COMPUTATIONAL PERCEIVED RISK MODELS IN SAE LEVEL 2 DRIVING AUTOMATION: FORMULATION, VALIDATION AND COMPARISON WITH EXPERIMENTAL DATA

Xiaolin He, Corresponding Author
Department of Cognitive Robotics
Delft University of Technology
Mekelweg 2, 2628 CD, Delft, the Netherlands.
Email: x.he-2@tudelft.nl

Riender Happee
Department of Cognitive Robotics
Delft University of Technology
Mekelweg 2, 2628 CD, Delft, the Netherlands.
Email: r.happee@tudelft.nl

Meng Wang
Chair of Traffic Process Automation
"Friedrich List" Faculty of Transport and Traffic Sciences
Technische Universität Dresden
Hettnerstraße 3,01069 Dresden, Germany.
Email: meng.wang@tu-dresden.de

Word Count: 6968 words + 2 table(s) × 250 = 7468 words

Submission Date: August ~~31~~, 2022

ABSTRACT

Perceived risk captures the risk during driving perceived by drivers and is crucial for the design and evaluation of driving automation systems. Inappropriate perceived risk may cause driver's failure to recognise dangerous situations that need driver's operation or take-over control in case of automated driving. However, real-time perceived risk models have rarely been reported in the literature. This study aims to formulate and validate three perceived risk models for drivers of SAE level 2 automated vehicles and compare them with multiple performance criteria. Regression perceived risk model (RPR), perceived probabilistic driving risk field model (PPDRF), and driving risk field model (DRF) were extended based on our previous work and literature. The models were calibrated using two different datasets. Results showed that all the models can capture risk variation in risky events and quantify perceived risk. The observation-based model RPR and DRF perform better in their original datasets. PPDRF has advantages in 2-D computation and prediction in the time domain and in explaining the mechanism of perceived risk. The motion prediction of other road users is essential for perceived risk computation.

Keywords: Percevied risk, computational models, experimental data

INTRODUCTION

Road crashes are a leading cause of injury and death worldwide, resulting in approximately 1.35 million deaths and 20-50 million non-fatal injuries each year (*1*). Traffic accidents are mainly caused by human misjudgements (*2*). Specifically, distorted perception of the driving risk by human drivers is one of the important causes of road accidents (*3*).

Perceived risk captures the level of risk experienced by drivers, which can differ from operational (or actual) risk (*4, 5*). A low perceived risk leads to feeling safe, relaxed, and comfortable, while a high risk perception results in cautious behaviour (*4*). With the advances of active safety and driving automation systems, the collision risk is generally reduced but driver's risk perception pattern is also changed. Inappropriate perceived risk may lead drivers to ignore the dangerous situations where they have to monitor the operation of the driving automation or take over the control, probably causing severe traffic accidents. Therefore, it is essential to understand and compute drivers' perceived risk in driving automation and in turn, use it to make driving automation perceived safer.

The zero risk theory postulates that all the motivations for actions are to maintain perceived risk below a specific threshold (*6*). We can accordingly predict human driver's behaviour by keeping perceived risk below a certain value (*7*). In automated vehicles (AV), human drivers will inversely perceive a high level of risk if the driving automation does not have appropriate driving behaviours, causing decreased trust and low acceptance, and even refusal to use AV (*8*). In this aspect, the design of AVs also needs to estimate human driver's perceived risk level.

A few attempts have been made to model and compute perceived risk. They are either phenomenological models based on observations or mechanistic models based on first-principles. In the first category, Kolekar et al. (*5*) established a driving risk field (DRF) model based on driver's subjective risk ratings and steering responses considering the probability of the event occurring and the event consequence. Ping et al. (*9*) used deep learning methods to model perceived risk in urban scenarios including factors related to the status of the subject vehicle and external environment information. Our previous study built a regression perceived risk model to explain and compute event-based perceived risk in highway driving reacting to merging and hard braking vehicles. Among other factors, the model captures the influence of relative motion to other road users on drivers' subjective perceived risk ratings (*10*).

Mechanistic perceived risk models are usually based on surrogate measures of safety (SMoS) or driving risk theory. For the classic SMoS, minimum time to collision (TTC) can show the drivers' threshold of perceived risk when they take braking actions (*11*). The inverse TTC represents drivers' relative visual expansion of the obstacle (*12*). There are more complicated driving risk models based on driving risk field theory. Wang et al. (*13*) use artificial field theory to model driving risk considering the influence of driver, vehicle, and road characteristics. Li et al. (*14*) use this theory to develop a warning strategy to prevent traffic accidents. Probabilistic driving risk field (PDRF) (*15*) using the paradigm of artificial field theory, considers the probability of motion predictions of other road users and the collision severity to estimate the collision risk in 2 dimensions. In fact, all above-mentioned models estimate the actual collision risk rather than perceived risk, although there is a connection between the actual collision risk and perceived risk.

The observation-based models can estimate perceived risk quite well in specific datasets but are not validated in various scenarios and are not fully explainable; the mechanistic models that are explainable can compute actual risk but the mapping between the actual collision risk and perceived risk is still unclear and the threshold of the SMoS has not been determined reasonably

based on the real driving data for perceived risk. Moreover, no comprehensive comparison has been conducted to analyse the different models' performance.

To fill the research gaps, this study has two main objectives: **Objective 1** is to formulate three perceived risk models based on different mechanisms and calibrate them using two different datasets; **Objective 2** is to analyse and comprehensively compare the three models in a systematic way. In this study, other road users' motion is focused on in all models since it accounts for most of driver's risk perception (*10, 16*). In recent years, personalised driver models that can describe human drivers' behaviour have gained increasing attention, the parameters of which can vary to adapt to individual differences. This study aims to develop **general models** that can be tuned to adapt to different scenarios and individual differences instead of direct personalised modelling. Our model parameters will be calibrated based on two different datasets.

The remainder of this paper is structured as follows. The formulation of three models is detailed in Section Model formulation. The methods including perceived risk data, model calibration approach and the performance indices are introduced in Section Model calibration and performance indices. The calibration and simulation results are represented in Section Results followed by Section Discussions, and the conclusions are given in Section Conclusions and future work.

MODEL FORMULATION

In this section, two models are formulated by extending our previous studies. A third model is adapted based on a state-of-art model from the literature.

Regression perceived risk model (RPR)

The regression perceived risk model (RPR) was an event-based perceived risk model established from a simulator experiment of us, where 18 merging events with various merging distance and braking intensity on a 2-lane highway were simulated. Event-based perceived risk ratings and continuous perceived risk measures were collected, based on which the event-based regression perceived risk model was established with the corresponding kinematic data (*10*). The details of the data will be introduced later.

The RPR model builds on several assumptions:

- Perceived risk stems from the vehicles directly in front of the subject vehicle, which means the merging vehicles cause perceived risk only after entering the current lane;
- Driver can accurately estimate the motion information (e.g., relative position, velocity, acceleration, etc.) with the human sensory system.
- Perceived risk in the forward driving direction is only influenced by the lane change vehicle or the leading vehicle, although the general perceived risk is caused by many objects (e.g., other traffic vehicles in the adjacent lane, infrastructure, etc.) during the drive.

The event-based regression perceived risk model was built based on the event-based perceived risk and the corresponding kinematic data in merging and braking events. The original model can predict event-based perceived risk (*10*), as shown in Equation (1)

$$perceived\_risk = 9.384 - 2.473 \cdot \ln(min\_gap) - 0.038 \cdot YDL - 0.201 \cdot max\_BI + 0.470 \cdot GEN \quad (1)$$

where *perceived_risk* is the event-based perceived risk ranging from 0-10; *min_gap* is the minimum relative gap to the leading vehicle during an event; *YDL* represents the years with a valid driving license; *max_BI* denotes the maximum braking intensity of the merging vehicle; *GEN*

represents the gender of the participants with $Female = 1$ and $Male = 0$.

The model can output event-based perceived risk after a certain event but cannot do real-time computation. Hence, a necessary extension is needed to compute real-time perceived risk. In this study, we replace the minimum gap between the leading and the subject vehicle and the average braking intensity of the leading vehicle in one certain event of the real-time values. YDL and GEN are neglected since they are always constant regarding a certain group of participants. In this way, RPR is formulated in the continuous time domain as

$$RPR(t) = C_0 + C_1 \cdot \ln(gap(t)) + C_2 \cdot BI(t) \quad (2)$$

where $gap(t)$ is the real-time gap between the subject vehicle and the leading vehicle (m); $BI(t)$ is the current braking intensity of the leading vehicle ($\mathrm{m/s^2}$); According to the simulator experiment settings (*10*), the validity range of the model is that $gap(t) < 33\,\mathrm{m}$ and $-8\,\mathrm{m/s^2} \leqslant BI(t) \leqslant -2\,\mathrm{m/s^2}$. The pattern of perceived risk needs to be verified outside this range. To get better performance, parameters C_0, C_1 and C_2 for different datasets will be calibrated later.

Perceived probabilistic driving risk field model (PPDRF)

Perceived probabilistic driving risk field model (PPDRF) is built based on the probabilistic driving risk field model (PDRF) (*15*), considering different traffic conditions and drivers' individual differences. The basis, artificial potential field, is a popular method that can be used to estimate the collision risk and control the vehicle motion in driving automation (*13, 14, 17*).

The original PDRF (*15*) includes two components to estimate the collision risk: the potential risk from non-moving objects and the kinetic risk caused by other road users. PDRF estimates the potential risk by the collision energy and the collision probability that increases with the decrease of relative distance to the non-moving obstacles. For the kinematic risk, PDRF computes the collision probability based on the overlap in space with objects at a single future time instant. Correspondingly, the predicted position of the subject vehicle according to the current motion state, a vehicle dynamics model, and the range of predicted positions of the neighbours are used. The risk is estimated by using a stochastic approach where the neighbour's longitudinal and lateral accelerations are treated as random variables, following a Gaussian distribution in stable highway driving (*18, 19*).

In real driving, human drivers perceive the driving risk by assessing the probability of a specified accident and how concerned they are with the possible consequences (*20, 21*), which is the same as the mechanism of PDRF. In this paper, we assume that human drivers' perception of other vehicles' motion such as the relative distance, velocity and acceleration is accurate but the motion uncertainties and behaviours deviate substantially from the real values and vary among drivers in different traffic conditions (*22–24*), leading to discrepancies between the objective risk and the perceived risk, and individual differences between drivers. Hence, we have several assumptions to extend the original PDRF as perceived probabilistic driving risk field (PPDRF) considering human drivers' risk perception:

- The longitudinal and lateral acceleration of the merging vehicle is independent of each other;
- The acceleration of merging vehicles follows Gaussian distribution with the real-time acceleration being the mean value of the distribution;
- The subject vehicle maintains the current state in a short prediction horizon;
- Different drivers estimate other road users' acceleration distribution differently, which can differ from the real statistics and vary among drivers;

- The risk gradient to a static object of human drivers varies among drivers.

The first three assumptions simplify other road users' motion and the last two consider human drivers' perception. Based on the discussions and the assumptions above, we can define the kinetic risk for PPDRF regarding moving objects as

$$R_{n,s}(t) = 0.5 M_s \beta^2 |\Delta v_{s,n}(t+\tau)|^2 \cdot \tilde{p}(n,s \mid t) \tag{3}$$

where $R_{n,s}(t)$ is the kinetic collision risk between the subject vehicle s and a neighbour vehicle n in Joules at time t. $\beta = \frac{M_n}{M_s+M_n}$ denotes the mass ratio. M_s and M_n are the mass of the subject vehicle and the neighbour vehicle. $\tilde{p}(n,s \mid t)$ is the collision probability to the neighbour vehicle estimated by drivers ranging on [0,1]. $\tilde{p}(n,s \mid t)$ is constructed as Equation (4).

$$\tilde{p}(n,s \mid t) = N\left(\frac{\Delta X(t) - \Delta V_X(t)\tau}{0.5\tau^2} \mid \mu_X(t), \tilde{\sigma}_X\right) \cdot N\left(\frac{\Delta Y(t) - \Delta V_Y(t)\tau}{0.5\tau^2} \mid \mu_Y(t), \tilde{\sigma}_Y\right) \tag{4}$$

where N is the collision probability density function that is assumed to be Gaussian distribution in this study. The acceleration distribution of the neighbour vehicle and the relative spacing between the subject vehicle and the neighbour vehicle can be visualised as Figure 1. Parameters $\mu_X(t)$ and $\mu_Y(t)$ denote the mean of the Gaussian distribution in longitudinal and lateral acceleration, which are set to the neighbour's real-time longitudinal and lateral acceleration at time t; $\tilde{\sigma}_X$ and $\tilde{\sigma}_Y$ are the respective standard deviations of the Gaussian distribution in the longitudinal and lateral directions, which are different from the real statistics and vary among drivers, representing drivers' different judgement of other road users' motion uncertainties. $\Delta X(t)$ and $\Delta Y(t)$ denote the relative spacing in X and Y direction between the subject vehicle and the neighbour vehicle; $\Delta V_X(t)$ and $\Delta V_Y(t)$ denote the longitudinal relative velocity and lateral relative velocity. τ is the prediction time horizon in second. $\tau = 3\,\mathrm{s}$ performs best in estimating the collision probability in rear-end conflict in stable highway driving up to 108 km/h (*15*). Note that the prediction horizon τ can vary among human drivers.

The potential risk taken by vehicle s due to a static object o can be modelled as

$$R_{o,s}(t) = 0.5 k M (V_{s,o}(t))^2 \cdot \max\left(e^{-|r_{s,o}|/D}, 0.001\right) \tag{5}$$

where $R_{o,s}(t)$ denotes the potential risk casued by the static object o; M denotes the mass of s; $|r_{s,o}|$ is the relative distance between the subject vehicle s and the non-moving object o; $V_{s,o}$ denotes the relative velocity; $0.5kM(V_{s,o})^2$ represents the expected crash energy scaled by the parameter k, with range $[0-1]$, which is set to 1 in this study representing the neighbour is immovable; the term $e^{-|r_{s,o}|/D}$ is the collision probability ranging between [0-1], where D determines the steepness of descent of the potential field.

Accordingly, at a given moment t, human drivers perceive a total risk consisting of kinetic risk and potential risk as follows

$$R(t) = R_{n,s}(t) + R_{o,s}(t) \tag{6}$$

$R(t)$ is an energy value that can reach 3×10^4 J in stable motorway driving (*15*). To compare the output with perceived risk ratings in specific ranges, $PPDRF(t)$ for moving objects can be normalised as Equation (7)

$$PPDRF(t) = C_{gain} \cdot R(t) / \max\left(R(t_0, t_{end})\right) \tag{7}$$

where $\max\left(R(t_0, t_{end})\right)$ represents the maximum $R_{n,s}(t)$ in a certain period from t_0 to t_{end}. C_{gain} depends on the range of perceived risk data, which is $0-10$ in this study.

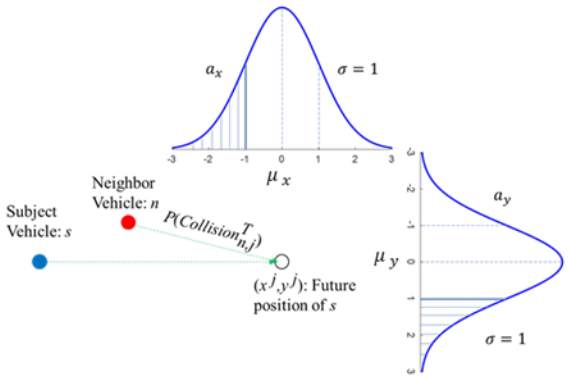


FIGURE 1: The acceleration probability distribution of neighbour vehicle

Driving risk field model (DRF)

DRF describes a 2D risk field around obstacles consisting of a 2-dimensional field to represent the driver's belief about the probability of an event occurring (probability field) and its consequence (severity field)(5), which are multiplied to provide an estimation of driver's perceived risk. The establishment of DRF is based on the following assumptions:

- Perceived risk is the product of the possibility of a hazardous event occurring estimated by drivers and the event severity;
- The risk field becomes wider as the longitudinal distance from the subject vehicle increases;
- The height of the risk field decays as the lateral and longitudinal distance from the vehicle increases;

The 2D probability field has a Gaussian cross-section laterally. The height of the Gaussian h and the width σ are separately modelled as a parabola and linear function of longitudinal distance x in front of the subject vehicle. Since there is only straight driving in this study, the probability field can be simplified as

$$p = h \cdot \exp\left(\frac{-y(t)^2}{2\sigma^2}\right) \tag{8}$$

$$h = s \cdot (x(t) - v(t) \cdot t_{la})^2 \tag{9}$$

$$\sigma = m \cdot x(t) + c \tag{10}$$

where the subject vehicle is at the original point (0,0); p is the probability of an event happening at position $(x(t), y(t))$; h and σ are the height and the width of the Gaussian at longitudinal position $x(t)$; s defines the steepness of the height parabola; t_{la} is the human driver's preview time (s); m defines the widening rate of the 2D probability field; c is the quarter width of the subject vehicle (m). At time t, $x(t)$ and $y(t)$ represent the longitudinal and lateral distance to the subject vehicle; $v(t)$ is the subject vehicle's velocity (m/s);

With the 2-D probability field, perceived risk can be computed if the event severity is defined. Hence, the severity field of the events in this study can be defined as

$$sev(t) = \begin{cases} 2500, & (x(t), y(t)) \in A^O, \\ 0, & (x(t), y(t)) \notin A^O. \end{cases} \tag{11}$$

where A^O represents a neighbour vehicle's spatial area. Note that the severity value 2500 can vary in different datasets.

According to the definition of DRF, the continuous perceived risk quantified by DRF can be modelled as

$$R_{DRF}(t) = \sum p(x(t), y(t)) \cdot sev(t) \tag{12}$$

In order to compare R_{DRF} with the scaled perceived risk data, $DRF(t)$ can be formatted as Equation (13).

$$DRF(t) = C_{gain} \cdot R_{DRF}(t) / \max\left(R_{DRF}\left(t_0, t_{\text{end}}\right)\right) \tag{13}$$

where $\max\left(R_{DRF}\left(t_0, t_{end}\right)\right)$ represents the maximum R_{DRF} in a certain period from t_0 to t_{end}; C_{gain} is 10 in this study.

Model properties

The risk field of three models for moving objects is detailed in Figure 2 based on different braking intensities ($-2\,\mathrm{m/s^2}$, $-5\,\mathrm{m/s^2}$ and $-8\,\mathrm{m/s^2}$) of the neighbour and Figure 3 shows the risk field for static objects. In this section, we qualitatively analyse and compare the properties of the three models based on the model formulation and the visualisation.

RPR is an observation-based model with a 1-D risk field using the relative gap to and braking intensity of the leading vehicle to compute perceived risk. Only when both variables are defined in the same lane, can RPR have output. RPR risk field has a width in Figure 2 and 3 since we considered the dimension of the vehicles or obstacles but RPR is still 1-D because the lateral threats are not included. With the braking intensity increasing, RPR has a larger risk area in front of the subject vehicle (Figure 2 (a-c)), meaning that human drivers perceived more risk of the object at the same position but with a stronger brake.

PPDRF is a mechanistic model with a 2-D risk field as it considers other road users' motion in two directions simultaneously and the potential risk in 2-D of the fixed objects. The shape of the risk field is different between moving and non-moving objects since the risk field is generated from the kinetic risk component in Figure 2(d-f)) but from the potential risk component in Figure 3 (b).

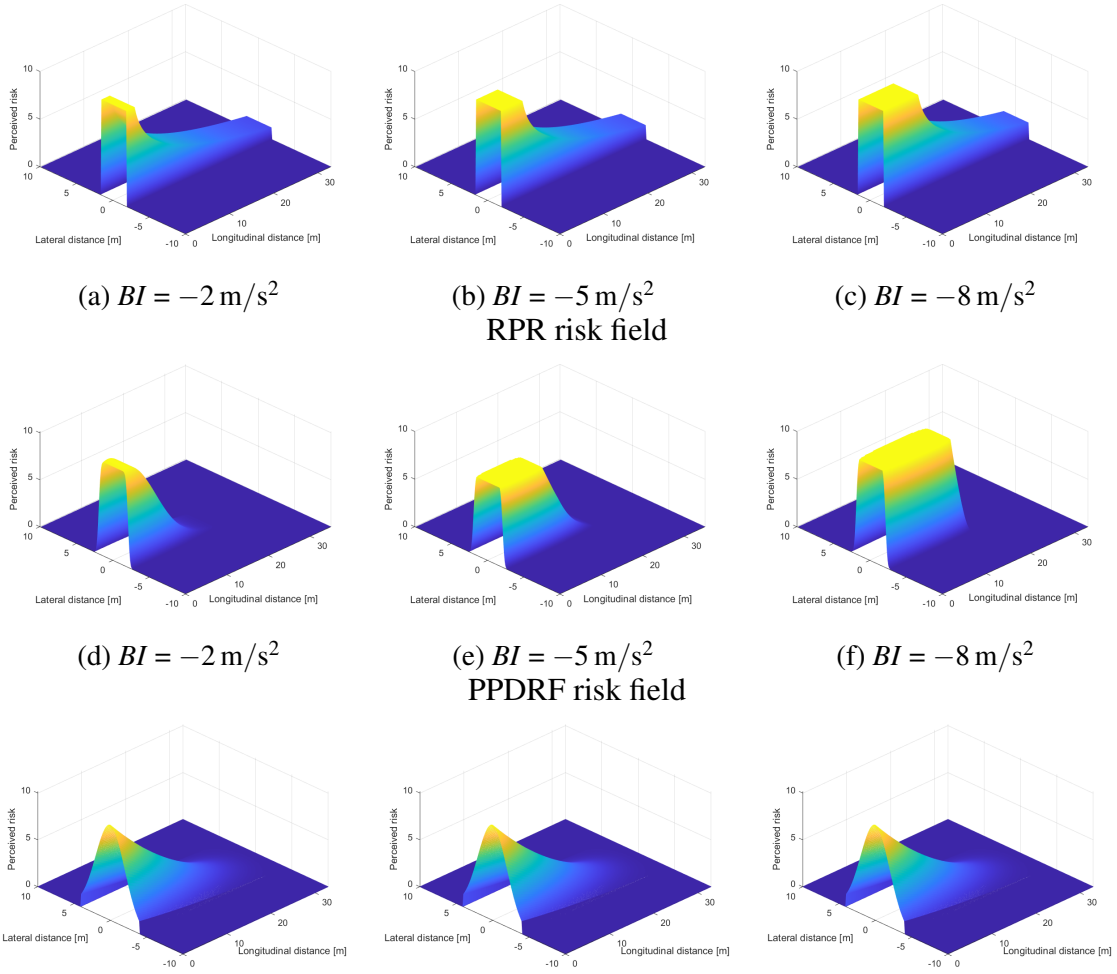
DRF is an observation-based model and it also has a 2-D risk field as its probability field and severity field are both 2-D. The risk field does not change with different braking intensities (Figure 2(g-i)) since no motion information of other road users is considered in the model. The risk area in front of the subject vehicle is larger with a longer preview time t_{la} of the driver and a higher subject vehicle's velocity.

Both PPDRF and DRF use the risk field theory to compute the driver's perceived risk but their formulation is different. For the non-moving objects, both PPDRF and DRF compute perceived risk based on a probability field and a severity field in 2-D. PPDRF considers the relative velocity to the object, changing the shape of the probability field and the severity field in real time but DRF's risk field is fixed once the parameters are pre-calibrated based on certain datasets. For other road users, PPDRF computes kinetic perceived risk based on motion prediction but DRF lacks the kinetic risk part. In other words, DRF is a static model heavily dependent on pre-known knowledge but PPDRF can adapt to different scenarios and datasets easier.

Based on the discussion above, the model features are summarised as follows:

- RPR: A regression model in highway driving reacting to merging and hard braking vehicles.
  - Attributes: Observation-based

- Dimension: 1-D
- PPDRF: A probabilistic driving risk assessment approach based on risk field theory considering collision probability and consequences
  - Attributes: Mechanism-based
  - Dimension: 2-D
- DRF: A perceived risk model consisting of a risk field and a severity field of consequences based on risk field theory
  - Attributes: Observation-based
  - Dimension: 2-D

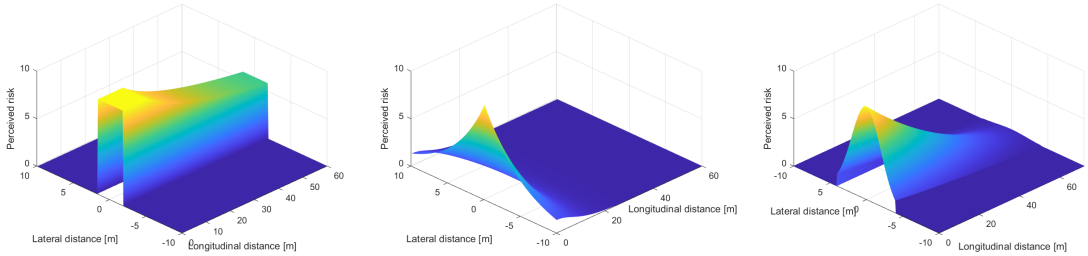


(a) $BI = -2\,\mathrm{m/s^2}$ (b) $BI = -5\,\mathrm{m/s^2}$ (c) $BI = -8\,\mathrm{m/s^2}$
RPR risk field

(d) $BI = -2\,\mathrm{m/s^2}$ (e) $BI = -5\,\mathrm{m/s^2}$ (f) $BI = -8\,\mathrm{m/s^2}$
PPDRF risk field

(g) $BI = -2\,\mathrm{m/s^2}$ (h) $BI = -5\,\mathrm{m/s^2}$ (i) $BI = -8\,\mathrm{m/s^2}$
DRF risk field

FIGURE 2: Risk field of human driver's longitudinal and lateral risk perception to traffic objects on a 2-D plane (with the subject and merging vehicle's velocity 27m/s). The subject vehicle is at the original point. Each row shows the risk field of one model with different braking intensity of leading vehicle. Model parameters are calibrated in Section Model calibration.



(a) RPR risk field (b) PPDRF risk field (c) DRF risk field

FIGURE 3: Risk field of human driver's longitudinal and lateral risk perception to traffic objects on a 2-D plane (with fixed obstacles and the subject velocity 25m/s). The subject vehicle is at the original point. Model parameters are calibrated in Section Model calibration.

MODEL CALIBRATION AND PERFORMANCE INDICES

In this section, the datasets, model calibration methods, and performance indices are introduced.

Dataset introduction

Two datasets from literature are used for the model calibration and validation in this study.

The first dataset (**Dataset Merging**) (*10*) is collected from our previous simulator experiment where 18 merging events with different merging distance and braking intensity on a 2-lane highway were simulated. Figure 4 shows an example of the simulated events during the experiment. The participants were asked to monitor the scenario as fall-back ready drivers for an SAE Level 2 driving automation. The participant used a pressure sensor to give perceived risk ratings from 0-10 continuously in the time domain (see the lower row in Figure 4), which are the continuous perceived risk data. After each event, the participants were also asked to give a verbal perceived risk rating from 0-10 regarding the previous event, which is the discrete event-based perceived risk data. The corresponding kinematic data (e.g. position, speed and acceleration of the subject vehicle and neighbouring vehicles) were collected in the meantime. It has been shown that the peak of continuous perceived risk in a specific event should not be much different from the corresponding verbal rating (*10*). With this criterion, the perceived risk data of 220 merging events (30 s of each event) with corresponding kinetic data were kept.

The second dataset (**Dataset Obstacle Avoidance**) (*16*) includes drivers' verbal perceived risk ratings and steering angle signals when the participants face static obstacles suddenly appearing in front with the subject vehicle's velocity of 25m/s in manual driving mode. The corresponding vehicle kinematic data and the position of the obstacles were recorded at the same time.

In model calibration, for Dataset Merging, we use the **event-based perceived risk** and the **peak of the continuous perceived risk** in a certain event as the references; For Dataset Obstacle Avoidance, we use the **event-based perceived risk**, and the **steering wheel angle peak** in a certain event as the references.

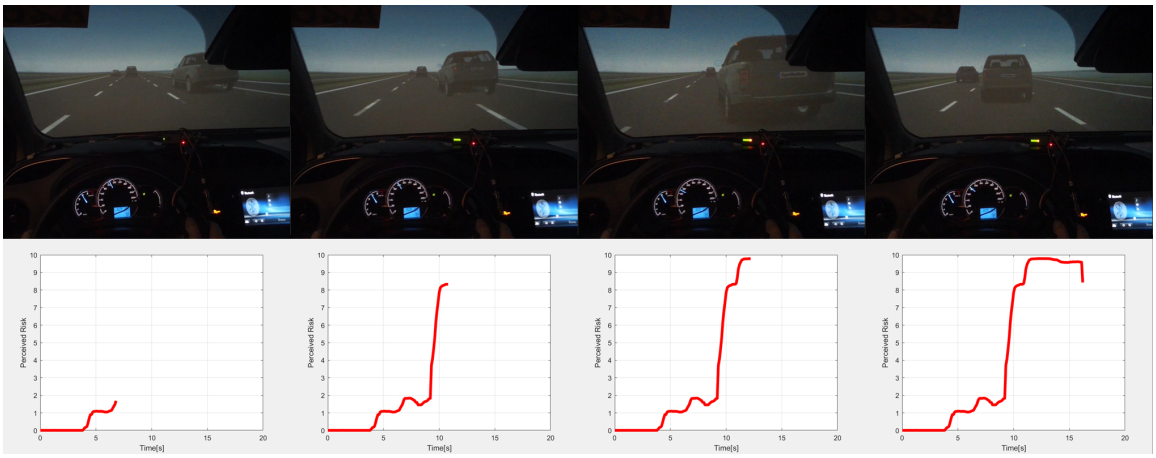


FIGURE 4: The upper row: Video stream of a merging with hard braking event simulated in the experiment. The lower row: Corresponding perceived risk values indicated by a participant with the pressure sensor.

Model calibration

Although we are developing general models considering the average characteristics of all participants, the influences of group features and scenarios cannot be ignored. Therefore, a dataset-level calibration of parameters is considered for all models to get the best performance. Here, we have $RMSE_i$ defined as

$$RMSE_i = \sqrt{\frac{\sum_{n=1}^{N} (\hat{y}_n - y_n)^2}{N}} \tag{14}$$

where $RMSE_i$ represents the root mean square error between the collected perceived risk data and the model output. For Dataset Merging, $i = event$ and $i = peak$ represent the $RMSE$ for event-based perceived risk and the peak of continuous perceived risk respectively; For Dataset Obstacle Avoidance, $i = event$ and $i = peak$ represents the $RMSE$ for event-based perceived risk and the maximum steering wheel angle separately. $\hat{y}_n$ is the model output; y_n is the referred perceived risk data. N is the number of events in different datasets with $N = 220$ for Dataset Merging and $N = 2496$ for Dataset Obstacle Avoidance.

The calibration tries to minimise $\sum RMSE_i$ for all models by tuning the key model parameters based on perceived risk and corresponding kinematic data. It should be pointed out that the first sample point of the kinematic data where the obstacle suddenly appears in Dataset Obstacle Avoidance is used for the calibration since the participants were asked to give a verbal perceived risk rating as soon as the obstacle appears.

Performance indicators

We use five indicators to evaluate the model performance: Prediction accuracy, Detection rate, Linear time scaling factor, Dynamic time warping and Time efficiency.

Prediction Accuracy

We use Root mean squared error (RMSE) to quantify the model Prediction accuracy, which is the same as the model calibration criterion. This indicator reflects the model's ability to compute the overall perceived risk in a certain event. A model with a smaller $RMSE$ can predict the overall perceived risk for a certain event more accurately.

Detection rate

The models cannot detect all dangerous events due to model limitations. The Detection rate represents the model's ability to detect a risky event that is also perceived as dangerous by human drivers. We defined detection rate as Equation (15)

$$R_{det} = \frac{N_{detected}}{N_{event}} \times 100\% \tag{15}$$

where $N_{detected}$ represents the number of events where the model manages to detect the risk with non-zero output; N_{event} is the total number of the events where human drivers gave perceived risk ratings in a certain dataset with $N_{event} = 220$ for Dataset Merging and $N_{event} = 2496$ for Dataset Obstacle Avoidance. A model with a higher detection rate can recognise more events that are also perceived as dangerous by human drivers.

Linear time scaling factor

The perceived risk signal has a time duration in certain events. Taking the continuous perceived risk signal as the reference, we can linearly scale the duration of the model output by a time factor to fit the duration of the continuous perceived risk. The linear time scaling factor (LTSF) can be defined as Equation (16).

$$LTSF = \frac{t_{10\%model}}{t_{10\%PR}} = \frac{n_{10\%model}}{n_{10\%PR}} \tag{16}$$

where $t_{10\%PR}$ is the duration when continuous perceived risk has more than 10 % of the peak in specific events; $t_{10\%model}$ is the duration when model has more than 10 % of the output peak. Equally, we can use the number of samples to define LTSF. $n_{10\%PR}$ is the number of sample points of continuous perceived risk, which is more than 10 % of the signal peak in specific events; $n_{10\%model}$ is the number of sample points of model output, which is more than 10% of the model output peak. Figure 5 shows the original two signals and the aligned signals by LTSF, where LTSF is larger than 1. LTSF reflects the model output range in the time domain. An LTSF that is closer to 1 indicates that the model can cover the whole process of human drivers' risk perception better.

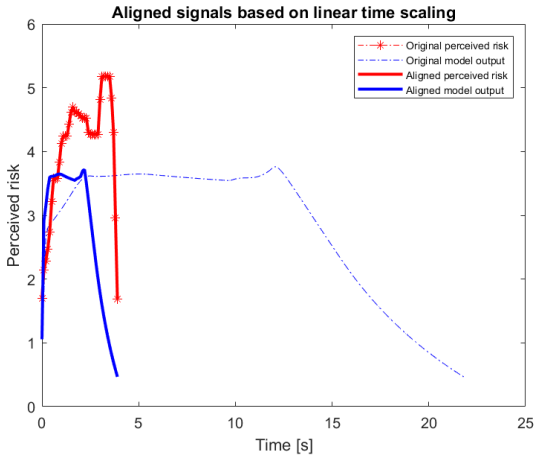


FIGURE 5: An example of linear time scaling where the duration of the original model output is longer than original perceived risk. The model output is compressed by a linear time scaling factor that is larger than 1.

Dynamic time warping (DTW) distance

As the duration of the model output and continuous perceived risk is different from each other, we cannot directly calculate the similarity between them. Hence, Dynamic time warping (DTW) algorithm is employed to quantify the similarity. There are two time series $\mathbf{X}$ and $\mathbf{Y}$ having M and N sampling points respectively. Define the distance matrix between the m th sample point of $\mathbf{X}$ and the n th sample point of $\mathbf{Y}$ as a Euclidean distance $d_{mn}(\mathbf{X},\mathbf{Y})$:

$$d_{mn}(\mathbf{X},\mathbf{Y}) = |x_m - y_n| \tag{17}$$

The DTW searches two path sequences ix and iy (of the same length) in the distance matrix d_{mn}, (m = 1,2,...,M, n = 1,2,...,N), which minimises the sum of the distances between the sampling points along the distance path. Figure 6 shows a pair of aligned signals by using DTW. Note that DTW is to align two signals to the same length and DTW distance finally represents the similarity of two signals with different duration. Two signals are more similar with a smaller DTW distance.

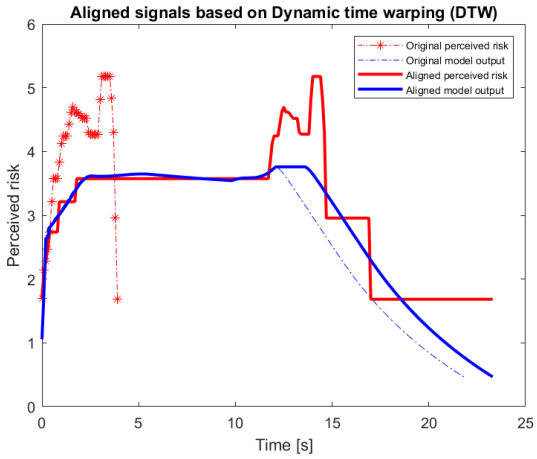


FIGURE 6: An example of Dynamic time warping. DTW align them by searching two path sequences to minimise distance sum. Both aligned signals have different time duration from the original signals.

Time efficiency

All the models should be used in real-time perceived risk computation so Time efficiency is critical. More complex models may have a better performance in Prediction accuracy but tend to take a longer time. We define Time Efficiency as the model's computation time cost per data sample. If the time consumption exceeds the computation time step, it means that the computation of perceived risk cannot be finished in real-time.

RESULTS

In this section, we illustrate the applicability of the three models and evaluate their performance with the performance indicators introduced previously regarding the two datasets.

Model calibrations and results

As the discussion above, we calibrate all models to get the best performance. According to the model structure and the dataset feature, the parameters to be calibrated are listed.

- RPR: C_0, C_1,C_2

- PPDRF: $\tilde{\sigma}_x$, $\tilde{\sigma}_y$ (Only for Dataset Merging); D (Only for Dataset Obstacle Avoidance)
- DRF: p, m, and t_{la}(only for Dataset Obstacle Avoidance)

Note that we re-scale the perceived risk linearly to 0-10 in Dataset Obstacle Avoidance to match the range in Dataset Merging. The calibration is conducted separately regarding two different datasets and Table 1 shows the calibration results for different models.

TABLE 1: Calibrated parameters for all models

Model	Parameters	Explanation	Values for Dataset Merging	Values for Dataset Obstacle Avoidance
RPR	C_0	The intercept in the regression model	11.30	16.30
	C_1	The coefficient of gap to the leading vehicle	-3.70	-2.50
	C_2	The coefficient of leading vehicle's braking intensity	-0.36	0
PPDRF	$\tilde{\sigma}_x$	The standard deviation of longitudinal acceleration distribution of neighbour vehicle	0.56	/
	$\tilde{\sigma}_y$	The standard deviation of lateral acceleration distribution of neighbour vehicle	0.11	/
	D	The steepness of descent of the potential field	/	0.46
DRF	p	The steepness of the height parabola of the risk field	0.15	0.15
	t_{la}	Human driver's preview time	/	4.87
	m	The rate of the risk field width expanding	3.98×10^{-8}	5.66×10^{-4}

RPR has three key parameters that are relevant to the initial risk, the gap to and the braking intensity of the leading vehicle. The participants have a positive initial perceived risk level and perceive more risk with a smaller gap to the object based on the signs of C_0 and C_1. According to C_2 for Dataset Merging, a stronger brake of the leading vehicle causes higher risk. Note that C_2 is 0 for Dataset Obstacle Avoidance since the objects in the experiment are static.

The calibrated $\tilde{\sigma}_x$ and $\tilde{\sigma}_y$ of the Gaussian distribution in PPDRF reflect the human drivers' estimation of other road users' motion uncertainties so the value may differ from the actual statistics. Only $\tilde{\sigma}_x$ and $\tilde{\sigma}_y$ are calibrated for Dataset Merging and only D for Dataset Obstacle Avoidance because the computation of potential risk and kinetic risk of PPDRF is different.

The parameter s of DRF is highly affected by the value in the severity field. The calibrated value is different from that in Kolekar et al. (*5*) since we have different perceived risk scales and considered more participants.

All three models' parameters are validated to be stable with sub-datasets covering 80% and 50% of all the events. We can conclude that the calibration is not easily affected by the data noise and we do improve the model performance by calibrating the parameters.

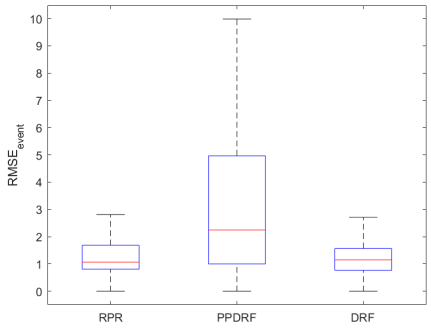
Quantitative analysis results

We use both datasets including the perceived risk data and the corresponding kinematic data to test the three models with the calibrated parameters in Table 1 and different aspects of performance are represented by the following sections.

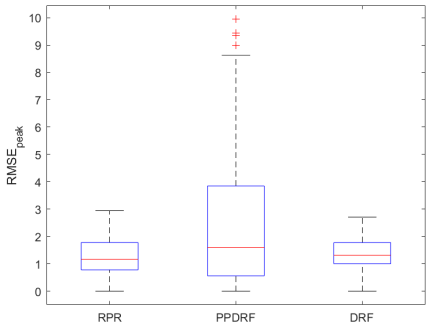
Prediction accuracy

As mentioned in Section Prediction accuracy, *RMSE* reflects the overall Prediction accuracy of the models in certain events where the continuous error in the time domain is not considered. Table 2 shows the *RMSE* for all three models regarding two different datasets. The *RMSE* distribution across all the events of the dataset is shown in Figure 7.

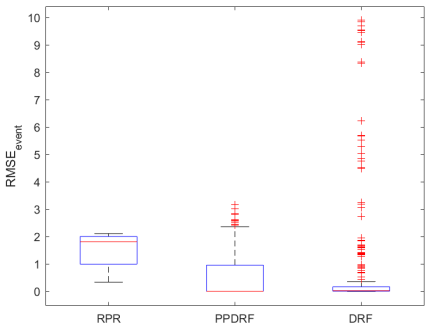
Two kinds of *RMSE* show that RPR and DRF perform better in Dataset Merging. A possible reason is that they are originally built for perceived risk computation based on observations but PPDRF is adapted from PDRF that is formulated to assess actual collision risk in traffic. In Dataset Obstacle Avoidance, $RMSE_{event}$ shows PPDRF and DRF have better performance since they consider lateral perceived risk as well, leading to a higher prediction accuracy regarding a 2-D dataset (Table 2). In Dataset Obstacle Avoidance, more than 80% of participants' verbal ratings are 0 but the steering response to the objects that suddenly appear is always non-zero. Hence, models in Figure 7(c) have more outliers than in Figure 7(d). Additionally, the risk field of PPDRF is steeper than DRF in the longitudinal direction because Exponential and Quadratic functions are used respectively, making DRF more sensitive than PPDRF in further areas regarding static objects. Consequently, DRF has more non-zero outputs in further areas, which can be outliers if the majority of the ratings is 0.



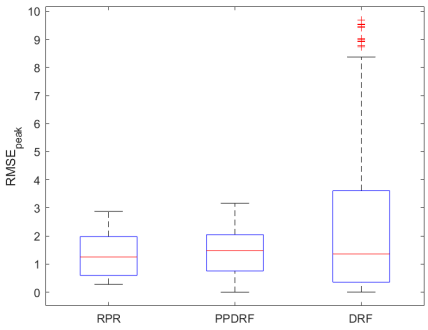
(a) $RMSE_{event}$ Dataset Merging



(b) $RMSE_{peak}$ for Dataset Merging



(c) $RMSE_{event}$ for Dataset Obstacle Avoidance



(d) $RMSE_{peak}$ for Dataset Obstacle Avoidance

FIGURE 7: *RMSE* distribution for all models in two datasets. Bars present RMSE at 5, 25, 50, 75, 95 percentile. The red '+' represent outliers. The same below.

Detection rate

According to Equation 15, the detection rate for three models regarding two datasets is shown in Table 2. In Dataset Merging, the merging vehicle causes only longitudinal risk in the same lane. Hence, all models can detect dangerous events no matter whether the model is 1-D or 2D. However, in Dataset Obstacle Avoidance, the obstacles were broadly distributed on a 2-D plane. Only the models that can capture lateral risk have the ability to detect dangerous objects outside the forward path. Hence, RPR has a low detection rate and PPDRF and DRF can recognise all dangerous events that are also perceived as risky by human drivers.

Linear time scaling factor

LTSF reflects the model's ability to cover human drivers' whole risk perception process in certain events. In this study, we only consider LTSF for Dataset Merging where the continuous perceived risk data is included. Figure 8 shows the LTSF distribution for the three models and the average is listed in Table 2. Generally speaking, the model can fit drivers' risk perception process better if LTSF is closer to 1. LTSF of RPR and DRF is larger than 1 indicating that the two models can cover the whole risk perception process of human drivers but can cause false positive errors. However, LTSF for PPDRF is smaller than 1, which means the model output comes later and ends earlier than human drivers' perceived risk, causing the false negative error. Note that we can tune the parameters of PPDRF to cover the whole risk perception process but that means the improvement of LTSF may be at the expense of other aspects of performance.

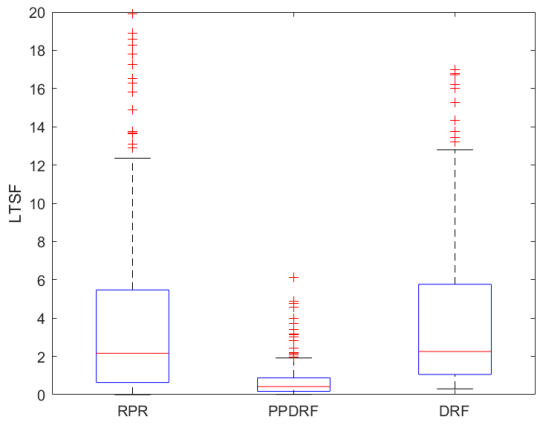


FIGURE 8: Linear time scaling factor distribution

Dynamic time warping (DTW) distance

DTW distance describes the similarity between the continuous perceived risk and the model output. Similar to LTSF, we only consider DTW for Dataset Merging because only this dataset contains continuous perceived risk data. According to Section Dynamic time warping (DTW) distance and Equation (17), the DTW distance is shown in Table 2 and the distribution is shown in Figure 9.

Generally speaking, two signals are more similar if DTW distance is smaller. 2 and Figure 9 show that PPDRF has the best performance indicating that the output of PPDRF can fit the curve of perceived risk best. The calibration of model parameters can improve the model performance but can hardly change the shape of the output curve. That means the mechanism of PPDRF reveals the pattern of human drivers' risk perception to some extent, which will be discussed later.

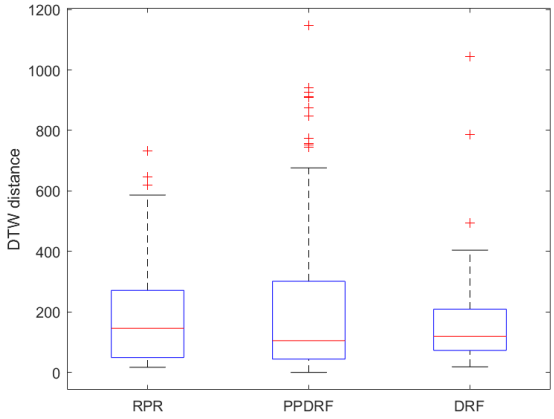


FIGURE 9: DTW distance distribution for five models

Time efficiency

For Time efficiency comparison, the simulation for different models should be conducted with the same device, which is a workstation with Intel Core i7-8665U 1.9Ghz and 8GB RAM. The computation time for different models is shown in Table 2. In general, models that consider more complicated situations took more time for computation. In both datasets, RPR took the least time since it only involves logarithmic calculation. In Dataset Merging, PPDRF is the most time-consuming model as it requires spatial overlap computations and multiple integrals in the overlap area. Although DRF also has multiple integrals, the overlap computations are much easier than PPDRF since the risk field and severity field are static and there is no motion prediction of neighbours. In Dataset Obstacle Avoidance, PPDRF took less time than DRF as PPDRF only computes potential risk, which is much simpler than the kinetic risk computation in Dataset Merging.

TABLE 2: Model performance represented by the performance indicators. The results are averaged values across all events that are different from the 50 percentile red dash lines in Figure 7-Figure 9

		RPR	PPDRF	DRF	p (K-W test)
Dataset Merging	$RMSE_{event}$	2.36	3.22	2.39	0.000
	$RMSE_{peak}$	2.61	3.37	2.70	0.000
	Detection rate	1.00	1.00	1.00	1.000
	DTW distance	145.72	105.08	118.97	0.001
	LTSF	2.16	0.42	2.25	0.000
	Time consumption (ms)*	0.0003	7.6292	0.4626	0.000
Dataset Obstacle Avoidance	$RMSE_{event}$	3.07	2.40	2.41	0.000
	$RMSE_{peak}$	3.31	3.61	3.42	0.000
	Detection rate	0.10**	1.00	1.00	1.000
	DTW distance	/	/	/	/
	LTSF	/	/	/	/
	Time consumption (ms) ***	0.0002	0.0141	0.5304	0.000

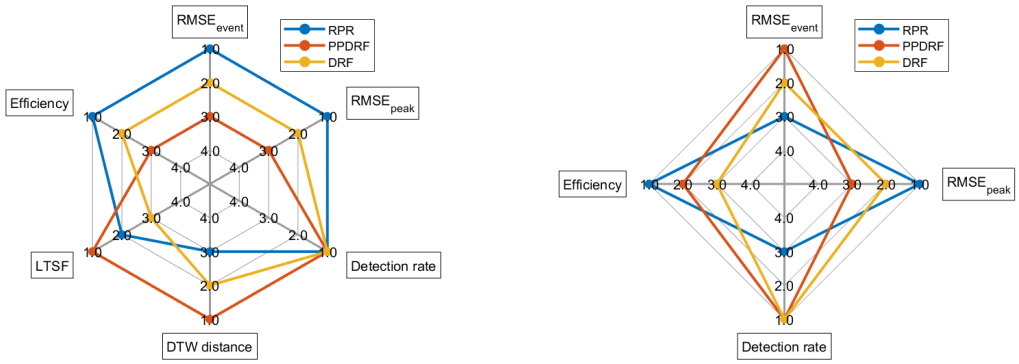
∗ The average value of computing 66220 samples.

∗∗ Only the objects directly in front of the vehicle can be detected by RPR, which leads to a low detection rate. See Kolekar et al. (*16*) for more experiment details

∗ ∗ ∗ The average value of computing 349440 samples.

Summary of quantitative model performance

According to the results presented in this section, we use radar charts to show the model performance in various aspects (Figure 10). Generally speaking, RPR has the best performance in overall prediction accuracy but PPDRF can fit the risk perception process best according to the DTW distance and LTSF performance in Dataset Merging. In Dataset Obstacle Avoidance, PPDRF and DRF have the advantages in Detection rate because they can capture lateral risk.



(a) Rankings for Dataset Merging (b) Rankings for Dataset Obstacle Avoidance

FIGURE 10: Rankings of models by performance indicators

DISCUSSIONS

Model mechanism

RPR and DRF are phenomenological models derived from observations, which means that the models depend on the scenarios and corresponding data heavily. They are originally built based on Dataset Merging and Dataset Obstacle Avoidance respectively and that is the possible reason why they compute perceived risk accurately in corresponding datasets. However, their parameters have to be carefully calibrated to get better performance for different scenarios and datasets. Although DRF is an observation-based model, it is still in the framework of risk field theory.

PPDRF is established based on the artificial field theory considering the motion of other road users and static obstacles so we treat it as a mechanism-based model. PDRF performs well in estimating actual collision risk (*15*) and PPDRF can easily compute perceived risk after calibration regarding the two datasets. We can conclude that PPDRF has the potential to be easily extended to fit different scenarios after calibrations if given the suitable probabilistic assumption of other traffic participants' motions. Mechanistic models are also helpful to explain the mechanism of human drivers' risk perception.

The relation between modelling and human driver's risk perception

The three models presented in this study can compute perceived risk accurately in two datasets, indicating that these models are promising to explain perceived risk.

The position of objects in interaction is vital for human drivers' perceived risk. All three models consider objects' position information. Specifically, RPR uses the gap to the leading vehicle as the position information in 1-D; PPDRF and DRF directly use objects' coordinates in 2-D as the position information. In other words, we have to know which object has threats and where it is

in dynamic driving, which is different from the concept studies where perceived risk is generally described as a feeling or a climate without any specific objects in conflict (*8, 25*).

The motion of the objects in interaction is essential to compute perceived risk, which is used by human drivers to predict the potential collision in a certain prediction horizon. RPR uses the braking intensity of the leading vehicle to predict the motion. The stronger the brake intensity is, the more likely the collision happens. PPDRF uses the relative velocity and acceleration to predict the position of other road users, which performs better in 2-D. DRF does not directly use the relative motion but the preview time provides a prediction horizon for human drivers.

Another essential part of perceived risk is the sequence of the potential events in the future. PPDRF and DRF prove that human drivers perceive risk by estimating not only the probability of the potential collision but also the severity of the collision, which corresponds to the results of Ping et al., Näätänen and Summala (*9, 26*). In the process of DTW, the duration of the model output is rescaled but the shape can be preserved, which means PPDRF provides the best shape of the perceived risk field. This is extra proof of this theory.

We can find several features of perceived risk according to the three models. In the longitudinal direction, human drivers perceive higher risk when driving closer to the objects in interaction and the increase is non-linear where perceived risk sharply in the near area. Logarithm, Gaussian, Exponential and Quadratic functions are used to describe this non-linear relation but further research is needed to find the best shape. In the lateral direction, the risk field's cross-section of PPDRF and DRF is coincidentally Gaussian, indicating that Gaussian could be a suitable function to describe the change of lateral perceived risk.

The detection rate of the three models proves that perceived risk is actually 2-D, but existing studies seldom discuss the underlying mechanism. We discuss several possible reasons to underpin the 2-D nature of perceived risk. Firstly, other road users may collide with the subject vehicle from the lateral direction or even any direction. This kind of collision can dramatically widen the risk field. Secondly, human drivers may have estimation errors of the subject vehicle's dimension, which can influence the width of the risk field in the nearby area. Thirdly, it is difficult for human drivers to judge whether an object is on the collision path or not, widening the risk field in a further area. Lastly, the motion uncertainties of other road users and the subject vehicle may widen the risk field. The latter uncertainty is relevant to human drivers' driving skills or the trust in driving automation.

Limitations

We use two datasets to calibrate and validate the models. Dataset Merging covers human drivers perceived risk data in highway automated driving, but the lateral risk is lacking. Dataset Obstacle Avoidance contains human drivers perceived risk data in 2-D, providing more information on lateral perceived risk. However, this dataset is collected from manual driving, which may cause bias in automated driving studies. Additionally, the objects in the experiment are fixed and suddenly displayed during the driving. The additional perceived risk caused by surprise cannot be neglected.

CONCLUSIONS AND FUTURE WORK

This study formulated or adapted three models to compute perceived risk in two datasets. All three models were calibrated and compared regarding their performance. Based on the simulation results and the analysis, we reach the following findings: (1) Perceived risk is 2-D coming from both longitudinal and lateral directions and increases non-linearly with the decrease of the distance to

surrounding vehicles; (2) It is proved that human drivers perceive risk by estimating the probability of potential collision and its severity; (3) Gaussian is suitable to describe the lateral change of perceived risk; (4) The observation-based model can compute perceived risk accurately regarding specific datasets but the extension to more scenarios and datasets is limited; (5) PPDRF has the potential to compute perceived risk accurately and explain its mechanism; (6) The prediction of other road users' motion is vital to estimate perceived risk.

In the next steps, we will collect more perceived risk data in 2-D to improve and validate the models. Specifically, we will add lateral perceived risk to RPR by calculating the safety metrics in 2-D; PPDRF will be improved by more accurate motion prediction of other road users considering different acceleration distributions or path planning.

ACKNOWLEDGEMENTS

The authors would like to thank Toyota Motor Europe NV/SA for the support for the simulator experiment. This research is also supported by the SHAPE-IT project funded by the European Union's Horizon 2020 research and innovation programme under the Marie Skłodowska-Curie grant agreement 860410.

REFERENCES

1. WHO, *Road traffic injuries*, 2020.
2. Nadimi, N., H. Behbahani, and H. R. Shahbazi, Calibration and validation of a new time-based surrogate safety measure using fuzzy inference system. *Journal of Traffic and Transportation Engineering (English Edition)*, Vol. 3, No. 1, 2016, pp. 51–58.
3. Eboli, L., G. Mazzulla, and G. Pungillo, How to define the accident risk level of car drivers by combining objective and subjective measures of driving style. *Transportation Research Part F: Traffic Psychology and Behaviour*, Vol. 49, 2017, pp. 29–38.
4. Griffin, W., N. Haworth, and D. Twisk, Patterns in perceived crash risk among male and female drivers with and without substantial cycling experience. *Transportation Research Part F: Traffic Psychology and Behaviour*, Vol. 69, 2020, pp. 1–12.
5. Kolekar, S., J. De Winter, and D. Abbink, Human-like driving behaviour emerges from a risk-based driver model. *Nature Communications*, Vol. 11, No. September, 2020, pp. 1–19.
6. Summala, H., Risk control is not risk adjustment: The zero-risk theory of driver behaviour and its implications. *Ergonomics*, Vol. 31, No. 4, 1988, pp. 491–506.
7. Kolekar, S., B. Petermeijer, E. Boer, J. de Winter, and D. Abbink, A risk field-based metric correlates with driver ' s perceived risk in manual and automated driving : A test-track study. *Transportation Research Part C*, Vol. 133, No. November, 2021, p. 103428.
8. Xu, Z., K. Zhang, H. Min, Z. Wang, X. Zhao, and P. Liu, What drives people to accept automated vehicles? Findings from a field experiment. *Transportation Research Part C: Emerging Technologies*, Vol. 95, No. February, 2018, pp. 320–334.
9. Ping, P., Y. Sheng, W. Qin, C. Miyajima, and K. Takeda, Modeling Driver Risk Perception on City Roads Using Deep Learning. *IEEE Access*, Vol. 6, 2018, pp. 68850–68866.
10. He, X., J. Stapel, M. Wang, and R. Happee, Modelling perceived risk and trust in driving automation reacting to merging and braking vehicles. *Transportation Research Part F: Psychology and Behaviour*, Vol. 86, No. February, 2022, pp. 178–195.
11. Kiefer, R. J., D. J. Leblanc, and C. A. Flannagan, Developing an inverse time-to-collision crash alert timing approach based on drivers' last-second braking and steering judgments. *Accident Analysis and Prevention*, Vol. 37, No. 2, 2005, pp. 295–303.
12. Lee, D., A theory of visual control of braking based on information about time-to- collision. *Perception*, Vol. 5, 1976, pp. 437–459.
13. Wang, J., J. Wu, X. Zheng, D. Ni, and K. Li, Driving safety field theory modeling and its application in pre-collision warning system. *Transportation Research Part C: Emerging Technologies*, Vol. 72, 2016, pp. 306–324.
14. Li, L., J. Gan, Z. Yi, X. Qu, and B. Ran, Risk perception and the warning strategy based on safety potential field theory. *Accident Analysis and Prevention*, Vol. 148, No. December, 2020.
15. Mullakkal-Babu, F. A., M. Wang, X. He, B. van Arem, and R. Happee, Probabilistic field approach for motorway driving risk assessment. *Transportation Research Part C: Emerging Technologies*, Vol. 118, No. July, 2020, p. 102716.
16. Kolekar, S., J. de Winter, and D. Abbink, Which parts of the road guide obstacle avoidance? Quantifying the driver's risk field. *Applied Ergonomics*, Vol. 89, No. July, 2020, p. 103196.
17. Ni, D., A Unified Perspective on Traffic Flow Theory Part I : The Field Theory. *Applied Mathematical Sciences*, Vol. 7, No. 39, 2013, pp. 1929–1946.

18. Wagner, P., R. Nippold, S. Gabloner, and M. Margreiter, Analyzing human driving data an approach motivated by data science methods. *Chaos, Solitons and Fractals*, Vol. 90, 2016, pp. 37–45.
19. Ko, J., R. Guensler, and M. Hunter, Analysis of effects of driver/vehicle characteristics on acceleration noise using GPS-equipped vehicles. *Transportation Research Part F: Traffic Psychology and Behaviour*, Vol. 13, No. 1, 2010, pp. 21–31.
20. Brown, I. D. and J. A. Groeger, Risk perception and decision taking during the transition between novice and experienced driver status. *Ergonomics*, Vol. 31, No. 4, 1988, pp. 585–597.
21. Sjöberg, L., E. Moen, and T. Rundmo, *Explaining risk perception. An evaluation of the psychometric paradigm in risk perception research*. Norwegian University of Science and Technology, 2004.
22. Gilinsky, a., *Gilinsky–Perceived Range and Distance in Visual Space.pdf*, 1951.
23. Hoffmann, E. R. and R. G. Mortimer, Scaling of relative velocity between vehicles. *Accident Analysis and Prevention*, Vol. 28, No. 4, 1996, pp. 415–421.
24. Saffarian, M., J. C. De Winter, and J. W. Senders, Measuring drivers' visual information needs during braking: A simulator study using a screen-occlusion method. *Transportation Research Part F: Traffic Psychology and Behaviour*, Vol. 33, 2015, pp. 48–65.
25. Osswald, S., D. Wurhofer, S. Trösterer, E. Beck, and M. Tscheligi, Predicting information technology usage in the car: Towards a car technology acceptance model. In *Proceedings of the 4th International Conference on Automotive User Interfaces and Interactive Vehicular Applications*, 2012, pp. 51–58.
26. Näätänen, R. and H. Summala, Road-user behaviour and traffic accidents. *Publication of: North-Holland Publishing Company*, 1976.