

See discussions, stats, and author profiles for this publication at: <https://www.researchgate.net/publication/343110045>

Driving risk assessment based on naturalistic driving study and driver attitude questionnaire analysis

Article in *Accident Analysis & Prevention* · September 2020

DOI: 10.1016/j.aap.2020.105680

CITATIONS

0

READS

119

6 authors, including:



Jianqiang Wang

Tsinghua University

170 PUBLICATIONS 2,601 CITATIONS

[SEE PROFILE](#)



Huang Heye

Tsinghua University

7 PUBLICATIONS 9 CITATIONS

[SEE PROFILE](#)



Jinxin Liu

Tsinghua University

8 PUBLICATIONS 1 CITATION

[SEE PROFILE](#)

Some of the authors of this publication are also working on these related projects:



Driving safety of ICVs [View project](#)



Test and evaluation of ICVs [View project](#)



Driving risk assessment based on naturalistic driving study and driver attitude questionnaire analysis

Jianqiang Wang^{a,*}, Heye Huang^a, Yang Li^a, Hanchu Zhou^b, Jinxin Liu^a, Qing Xu^a

^a State Key Laboratory of Automotive Safety and Energy, Tsinghua University, Beijing, China

^b School of Traffic and Transportation Engineering, Central South University, Changsha, Hunan, China



ARTICLE INFO

Keywords:

Intelligent vehicles
Driving risk coupling model
Internal and external field
Naturalistic driving study
Driver attitude questionnaire

ABSTRACT

Traffic accident statistics have shown the necessity of risk assessment when driving in the dynamic traffic environment. If the risk associated with different traffic elements (*i.e.*, road, environment and vehicles) could be evaluated accurately, potential accidents could be significantly avoided or mitigated. This paper proposes a driving risk assessment model that can quantitatively evaluate the driving risk associated with intelligent vehicles via the coupled analysis of different traffic elements. First, we present a concept of the internal field and external field for establishing the driving risk coupling model, through employing the internal field to define the risk range of driver's perspective and the external field to calculate the risk coefficients of those traffic elements. Then, the relative risk coefficients are computed by incorporating both naturalistic driving study (NDS) and driver attitude questionnaire (DAQ) using a multinomial logit model. Specifically, we perform a large-scale naturalistic driving study to investigate the objective driving risks. Typical driver behavior parameters, such as velocity, time headway, and acceleration, are analyzed. Besides, a self-reported survey of 364 drivers is conducted to subjectively evaluate the potential risks that drivers may face in various situations. Finally, validation of the model is conducted by comparing the accuracy with the typical risk assessment index, *i.e.*, TTC and THW. Results demonstrate that the proposed approach is effective in evaluating the comprehensive driving risks by quantifying the influence factors of driving risks in dynamic environments.

1. Introduction

Transportation systems comprise various elements including drivers, vehicles, and roads. However, a generalized instability of the driver-vehicle-road closed-loop system can lead to a driving risk owing to these potential dynamic factors, such as driver error, vehicle failure, road condition, and environment state (Laugier et al., 2011; Rolison et al., 2018; Wang et al., 2016). Previous studies have suggested that drivers' perceptions and reactions can change in different conditions (Martensen and Dupont, 2013; Morgan and Mannering, 2011; Saifuzzaman and Zheng, 2014). A previous study (Brackstone et al., 2009) has found that the involvement of multiple trucks, buses, cars, or motorcycles can cause potential traffic accidents. Environmental elements (Ahmed et al., 2014) have also been proven to be the main factors affecting driving safety. Several studies (Hosseinpour et al., 2014; Jafari Anarkooli and Hadji Hosseiniou, 2016; Morgan and Mannering, 2011) have reported that the possible effect of road features on risk levels, providing the importance of studying the influence of

traffic factors on driving risk to improve driving safety. In the past decades, several studies (Goerlandt and Reniers, 2016; Ni et al., 2010) have assessed various influencing factors that lead to driving risks, which can be divided into objective risk analysis, subjective risk analysis, and combined ones.

One line of research has focused on objective driving risk analysis, which can also be divided into two aspects, namely macroscopic traffic data and microscopic operating parameters of the driving process to analyze risk factors. Specifically, macro accident data analysis includes the absolute number method, the accident rate method, and the accident severity analysis method. The analysis addresses attributes that are related to accidents. To study the relation between accident probability and contributing factors, macro accident data analysis involves extracting and analyzing traffic accident data (Çelik and Oktay, 2014; Elvik, 2009) to study the relation between accident probability and influence factors. A previous study (Al-Ghamdi, 2002) has developed the logistic regression model to effectively combine the accident factors with the possible influence characteristic variables, designing the

* Corresponding author.

E-mail addresses: wjqlws@tsinghua.edu.cn (J. Wang), hhy18@mails.tsinghua.edu.cn (H. Huang), lyxc56@gmail.com (Y. Li), hanchzhou@csu.edu.cn (H. Zhou), liujx17@mails.tsinghua.edu.cn (J. Liu), qingxu@tsinghua.edu.cn (Q. Xu).

characterization scale to analyze the factors influencing the accidents. Although these macro data can reflect the overall traffic system risks, while the evaluation results cannot reflect the dynamic driving risk of each driver. Moreover, the detailed driving data are limited, which include vehicle kinematics parameters and road user moving characteristics, reflecting the objective driving process and risk generation mechanism associated with various factors. From the viewpoint of microscopic operating parameters, the existing driving risk evaluation parameters infer the longitudinal driving risk, the collision time (TTC) (Archibald et al., 2008), the headway distance (THW) (Allen et al., 2013), and the horizontal driving risk, such as the current location of the vehicle (CCP) (Khatib, 1990), and time of lane change (TLC). These evaluation indicators can effectively estimate the risk in a certain driving process. Vehicle kinematics and kinetics are considered when evaluating the driving risks from the perspective of time or distance. However, the potential risk that underlies the driver's characteristics is ignored.

Evaluating the subjective risk of drivers is also necessary to help analyze safety levels in the driving process. Statistics show that drivers are the direct object of risk perception in the driving process and their impairments or errors may provoke 94 % of motor accidents (Kinnear et al., 2015; Wang et al., 2014). Many studies have been carried out to study the driver's risk perception and driving behavior. Worrall et al. (2010) evaluated risk using the safety operations rule defined by safety management. Musicant et al. (2014) proposed a driving behavior statistical model to represent the potential risk caused by driver manipulation, the model was established by collecting and screening the data of drivers' actual manipulation. Besides, drivers' self-report, including interviews and questionnaires, can be regarded as one of the main subjective analysis methods (Martinussen et al., 2014; Reimer et al., 2005). For example, we can use driving behavior questionnaire (DBQ) and driving style questionnaire (DSQ) to explore the relationship between drivers' driving behaviors and driving risk. These above-mentioned studies have extracted the characteristics of drivers' psychophysiological state, cognition, driving skills, and traffic rule consciousness, among others. These characteristics have helped in developing a new definition of risk from the driver's perspective; however, most subjective risk analyses are with qualitative description, making it challenging to accurately quantify the real risk.

The accurate identification of driving risk needs not only objective risk identification but also combining subjective risk. While exploring the combination of subjective and objective risks, some researches introduced artificial potential field theory to describe the relation between vehicles and driving environment, including pedestrians and cyclists (Gerdes and Rossetter, 2001; Moreau et al., 2017; Tu et al., 2016). Gravitational field concept has been adopted to establish the gravity model based on the attraction generated by the vehicle safety clearance and repulsive force generated by other surrounding vehicles (Hsu et al., 2012). However, properly incorporating the driver's

characteristics and traffic environment into a unified model is still challenging. Another study (Wang et al., 2016) also applied the thought of field theory, incorporated the driver behavior factor, and proposed a unified safety evaluation method based on the coupling relation of the "diver-vehicle-road," which was the safety field theory. This method attributes a full consideration to the driver, vehicles, and road for good expansion, and combines subjective and objective risks. However, reducing the dependence on the parameters calibration model applications remains challenging.

Our study aims to better assess the driving risk in a dynamic traffic environment. To fill in these research gaps, we develop and examine a comprehensive analysis of driving risk influencing factors combining naturalistic driving study and subjective questionnaire analysis. The contributions can be concluded as follows: 1) we combined objective and subjective data quantitatively to reveal the potential risk associated with different traffic element; 2) we employed the multinomial logit model to obtain the relative risk coefficient to support the driving risk assessment, and 3) we proposed a concept of the internal field and external field to establish the driving risk coupling model and output dynamic risk assessment of the traffic environment. This paper is organized as follows. Section 2 describes the concept of the internal and external field and establishes the coupled driving risk assessment model. Section 3 gives detailed information about the analysis of the NDS and DAQ. Section 4 presents the analysis results with the corresponding discussion.

2. Methodology

To explore the potential risk associated with different traffic elements in the dynamic environment, a coupled driving risk assessment model based on the concept of the internal and external field was employed. By combining the internal field centered by the driver with the external field formed by the traffic environment, the influence range and degree of different elements can be calculated in real-time.

2.1. A framework of driving risk assessment

Traffic system is a complex dynamic system composing dynamic road users (including drivers, vehicles, pedestrians, cyclists and etc.) and static traffic elements (including roads, traffic lights, static obstacles, etc.). Risk can be described as the combination of the probabilities and the severity of events (Aven, 2011; Goerlandt and Kujala, 2014). Accidents can take place when a problem occurs during the coordination of the traffic system composing of drivers, vehicles.

Fig. 1 illustrates the entire framework of driving risk assessment. Herein, we classify driving risks into subjective risks and objective risks. To analyze the subjective risks, different drivers display different perceptions of current driving risks when facing the same traffic incidents. Drivers' physiological and psychological factors, such as their driving

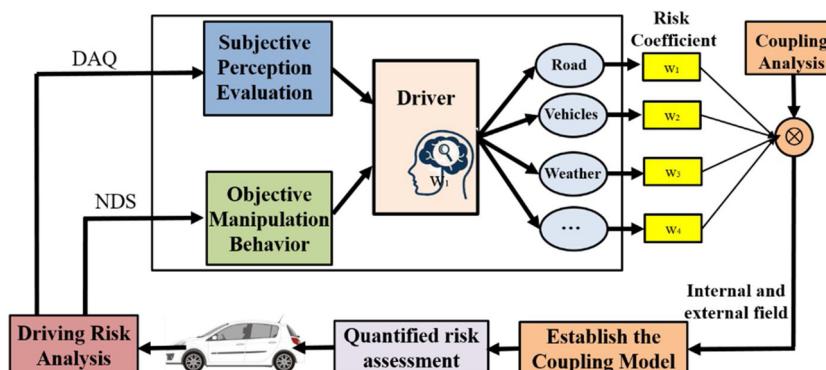


Fig. 1. A framework of driving risk assessment based on objective and subjective data.

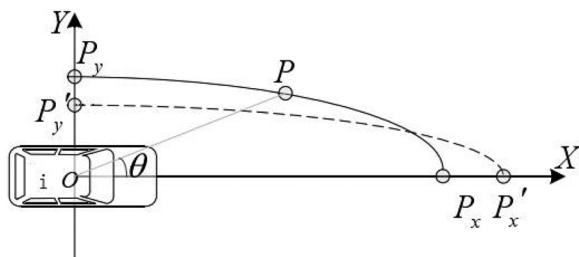


Fig. 2. Schematic diagram of internal field potential change.

skills, personalities, attitudes, emotions, and states, may affect their information reception and processing processes. They may lead to the difference in the subjective measurement of the degree of risk. When identifying the objective risks, the vehicle-road system characteristics are extracted by taking each road user in the system as the source of the objective risks, which include the vehicle performance, mixed traffic characteristics, pavement characteristics, and traffic signs. Considering the mechanism influencing driving risk, the physical characteristics between the vehicle-road system and driving risk are excavated. Then, the risk measurement values of objective risk attributes are established. The driver's behaviors in the driving process are disturbed by several external factors, such as the road environment and dynamic characteristics of the vehicle. Finally, a driving risk coupling model is established to evaluate and quantify the driving risk. Therefore, the combination of subjective and objective risks can more accurately realize the comprehensive situation assessment under complex scenarios. In conclusion, the subjective risk is defined as the risk brought by the individual attributes of the driver; objective risk is defined as the objective potential impact brought by dynamic changes of surrounding traffic elements. The comprehensive risk perception is to describe the psychological projection of drivers' risk perception from the perspective of drivers by various factors in the dynamic traffic environment.

2.2. Internal field model

2.2.1. Concept of an internal model

According to Lewin's psychological principle (Householder, 1939; Lewin, 1936), the generation of driver constraints and the execution of actions are the common results of driver attributes and environment. Drivers, vehicles, and roads are correlated in the real traffic environment. Therefore, it is concluded that the driver's manipulation behavior is constrained and influenced by the external environment. Therefore, the entire system needs to be modeled. Herein, the influence of the traffic environment on the drivers is regarded as the mechanism of the driver's internal effect. The extended time-distance measures the distribution of drivers' internal presence intensity, and the distribution is mathematically described.

The physical field form, $E = kQ/r^2$, and the change of the space between a certain point and the target vehicle in the driver's in-field space has a certain functional relation with the intensity of the field. We introduce the polynomial of the velocity's power function term to express the influence of the target vehicle velocity on the field intensity based on other studies (Wang et al., 2016). Since the internal field intensity is also affected by the characteristics of driver, road environment, and vehicle, the expression of field intensity of the internal field is defined as follows:

$$E_D = f(D, R, V) \frac{v_i^\delta}{r^k} \vec{n} \quad (1)$$

where E_D is the internal field intensity, v_i is the velocity of the target vehicle, and r is the distance from the origin of the target vehicle along the direction of the target vehicle. For a field vector, a unit vector \vec{n} is introduced to describe its direction. $f(D, R, V)$ is a comprehensive risk function, which can be expressed as a functional relation of driver risk

coefficient D_a , road environmental risk coefficient R_a , and vehicle physical risk coefficient V_a .

2.2.2. Definition of driver visual characteristics

The driver plays the role of the receiver and output decision-maker during the entire driving process. Therefore, it is necessary to define the effective scope of the risk coupling model. According to the main channels of driver information acquisition in the traffic system, the visual characteristics of drivers are closely related to their safe driving process (Zhang et al., 2016). In the driving process, visual discrimination ability is greatly affected, and the visual field angle is narrowed with an increase in speed. Therefore, the scope of the internal presence is defined by the driver's visual characteristics as shown in Fig. 2, and the object field intensity is determined by calculating the curvilinear integral in the field of the object projected line segment within this range. The driving direction of the target vehicle i is defined as the X -axis. Within the driving range centered on the target vehicle i , a range of areas with the equal potential field are constructed according to the principle of visual characteristics. Specifically, the radius of the equal potential field can be defined as:

$$l = r - (1-\delta)r |\sin\theta| \quad (2)$$

where l is the distance from the point of the equipotential line to the origin O of the vehicle i , and θ is the angle of the connection between OP at a point P on the equipotential line and the direction of the vehicle, i.e., the X -axis in the figure. $\delta = OP_x/OP_y$ is related to the velocity of i . The faster the speed, the narrower will be the line of sight, which is expressed as follows:

$$r = \frac{l}{1 - (1-\delta)|\sin\theta|} \quad (3)$$

According to the driver's visual characteristics, the change of the vehicle's relative speed will affect the perspective of the front vehicle. The change rate of the perspective is related to the perception threshold k as follows based on the research of (van Winsum, 1999):

$$k = \frac{d\theta}{dt} \approx \frac{\Delta v}{D_p^2} \quad (5)$$

where θ is the angle; Δv is the relative speed of the vehicle; D_p is the distance between vehicles. Drivers always keep a certain time headway as a safe margin for the following. Based on Winsum, the space between vehicles can be expressed as:

$$D_p = t_p v_i \quad (6)$$

where t_p is the time interval between the front and rear vehicles (s); v_i represents vehicle speed under the stable following state.

Therefore, the faster the speed and the smaller the angle θ , the more the driver's field of vision focuses on the longitudinal direction and ignores the lateral effect. Then the driver's basic field intensity at any point in the space can be transferred as follows:

$$E_D = f(D, R, V) \frac{v_i^\delta}{l_{(x,y)}^k} [1 - (1 - \delta(v_i))|\sin\theta_{(x,y)}|]^k \vec{n} \quad (4)$$

2.2.3. The comprehensive function of driving risk

Driver characteristics comprise personality characteristics, such as gender differences, age difference, temperamental pilot differences, and driver's personality differences; visual cognition characteristics, such as vision, visual adaptation, and blinding; and behavior characteristics, such as information processing, reaction characteristics, and psychological characteristics. Therefore, we select its main characteristics to construct the driver risk coefficient D_a to describe and analyze the influence of driver characteristics on the coupling model. After referring to (Knoefel et al., 2018; Zhang et al., 2009), the expression of driver risk coefficient D_a is obtained as follows:

$$D_a = \gamma e^a \quad (7)$$

where γ and a are mainly related to the driver's personality characteristics.

The determination of road environmental risk factor, R_a , mainly refers to the definition of road environmental risk factor (Wang et al., 2016), which can be expressed as follows:

$$R_a = \Psi_\mu(\mu_i) \cdot \Psi_\rho(\rho_i) \cdot \Psi_\tau(\tau_i) \cdot \Psi_\delta(\delta_i) \quad (8)$$

where Ψ_μ , Ψ_ρ , Ψ_τ , and Ψ_δ are the tire-road friction, the road curvature, and slope and visibility of risk evaluation function, respectively.

Finally, we define the vehicle physical risk coefficient V_a to describe the potential influence on the performance of the vehicle. M represents the vehicle mass, L represents the vehicle volume parameter, and X represents the vehicle design performance index; the calibration process of parameters can be referred in (Li et al., 2017; Wang et al., 2016). Therefore, the formula is described as follows:

$$V_a = V_a(M, L, X) = M \cdot L \cdot X \quad (9)$$

Therefore, the comprehensive risk function $f(D, R, V)$ can be expressed as follows:

$$f(D, R, V) = D_a \cdot R_a \cdot V_a \quad (10)$$

This formula means that the comprehensive risk function $f(D, R, V)$ can combine the potential influence brought by the driver, vehicle and road. When the state of any loop in the closed-loop system changes, the whole risk function will change and the risk value will change dynamically in traffic scene.

2.3. External field model

Herein, all objects, in motion or stationary, in the surrounding environment of driving target vehicles are defined as obstacles, including vehicles, lane markings, and pedestrians. Since all the obstacles may pose a threat to the driver's safety (Yoshitake and Shino, 2018), the threat generated by the obstacle is abstracted as a repulsive force field generated around it, namely, the external field. The external field produced by different objects is determined by its properties. Besides, based on the visual characteristics of the driver, the objects felt by the driver within the range of vision are divided into five categories.

The first type is static obstacles. Risk factors generated by such static obstacles are mainly determined by their own volume, which is reflected by the volume size projected in the driving safety field. Therefore, we define the risk coefficient as S_o . Constrain obstacles are considered as the second type. Risk factors generated by constrained obstacles, such as lane lines, guardrails, road edges, and warning signs for road construction, are determined by constraint types and width, among other factors. The risk coefficient is defined as S_c , reflected in terms of the degree of psychological pressure on drivers for driving safety. Further, instead of these static obstacles or constraints, the dynamic characteristic of traffic elements can bring the uncertainty of surrounding environment, which will influence the external field. Therefore, we first consider motor vehicles, which are usually in motion and have a sense of autonomous driving. The risk factors generated are mainly determined by the driving speed, volume size, location, and other factors of the motor vehicle, mainly manifested in the influence of these attributes on the decision-making process of the driving target vehicle (Malta et al., 2009). The risk coefficient can be defined as S_v . In addition, pedestrians are easily driven by their own conditions, affecting driving behavior (Yoshitake and Shino, 2018). The risk factors generated are determined by the free walking range, type, and code of conduct, among other factors, manifested by the influence of the driving target vehicles in the coupling model on the decision-making and judgment. And the risk coefficient is defined as S_p . Finally, non-motor vehicle's behavior is influenced by people's subjective consciousness (Greene et al., 2011), smaller in size and more convenient,

which is irregular to walk on the road. The risk factors generated are mainly determined by the free walking range and flexibility, among other factors, which are manifested as the influence of the driving safety field on the decision-making and judgment of the driving target vehicle. We define its risk coefficient as S_n . To conclude, the attributes of different objects within the visual range of drivers are distinguished and the risk coefficient S has been obtained from the comprehensive risk analysis.

2.4. Driving risk coupling model

The correlation between the internal and external physical fields is noted. In the inner presence, people are similar in the field source, and the occurrence source in the physical field has an effect on the objects in the field. Moreover, people in the inner presence will react to the environment under environmental constraints, affecting the objects in the surrounding field owing to their own behavior. For a complete traffic system, the coupling model of total driving risk influencing factors is constructed as follows:

$$E_s = f(D, R, V) S_i \frac{(v_i - v_j)^\delta}{l_{(x,y)}^k} [1 - (1 - \delta(v_i)) |sin\theta_{(x,y)}|]^k \vec{n} \quad (11)$$

Meanwhile, it should be noted that when there are more than one external driving risk influencing factors, that is, multiple influencing factors interact with each other, the external risk field S_i is superimposed in the form of multiplication, and finally acts on the driver's internal field model to jointly output the driving coupling risk.

3. Data processing

3.1. Objective manipulation data analysis

To lay a solid foundation for assessing driving risk and improving driving safety, two aspects are necessary, namely reasonable experimental design and real and detailed driving data. Therefore, we selected a large number of naturalistic driving data of excellent drivers, extracted typical car-following scenarios, and analyzed driver characteristic parameters, including vehicle speed, distance, acceleration and deceleration. Through objective manipulation analysis, we can obtain the common risk cognition rules in real traffic environment.

3.1.1. Naturalistic driving test platform

We establish a driving-data-acquisition platform to collect driving data involving the vehicle, driver, and surrounding environment. As shown in Fig. 3, the experimental vehicle is equipped with the onboard data recorders, cameras, CalmCar vision system, and industrial PC (IPC), among other equipment.

The onboard vehicle data recorder mainly collects vehicle state parameters, including the speed of ego vehicle, the relative speed between ego vehicle and other vehicles, and the relative distance among others. The state information between vehicles is derived from the front-mounted radar onboard the vehicle. Besides, the vehicle is equipped with five cameras, including one forward telephoto camera, two forward blind spot detection cameras, and two side cameras. These cameras can achieve complete coverage of 360° to collect the surroundings' and drivers' operational information. Meanwhile, CalmCar vision system is a driving analysis tool that can fully cover the cognition of complex road conditions. This tool performs data recording and conservation functions. The IPC is equipped with the image acquisition program to save the image information collected by the camera in real-time. The NDS data collected during the driving process are mainly Controller Area Network data (10 Hz), GPS (20 Hz), and LiDAR signal (10 Hz).

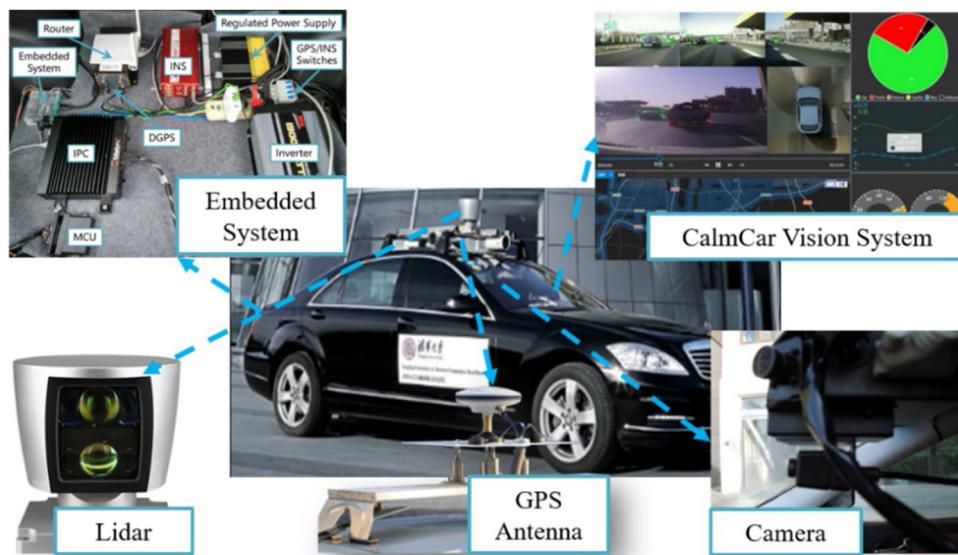


Fig. 3. The architecture of the test platform.

3.1.2. Experiment design

In the process of collecting manipulation behavior data to reveal the universal objective risk rule, it is necessary to reduce the impact of excessive subjective individual differences in risk assessment. Therefore, we selected several experienced and excellent drivers for naturalistic driving data collection. They all have more than 10 years of driving experience, and accumulated mileage exceeds 100,000 km. Further, they have a stable driving style (neither too aggressive nor too conservative), and a low traffic accident rate. All the participants recruited hold valid driving licenses, have normal vision or are corrected to normal, and meet the standard of excellent drivers. Through the analysis of a large number of naturalistic driving data collected, we can more accurately reveal the universal objective risk rule.

In addition, the experiment route is the Beijing highway, including four types of roads, namely urban highway (usually with low traffic volume, over 90 km/h), city ring road (a little congestion, approximately 60–80 km/h), national road (mixed traffic environment, 60–80 km/h, mixed traffic conditions), and inner-city road (mixed traffic environment, within 60 km/h). The naturalistic driving data are collected by 16 drivers with eight data acquisition vehicles; the amount is approximately 100,000 km in total, as shown in Fig. 4. The statistical characteristics of the sampled data are consistent with the traffic rules to reflect the real traffic laws in the city. Drivers are allowed to drive the vehicle based on their habits.

3.1.3. Extraction of scenarios and traffic elements

3.1.3.1. Extraction of car-following data. In this paper, the most dangerous working conditions that provoke most accidents in China are first counted to screen possible driving conditions for risk generation. Hazardous conditions are screened out from the naturalistic driving data, wherein the vehicle rear-end condition accounts for the highest proportion of 43.83 % (X. Wang et al., 2016). Therefore, the car-following scenario is selected. Continuous and effective car-following time series are extracted from the original data series. In extracting the car-following time series, we adopt the extracting rule as described in (Wang et al., 2013), as follows:

$$\begin{cases} 5m < \Delta d(k) \leq 2.5V_h(k) + 10m \\ |d_y(k) - d_y(k-1)| < 5m \\ V_h(k) > \frac{30}{3.6} m/s \\ |\Delta V_j(k)| < 5m/s \end{cases} \quad (12)$$

where $\Delta d(k)$ is the relative distance; $V_h(k)$ describes the velocity of ego vehicle; $\Delta V_j(k)$ is relative velocity; $d_y(k)$ is the longitudinal position. When the relative distance between the two vehicles is greater than 5m and less than $2.5V_h(k) + 10m$, it conforms to the defined car-following range, which can be customized according to the traffic flow density of actual natural driving data (Treiber and Kesting (2013)). When the



Fig. 4. The selected driving sections and sensor acquisition ranges.

Table 1
Traffic elements extraction.

Road users	Car	Bus	Truck	Pedestrian	Two-wheels
Road					
	Flat	Rough	Free	Congestion	Complex
Environment					
	Sunny	Rain	Foggy	Dusk	Dark

Note: congestion represents a traffic flow with a speed less than 20km/h and lasting more than 5 min. Free flow represents that the average speed of traffic flow is higher than 65km/h.

velocity of ego vehicle is larger than $\frac{30}{3.6}$ m/s, it will not be too slow to accurately obtain the dynamic changes of the driver's control of ego vehicle.

3.1.3.2. Extraction of different influence parameters. During the test process, we classify and extract the independent variables that may affect the driving risk under the following conditions. The elements marked by the data collection process in the driving scene fall into the following categories listed in Table 1. We consider the main factors that influence driver manipulation level, and design variables mainly from two aspects: objective stability of driver manipulation and psychological comfort of driver manipulation. Therefore, in the driver-vehicle-road-environment system, the influence of potential variables such as vehicle, road and environment on the driver's manipulation level is mainly considered. Therefore, the variables in the designed scenarios include driving time (day/night), road geometry (flat/rough), road congestion state (congestion/free), light condition (luminosity/dusk/darkness), weather (rain/foggy/ sunny), and object (cars/buses/trucks/two-wheels/pedestrians).

3.1.4. Data preprocessing and result analysis

Based on the speed signal obtained in the process of the natural driving test, we utilize Kalman filter (Faragher, 2012) to further estimate and smoothen the vehicle velocity, widely applied in signal processing. Therefore, we employ Kalman filter to eliminate noise from the radar data of the longitudinal velocity in the ego vehicle. The basic assumption in the car-following condition is that each driver tries to maintain the desired position in the back lane behind the leading vehicle by adjusting the throttle and brake. Herein, five variables are mainly used for considering different object working conditions, namely the speed of ego vehicle, relative distance, time headway, lateral offset, and transverse overlap rate. After the statistical analysis of the parameters, statistical results present the mean value, standard deviation, maximum and minimum values of sample data, 5% cumulative frequency value (5% C.F.), 50 % cumulative frequency value, and 95 % cumulative frequency. Fig. 5 shows the statistical analysis of indexes, i.e., velocity, relative distance, and time headway, when tracing different objects, i.e., car, bus, truck, and pedestrian. Table 2 lists the statistical results of various statistical tests.

When analyzing the velocity of ego vehicle and tracking objects of a vehicle, preliminary analysis indicates that the safety degree is higher than that of the tracking pedestrian. The velocity of tracking pedestrians is relatively scattered, indicating that the complexity is higher. This condition has a higher likelihood to contribute to collisions (Fig. 5; Table 2). Owing to the heavy psychological pressure on the driver for different types of vehicles, a low speed is remarked with the large truck, whereas the differences between tracking bus and car are insignificant.

The relative distance is the distance between the front vehicle and

the ego vehicle. Through comparison, the results indicate that when tracking different types of vehicles, the following relation exists with the car-following distance: truck > bus > car. When encountering pedestrians, the vehicle velocity turns out to be relatively slow, and the relative distance is concentrated, indicating that the vehicle is changing its state steadily in the following movement.

The time headway (THW) is also selected as the main indicator of driver behavior. To a certain extent, the smaller its value, the more dangerous it is. As shown in Fig. 5 and Table 2, the THW trend is consistent with that of relative distance. During the car-following process, drivers are more likely to believe that tracking the truck will bring more psychological pressure, and the risk cognition value tends to be truck > bus > car.

In addition to the above three direct parameters, we also extracted other potential parameters such as transverse, overlap rate, and lateral offsets, to reflect the driving behavior of a driver facing different road users. The transverse overlap rate is obtained by dividing the transverse overlap between the ego vehicle and the front vehicle based on ego vehicle's width. The lateral offset refers to the distance deviation between the center point of the ego vehicle and the center point of the target vehicle (front vehicle). When the centerline of the car in front is flush with the centerline of the ego car, the lateral deviation is zero. As shown in Fig. 6, when the current vehicle is a car, the skewness is the minimum, the lateral offset is concentrated, and the overlap rate is large, reflecting that the driver often adopts the stability following velocity when tracking the front vehicle. When the front vehicle is a bus, the overlap rate is evenly distributed, indicating that the lateral direction of the ego vehicle is adjusted at every moment. In the case of tracking a truck, the overlap rate is discrete and the proportion stability of the vehicle is small, reflecting an unstable car-following state.

The findings indicate that the potential risk of buses and trucks are associated with their scales and driving characteristics in comparison with cars. In the following scenario, drivers are more likely to change the unsafety state, such as changing lanes or keeping a longer distance, to avoid the potential risks. Besides, the uncertainty of pedestrians is more likely to change the state of vehicles, which may be associated with higher risks.

3.2. Subjective questionnaire data analysis

For the evaluation of subjective risk, it is necessary to cover enough driver samples of differentiated characteristics. The differentiation of drivers' driving style, driving skills and experience is more conducive to obtaining more reasonable subjective risks caused by individual characteristics. Therefore, in our experiment, 461 participants aged between 18 and 55 were recruited through the existing contact group and network (WeChat, weibo, QQ and other social media), and we employ SPSS version 24 to analyze the reliability and validity of the designed questions to screen out 364 valid questionnaires from 461 questionnaires collected for analysis and to quantitatively output the subjective risk identification of different drivers.

3.2.1. Design of DAQ

DAQ is designed to reflect the driver's subjective risk identification level. The main design questions are in the form of text, dialog, picture, and video. The entire questionnaire takes approximately 20 min to complete; the entire process is shown in Fig. 7. In the questionnaire survey, participants are asked to fill in their basic information. They subsequently complete the scoring of their subjective risk factors based on the real driving scene pictures and video information. Finally, they describe their feelings in a question-and-answer style. Particularly, the survey questions are in four parts:

- Individual information (9 questions), including gender, age, occupation and driving experience.
- Driving information (18 questions), including the types of risks

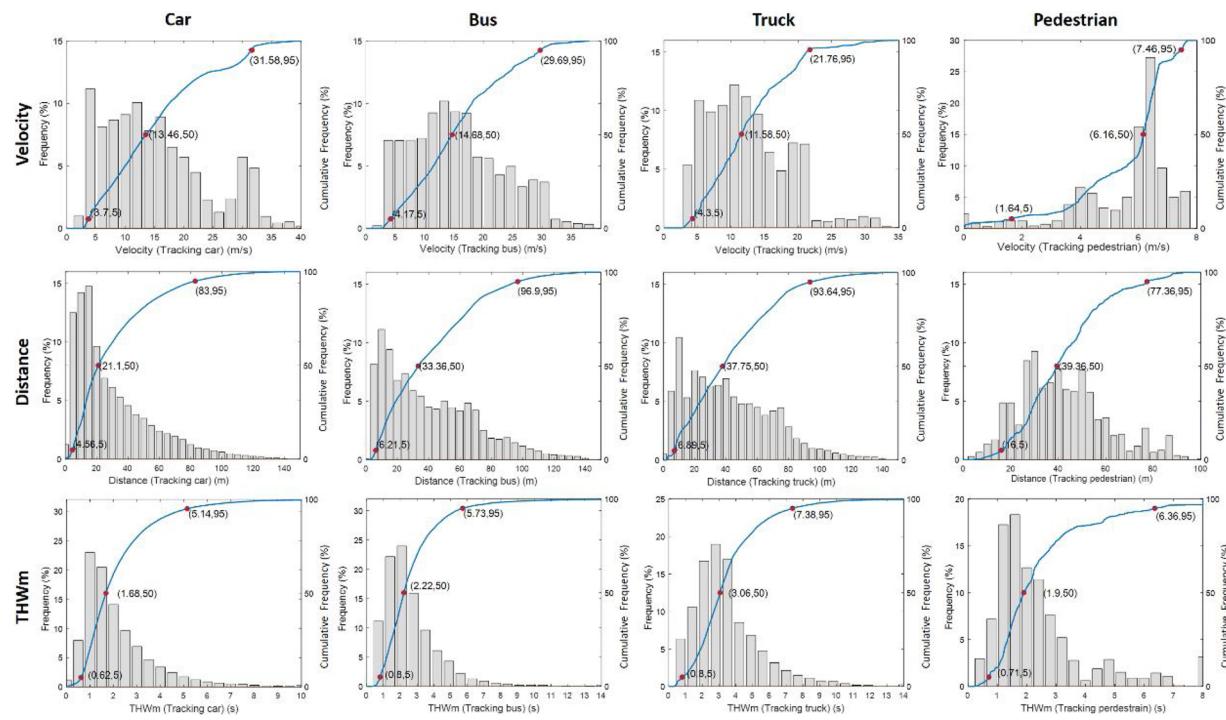


Fig. 5. Statistical analysis of manipulation parameters when tracing different objects.

encountered during driving, number of accidents, general velocity under different working conditions, and the evaluation criteria of different risks.

- Experimental information (8 scenarios with 24 pictures). Herein, the driver ranks the risk of the experimental scene according to the effect of different driving scenarios shown in the picture from the viewpoint of their feelings.
- Video information (6 videos), including the dynamic description and materialization of the above pictures using video information to assist drivers to have an intuitive feeling and judge the risks of different tracking objects (refer to Fig. 8)

3.2.2. Participants information

A sample of 364 participants with full China driving licenses is recruited to finally complete the whole questionnaire. The participants' demographic information shows in Table 3, which specifically gives a detailed description of the recruited drivers in our experiments. Other information can be found in Appendix.

3.2.3. Statistical data analysis

Statistically, reliability and validity are the major indicators that determine the efficiency and accuracy of self-reported questionnaire results. Cronbach's alpha reliability coefficient method (Peterson, 1994) is the most commonly used method to reflect the internal consistency of the scale. If the coefficient is greater than 0.6, then this ratio

is convincing. In our questionnaire, this coefficient is 0.86, proving that it is convincing. Besides, part of the detailed information of DAQ can be obtained from Table A1 in Appendix.

We investigate the main factors of driving risk by drawing on multiple sources. In all experiments, we set the first variable (good light, sunny days, and flat road) as the standard scenario in Table 1. Comparison 1 is the second one, whereas comparison 2 is the last one. Drivers' subjective risk assessment scores range from 0 to 10, where 0 and 10 represent absolute security and risk, respectively. In experiments 1 and 2, there are differences in exploring the geometric size, handling performance, and the psychological influence of different objects of neighboring vehicles. As revealed in Fig. 9 (a), drivers' subjective risk assessment scores of encountering car, bus, large truck, two-wheels, and pedestrian are 3.33, 4.95, 6.22, 5.17, and 6.24, respectively. While the complex scenario with various elements (cars, buses, trucks, among others) is 7.92.

In experiment 3, 4, and 5, we investigate whether diverse weather, road, and light influenced driver's judgment of risk level. Results show that different weather conditions and variations in rain and fog are larger than those of sunny days. For sunny, rainy, and foggy days, drivers' subjective risk assessment scores are 2.68, 6.23, and 7.95, respectively. When focusing on road conditions, the subjective scores are 2.82, 5.26, and 5.65, respectively, in the three cases of free flow and the good road surface, free but rough road surface, and congested but good road surface. Regarding lighting conditions, the distances in the dusk

Table 2

Characteristics of velocity, relative distance and time headway when tracing different objects.

Index	Ego vehicle' velocity (m/s)		Relative distance (m)		Time headway [s]	
	Object	F.	Mean ± S. D.	C.F.	Mean ± S. D.	C.F.
Car	948,999	15.11 ± 8.78	13.46 (3.70, 31.58)	30.11 ± 25.38	21.1 (4.56, 83.00)	2.13 ± 1.56
Bus	69,521	15.57 ± 7.84	14.68 (4.17, 29.69)	33.36 ± 29.28	33.36 (6.21, 96.90)	2.64 ± 1.80
Truck	30,998	12.45 ± 6.06	11.58 (4.30, 21.76)	42.21 ± 28.17	37.75 (6.89, 93.64)	3.44 ± 2.05
Pedestrian	1113	5.53 ± 1.68	6.16 (1.64, 7.46)	41.56 ± 18.2	39.36 (16.00, 77.36)	2.13 ± 1.56

Note: F. denotes frequency, S.D. denotes standard deviation, and C.F. denotes cumulative frequency. Inside the C.F., the median is in front, and first and second number in the bracket is 5% and 95%, respectively.

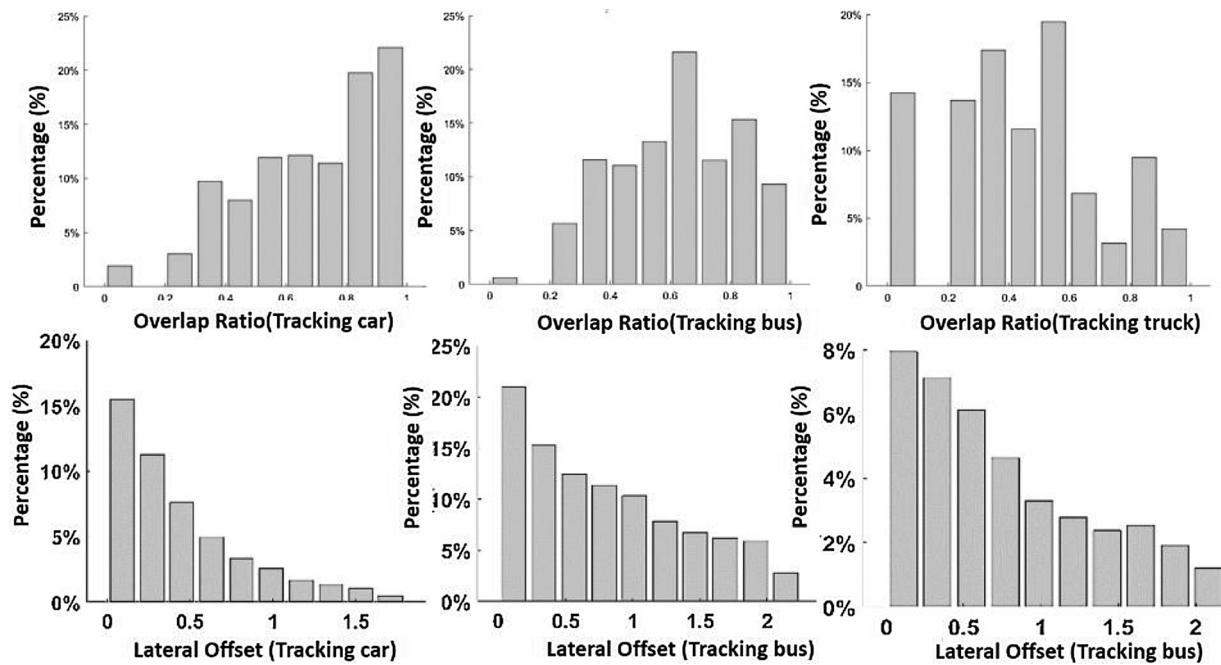


Fig. 6. Transverse overlap rate and lateral offset.

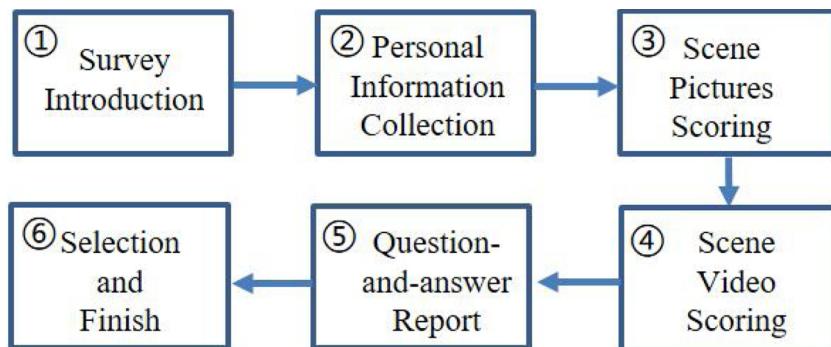


Fig. 7. The experimental procedures of driver subjective evaluation.



Fig. 8. Examples of videos used in the study.

Table 3

Demographic information for the driver participants.

Item	Number	%	Item	Number	%
Gender					
Male	190	52.19	High school or below	68	18.68
Female	174	47.81	Bachelor degree or above	296	81.32
Age					
< 20	9	2.47	< 1	44	12.09
20–30	180	49.45	1–2	65	17.86
30–39	123	33.79	3–5	149	40.93
40–50	40	10.99	6–10	77	21.15
> 50	12	3.30	> 10	29	7.97
Occupation					
Public institution	49	13.46	Very conservative	29	7.97
Company employee	152	41.76	Relative conservative	142	39.01
Professional	25	6.87	Normal	141	38.74
Civil servant	48	13.19	Relative aggressive	49	13.46
Student	90	24.73	Very aggressive	3	0.82
Frequently driving type					
Truck	6	1.65	Urban road	320	87.91
Car	278	76.37	Highway	172	47.25
SUV	76	21.7	Backroad	101	27.75
Bus	4	1.1	City loop	189	51.92

and dark are larger than that of good light days. The subjective scores are 2.79, 5.55, and 6.16 in the three cases of luminosity (good light conditions), dusk (dim light, no street light), and night (no natural light, no street light), respectively. These findings are shown in Fig. 9 (b), indicating that the driver's attitude could have a dynamic changing tendency while keeping a relatively stable evaluation of the various conditions.

In summary, the relation between driving risk and influencing factors, such as road users, weather, and road conditions are complex. From the potential influencing factors of driver's subjective perception, results identify that drivers' physiological and psychological factors, such as driving skills, personalities, ages, genders, emotions, and states, could provoke driving risk. The variation of road users shows a significant difference in the degree of risk perception. Road type and road characteristics may bring driving comfortableness, resulting in high subjective risks. Environmental variables, such as weather conditions and light conditions, are also the possible risk factors.

Table 4

The correlation coefficient between parameters.

	Index	LO	TOR	V	RD	THW
Index	1	0.146	-0.189	0.264	-0.443	-0.304
LO	0.146	1	0.871	0.042	-0.075	-0.105
TOR	-0.189	0.871	1	0.031	0.025	0.203
V	0.264	0.042	0.031	1	0.439	0.493
RD	-0.443	-0.075	0.025	0.437	1	0.944
THW	-0.304	-0.105	0.203	0.493	0.944	1

Note: In these tables, LO represents lateral offset; TOR means transverse overlap rate; V is the velocity of ego vehicle; RD is relative distance.

Table 5

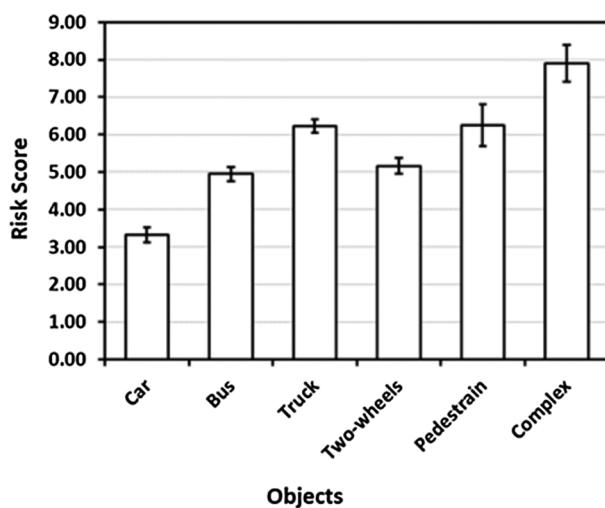
Risk coefficient value of different influencing factors.

Items	Risk value	Rank
Road users	Car	1.000
	Bus	1.363
	Truck	1.742
	Two-wheels	1.641
	Pedestrian	1.825
Environment	Sunny	1.000
	Rain	2.325
	Foggy	2.967
	Luminosity	1.000
Light	Duck	1.989
	Dark	2.210
Road	Flat	1.000
	Rough	1.865
	Free	1.000
Traffic	Congestion	2.004

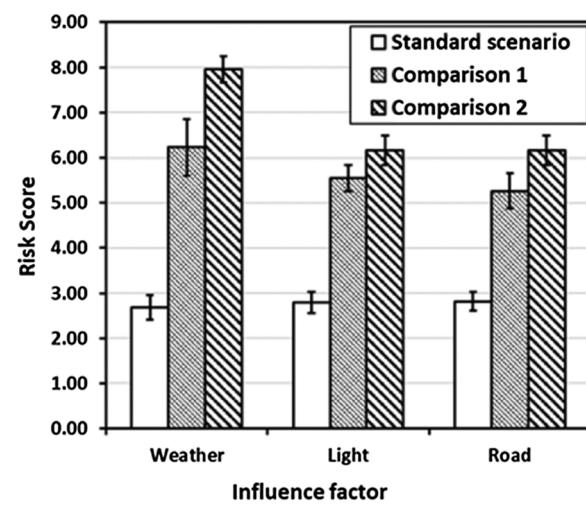
3.3. Comprehensive risk coefficient

3.3.1. Single-feature correlation analysis

We employ the single-feature correlation to analyze the parameters to attribute a greater impact on driving risk. The whole process can be described as: (1) extract the corresponding risk level from each segment of natural driving data and driver's subjective score. (2) Integrate the output score, evaluation index, and other data of subjective and objective data under the same type. (3) Synthesize the risk coefficient of each scenario (including different vehicle types, traffic participants, weather, lighting, and road conditions) that show the impact of a single factor on driving risk. In this method, the evaluation index of the risk level is calculated and the possible accident severity caused by each



(a) Risk scores of different objects



(b) Risk scores of different conditions

Fig. 9. Subjective risk scores of different objects and conditions.

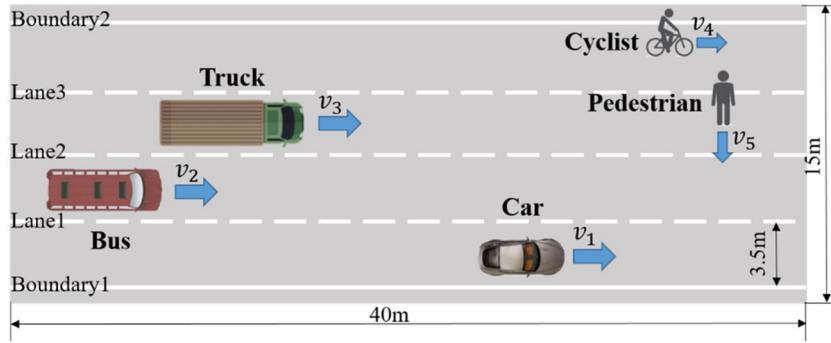


Fig. 10. Experimental setting diagram of traffic scenario.

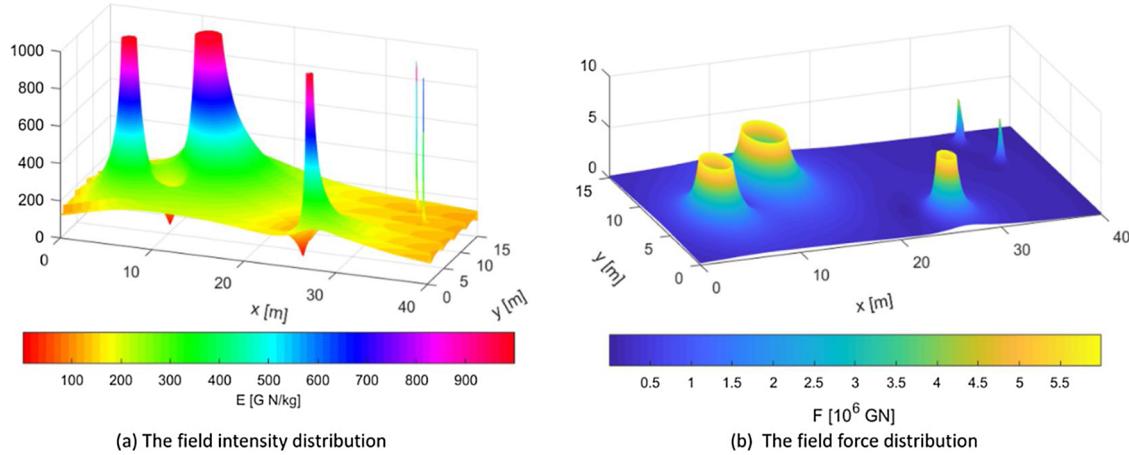


Fig. 11. The distribution of field intensity and force in a complex traffic scenario.

single factor variable.

In the single-feature correlation analysis method, the formula to calculate the correlation coefficient between target value t and feature x_j is as follows:

$$R(b, x_j) = \frac{\sum_i^N (b_i - \bar{b})(x_{ij} - \bar{x}_j)}{[\sum_{i=1}^N (b_i - \bar{b})^2 \sum_{i=1}^N (x_{ij} - \bar{x}_j)^2]^{\frac{1}{2}}} \quad (13)$$

where i is the sample number; b_i represents the object value of i th sample; x_{ij} is the j th characteristic value of sample i and \bar{b} , \bar{x}_j are the mean value respectively.

We also apply the single-feature correlation analysis to screen out the significant variables in the parameters obtained from NDS and to construct a factor analysis model, as shown in Table 4. Then, the three parameters of time headway, the speed of ego vehicle, and relative distance are selected as the modeling basis. Besides, we normalize the three parameters with the normalization formula:

$$Cox_{xj} = \frac{x_{ij} - minx_j}{maxx_j - minx_j} \quad (14)$$

where Cox_{xj} is the normalized parameter value.

3.3.2. Multinomial logit model analysis

When combining the data acquired subjectively and objectively, a probabilistic model is developed to evaluate the driving risk associated with each contributing factor of a potential accident. Previous studies have proven that the multinomial logit model (MLM) (Çelik and Oktay, 2014) can facilitate the coefficient of different variables to vary among individuals involved in driving risk. Besides, this model can analyze the relative risk by the utility function. Therefore, we assume that P_{di} represents the probability of driving risk, and U_{di} is a linear function that determines the value of the driving risk. The definition of P_{di} is

described as:

$$P_{di} = P(U_{di} \geq U_{di'}), \forall i' \in I, i' \neq i \quad (15)$$

where I denotes a set of risk levels. The function of U_{di} can be given as follows:

$$U_{di} = \beta_i x_d + \varepsilon_{di} \quad (16)$$

where β_i is a vector of estimable indexes; the polynomial logit model uses the linear function U_{di} to represent the relationship between the observable influencing factors x_d that directly cause risks (objective risks) and the non-observable random influencing factors ε_{di} that affect the correct behaviors (subjective risks). Thus MLM is described as follows:

$$P_{di} = \exp(\beta_i x_d) / \sum_{\forall i' \in I} \exp(\beta_i' x_d) \quad (17)$$

Analyses are conducted out for the first categories (road users, road condition, and environment), and they are subdivided into sub-categories of each major category (e.g., light condition, weather condition). Owing to the nonlinear characteristics of MLM, the estimated coefficients of independent variables cannot reflect the influence of independent variables directly (Çelik and Oktay, 2014). Therefore, it is necessary to set a standard condition. The risks associated with all contributing factors are then evaluated through a comparison with the set standard scenario to explain the differences clearly. Therefore, we set the standard category as the basis, defined as the risk coefficient as 1; relative risk (S_i) represents the risk of different traffic variables relative to the standard, which is defined as follows:

$$S_i = \frac{\Pr(i = k)}{\Pr(i = 1)} \quad (18)$$

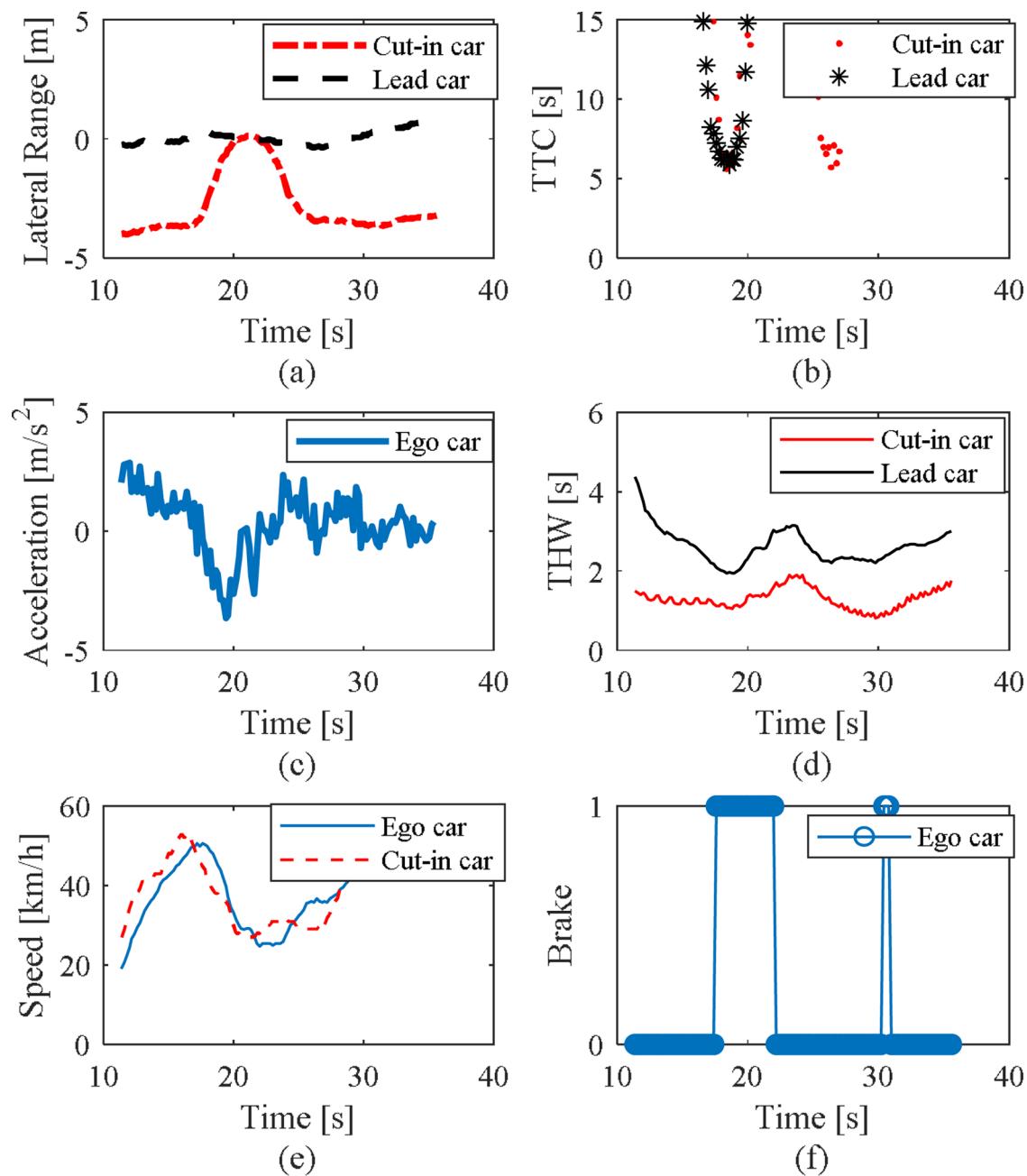


Fig. 12. TTC and THW in a critical cut-in scenario.

where Pr is defined as the relative probability of a potential risk ($i = k$) to the standard category ($i = 1$). When $S_i > 1$, this indicates that the risk of this variable is greater than the standard case, otherwise, the compared condition is safer than the standard. Finally, the degree of influence of different influencing factors on driving risk, namely the risk coefficient, is shown in Table 5.

From the obtained results, we can conclude that the discrimination of the risk score obtained by subjective factors is more obvious than that from objective data. Results show that when the drivers are influenced by more than one factor in the subjective judgment process, they fail to focus on one factor, put them in a more dangerous position and make them obtain higher risk scores. Besides, the evaluation results reveal that in the car-following scenario, road users, road conditions, and the environment have a direct impact on driving risk.

For different road users, their geometric size and handling performance have different psychological effects on drivers. Existing studies

(Brackstone et al., 2009) show that when at a distance velocity limit, drivers are at the highest risk when the front object is a truck and running at high speed. The value is approximately 1.7 times than that of a normal tracking car. Because in the car-following states, the following range is considerably short, which has a negative impact on the driver's visual effect, resulting in a large psychological threshold. Therefore, drivers tend to stay away from trucks and buses, which is consistent with our findings. For different road conditions, the horizontal curvature, terrain type, and traffic congestion of the road are closely related to the accident rate (Hosseinpour et al., 2014). When controlling for a single variable, traffic congestion is more likely to bring potential risks than uneven roads. For a diverse environment, under different weather conditions, the risk of the foggy day is the highest. Rainy days are 2.3 times riskier than sunny days. Under different light conditions, the light at dusk and night has a negative effect on the subjective feeling and objective behavior of the driver.

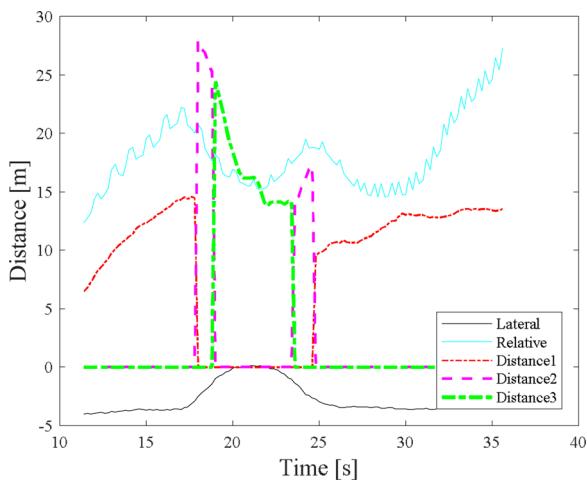


Fig. 13. The distance limits at different stages in the cut-in scenario.

4. Results and discussion

This study aims to accurately evaluate the driving risk in the dynamic traffic environment, output the relative risk value of each traffic factor, realize early warning of driving and improve travel safety. Therefore, based on the concept of driving safety field put forward by (J. Wang et al., 2016), the coupled driving risk assessment model is established. Specifically, risk coefficients of different influencing factors output in Table 5 are obtained through subjective and objective data analysis, which are input into the external field formed by traffic environment. The internal field is established by analyzing the driver's visual characteristics. Finally, the accuracy of the proposed model is verified through applying in the dynamic traffic environment and comparing with other risk indexes.

4.1. Risk map of traffic environment

The field intensity of the driving risk coupling model refers to the coupling field intensity generated by the internal field of the driver and the external field of other traffic elements, as the entire traffic scenario setting is shown in Fig. 10. In this traffic scenario, from 0 to 15 on the y-axis, moving cars, buses, trucks, bicycles, and pedestrians are in order.

In the complex traffic scene with multiple lanes and multiple traffic users, the field intensity distribution of the entire traffic environment is obtained by calculating the field intensity generated by each traffic element (refers to Fig. 11 (a)). The distribution diagram of the field force is obtained by taking the distance derivative of the field intensity, as shown in Fig. 11 (b). In Figs. 11 (a), (b), the difference in the influence range and intensity of field among different traffic factors are revealed. Besides, through the field intensity distribution, the potential risks brought by different traffic factors are evaluated.

4.2. TTC and THW

To verify the effect of the field model, this study simultaneously compares the field model and the traditional driving risk assessment indexes TTC and THW in the dangerous cut-in scenario. Therefore, we define the entire driving process as stage 1: completely in adjacent lanes, stage 2: cut-in process, and stage 3: car-following process. Besides, the cut-in scenario includes ego vehicle, the front vehicle in the same lane, and the cut-in vehicle in the next lane. From Fig. 12 (a), the entire cut-in process starts at 12 s and ends at 25 s. Although TTC signal is not enough to trigger collision warning in this process, it is seen in Fig. 12 (b) that the TTC curve drops sharply, whereas THW signal fluctuates slowly Fig. 12 (d). Similarly, in the cut-in process, the driver performs emergency braking and speed reduction (refer to Figs. 12 (c), (e),(f)), consistent with the TTC signal trend, which shows that the TTC signal is more sensitive to the driver.

4.3. Field-based coupling model

The entire cut-in process in Fig. 13, the changes in the relative lateral position of the cutting vehicle (lateral), the relative distance between the ego vehicle and the cut-in vehicle (Relative), and the distance in the first, second and third stages of the cut-in process (Distance 1, 2, and 3) are remarked.

As seen in Fig. 14, the positive field intensity is an attractive effect, whereas the negative field intensity is represented as a repulsive effect. When the driver suffers a significant amount of driving risk and the repulsive effect is greater than the driving effect, the driver will reduce the acceleration to weaken the risk. On the contrary, when the current environment is safe, the driving effect is greater than the repulsive effect, and the driver will accelerate. In Fig. 14, when the driver

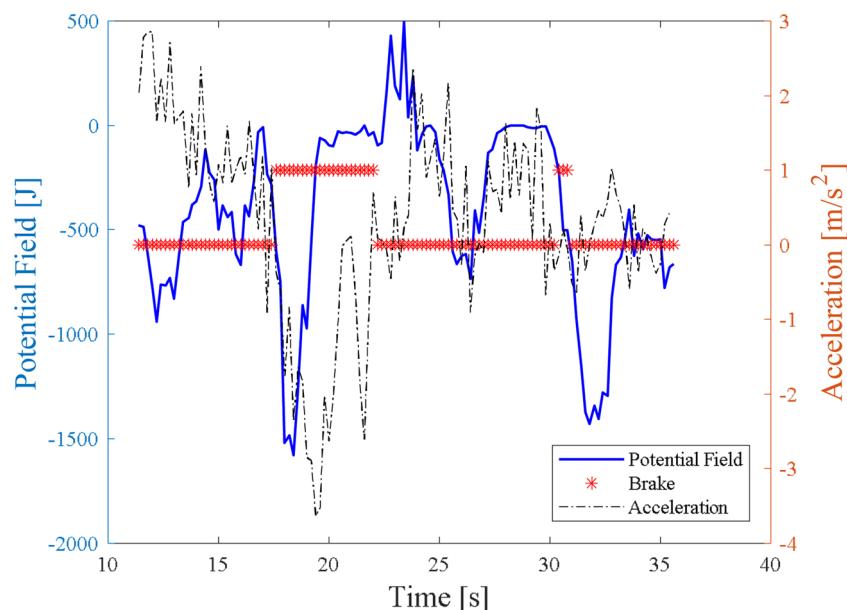


Fig. 14. The potential field of the cut-in vehicle in the cut-in scenario.

accelerates in the first stage, the acceleration decreases as it gradually approaches other vehicles. The driver perceives the risk, and the potential field becomes negative. During the second stage (18–21 s), the driver cuts another lane and the repulsive effect of the potential field increases rapidly. When emergency brakes are applied, the collision risk is reduced and repulsive effect disappears at the end of the second stage (21–24 s), and the field strength reaches zero. In the third stage, the cut-in process ends. At this time, the potential risk is caused by the distance between the front and rear vehicles. If the driver is too close to the front car, the potential energy field will show the repulsive effect. As the driver applies brake, the repulsion diminishes.

In conclusion, we verify the strengths of the proposed algorithm by comparing the proposed driving risk coupling field model with TTC and THW methods in the cut-in scenario. TTC and THW metrics are very classical and effective risk assessment indicators, which are very common in the comparison and verification of risk assessment. Specifically, TTC is a method based on time logic, and THW is a method based on distance logic. Both methods are simple and effective, consistent with human drivers' intuitive perception of a collision or potential risk. Therefore, the use of these two safety measures as a comparative verification method is persuasive. Results show that TTC is more sensitive than THW in the cut-in scene. However, its longitudinal direction signals change rapidly, whereas the lateral direction is insensitive. Moreover, the proposed potential field reflects the change of collision risk in advance, which is also consistent with the operation of the driver. The proposed potential field is a comprehensive risk assessment index, which can more sensitively reflect the directional risks in the driving process. This is more accurate than TTC, THW, and other one-way risk assessment indexes. For example, it responds to lateral collision risks faster than TTC, as shown in Fig. 12 (b).

5. Conclusion

This paper presented a practical methodology for assessing driving risk based on an integrated driving risk model. In particular, the proposed model helps in evaluating the driving risk of traffic elements, with the internal field to define the risk range and the external field to calculate the risk level. Furthermore, to quantify the relative risk coefficient of the external field, a multinomial logit model (MLM) is proposed to incorporate the objective analysis obtained from NDS and subjective evaluation through DAQ. In particular, a large-scale

naturalistic driving test has been carried out to extract the driver's behavior parameters. The objective risk is analyzed based on significant driver characteristics. The DAQ influencing factors are designed by considering the driver's subjective identification. Results indicate that the proposed model is more accurate and sensitive in evaluating driving risk combining internal with the external field. Moreover, the rationality and strengths of the proposed coupling model are verified by different scenarios in comparison with TTC and THW. The calculated value describes the relative potential risk in various scenarios to support drivers for safer decisions.

Some interesting extensions to this study are worth exploring in future research. We can develop driver assistance systems employing the proposed model to better ensure driving safety. In addition, some potential risk influencing factors have not been extracted due to the complex environment; therefore, future work will focus on extracting more data from objective and subjective experiments to improve the accuracy of the algorithm, and apply the proposed method to the risk assessment of complex traffic scenes and accident prevention of road sections with frequent accidents.

Author contributions

Jianqiang Wang, Heye Huang: Methodology, Simulations, Data analysis, Writing - Original draft preparation.

Yang Li, Hanchu Zhou, Jinxin Liu: Discussion, Investigation, Validation, Visualization, Improvement.

Qing Xu: Conceptualization, Writing - Reviewing and Editing.

Declaration of Competing Interest

None.

Acknowledgments

This research was supported by National Natural Science Foundation of China (Grant No. 51625503 and 61790561) and the Joint Laboratory for Internet of Vehicle, Ministry of Education - China Mobile Communications Corporation. Natural Data source comes from China Automotive Engineering Research Institute Co., Ltd.

Appendix A

Table A1
Driving behavior of driver participants (Partial questionnaire information).

Q1		Have you ever had an accident while driving or riding in a car?									
		1	rear-end	2	head-on collision	3	Touching the side of ego car	4	Touching the side of other car	5	The side scratches
N/%	215 (59.07 %)			42 (11.54 %)		162 (44.51 %)		74 (20.33 %)		237 (65.11 %)	
6	Collision of VRUs	6	7	other _____		8	No				
N/%	75 (20.6 %)			4 (1.1 %)		45 (12.36 %)					
Q2	What do you think are the main causes of the above traffic accidents?										Vehicle factors (e.g., driver assistant system interfering with judgment)
N/%	1 Human factors (e.g., driver distraction)	1	303 (83.24 %)		133 (36.54 %)						
N/%	3 Road factors (e.g., unclear traffic signs)	3	177 (48.63 %)		4	Weather factors (e.g., rain and snow days)	5	Other _____	205 (56.32 %)	12 (3.3 %)	
Q3	What is your normal driving speed on the highway?										100 – 120 km/h
N/%	60 – 80 km/h	67 (21.15 %)		80 – 100 km/h		206 (56.59 %)		81 (22.25 %)			
Q4	The safety level driving in the following types of roads (e.g., Beijing) (from 0 to 10; most safe:0,most danger:10)										Municipal mixed road
Mean	City expressway	3.37		Highway	1.55		2.37				Suburban country road
Q5	The safety level driving in the following weather conditions										Rain
Mean	Sunny	2.68		Rain	6.23		7.95				Foggy(visibility < 100 m)
Q6	The safety level driving in the following light conditions										Night (no natural light, no street light)
Mean	Luminosity (good light conditions)	2.79		Dusk (dim light, no street light)	5.55		5.13				Night (no natural light, with street light)
Q7	The degree to which you perceive the danger of the front object in the following process										6.16
Mean	Car	3.33		Large truck	6.22		5.17				Two-wheels
Q8	When you see a large truck on the road ahead, which of the following driving options will you choose										Pedestrian
Mean	Slow down, pull away	245 (67.31 %)		Speed up, change lanes	101 (27.75 %)		13 (3.57 %)				6.24
Q9	When you see a bus on the road ahead, which of the following driving options will you choose										Complex
Mean	Slow down, pull away	193 (53.02 %)		Speed up, change lanes	112 (30.77 %)		56 (15.38 %)				7.92
Q10	When you see a car on the road ahead, which of the following driving options will you choose										Other manipulation
Mean	Slow down, pull away	83 (22.8 %)		Speed up, change lanes	58 (15.93 %)		217 (59.62 %)				5 (1.37 %)
Q11	The following questions are analyzed based on the normal following scene on the highway:										Night (no natural light, no street light)
Q11	When driving in the highway with no block while limit in lane change, following with the front car, how much km/h is your driving speed at the following stage? What is the average distance from the car?										90 – 100
Speed	60 – 70	70 – 80		80 – 90		90 – 100		100 – 120			100 – 120
N/%	138 (37.91 %)	98 (26.92 %)		89 (24.45 %)		77 (21.15 %)		19 (5.22 %)			
Distance	0 – 10	10 – 40		40 – 70		40 – 70		> 100			
N/%	19 (5.22 %)	84 (23.08 %)		109 (29.95 %)		94 (25.52 %)		84 (23.08 %)			
Q12	When driving in the highway with no block while limit in lane change, following with the front truck, how much km/h is your driving speed at the following stage? What is the average distance from the truck?										90 – 100
Speed	60 – 70	70 – 80		80 – 90		90 – 100		100 – 120			100 – 120
N/%	138 (37.91 %)	91 (25.00 %)		83 (22.80 %)		83 (22.80 %)		15 (4.12 %)			
Distance	0 – 10	10 – 40		40 – 70		40 – 70		> 100			
N/%	8 (2.20 %)	45 (12.36 %)		149 (40.93 %)		104 (29.94 %)					
Q13	When driving in the highway with no block while limit in lane change, following with the front bus, how much km/h is your driving speed at the following stage? What is the average distance from the bus?										90 – 100
Speed	60 – 70	70 – 80		80 – 90		90 – 100		100 – 120			100 – 120
N/%	88 (24.18 %)	104 (28.57 %)		112 (30.77 %)		112 (30.77 %)		12 (3.30 %)			
Distance	0 – 10	10 – 40		40 – 70		40 – 70		> 100			
N/%	8 (2.20 %)	61 (16.76 %)		165 (45.33 %)		165 (45.33 %)		39 (10.71 %)			
Q14	When driving in the highway with no block at night, following with the front car, how much km/h is your driving speed at the following stage? What is the average distance from the car?										90 – 100
Speed	60 – 70	70 – 80		80 – 90		80 – 90		100 – 120			100 – 120
N/%	97 (26.65 %)	88 (24.18 %)		88 (24.18 %)		88 (24.18 %)		10 (2.75 %)			

(continued on next page)

Table A1 (continued)

Q1	Have you ever had an accident while driving or riding in a car?	10–40	40–70	70–100	> 100
Distance	0–10	52 (14.29 %)	153 (42.03 %)	99 (27.20 %)	46 (12.64 %)
N/%	14 (3.85 %)				
Q15	When driving in the highway with no block in a rainy day, following with the front car, how much km/h is your driving speed at the following stage? What is the average distance from the car?	70–80	80–90	90–100	100–120
Speed	60–70	104 (28.57 %)	112 (30.77 %)	48 (13.19 %)	12 (3.30 %)
N/%	88 (24.18 %)				
Distance	0–10	10–40	40–70	70–100	> 100
N/%	8 (2.20 %)				
Q16	When driving in the highway with block and limit in lane change, following with the front car, how much km/h is your driving speed at the following stage? What is the average distance from the car?	61 (16.76 %)	165 (45.33 %)	61 (16.76 %)	69 (18.96 %)
Speed	60–70	70–80	80–90	90–100	100–120
N/%	273 (75.00 %)	58 (15.93 %)	28 (7.69 %)	3 (0.82 %)	2 (0.55 %)
Distance	0–10	10–40	40–70	70–100	> 100
N/%	64 (17.58 %)				
Q17	When driving in the highway with no block while limit in lane change, following with the front car on uneven road, how much km/h is your driving speed at the following stage? What is the average distance from the car?	70–80	80–90	90–100	100–120
Speed	60–70	100 (27.47 %)	50 (13.74 %)	18 (4.95 %)	3 (0.82 %)
N/%	193 (53.02 %)	10–40	40–70	70–100	> 100
Distance	0–10				
N/%	26 (7.14 %)				

References

- Ahmed, M.M., Abdel-Aty, M., Lee, J., Yu, R., 2014. Real-time assessment of fog-related crashes using airport weather data: a feasibility analysis. *Accid. Anal. Prev.* 72, 309–317.
- Al-Ghamdi, A.S., 2002. Using logistic regression to estimate the influence of accident factors on accident severity. *Accid. Anal. Prev.* 34 (6), 729–741.
- Allen, C., Ewing, M., Keshmiri, S., Zakharov, M., Florencio, F., Niakan, N., Knight, R., 2013. Multichannel sense-and-avoid radar for small UAVs. In: Presented at the 2013 IEEE/AIAA 32nd Digital Avionics Systems Conference (DASC). pp. 6E2-1-6E2-10.
- Archibald, J.K., Hill, J.C., Jepsen, N.A., Stirling, W.C., Frost, R.L., 2008. A satisficing approach to aircraft conflict resolution. *IEEE Trans. Syst. Man Cybern. C Appl. Rev.* 38 (4), 510–521.
- Aven, T., 2011. A risk concept applicable for both probabilistic and non-probabilistic perspectives. *Saf. Sci.* 49 (8), 1080–1086.
- Brackstone, M., Watson, B., McDonald, M., 2009. Determinants of following headway in congested traffic. *Transp. Res. Part F Traffic Psychol. Behav.* 122, 131–142.
- Celik, A.K., Oktay, E., 2014. A multinomial logit analysis of risk factors influencing road traffic injury severities in the Erzurum and Kars Provinces of Turkey. *Accid. Anal. Prev.* 72, 66–77.
- Elvik, R., 2009. An exploratory analysis of models for estimating the combined effects of road safety measures. *Accid. Anal. Prev.* 41 (4), 876–880.
- Faragher, R., 2012. Understanding the basis of the kalman filter via a simple and intuitive derivation [Lecture notes]. *IEEE Signal Process. Mag.* 29 (5), 128–132.
- Gerdes, J.C., Rossetter, E.J., 2001. A unified approach to driver assistance systems based on artificial potential fields. *J. Dyn. Syst. Meas. Control* 123 (3), 431. <https://doi.org/10.1115/1.1386788>.
- Goerlandt, F., Kujala, P., 2014. On the reliability and validity of ship-ship collision risk analysis in light of different perspectives on risk. *Saf. Sci.* 62, 348–365.
- Goerlandt, F., Reniers, G., 2016. On the assessment of uncertainty in risk diagrams. *Saf. Sci.* 84, 67–77.
- Greene, D., Liu, J., Reich, J., Hirokawa, Y., Shinagawa, A., Ito, H., Mikami, T., 2011. An efficient computational architecture for a collision early-warning system for vehicles, pedestrians, and bicyclists. *IEEE Trans. Intell. Transp. Syst.* 12 (4), 942–953.
- Hosseinpour, M., Yahaya, A.S., Sadullah, A.F., 2014. Exploring the effects of roadway characteristics on the frequency and severity of head-on crashes: case studies from Malaysian Federal Roads. *Accid. Anal. Prev.* 62, 209–222.
- Householder, A.S., 1939. (Lewin, Kurt. Principles of topological psychology. Translated by Fritz and Grace Heider. New York: McGraw-Hill, 1936. Pp. 231.). *Pedagog. Semin. J. Genet. Psychol.* 54 (1), 249–259.
- Hsu, T.-P., Wang, G.-Y., Lin, Y.-J., 2012. conceptual structure of a novel car-following model upon gravitational Field concept. Presented at the 19th ITS World CongressERTICO - ITS EuropeEuropean CommissionITS AmericaITS Asia-Pacific.
- Jafari Anarkooli, A., Hadji Hosseiniou, M., 2016. Analysis of the injury severity of crashes by considering different lighting conditions on two-lane rural roads. *J. Safety Res.* 56, 57–65.
- Khatib, O., 1990. Real-time obstacle avoidance for manipulators and Mobile robots. In: Cox, I.J., Wilfong, G.T. (Eds.), *Autonomous Robot Vehicles*. Springer, New York, New York, NY, pp. 396–404.
- Kinnear, N., Helman, S., Wallbank, C., Grayson, G., 2015. An experimental study of factors associated with driver frustration and overtaking intentions. *Accid. Anal. Prev.* 79, 221–230.
- Knoefel, F., Wallace, B., Goubran, R., Marshall, S., 2018. Naturalistic driving: a framework and advances in using big data. *Geriatrics* 32, 16.
- Laugier, C., Paromtchik, I.E., Perrollaz, M., Yong, M., Yoder, J., Tay, C., Mekhnacha, K., Nègre, A., 2011. Probabilistic analysis of dynamic scenes and collision risks assessment to improve driving safety. *IEEE Intell. Transp. Syst. Mag.* 34, 4–19.
- Lewin, K., 1936. *Principles of Topological Psychology*, *Principles of Topological Psychology*. McGraw-Hill, New York, NY, US.
- Li, Y., Wang, J., Wu, J., 2017. Model calibration concerning risk coefficients of driving safety field model. *J. Cent. South Univ.* 24 (6), 1494–1502.
- Malta, L., Miyajima, C., Takeda, K., 2009. A study of driver behavior under potential threats in vehicle traffic. *IEEE Trans. Intell. Transp. Syst.* 10 (2), 201–210.
- Martensen, H., Dupont, E., 2013. Comparing single vehicle and multivehicle fatal road crashes: a joint analysis of road conditions, time variables and driver characteristics. *Accid. Anal. Prev.* 60, 466–471.
- Martinussen, L.M., Möller, M., Prato, C.G., 2014. Assessing the relationship between the driver behavior questionnaire and the driver skill inventory: revealing sub-groups of drivers. *Transp. Res. Part F Traffic Psychol. Behav.* 26, 82–91.
- Moreau, J., Melchior, P., Victor, S., Aïoun, F., Guillemand, F., 2017. Path planning with fractional potential fields for autonomous vehicles. *IFAC-Pap.* 50 (1), 14533–14538.
- Morgan, A., Manning, F.L., 2011. The effects of road-surface conditions, age, and gender on driver-injury severities. *Accid. Anal. Prev.* 43 (5), 1852–1863.
- Musicant, O., Bar-Gera, H., Schechtman, E., 2014. Temporal perspective on individual driver behavior using electronic records of undesirable events. *Accid. Anal. Prev.* 70, 55–64.
- Ni, H., Chen, A., Chen, N., 2010. Some extensions on risk matrix approach. *Saf. Sci.* 48 (10), 1269–1278.
- Peterson, R.A., 1994. A meta-analysis of Cronbach's coefficient alpha. *J. Consum. Res.* 21 (2), 381–391.
- Reimer, B., D'Ambrosio, L.A., Gilbert, J., Coughlin, J.F., Biederman, J., Surman, C., Fried, R., Aleardi, M., 2005. Behavior differences in drivers with attention deficit hyperactivity disorder: the driving behavior questionnaire. *Accid. Anal. Prev.* 37 (6), 996–1004.
- Rolison, J.J., Regev, S., Moutari, S., Feeney, A., 2018. What are the factors that contribute

- to road accidents? An assessment of law enforcement views, ordinary drivers' opinions, and road accident records. *Accid. Anal. Prev.* 115, 11–24.
- Saifuzzaman, M., Zheng, Z., 2014. Incorporating human-factors in car-following models: a review of recent developments and research needs. *Transp. Res. Part C Emerg. Technol.* 48, 379–403.
- Treiber, M., Kesting, A., 2013. Car-following models based on driving strategies. *Traffic Flow Dynamics*. Springer, Berlin, Heidelberg.
- Tu, Q., Chen, H., Li, J., 2016. A potential field based lateral planning method for autonomous vehicles. *SAE Int. J. Passeng. Cars - Electron. Electr. Syst.* 10, 1.
- van Winsum, W., 1999. The human element in car following models. *Transp. Res. Part F Traffic Psychol. Behav.* 2 (4), 207–211.
- Wang, J., Zhang, L., Zhang, D., Li, K., 2013. An adaptive longitudinal driving assistance system based on driver characteristics. *IEEE Trans. Intell. Transp. Syst.* 14 (1), 1–12.
- Wang, W., Xi, J., Chen, H., 2014. Modeling and recognizing driver behavior based on driving data: a survey. *Math. Probl. Eng.*
- Wang, J., Wu, J., Zheng, X., Ni, D., Li, K., 2016. Driving safety field theory modeling and its application in pre-collision warning system. *Transp. Res. Part C Emerg. Technol.* 72, 306–324.
- Yoshitake, H., Shino, M., 2018. Risk assessment based on driving behavior for preventing collisions with pedestrians when making across-traffic turns at intersections. *IATSS Res.* 42 (4), 240–247.
- Zhang, L., Wang, J., Yang, F., Li, K., 2009. A quantification method of driver characteristics based on driver behavior questionnaire. *2009 IEEE Intelligent Vehicles Symposium*. Presented at the 2009 IEEE Intelligent Vehicles Symposium 616–620.
- Zhang, W., Dai, J., Pei, Y., Li, P., Yan, Y., Chen, X., 2016. Drivers' visual search patterns during overtaking maneuvers on freeway. *Int. J. Environ. Res. Public Health* 13 (11), 1159.