomaly floating data set (Wu and Jovanis,

2012), named as "AFDS". As can be seen from Table 3, the AFDS for the study variables ranged in size between 688 and 696 of the 6871 observations (i.e. roughly 10 %). Comparison with Table 4 shows that both the mean and minimum values of the AFDS were markedly higher than the original dataset. Therefore, there was reason to believe that other road-related factors have caused the vehicle kinetic response to produce abnormal fluctuations on these road segments.

In particular, some other methods of reducing the size of the data set have been tried during the preprocessing of kinetic data, such as ① selecting the extreme value; ② selecting the 75 % quantile; ③ selecting the median; ④ defining the data outside the range of $\bar{X} \pm n\sigma$ ($n = 1, 2, 3$) used for research; ⑤ taking the average of a certain time unit. After several attempts, this study finally adopted the combination of " (see 2.4 section) " + "⑤ (the time unit is 0.1 s, see 2.3 section) ". The advantage of this approach is as follow:

- (1) The time interval of 0.1 s is relatively small. Taking 0.1 s as the time unit for averaging can not only reduce the data set, but reflect the impact of extreme values on the mean. If the average value is calculated directly in units of 1 s, some extreme values of kinetic data may be easily overwritten.
- (2) Considering that the risk of traffic crashes is likely to be related to extreme values, the maximum 10 % of the change rate of kinetic data is selected as the AFDS.
- (3) If the extreme value is directly selected for analysis without the mean value processing, it may be misled by extreme values of "false positives". The distance between adjacent data points in the original data set is only about 20 cm, and the risk of road crashes is often the result of the combined action of a certain long-distance road segment. Therefore, it is often accidental if the data of just a certain location or a few locations is abnormal. With the purpose of avoiding similar interference, the mean value is first obtained within a certain range, and then the extreme value processing is performed, which can weaken the effect of "false positive" data.

3. Definition of crash risk and kinetic variables

3.1. Defining road crash risk

The reason for defining road crash risk as Eq. (3) is as follows:

- Due to the large range of original equivalent crashes, it is difficult to judge the prediction accuracy of the model in Section 5.
- It is also determined by the structure of the model used in Section 4. It can reduce the sensitivity of the estimation results to the intercept term (logarithm operation), and map the results of all models and historical crash to the same dimension for comparison (divided by the mean of the overall risk).

$$R_i = round\left(\frac{\ln(N_{ei})}{\sum \ln(N_{ei})}, 1\right) \quad (3)$$

Where N_{ei} is the equivalent crash data of the i -th road segment, R_i is the risk value of the i -th road segment; " \ln " is the logarithm; "round" means round the result; "1" means take one decimal place.

Table 4
Correlation analysis of road segment R value and vehicle kinetic anomaly data.

Variables	R value	CRFx_1	CRFy_1	CRFz_1	CRTx_1	CRTy_1	CRTz_1	CRV_1
R value	1.000	0.157	0.164**	0.166**	0.145**	0.286**	-0.026	0.326**
Variables	R value	CRFx_2	CRFy_2	CRFz_2	CRTx_2	CRTy_2	CRTz_2	CRV_2
R value	1.000	0.189**	0.202**	0.202**	0.209**	0.305**	0.048	0.329**
Variables	R value	CRFx_3	CRFy_3	CRFz_3	CRTx_3	CRTy_3	CRTz_3	CRV_3
R value	1.000	0.193**	0.203**	0.203**	0.207**	0.306**	0.050	0.330**

** shows significance at the 0.05 level.

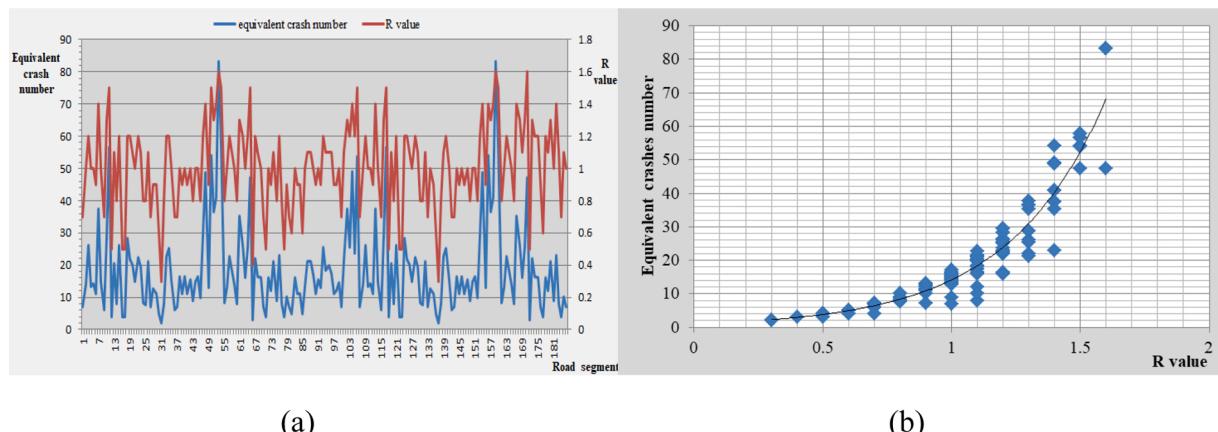


Fig. 5. The relationship between the *R* value and equivalent crashes (a- Trend distribution plot; b- Scatter plot).

The value of crash risk of 186 valid segments was calculated according to Eq. (3). Fig. 5 provides the distribution of *R* value and the number of equivalent crashes for each road segment. It can be seen that the trends in the two curves are basically consistent. That is, the *R* value of a road segment is a reasonable reflection of the historical crash data (weighted by severity) for that segment.

3.2. Define outlier for kinetic data

So as to more clearly explore the relationship between the kinetic data of different anomaly floating levels and crashes, and to match these kinetic data with road segments reasonably, the AFDS of the kinetic data was divided into three grades according to the following rules:

- Take the minimum value as the lower bound and the maximum value as the upper bound. This paper recategorises each AFDS into 3 subsets, corresponding to 3 nested anomaly grades. Note that the anomaly grade I is “CRFx_1”, “CRFy_1”, “CRFz_1”, “CRTx_1”, “CRTy_1”, “CRTz_1”, “CRV_1”; the anomaly grade II is “CRFx_2”, “CRFy_2”, “CRFz_2”, “CRTx_2”, “CRTy_2”, “CRTz_2”, “CRV_2”; the anomaly grade III is “CRFx_3”, “CRFy_3”, “CRFz_3”, “CRTx_3”, “CRTy_3”, “CRTz_3”, “CRV_3”.

$$x_{mn} \in \begin{cases} \text{Anomaly I, if } \min_{x_n} \leq x_{mn} < \min_{x_n} + \frac{\max_{x_n} - \min_{x_n}}{3} \\ \text{Anomaly II, if } \min_{x_n} \leq x_{mn} < \min_{x_n} + \frac{2(\max_{x_n} - \min_{x_n})}{3} \\ \text{Anomaly III, if } \min_{x_n} \leq x_{mn} < \max_{x_n} \end{cases} \quad (4)$$

Where x_{mn} is the value of a certain type of kinetic change rate at position m , where $m \in i$, \min_{x_n} and \max_{x_n} are the minimum and maximum values of a certain type of kinetic change rate at 186 segments.

- If a certain type of vehicle kinetic change rate at a certain position falls within a range of anomaly grade, the data is recorded as 1,

otherwise it is recorded as 0. As shown in Eq. (5), where a is the anomaly grade, $a = \text{I, II, III}$.

$$x_{mn}^a = \begin{cases} 0, \text{if } x_{mn} \notin \text{Anomaly level } a; \\ 1, \text{if } x_{mn} \in \text{Anomaly level } a; \end{cases} \quad (5)$$

- Calculate outliers at various grades of various types of kinetic change rates for each road segments as shown in Eq. (6), where A_{in}^a represents a certain class of outlier of a certain type of kinetic data in the road segment i ($a = \text{I or II or III}$). It is assumed that in the “CRFx” abnormal floating data set of segment i , 6 data fall into the range of abnormal grade I, 9 data fall into the range of grade II, and 12 data fall into the range of grade III. That is, for “CRFx”, outlier for grade I is 6, the outlier for grade II is 9, and the outlier for grade III is 12, which are recorded as “CRFx_1 = 6”, “CRFx_2 = 9”, and “CRFx_3 = 12”. Therefore, when the kinetic of certain types in a road segment fluctuate more strongly, the corresponding high-level outliers also increase.

$$A_{in}^a = \sum_{m \in i} x_{mn}^a \quad (6)$$

Table 4 summarizes the results of the analysis of the correlation between the *R* value of the road segments and the outlier for each anomaly grade.

- The rate of change of returning moment “CRTz” had the lowest correlation with the *R* value, and the rate of change of speed “CRV” had the greatest correlation with the risk level of the road segment.
- “CRFx”, “CRFy”, “CRFz”, “CRTx”, “CRTy” and “CRV” were all highly correlated with the risk level of the road segment with 95 % confidence interval.
- As the anomaly grade of kinetic data increased, the correlation between outlier for various kinetic data and the *R* value increased.

Table 5
Correlation analysis.

Variables	CRFx_3	CRFy_3	CRFz_3	CRTx_3	CRTy_3	CRTz_3	CRV_3
CRFx_3	1.000						
CRFy_3	0.385**	1.000					
CRFz_3	0.482**	0.744**	1.000				
CRTx_3	0.418**	0.826**	0.798**	1.000			
CRTy_3	0.137	0.054	0.087	0.117	1.000		
CRTz_3	0.348**	0.325**	0.286**	0.344**	-0.089	1.000	
CRV_3	0.284**	0.271**	0.331**	0.348**	0.559**	0.080	1.000

** shows significance at the 0.05 level.

Table 6
Factor rotation.

Factor	Variance	Difference	Proportion	Cumulative
Factor1	2.620	1.656	0.768	0.768
Factor2	0.964	0.621	0.282	1.050
Variable	Factor1	Factor2	Uniqueness	
CRFx_3	0.426	0.184	0.624	
CRFy_3	0.864	0.043	0.247	
CRFz_3	0.832	0.127	0.270	
CRTx_3	0.897	0.128	0.171	
CRTy_3	0.033	0.668	0.552	
CRTz_3	0.345	-0.080	0.733	
CRV_3	0.273	0.666	0.476	

LR test: independent vs. saturated: chi2 (21) = 595.74 Prob > chi2 = 0.0000.
Rotated factor loadings (pattern matrix) and unique variances.

- All the correlation coefficients were small, probably because traffic crashes were most likely the result of the coupling of multiple kinetic response.

Because of its higher correlations, the third-grade outlier classification was selected as the basic variables for building the model, namely “CRFx_3”, “CRFy_3”, “CRFz_3”, “CRTx_3”, “CRTy_3”, “CRTz_3” and “CRV_3”.

4. Construction of crash risk identification model

4.1. Correlation analysis of kinetic variables

The tire is an elastic body. In the course of vehicle operation, the tire is always subjected to uneven forces due to road alignment, elevation, braking and acceleration. These forces interact with each other to result in irregular deformation of the ground contact area of the tire. Consequently, in order to determine whether there was a highly correlated interference between various types of kinetic variables, a Pearson correlation analysis was undertaken for the variables “CRFx_3”, “CRFy_3”, “CRFz_3”, “CRTx_3”, “CRTy_3”, “CRTz_3” and “CRV_3”.

According to the analysis results in Table 5, there is a high degree of correlation between the variables. When the correlation between explanatory variables is high, factor analysis can be undertaken to reduce the dimensionality of independent variables. By combing the interactions between variables, the principle of least information loss is used to extract unobservable factors affecting the explanatory variables. The highly correlated original variables are transformed into linear combinations of independent factors. The model structure is provided in Eq. (7):

$$\begin{aligned} X_1 - \mu_1 &= \ell_{11}F_1 + \ell_{12}F_2 + \dots + \ell_{1q}F_q + \varepsilon_1 \\ X_2 - \mu_2 &= \ell_{21}F_1 + \ell_{22}F_2 + \dots + \ell_{2q}F_q + \varepsilon_2 \\ \vdots &\vdots \vdots \vdots \\ X_p - \mu_p &= \ell_{p1}F_1 + \ell_{p2}F_2 + \dots + \ell_{pq}F_q + \varepsilon_p \end{aligned} \quad (7)$$

Its matrix expression is:

$$(X - \mu)_{p \times 1} = L_{p \times q} F_{q \times 1} + \varepsilon_{p \times 1} \quad (8)$$

Where F_q is the extracted new factor and ℓ_{pq} is the factor load, the value of which is between 0 and 1. The closer to 1, the greater the degree of influence of the original variable (Washington et al., 2010). ε_p is a random error term related to X_p only.

Next, factor analysis was used to mine unobservable events of various kinetic variables. Table 6 summarizes the preliminary results of factor analysis after rotating. From the p value, it could be seen that the use of factor analysis in this example being significant statistically significant. Rotating the factor loading matrix can further simplify the factor structure and facilitate the mining of the representative meaning of each factor. Among them, factor 1 had high loadings from “CRFy_3”,

Table 7
Factor analysis results after trimming variables.

Factor	Variance	Difference	Proportion	Cumulative
Factor1	2.332	1.412	0.815	0.815
Factor2	0.920	.	0.322	1.137
Variable	Factor1	Factor2	Uniqueness	
CRFy_3	0.867	0.056	0.245	
CRFz_3	0.836	0.127	0.285	
CRTx_3	0.899	0.144	0.170	
CRTy_3	0.027	0.658	0.566	
CRV_3	0.268	0.669	0.481	

LR test: independent vs. saturated: chi2(10) = 503.72 Prob > chi2 = 0.0000.
Rotated factor loadings (pattern matrix) and unique variances.

Table 8
High loading factor analysis.

Factor and label	High loading variables	Variable name
Factor1: Lateral stability indicator	CRFy	Rate of change of lateral force
	CRFz	Rate of change of vertical force
	CRTx	Rate of change of overturning moment
Factor2: Longitudinal stability indicator	CRTy	Rate of change of rolling moment
	CRV	Rate of change of speed

Table 9
The result of NBR model.

Variables	Coef.	St.Err.	t-value	p-value	Sig
Lateral stability indicator	0.097	0.050	1.93	0.054	*
Longitudinal stability indicator	0.223	0.059	3.80	0.000	***
Constant	2.898	0.047	61.43	0.000	***
Akaike crit. (AIC)	1398.313			Prob > chi2	0.000

*** p < 0.01, ** p < 0.05, *p < 0.1.

“CRFz_3”, and “CRTx_3”; factor 2 had high loadings from “CRTy_3” and “CRV_3”.

Both “CRFx_3” and “CRTz_3” did not load strongly onto factors 1 and 2, so they were excluded and factor analysis was conducted on the remaining 5 variables. The analysis process is shown in Table 7, revealing a clearer factor structure. Table 8 shows the two factors finally extracted. In accordance with the variables that loaded onto them, they were labeled “lateral stability indicator” and “longitudinal stability indicator”.

4.2. Negative binomial regression model

The NBR model was constructed to predict the R value of the road segments in terms of the two indicators (factors). The NBR model is widely used in road crash research to predict crash frequency (Abdel-Aty and Radwan, 2000; Milton and Mannerling, 1998; Sarhan and Hassan, 2008). In this example, the lateral stability indicator and longitudinal stability indicator were used as explanatory variables in the NBR model, and the number of equivalent crashes was the dependent variable. Table 9 displays the output of the NBR model. It can be seen from the p value that the model's parameters were statistically significant. Furthermore, the longitudinal stability indicator had a stronger impact on the crash than the lateral stability indicator, and the lateral stability indicator was only meaningful at a significance level of 0.1.

The NBR model does not take into account the random effects among the sample data of each group, i.e. the NBR model assumes that the mean attribute effect for each segment is fixed. However, there may be some

Table 10

Random-effects negative binomial regression.

Variables	Coef.	St.Err.	t-value	p-value	Sig
Lateral stability indicator	0.116	0.043	2.69	0.007	***
Longitudinal stability indicator	0.201	0.056	3.59	0.000	***
Constant	1.000	0.115	8.67	0.000	***
Akaike crit. (AIC)	1405.568		Prob>chi2		0.000

*** p < 0.01, ** p < 0.05, * p < 0.1.

Table 11

Prediction results of R value.

Segment	R value	Prediction result of NBR model	Prediction result of RENBR model
1	0.7	1.1	1.2
2	1.0	0.9	0.8
3	1.2	1.0	1.1
4	1.0	1.0	1.2
5	1.0	1.1	1.3
...
182	1.0	0.9	0.8
183	1.4	0.9	0.9
184	1.1	1.0	1.1
185	0.7	0.9	0.8
186	1.1	1.0	1.0

non-observed factors independent of the observed variables that have an impact on the results (Hou et al., 2018; Ma et al., 2017), i.e. the mean attribute effect in a segment may vary for extraneous reasons, so that each measure of an attribute is drawn from a distribution which is not fixed but varies over time. This problem can be addressed by using a random effects regression model (RENBR model) (Naznin et al., 2016).

4.3. Random effects negative binomial regression model

The vehicle lateral stability indicator and longitudinal stability indicator (factors 1 and 2) were again used as explanatory variables of the RENBR model. Table 10 presents the parameter estimation results. The p value confirmed that the model and its parameters were statistically significant, and that all explanatory variables were significant at level of 0.01. Moreover, it can be seen from the model that both the longitudinal stability indicator and the lateral stability indicator had positive effects on the crash, consistent with the characteristics of the attributes of the original data set.

The estimated results of RNB model and RENBR model is a good proof of our hypothesis before conducting the research, that although some crashes occurred in bad weather such as rain and snow, they are closely related to the design defects of the road itself and these are intuitively reflected in the abnormal excitation of the tires.

Table 12

Road risk prediction results.

	Model results		Accuracy rate (%)
	Low-risk	High-risk	
Historical crash risk	Low-risk	40	29
	High-risk	33	84
			71.8

5. Results analysis

The number of equivalent crashes, which would convert to risk value according to Eq. (3), was predicted based on the estimation results of the model. Table 11 present the comparison of historical crash risk value, the risk value predicted by NBR model and the risk value predicted by RENBR model.

In order to quantitatively analyze the prediction results, the road segments were divided into two sets of high risk and low risk according to the historical crash data, and the output of the model was analyzed by checking the identification accuracy of the high/low risk road segment.

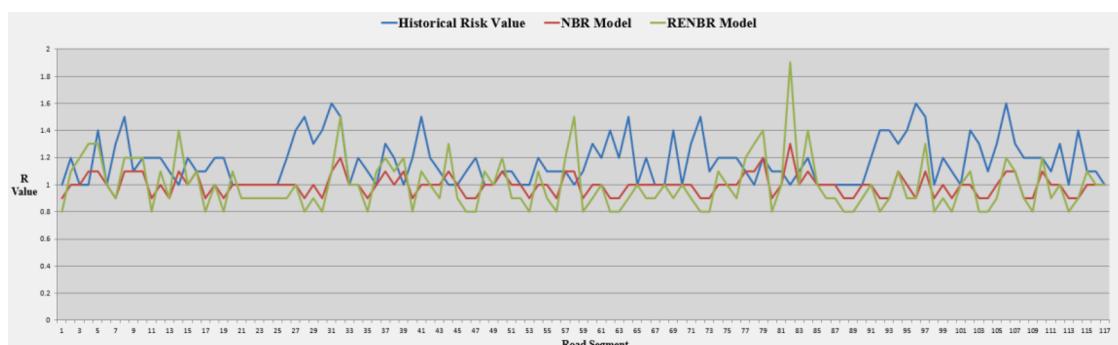
With reference to another literature (Liu et al., 2020), considering that the average value of historical crash risk value is 1, "R = 1" is positioned as the threshold of high and low risk. If the crash risk value of the road segment is less than the threshold, the road segment is "the road segment with relatively low crash risk" (note as low-risk); otherwise it is "the road segment with relatively high crash risk" (note as high-risk).

As shown by the curve in Fig. 6, the prediction range of the NBR model is relatively concentrated, and the prediction range of the RENBR model is relatively large. That is to say the NBR model has a better performance on some road segments with moderate risk values, and on the road segments with higher risk values, the RENBR model has a better performance.

Therefore, the output of the two models are coupled to identify segments with high-risk by combining the performance advantages of the two (ie: if a certain segment is the "road segment with relatively high crash risk", then the prediction fails only when the prediction results of the NBR model and the RENBR model are both <1). Table 12 illustrates the prediction results of this method.

It can be seen from the results that 28.2 % of the high-risk road segments have been missed. This may be because the road design on these segments is favorable, or the road design defects are not the main cause of the crashes. As such, when we aim to identify roads with design flaws that are related to crash risks, its actual efficiency is significantly higher than the recall of the model in theory, as shown in Fig. 7.

Although the overall accuracy rate needs to be further improved, in terms of engineering applications, the "recall rate" is more effective than overall accuracy in the context of the lack of crash statistics on many roads. On these roads, it is extremely difficult to identify road design defects that may cause crashes. In addition, the identification of low-risk segments as high-risk may lead to a certain waste of resources, but it does not increase the risk of crashes. The two kinetic indicators in the

**Fig. 6.** Comparison of prediction performance on high-risk segments.

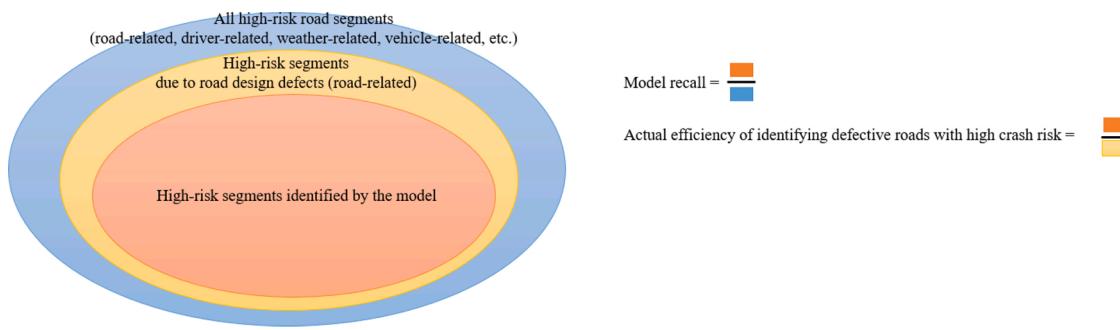


Fig. 7. Actual efficiency of identifying defective roads with high crash risk.

real driving conditions proposed in this paper have been revealed to be significantly related to crash risk. This also provides a reference for crash prevention.

It is worth mentioning that there are many low-risk segments identified as high-risk segments. One possible explanation is that there are factors that lead to abnormal kinetic response on these road segments, but do not increase the risk of crashes. This interesting finding merits further investigation, because the evolution law of abnormal kinetic indicators caused by these unrelated to crash may be different from the crash-related factors. If we can distinguish the abnormal kinetic response in these two situations, it is very likely to make the output of the model more in line with the real crash risk.

In summary, this section discussed the prediction method of crash risk. The NBR model and the RENBR model are jointly applied, which can combine the advantages of the two models to improve identification performance. We found that the accuracy rate for identifying segments with high-risk is at least 71.8 %. As such, for new-built roads or roads lacking crash statistics, relevant departments can deploy experimental methods developed by this study to analyze the characteristics of kinetic data and to make a preliminary judgment on the high-risk segments of crash.

6. Conclusion

On the one hand, this paper picked Wenli Freeway (K117 ~ K189) as the experimental segments, which was divided in units of 1 km. Afterwards, segments with horizontal curvature were selected. Then based on the crash data of equivalent, the value of crash risk on each segment with horizontal curvature was calculated.

On the other hand, a small car equipped with the automotive instrumentation was used to collect the vehicle longitudinal force, lateral force, vertical force, overturning torque, rolling moment, return moment and speed per 0.01 s on the experimental segment at a speed of 100 km / h. By filtering, analyzing and processing the vehicle kinetic data, the indicators for the risk analysis of the segments with horizontal curvature linearly expressed by multiple types of kinetic response were gained—the longitudinal stability indicator and the lateral stability indicator. Based on this, the NBR model and the RENBR model were constructed. Both models showed that the vehicle's lateral stability indicator and longitudinal stability indicator had remarkable statistically significant with crash risk.

Finally, the value of crash risk predicted based on the model was compared with the historical risk value. All road segments are classified into "the road segments with relatively high crash risk" and "the road segments with relatively low crash risk" according to historical crash risk values. Based on the performance advantage analysis of the NBR model and the RENBR model, the two models were jointly employed to predict the risk of crashes. This measure showed good performance (at least 71.8 %) for high-risk road segments, but it would be accompanied by more samples showing "false positives".

Hence, this study considered the impact of the six-point force of the

vehicle, and proposed a vehicle's lateral stability indicator and longitudinal stability indicator with multiple types of kinetic parameter effects that could be used for the high-risk identification. In addition, given the strong adaptability to high-risk road segments, vehicle longitudinal stability indicators and lateral stability indicators could also be considered as "crash surrogate events". Although it was difficult to identify akin events from in-vehicle data, the great role of in-vehicle data in road safety analysis had been demonstrated by this research. Using the kinetic parameters available from the experiment in real world may help in determining whether certain safety measures need to be undertaken without having to wait for sufficient crash data to be accumulated.

We proposed the use of vehicle kinetic response indicators for road safety analysis, but there is still a lot of work that has not been carried out. Although using only one driver in the experiment has certain advantages, it also limits the scope of this method. Generally, drivers could be divided into several clusters, including defensive, aggressive, dissociative, anxious, and patient (Freuli et al., 2020; Sagberg et al., 2015). Given different drivers will adopt various driving behaviors on dangerous road segments, which will lead to different fluctuations in vehicle kinetic indicators. In other words, due to the different types of drivers, there is unobservable heterogeneity in kinetic indicators when modeling. Therefore, after this study reveals the effectiveness of the indicators, it is necessary to further investigate the heterogeneous effects of different drivers, and try to build random parameter models or machine learning models to mine these unobservable laws.

When we can capture these heterogeneities, we can identify high-risk areas sensitive to different types of drivers. Note that the method developed in this study can only estimate the overall risk propensity, and different types of drivers may have different emphasis on these high-risk areas.

We hope to conduct more in-depth research on traffic safety from the perspective of microscopic kinetic parameters and provide greater assistance for the prevention and management of crashes.

CRediT authorship contribution statement

Changjian Zhang: Conceptualization, Methodology, Software, Formal analysis, Writing - original draft. **Jie He:** Conceptualization, Investigation, Writing - original draft, Supervision, Project administration. **Mark King:** Formal analysis, Writing - original draft. **Ziyang Liu:** Data curation, Validation. **Yikai Chen:** Software, Investigation. **Xintong Yan:** Methodology, Writing - review & editing. **Lu Xing:** Conceptualization, Supervision. **Hao Zhang:** Investigation, Project administration.

Declaration of Competing Interest

The authors report no declarations of interest.

Acknowledgments

This research was funded by National Natural Science Foundation of China (Grant No. 51778141, 52072069 and 71871078), Henan science and technology project (Grant No.182102310733), and Jiangsu Creative PHD student sponsored project (KYCX20_00138). Their assistance is gratefully acknowledged.

References

- Abdel-Aty, M.A., Radwan, A.E., 2000. Modeling traffic accident occurrence and involvement. *Accid. Anal. Prev.* [https://doi.org/10.1016/S0001-4575\(99\)00094-9](https://doi.org/10.1016/S0001-4575(99)00094-9).
- Antin, J., Lee, S., Hankey, J., Dingus, T., 2011. Design of the In-Vehicle Driving Behavior and Crash Risk Study: In Support of the SHRP 2 Naturalistic Driving Study. <https://doi.org/10.17226/14494>.
- Dadashova, B., Ramírez, B.A., McWilliams, J.M., Izquierdo, F.A., 2016. The identification of patterns of interurban road accident frequency and severity using road geometry and traffic indicators. *Transp. Res. Procedia* 14, 4122–4129. <https://doi.org/10.1016/j.trpro.2016.05.383>.
- Dahbir, B., Hassan, Y., 2019. Probabilistic, safety-explicit design of horizontal curves on two-lane rural highways based on reliability analysis of naturalistic driving data. *Accid. Anal. Prev.* 123, 200–210. <https://doi.org/10.1016/j.aap.2018.11.024>.
- Dingus, T.A., Klauer, S., Neale, V.L., Petersen, A., Lee, S.E., Sudweeks, J., Perez, M.A., Hankey, J., Ramsey, D., Gupta, S., Bucher, C., Doerzaph, Z.R., Jermeland, J., Knippling, R., 2006. The 100-car naturalistic driving study phase II – results of the 100-car Field experiment. United States. Department of Transportation. National Highway Traffic Safety Administration. <https://doi.org/10.1211/pj.2017.20203911>.
- Dixit, V., Harrison, G.W., Rutström, E.E., 2014. Estimating the subjective risks of driving simulator accidents. *Accid. Anal. Prev.* 62, 63–78. <https://doi.org/10.1016/j.aap.2013.08.023>.
- Ellison, A.B., Greaves, S.P., Bliemer, M.C.J., 2015. Driver behaviour profiles for road safety analysis. *Accid. Anal. Prev.* 76, 118–132. <https://doi.org/10.1016/j.aap.2015.01.009>.
- Freuli, F., De Cet, G., Gastaldi, M., Orsini, F., Tagliabue, M., Rossi, R., Vidotto, G., 2020. Cross-cultural perspective of driving style in young adults: psychometric evaluation through the analysis of the Multidimensional driving Style Inventory. *Transp. Res. Part F Traffic Psychol. Behav.* 73, 425–432. <https://doi.org/10.1016/j.trf.2020.07.010>.
- Geng, C., Peng, Y., 2018. Identification method of traffic accident black spots based on dynamic segmentation and DBSCAN algorithm. *J. Chang'an Univ. (Nat. Sci. Ed.)* 38 (5), 131–138.
- Gettman, D., Pu, L., Sayed, T., Shelby, S., 2008. Surrogate Safety Assessment Model and Validation: Final Report. In Publication No. FHWA-HRT-08-051 (Issue June).
- Geurts, K., Wets, G., 2003. Black spot analysis methods: literature review. *Onderzoekslijn Kennis Verkeersveiligheid* 1, 32.
- Geurts, K., Wets, G., Brijs, T., Vanhoof, K., 2004. Identification and ranking of black spots: sensitivity analysis. *Transp. Res. Rec.* 1897, 34–42. <https://doi.org/10.3141/1897-05>.
- Ghasemzadeh, A., Ahmed, M.M., 2019. Quantifying regional heterogeneity effect on drivers' speeding behavior using SHRP2 naturalistic driving data: a multilevel modeling approach. *Transp. Res. Part C Emerg. Technol.* 106, 29–40. <https://doi.org/10.1016/j.trc.2019.06.017>.
- Gitelman, V., Bekhor, S., Doveh, E., Pesahov, F., Carmel, R., Morik, S., 2018. Exploring relationships between driving events identified by in-vehicle data recorders, infrastructure characteristics and road crashes. *Transp. Res. Part C Emerg. Technol.* 91, 156–175. <https://doi.org/10.1016/j.trc.2018.04.003>.
- Guo, F., Klauer, S.G., Hankey, J.M., Dingus, T.A., 2010. Near crashes as crash surrogate for naturalistic driving studies. *Transp. Res. Rec.: J. Transp. Res. Board* 2147 (1), 66–74. <https://doi.org/10.3141/2147-09>.
- Hallmark, S., McGehee Dan Bauer, K.M., Hutton, J.M., Davis, G.A., Hourdos, J., Dozza, M., 2013. Initial Analyses From the SHRP 2 Naturalistic Driving Study: Addressing Driver Performance and Behavior in Traffic Safety(No. SHRP 2 Safety Project S08) (Issue 282). The National Academies Press. <https://doi.org/10.17226/22621>.
- Hou, Q., Tarko, A.P., Meng, X., 2018. Investigating factors of crash frequency with random effects and random parameters models: new insights from Chinese freeway study. *Accid. Anal. Prev.* 120, 1–12. <https://doi.org/10.1016/j.aap.2018.07.010>.
- Hydén, C., 1987. The development of a method for traffic safety evaluation: the Swedish traffic conflict technique. *Transp. Res. Rec.* 133–139. <https://doi.org/10.1002/2016GC006399>.
- Kluger, R., Smith, B.L., Park, H., Dailey, D.J., 2016. Identification of safety-critical events using kinematic vehicle data and the discrete fourier transform. *Accid. Anal. Prev.* 96, 162–168. <https://doi.org/10.1016/j.aap.2016.08.006>.
- Li, P., He, J., 2016. Geometric design safety estimation based on tire-road side friction. *Transp. Res. Part C Emerg. Technol.* 63, 114–125. <https://doi.org/10.1016/j.trc.2015.12.009>.
- Li, L., Shrestha, S., Hu, G., 2017. Analysis of road traffic fatal accidents using data mining techniques. In: Proceedings - 2017 15th IEEE/ACIS International Conference on Software Engineering Research, Management and Applications, SERA, 2017, pp. 363–370. <https://doi.org/10.1109/SERA.2017.7965753>.
- Liu, Z., He, J., Zhang, C., Xing, L., Zhou, B., 2020. The impact of road alignment characteristics on different types of traffic accidents. *J. Transp. Saf. Secur.* 12 (5), 697–726. <https://doi.org/10.1080/19439962.2018.1538173>.
- Lv, C., Shi, D., Gao, S., Liu, S., 2018. The safety research of road alignment based on virtual simulation technology. *Int. J. Eng. Syst. Model. Simul.* 10 (4), 197–206. <https://doi.org/10.1504/IJESMS.2018.095966>.
- Ma, X., Chen, S., Chen, F., 2017. Multivariate space-time modeling of crash frequencies by injury severity levels. *Anal. Methods Accid. Res.* 15, 29–40. <https://doi.org/10.1016/j.amar.2017.06.001>.
- Milton, J., Manning, F., 1998. The relationship among highway geometrics, traffic-related elements and motor-vehicle accident frequencies. *Transportation* 25 (4), 395–413. <https://doi.org/10.1023/A:1005095725001>.
- Naznin, F., Currie, G., Logan, D., Sarvi, M., 2016. Application of a random effects negative binomial model to examine tram-involved crash frequency on route sections in Melbourne, Australia. *Accid. Anal. Prev.* 92, 15–21. <https://doi.org/10.1016/j.aap.2016.03.012>.
- Pande, A., Chand, S., Saxena, N., Dixit, V., Loy, J., Wolshon, B., Kent, J.D., 2017. A preliminary investigation of the relationships between historical crash and naturalistic driving. *Accid. Anal. Prev.* 101, 107–116. <https://doi.org/10.1016/j.aap.2017.01.023>.
- Sagberg, F., Selpi Bianchi Piccinini, G.F., Engström, J., 2015. A review of research on driving styles and road safety. *Hum. Factors* 57 (7), 1248–1275. <https://doi.org/10.1177/0018720815591313>.
- Sarhan, M., Hassan, Y., 2008. Safety explicit design of freeway speed change lanes safety performance of freeway sections and relation to length of speed change lanes. *Can. J. Civ. Eng.* 5 (35), 531–541. <https://doi.org/10.1139/L07-135>.
- Urie, Y., Velaga, N.R., Maji, A., 2016. Cross-sectional study of road accidents and related law enforcement efficiency for 10 countries: a gap coherence analysis. *Traffic Inj. Prev.* 17 (7), 686–691. <https://doi.org/10.1080/15389588.2016.1146823>.
- Wang, Y., Wang, L., 2019. An identification method of traffic accident black point based on street-network spatial-temporal kernel density estimation. *Sci. Geogr. Sin.* 39 (8), 1238–1245.
- Wang, H.W., Wen, H.Y., Niu, L.Z., Guo, X.H., Zhang, R.H., Pei, Y.L., 2015. A new nonlinear car-following model for mountain highways based on the traffic safety analysis. *Proceedings of the 2015 International Conference on Electronics, Electrical Engineering and Information Science (EEEIS2015)* 823–832. https://doi.org/10.1142/9789814740135_0083.
- Washington, S.P., Karlaftis, M.G., Manning, F.L., 2010. *Statistical and Econometric Methods for Transportation Data Analysis*, second edi. Chapman and Hall/CRC.
- Wu, K.F., Jovanis, P.P., 2012. Crashes and crash-surrogate events: exploratory modeling with naturalistic driving data. *Accid. Anal. Prev.* 45, 507–516. <https://doi.org/10.1016/j.aap.2011.09.002>.
- Wu, K.F., Jovanis, P.P., 2013. Defining and screening crash surrogate events using naturalistic driving data. *Accid. Anal. Prev.* 61, 10–22. <https://doi.org/10.1016/j.aap.2012.10.004>.
- Xing, L., He, J., Abdel-Aty, M., Wu, Y., Yuan, J., 2020. Time-varying analysis of traffic conflicts at the upstream approach of toll plaza. *Accid. Anal. Prev.* 141, 1–14. <https://doi.org/10.1016/j.aap.2020.105539>.
- Yang, H., Ozbay, K., Bartin, B., 2019. Application of simulation-based traffic conflict analysis for highway safety evaluation. Selected Proceedings of the 12th World Conference on Transport Research 11–15. <http://intranet.imet.gr/Portals/0/Useful/Documents/documents/02303.pdf>.
- Yuan, Q., Li, Y., Lu, G., 2010. Identification analysis model of traffic accident-prone locations based on geographical view angle. *J. Traffic Transp. Eng.* 10 (1), 101–105.
- Zhang, C., Ma, R., 2010. Identifying method of traffic accident black spots on mountain freeway. *J. Chang'an Univ. (Nat. Sci. Ed.)* 30 (6), 76–80.
- Zhou, Yaoqun, Zhang, Weigong, Liu, Guangfu, Li, Guozhong, 2007. Research and development on the vehicle roadway test system based on a new six-component wheel force transducer. *China Mech. Eng.* 18 (20), 2510–2514. U467.11;TP212.9.