A risk field-based metric correlates with driver's perceived risk in manual and automated driving
A test-track study
Kolekar, Sarvesh; Petermeijer, Bastiaan; Boer, Erwin; de Winter, Joost; Abbink, David

DOI
10.1016/j.trc.2021.103428

Publication date
2021

Document Version
Final published version

Published in
Transportation Research Part C: Emerging Technologies

Citation (APA)

Important note
To cite this publication, please use the final published version (if applicable).
Please check the document version above.
A risk field-based metric correlates with driver’s perceived risk in manual and automated driving: A test-track study

Sarvesh Kolekar*, Bastiaan Petermeijer, Erwin Boer, Joost de Winter, David Abbink

Department of Cognitive Robotics, Faculty of Mechanical, Maritime and Materials Engineering (3mE), Delft University of Technology, The Netherlands

Abstract

Quantifying drivers’ perceived risk is important in the design and evaluation of the behaviour of automated vehicles (AVs) and in predicting takeovers by the driver. A ‘Driver’s Risk Field’ (DRF) function has been previously shown to be able to predict manual driving behaviour in several simulated scenarios. In this paper, we tested if the DRF-based risk estimate ($\hat{r}$) could predict manual driving behaviour and the driver’s perceived risk during automated driving. To ensure that the participants perceived realistic levels of risk, the experiment was conducted in a test vehicle. Eight participants drove five laps manually and experienced 12 different laps of automated driving on a test track. The test track consisted of three sections (which were sub-divided into 12 sectors): curve driving (9 sectors), parked car (1 sector), and 90-degree intersections (2 sectors). If the driver verbally expressed risk or performed a takeover, that particular sector was labelled as risky. The results show that the DRF risk estimate ($\hat{r}$) predicted manual driving behaviour ($\rho_{\text{steering}} = 0.69$, $\rho_{\text{speed}} = 0.64$), as well as correlated with the driver’s perceived risk in curve driving ($r^2 = 0.98$) and while negotiating a car parked outside the lane boundary ($r^2 = 0.59$). In conclusion, the DRF-based risk estimate ($\hat{r}$) is predictive of manual driving behaviour and perceived risk in automated driving. Future research should include tactical and strategic components to the driving task.

1. Introduction

Recent years have seen a surge in the popularity of automated vehicles (AVs). Although AVs may solve several problems, they also introduce new ones. What is acceptable driving behaviour for an AV, and when do humans takeover the AV’s control? To answer these questions, it may be useful to have an estimate of the driver’s perceived risk in the AV (Zhang et al., 2019a). AVs will be used by drivers and passengers, and how these drivers and passengers perceive their system while driving is hardly known, yet essential to know. If drivers perceive high levels of risk, and the AV does not react appropriately and reliably, they could lose trust (Azevedo-Sa et al., 2021) and take over control, or reject the use of the AV altogether. Accordingly, it can be argued that new AVs need an assessment of how end users perceive their behaviour.

Perceived risk is influenced by a variety of factors, including the environment and traffic situation, the mental state of the driver, and prior experiences. Several studies have shown a correlation between personality characteristics and risky driving behaviour (Jonah, 1997; Ulleberg and Rundmo, 2003; Dahlen and White, 2006; Machin and Sankey, 2008). In this study, in which we provide a computational estimate of perceived risk, we consider the effect of the road environment and traffic situation only.
Summala (2007) proposed four factors that need to be maintained above a certain threshold to keep drivers within their “comfort zone”. These are safety margins (to road edges, obstacles or other vehicles), vehicle–road system (accelerations, road geometry), rule-following (obeying traffic laws, maintaining speed limits), and good progress of the trip (meeting one’s expectations for the pace or progress of the travel). Siebert et al. (2013) noted that the rule-following factor for comfort is redundant, as automated vehicles (AVs) will almost certainly follow the rules, and that good progress of the trip depends on traffic conditions, rather than automation state in itself. Therefore, in this paper, we focus specifically on factors that affect the safety margins and vehicle–road system for manual and automated driving. Other researchers have investigated the acceleration-related comfort aspect of automated driving, which we do not investigate in this paper (Hartwich et al., 2018). Specifically, we will focus only on the perceived risk that arises due to the presence of physical objects (road curb, obstacles, cars, etc.) and exclude risks that arise due to other factors related to vehicle dynamics (loss of traction, roll over, etc.).

Several studies have attempted to predict manual driving behaviour for scenarios involving road boundaries and obstacles. For example, longitudinal behaviour has been modelled using the optical edge rate on open roads (Denton, 1980), the time to extended tangent point in curves (Gruppelaar et al., 2018), time to collision (TTC) while approaching obstacles (Lee, 1976; Kondoh et al., 2008), and time headway (THW) during car-following (Van Winsum and Heino, 1996). Lateral behaviour has been modelled using the two-point model (Salvucci and Gray, 2004), in which lateral and heading errors were used as signals. Other models have used features such as the ‘angle to the tangent point’ (Boer, 1996) and time to lane crossing (TLC) as signals to steer the vehicle (Van Winsum and Godthelp, 1996; Boer, 2016). Some studies have also utilised a combination of signals such as speed and acceleration (executed by the drivers) to classify manual driving behaviour into different levels of risk-taking behaviour (Éboli et al., 2016, 2017).

All the above models used a particular measure (TLC, TTC, THW, etc.) to predict the actions of a human driver. If we abide by Näätänen and Summala’s (Näätänen and Summala, 1976) risk threshold theory, which proposes that all motivations for actions are to reduce the perceived risk and maintain it below a threshold level, these measures would represent the driver’s perceived risk. However, using different measures to quantify perceived risk in different scenarios leads to a fragmented model. For example, Basu et al. (2017) used ten different measures for assessing the behaviour in 14 different scenarios.

The problem with having separate measures for separate scenarios is that: first, it is difficult to compare the estimated risk between scenarios; second, if two or more scenarios occur simultaneously, it is difficult to estimate the combined risk; third, the operational level driving behaviour is often influenced by higher-level factors such as a driver’s familiarity with a particular road, his/her mood, etc. In the case of a fragmented model, one would have to tune the parameters of several models in a coherent manner. On the other hand, with a unified model, one has to tune the parameters of only one model. Accordingly, a unified model that can estimate driving behaviour or generate a risk estimate is needed.

The first attempt of making a unified model for explaining driver behaviour in different scenarios was made by Gibson and Crooks, using their ‘Field of Safe Travel’ concept (Gibson and Crooks, 1938). However, this was a qualitative description of how humans drive in different scenarios. There have been several models that utilise this concept of risk field (Mullakkal-Babu et al., 2020; Li et al., 2020; Guofa et al., 2021). For example, Wang et al. (2015, 2016) tested the Driver Safety Index (DSI) in real cars on straight sections of the highway during car-following and cut-in scenarios. Rasekhipour et al. (2016) proposed a path planning approach by combining the potential field approach with an optimal controller. More recently, Kolekar et al. (2020a) proposed a driver model based on the ‘Driver’s Risk Field’ (DRF) which could predict driving behaviour in a unified manner in 7 different scenarios (curve driving, obstacle avoidance, car following, etc.). Hence, we hypothesise that the DRF risk estimate (r) is a good candidate for evaluating perceived risk during automated driving.

Kolekar et al. (2020a) validated their driver model by comparing its output (i.e., steering angle and speed) to manual driving behaviour. They did not test if the risk estimate (r) that they calculated (which was used as a ‘cost function’ in the driver model) corresponded to the perceived risk of human drivers. Testing this is important since, if we can formulate a signal that correlates with the perceived risk by the driver, it can be a valuable tool to assess the behaviour and acceptance of automated driving systems. Also, Kolekar et al. (2020a) compared their simulation results to manual driving behaviour results published in the literature. Here, since we aim to quantify the driver’s perceived risk, it becomes imperative that an experiment is conducted in a real vehicle to generate a realistic feeling of risk.

In this paper, we test if the DRF-based risk estimate (r) correlates with the drivers’ perceived risk, in a test vehicle. We perform two main analyses. First, as an initial check, we tested if the DRF risk estimate (r) predicts the driver’s behaviour during manual driving. This is a necessary validation before performing the second step, where we compared the DRF risk estimate (r) to the driver’s comments and the takeovers that they performed, while being driven in an automated manner around a test track. The test track consisted of three scenarios: curves with different curvatures, a parked car (placed outside the lane), and intersections. The parked car being placed outside the lane boundary was done deliberately to examine if objects outside the lane boundary affected the perceived risk or not.

2. Methods

2.1. DRF risk estimate (r)

The DRF risk estimate (r) is a product of the ‘consequence of an event’ (e.g., hitting a tree is worse compared to being on the road) and the driver’s subjective belief about the ‘probability of the event occurring’ (Fig. 1d).
2.1.1. Consequence of an event

In the DRF model, the ‘consequence of an event’ is represented by a cost map that provides relative costs of colliding with different objects in the environment (e.g., road = 0, grass off-road = 500, tree = 2000, etc.). The DRF model can generate the $\hat{r}$ in real-time by using the environment generated by lidar and camera systems on-board the test vehicle. However, in this study, we calculate the $\hat{r}$ offline. We acquired a high definition map of the test track, converted it into a cost map, and tracked the location of the car using GPS (± 5 cm accuracy). The cost of being on the road ($C_{\text{road}}$), cost of being off-road ($C_{\text{off-road}}$), and cost of hitting the parked car ($C_{\text{parked car}}$) were determined using manually driven trajectories (Table 1).

2.1.2. Probability of the event occurring

The ‘probability of the event occurring’ is represented by the Driver’s Risk Field (Fig. 1a), which was empirically determined by Kolekar et al. (2020b) and is mathematically represented by a modified torus with a Gaussian cross-section, the height of which decreases and width increases, as you go further away from the ego-car (Fig. 1b). The formulation begins by predicting the vehicle path, which is calculated using a kinematic car model. The position ($x_{\text{car}}, y_{\text{car}}$), heading ($\phi_{\text{car}}$), and steering angle ($\delta$) determine the radius of the arc ($R_{\text{car}}$) in which the car is predicted to travel, assuming a constant steering angle (Eq. (1)).

$$R_{\text{car}} = \frac{L}{\tan(\delta)}$$

where $L$ is the wheel-base of the car. Using $x_{\text{car}}, y_{\text{car}}, \phi_{\text{car}},$ and $R_{\text{car}},$ the centre of the turning circle ($x_c, y_c$) is determined, which is used to calculate the arc-length ($s$) along the predicted path. The DRF is modelled as a torus with a Gaussian cross-section (Eq. (2)). The height ($a$) and width ($\sigma$) of the Gaussian are a function of the arc length ($s$) (Fig. 1b).

$$z(x, y) = a \exp\left(-\frac{\sqrt{(x-x_c)^2 + (y-y_c)^2 - R_{\text{car}}^2}}{2\sigma^2}\right)$$
The height of the Gaussian ($\sigma$), is modelled as a parabola (Eq. (3)).
\[ a(s) = \rho \left( s - v \, t_{la} \right)^2 \]  
(3)

With a fixed look ahead time ($t_{la}$), the look-ahead distance is assumed to increase linearly with speed ($v$). Parameter ($\rho$) defines the ‘steepness’ of the parabola. The width of the Gaussian ($\sigma$) is modelled as a linear function of arc length ($s$) (Eq. (4)) and includes the following parameters: first, $c$ defines the width of DRF at the location of the vehicle and is related to the car-width. In this paper, $c$ is equal to car-width/4. Second, $m$ defines the slope of widening (or narrowing for negative values) the width of the DRF proportional to the (absolute) steering angle ($|\delta|$). This relationship between the DRF and $\delta$ is based on the rationale that the variability in steering angle increases linearly with the (mean) steering angle (Van Winsum and Godthelp, 1996; Boer, 2016; Eboi et al., 2016, 2017; Näätänen and Summala, 1976; Basu et al., 2017; Gibson and Crooks, 1938; Mullakkal-Babu et al., 2020; Li et al., 2020; Guofa et al., 2021; Wang et al., 2015, 2016; Rasekhipour et al., 2016; Kolekar et al., 2020a,b, 2018) and is similar to the empirically confirmed signal-dependent noise present in the human sensorimotor system (Harris and Wolpert, 1998; Clamann, 1969). $k_1$ and $k_2$ represent the parameters for the inner and outer edges of the DRF, respectively, and allow for an asymmetric DRF. The expansion of DRF proportional to $\delta$ results in the accumulation of a higher risk for a curve with a smaller radius. The asymmetric expansion ($k_1$ and $k_2$) provides flexibility to exhibit curve-cutting ($k_1 < k_2$), centreline ($k_1 = k_2$), or curve overshooting ($k_1 > k_2$) behaviour.
\[ \sigma_i = (m + k_1 |\delta|) \, s + c \]  
(4)

In short, the Driver’s Risk Field (DRF) is parameterised by $\rho$, $t_{la}$, $m$, $c$, $k_1$, $k_2$, and is only dependent on the driver’s state, not the environment. Although these DRF parameters are specific to each individual, in this study, we decided to use the same parameter values (Table 1) for the eight participants. This was done so that we could combine/average the results over all the participants.

The values in Table 1 were estimated using a grid search algorithm which tried to minimise $\sum_{i=1}^{3} \left( \hat{y}_i \text{ model} - \hat{y}_i \text{ experiment} \right)^2$, where $i = 1$: steering angle, $i = 2$: speed, and $i = 3$: lateral deviation from the lane centre. Where, $y_i$ model is the trajectory generated by the DRF model (page 10 of Kolekar et al., 2020a), and $y_i$ experiment is the average of five laps manually driven by a test driver.

### 2.1.3. DRF risk estimate for each trajectory

The DRF risk estimate ($\hat{r}$) is a function of the vehicle’s trajectory, which is defined by the vehicle’s position ($x$, $y$), heading ($\phi$), steering angle ($\delta$), and speed ($v$). We calculated the $\hat{r}$ for each manual and automated lap of every participant.

### 2.2. Test track and test vehicle

The experiment was conducted at a test facility in collaboration with Nissan, Japan. The test vehicle was a 1st generation Nissan Leaf equipped with an automated driving system. The participants sat on the right side of the vehicle and drove on the left hand side of the road.

To test different scenarios, we divided the test track into three sections. The first section consisted of curves with varying curvatures and road widths. The second section (parked car section) consisted of a Nissan NV500 van parked along the left shoulder of the road (outside the lane boundary). The third section consisted of two 90-degree intersections.

We further subdivided these three sections into 12 sectors. The curve driving section was divided into 9 sectors (S1 to S9), the parked car was its own sector (PC), and the intersection segment consisted of two sectors (I1 and I2), as shown in Fig. 2. One of the intersections (I1) had a zebra crossing, which the participants were asked to ignore for this experiment.

In Fig. 2, we show a snapshot of the signals recorded during the second manually driven lap by Participant 2. The dotted green and red lines indicate the start and end of the trajectory used for analysis, respectively. The driver started every lap from the green dot and stopped at/after the red dot. In the automated driving condition, the car would need some initialisation time, and hence it had to be driven manually from the green dot, and it would get into automated mode by the time it crossed the green line. Hence, the green line is used as the start of the analysis for both the manual and automated trajectories.

<table>
<thead>
<tr>
<th>Table 1</th>
<th>DRF and cost map parameter values.</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>DRF parameters</strong></td>
<td>$p$</td>
</tr>
<tr>
<td>0.04</td>
<td>3</td>
</tr>
<tr>
<td><strong>Cost map parameters</strong></td>
<td>$C_{\text{road}}$</td>
</tr>
<tr>
<td>500</td>
<td>250</td>
</tr>
</tbody>
</table>

In Fig. 2, we show a snapshot of the signals recorded during the second manually driven lap by Participant 2. The dotted green and red lines indicate the start and end of the trajectory used for analysis, respectively. The driver started every lap from the green dot and stopped at/after the red dot. In the automated driving condition, the car would need some initialisation time, and hence it had to be driven manually from the green dot, and it would get into automated mode by the time it crossed the green line. Hence, the green line is used as the start of the analysis for both the manual and automated trajectories.
2.3. Participants and test procedure

Eight male Nissan test drivers, certified to operate the automated test vehicle, participated in the experiment (mean age = 44.1 and SD = 13.4 years). Each driver drove five laps of the test track manually on one day and on another day they were driven in the same car, along the same track, in an automated manner for twelve laps. The experimenter (first author) sat behind the driver’s seat and an interpreter (Japanese ↔ English) sat in the passenger’s seat.
During the automated driving condition, the participants were requested to comment on how they felt about the driving behaviour of the vehicle. They were encouraged to specify the ‘why and where’ of their feedback. This was done because, first, the driver’s steering or speed control input was unavailable since the car was being driven in an automated manner. Second, we could not provide the drivers with a hand-held or foot-operated device to indicate their perceived risk level. This was in accordance with Nissan’s safety committee’s guidelines which required the drivers to always be able to takeover the vehicle’s controls. Third, physiological measures are reliable in a laboratory setting, but in the dynamic environments of the test course, it is challenging to acquire reliable physiological data. Hence, the decision was made to ask the participants to think aloud as they were driven around the lap. The audio and video of all the laps were recorded, and the participants could take over the vehicle’s control whenever they wanted.

2.4. Data exclusion

For the automated driving conditions, participants experienced twelve different trajectories (speed and position profiles). The initial plan was to have the driver’s experience two repetitions of six different controllers (1: centreline follower, 2: manual replay, 3: manual replay with 90% speed, 4: DRF model, 5: DRF model safer (90% risk threshold) from Kolekar et al. (2020a), and 6: ‘safe’ driver’s trajectory). However, we could not replicate these trajectories on the track. Hence, we decided not to analyse the results per controller and attribute the trajectories to any controller type. In other words, we now analyse them purely based on the trajectory that the driver’s experienced (not what the car was supposed to implement). After every lap, the participants filled up a questionnaire that evaluated the characteristics of the controllers. Since we do not draw or attribute any conclusions to the type of controller, we did not analyse the questionnaires. Additionally, due to experimenter error, the video for Lap 4 of participant 6 was not recorded and hence this lap has been excluded from the analysis.

3. Results

3.1. Manual driving

For the manual driving conditions, the eight participants drove five laps of the test course. They were instructed to drive as they normally would. To test if the DRF model could predict the driver’s behaviour during manual driving we compared the predictions of the model to the steering and speed control actions performed by the driver.

3.1.1. Driver actions

To mitigate a rise in risk, the driver could perform steering corrections and speed reductions. Hence, we calculated the following signals to evaluate the DRF prediction estimates.

- Absolute steering angle (|\(\delta\)|): Since both positive (left turn) and negative (right turn) steering actions indicate the driver’s intent to reduce risk, we calculate the absolute value of the steering angle to quantify the driver’s action.

- Speed reduction (\(v_{\text{reduction}}\)): We assume that speed reduction is an indication of the driver’s intent to reduce the risk. The reduction in speed is calculated with respect to the maximum speed at which the driver drove during that particular lap (\(V_{\text{max}}\)). Hence \(v_{\text{reduction}} = V_{\text{max}} - v\). We calculate the \(v_{\text{reduction}}\) with respect to \(V_{\text{max}}\), because we think that \(V_{\text{max}}\) represents the driver’s max willingness to incur risk, in the context of this track.

3.1.2. DRF model predictions

The DRF model by Kolekar et al. (2020a) generates a single risk estimate. However, as mentioned earlier, the driver’s actions (to mitigate the rise in risk) are split mainly into two components: |\(\delta\)| and \(v_{\text{reduction}}\). To quantify which part of the \(\hat{r}\) corresponds to a reduction in risk due to |\(\delta\)| and which part corresponds to a reduction in risk due to \(v_{\text{reduction}}\), we calculate the DRF steering risk potential (\(\hat{p}_{\text{steering}}\), Fig. 3b) and DRF speed risk potential (\(\hat{p}_{\text{speed}}\), Fig. 3c), respectively.

\[
\hat{p}_{\text{steering}} = \hat{r}_{\delta = 0, v, \text{max}} - \hat{r}_{v, \text{max}} \\
\hat{p}_{\text{speed}} = \hat{r}_{\delta = 0, v, \text{max}} - \hat{r}_{\delta, 0}
\]

The \(\hat{r}_{\delta = 0}\) is the \(\hat{r}\) when \(\delta = 0\) and hence, indicates the \(\hat{r}\) if the driver had not taken any steering actions. Similarly, \(\hat{r}_{v, \text{max}}\) is the \(\hat{r}\) when \(v = V_{\text{max}}\) (\(V_{\text{max}}\) of that particular lap) and hence, it indicates the \(\hat{r}\) if we were to disregard the speed reductions performed by the driver. \(\hat{r}_{\delta = 0, v, \text{max}}\) is the \(\hat{r}\) when \(\delta = 0\) and \(v = V_{\text{max}}\) and hence, it represents the risk estimate if the driver had not made any speed or steering corrections.

From Fig. 3a we can see that, for most parts, \(\hat{r}_{\delta = 0, v, \text{max}} > \hat{r}_{\delta, 0}\), \(\hat{r}_{v, \text{max}} > \hat{r}\). It indicates that, as expected, most of the steering and speed corrections performed by the driver reduced the risk. In some instances (in the neighbourhood of 300 m) the \(\hat{r}_{v, \text{max}}\) is higher than the \(\hat{r}_{\delta = 0, v, \text{max}}\). The difference indicates that the model thinks that the steering implemented by the diver at this point was ‘incorrect’. Interestingly, it can be seen that the \(\hat{r}_{\delta = 0}\) is lower compared to the \(\hat{r}_{\delta = 0, v, \text{max}}\), which indicates that the driver compensated for this ‘steering error’ by reducing the speed more than the model have expected him/her to.

The predictions of the model and the driver’s steering behaviour showed a moderate-to-strong correlation (\(\rho_{\text{steering}} = 0.69 \pm 0.04\) (Mean \(\pm\) SD), Table 2). The peaks of the \(\hat{p}_{\text{steering}}\) and the |\(\delta\)| signals also align, indicating the steering timing also matches. In the
Fig. 3. Manual driving predictions: The figure shows DRF risk estimates, DRF steering risk potential ($\hat{p}_{\text{steering}}$), and DRF speed risk potential ($\hat{p}_{\text{speed}}$) as a function of the distance covered along the lane centre, for Participant 2. All the signals are the mean of the five repetitions of the manual trajectories. (a) The four different DRF Risk Estimates: $\hat{r}_{\|\delta\|=0}$, $\hat{r}_{v_{\text{max}}}$, $\hat{r}_{\|\delta\|=0}$, and $\hat{r}$ are plotted. The $\hat{r}_{\|\delta\|=0}$ has the highest value, and the $\hat{r}$ has the lowest value. In between these two are the $\hat{r}_{\|\delta\|=0}$ and $\hat{r}_{v_{\text{max}}}$ plots. (b) We compare the $\hat{p}_{\text{steering}}$, which is the model’s prediction about how much steering the driver should have implemented, to the steering angle implemented by the driver ($|\delta|$). (c) We compare the $\hat{p}_{\text{speed}}$, which is the model’s prediction about how much speed reduction is needed, to the speed reduction implemented by the driver ($v_{\text{reduction}}$). The driver’s speed reduction can be seen to have ‘slower dynamics’ compared to the model predictions. $\rho$ is the Pearson’s correlation coefficient.

Table 2
Correlation coefficients between DRF model predictions and manual driving behaviour.

<table>
<thead>
<tr>
<th>Participant No.</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>Mean ± SD</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\rho_{\text{steering}}$</td>
<td>0.69</td>
<td>0.73</td>
<td>0.75</td>
<td>0.62</td>
<td>0.70</td>
<td>0.67</td>
<td>0.74</td>
<td>0.67</td>
<td>0.69 ± 0.04</td>
</tr>
<tr>
<td>$\rho_{\text{speed}}$</td>
<td>0.77</td>
<td>0.63</td>
<td>0.63</td>
<td>0.61</td>
<td>0.67</td>
<td>0.64</td>
<td>0.65</td>
<td>0.52</td>
<td>0.64 ± 0.07</td>
</tr>
</tbody>
</table>

neighbourhood of 300 m, the $|\delta|$ is higher than the $\hat{p}_{\text{steering}}$, which means that the driver steered more than the model expected him/her to.

Fig. 3 shows that the driver’s speed reduction correlates moderately with the model predictions ($\rho_{\text{speed}} = 0.64 \pm 0.07$ (Mean ± SD), Table 2). The peaks of the $v_{\text{reduction}}$ signal occur earlier than the peaks in the $\hat{p}_{\text{speed}}$, indicating that the drivers reduced their speed earlier than the model expected them to. This, we think, is due to the parameter values identified for the DRF. Specifically, the look-ahead-time ($t_{la}$) parameter of the DRF was set to 3 seconds (Table 1). Increasing this value will lead to the model predicting an earlier speed reduction.

Ideally, for the speed control case, we should have compared the model predictions to the brake and accelerator pedal activity since the perceived risk estimate of the driver is more directly related to the actions of the driver. However, the pedal activity was recorded as a binary (on/off) signal. Hence, we decided to use the speed signal, even though the car dynamics would entail that it would have ‘slower dynamics’ compared to the risk metric, as can be seen in Fig. 3c.

The results hence verify that DRF risk estimate correlates well with the steering and speed reduction actions performed by the drivers during manual driving condition.
3.2. Automated driving

Investigating if the DRF risk estimate (\(\hat{r}\)) can predict the driver’s subjective feeling of risk during an automated drive is the main aim of this study. To test this we compared the DRF risk estimate (\(\hat{r}\)) to the comments provided by the driver, and the takeovers they performed during each lap. Essentially, we expected a Risky comment or a takeover when the \(\hat{r}\) value was high and a Non-risky comment, no comment, or no takeover when the \(\hat{r}\) value was low.

3.2.1. Driver’s comments and takeovers

We segregated the driver’s comments into two categories: Risky comments and Non-risky comments. Risky comments were those in which the driver indicated a sense of danger, discomfort, or risk by using words or phrases such as ‘close to the curb’, ‘scary’, etc. Non-risky comments are those that indicated that the drivers liked the drive, were ‘OK’ with it, found it too safe (boring), or did not comment at all. Table 3 shows a few words that, for illustration purposes, we categorized into (i) the reasoning (e.g., close, near), (ii) the object (e.g., curb), (iii) the noises participants made (e.g., woh, aaahhh), (iv) how they described the sector (e.g., risky, scary), (v) the takeovers (e.g., overtook, (steering) correct(ion)). The complete list of Risky and Non-risky comments can be found in the supplementary material.

As mentioned in the methods section, we had segregated the track into 12 sectors (Curve driving: S1–S9, Parked car: PC, Intersections: I1, and I2) because participants usually commented while referring to these salient features, after they had experienced the sector. Hence using time stamps of the comments to attribute the comments to the location of the vehicle on the track would not make sense. The comments were hence attributed to each sector. So, if a driver complained about a particular sector being Risky, the trajectory for the entire sector was ‘marked red’ (Fig. 4). The takeovers by the drivers were measured as a binary signal (black line in Fig. 4).

3.2.2. DRF model predictions

We compared the driver’s comments and takeovers to the \(\hat{r}\) calculated during the automated trajectory. Similar to that calculated for manual driving condition, the \(\hat{r}\) in the automated driving condition is unique to the trajectory of each of the 12 laps, for each of the eight participants.

**DRF risk estimate trajectories:** The \(\hat{r}\) trajectories for each participant are plotted in Fig. 4. Visual inspection of the plot reveals that the participants made a Risky comment when the \(\hat{r}\) was also calculated to be high. This points towards a good match between the driver’s subjective feeling of risk and the objective risk estimate calculated by the DRF model. However, there are some instances in the I1 and I2 sectors (intersections) where red sectors are marked at very low \(\hat{r}\) values.

Another interesting point is that the \(\hat{r}\) indicates a high-risk level in S1 and S7, but a low-risk level in S9. However, in Fig. 2c, we can find that the curvature of S9 is even higher than that of S1 and the curvature in S7 is relatively small. The reason for this high risk is due to the lateral positioning of the vehicle. The AV drove very close to the curb in S7 and S1 and hence led to a high value for \(\hat{r}\). Several Risky comments from the participants also corresponded to the vehicle being too close to the curb. In S9, the AV would slow down considerably more compared to that in S1 and S7, and hence the \(\hat{r}\) is low despite the curvature being more than that in S1 and S7.

**Maximum of \(\hat{r}\) per sector:** We assume that the maximum risk experienced by the driver is a more valid index of perceived risk as compared to the mean risk experienced during a sector. Hence, we calculate the max(\(\hat{r}\)) in each sector, as a metric, to quantify the ‘riskiness’ of that particular sector. This metric (max(\(\hat{r}\)) per sector) was labelled either Risky or Non-risky based on the comments, and takeover or no takeover depending on whether the participant took over the vehicle’s control or not (Fig. 5).

**Counting the number of Risky and Non-risky sectors:** We plotted a histogram for each participant (Fig. 6) using the max(\(\hat{r}\)) for each sector (from Fig. 4). We compared the qualitative response (the comments provided by the driver) with a quantitative response (the takeovers performed by the drivers) in a stacked histogram. Fig. 6 shows that in the curve driving and the parked car sections, none of the participants made any Risky comments or performed takeovers in the lowest risk estimate (bin). However, there are

<table>
<thead>
<tr>
<th>Table 3</th>
<th>Words mentioned in drivers’ comments (8 participants × 12 laps × 12 sectors)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Risky comments</strong></td>
<td>Count</td>
</tr>
<tr>
<td>close, left, near, closer</td>
<td>105</td>
</tr>
<tr>
<td>curb, stone</td>
<td>37</td>
</tr>
<tr>
<td>woh, aahhh, …(noises)</td>
<td>40</td>
</tr>
<tr>
<td>risk, risky, scary, hit, aggressive, immediately, wrong, bad, dangerous, harshly, intensely, jumps, rush, scared, sharp, sharply, steep, suddenly, surprised, weird</td>
<td>45</td>
</tr>
<tr>
<td>override, overrode, overshoot, correct</td>
<td>10</td>
</tr>
</tbody>
</table>
The shaded rectangles indicate the 12 different sectors. S1–S9 for the curve driving segment, PC is the parked car segment, and I1 and I2 indicate the two 90 degree intersections. The plots show the DRF risk estimate ($\hat{r}$) for each lap. The blue parts indicate Non-risky sectors and the red parts indicate the Risky sectors. The black 'steps' indicate the steering takeover by the driver. It can be seen that, in general, the risk estimate is marked red when the $\hat{r}$ is high. Also, the takeovers occur when the $\hat{r}$ peaks.

several instances in the intersections when the participants (P1, P5, P6, P7, and P8) either made a Risky comment or performed a takeover.

Naturalistic driving is majorly comprised of non-risky situations and a very few risky situations. A similar trend was shown by the DRF risk estimate ($\hat{r}$), where the number of sectors decreased as the risk estimate increased (Fig. 6). Also, most of the takeovers are accompanied by a Risky comment.

**DRF risk estimate as a decision variable:** The main aim of this paper was to test if the DRF risk estimate could be used to detect Risky and Non-risky sectors. Since the participants can express their 'feeling of risk' by either commenting or by taking over, we combine the two modes (using the 'or' function). Hence a sector is marked Risky if the participant either made a Risky comment, or a takeover, or both. The remaining sectors are labelled Non-Risky.

We calculated the proportion of Risky sectors ($\frac{\text{Risky}}{\text{Risky} + \text{Non-Risky}}$) for each $\hat{r}$ value (mid point of each bin) and averaged it over all the eight participants. The mean (circular markers) and standard deviation (error bars) are plotted in Fig. 7. The curve driving plot (Fig. 7a), all the 14 points together, are calculated from 855 points (8 participants $\times$ 12 laps $\times$ 9 sectors - 9 sectors of lap four of Participant six, which the experimenter failed to record). The parked car section plot (Fig. 7b), all 8 points combined, represent...
Fig. 5. Steps in automated driving analysis: Step 1: The comments were classified into ‘Risky’ or ‘Non-risky’ (Table 3). The takeovers were recorded by the vehicle’s data collection system. Step 2: The trajectory of the entire sector the comment was attributed to was tagged ‘Risky’ or ‘Non-risky’. Similarly, the entire sector in which the takeover was initiated was tagged ‘takeover’ or ‘no takeover’. If a takeover was initiated or comment was attributed to parts between two sectors, both sectors were tagged. Step 3: The maximum value of \( \hat{r} \) for each sector was chosen as a metric. The number sectors with different levels of \( \hat{r} \) were plotted in a stacked histogram. The two histograms, comments and takeovers, are combined and presented in Fig. 6. Step 4: The proportion of risky comments or takeovers was calculated for each participant in each of the histogram bins, and then averaged over all eight participants. The red circular markers indicate the mean, and the bars indicate ± standard deviation (SD). The plot shown here is for the curve driving section (S1–S9). An identical analysis was performed for the parked car (PC) and the intersections (I1, I2) and the results are shown in Fig. 7.

The mean and standard deviation over 95 points (8 participants × 12 laps × 1 sector - 1 sector of lap four of Participant six). The intersection plot (Fig. 7c), all 9 points combined, represent the mean and standard deviation over 190 points (8 participants × 12 laps × 2 sectors - 2 sectors of lap four of participant six). We think that this difference in the number of data points explains the larger uncertainty in the parked car section and intersections compared to the curve driving section.

A logistic function was fit to the (mean) data points in all three sections (Fig. 7). As can be seen from the figure, instances with \( \hat{r} \) higher than 2250 and 1000 did not occur (Fig. 6, 7) and the proportions of ‘Risky’ comments were low at low \( \hat{r} \) values. It is unknown how the results extrapolate to higher values of \( \hat{r} \). However, the fact that the current version of the DRF only accounts for the physical presence of objects suggests that the \( \hat{r} \) at intersections may be a limitation on this model. This is because at an intersection the DRF sees a larger patch of road (‘safe zone’) and hence reduces the \( \hat{r} \), which may result in speeding rather than braking at an intersection. Humans slow down at an intersection in anticipation of traffic. This kind of learned conscious behaviour will have to be integrated in future versions of the DRF.

In this study, we derived the perceived risk (indicated by the driver) as a measure of it being ‘Risky’ or ‘Non-risky’. This was based only on the driver’s comments and whether they took over the controls of the vehicle or not. However, while analysing the videos we found several interesting behaviours where the drivers would hover their hands near the steering wheel (presumably when they felt a high risk) (De Diego et al., 2013). This kind of bio-mechanical preparedness to takeover also contains valuable information about the risk perceived by the driver (Zhang et al., 2019b) and could be analysed in detail in future studies.

4. Discussion

The aim of this paper was to test if the DRF-based risk estimate (\( \hat{r} \)) correlates with drivers’ perceived risk. The results show that \( \hat{r} \) predicted driver’s manual driving behaviour — speed and steering actions. Furthermore, in automated driving, the risk estimate, correlated with the comments and takeovers performed by the drivers in curve driving and avoiding a car parked outside the lane boundary. In other words, the DRF, a field-based risk metric, captured the perceived risk in curve driving and with a static obstacle.

The results for 90 degree intersections were inconclusive. Events with high \( \hat{r} \) did not occur (Fig. 6, 7) and the proportions of ‘Risky’ comments were low at low \( \hat{r} \) values. It is unknown how the results extrapolate to higher values of \( \hat{r} \). However, the fact that the current version of the DRF only accounts for the physical presence of objects suggests that the \( \hat{r} \) at intersections may be a limitation on this model. This is because at an intersection the DRF sees a larger patch of road (‘safe zone’) and hence reduces the \( \hat{r} \), which may result in speeding rather than braking at an intersection. Humans slow down at an intersection in anticipation of traffic. This kind of learned conscious behaviour will have to be integrated in future versions of the DRF.

In this study, we derived the perceived risk (indicated by the driver) as a measure of it being ‘Risky’ or ‘Non-risky’. This was based only on the driver’s comments and whether they took over the controls of the vehicle or not. However, while analysing the videos we found several interesting behaviours where the drivers would hover their hands near the steering wheel (presumably when they felt a high risk) (De Diego et al., 2013). This kind of bio-mechanical preparedness to takeover also contains valuable information about the risk perceived by the driver (Zhang et al., 2019b) and could be analysed in detail in future studies.

\[ x = \frac{a_1}{1 + e^{-a_2(y-a_3)}} - a_4 \]
Fig. 6. Histogram of comments and takeovers: P1–P8 indicate the eight participants. The 3 columns indicate the segregation of sectors into the three segments: curve driving (S1–S9), parked car (PC), and intersections (I1–I2). The x-axis represents the number of sectors and the y-axis represents the DRF risk estimate. We plot the Risky and Non-risky comments stacked on top of each other for comparison. Simultaneously, we also plot the takeovers and non-takeovers in a stacked manner. Hence the height of two adjacent bars is equal. The bar plots have 14 bins (y axes) with minimum value of 0 and maximum value of 3500, and a width of 250.
Fig. 7. Proportion of risky sectors: This plot shows the mean (circular markers) and ±SD (bars) of the proportion of risky sectors (either the comment was Risky or there was a takeover) over the 12 laps of all the 8 participants (95 laps in total, since 1 lap’s video was accidentally not recorded). The sectors for the three sections: curve driving (S1–S9), parked car (PC), and intersections (I1, I2) contained 855, 95, and 190 data points respectively. A logistic function was fit to the mean (circular markers) data points (Eq. (5)). The goodness of fit is indicated by the $r^2$ metric.

To put our results in the context of literature, we compare it to ‘time to lane crossing’ (TLC) (Van Winsum and Godthelp, 1996). There are several methods for calculating TLC (Van Winsum et al., 2000), such as the ‘TLC swath’ approach proposed by Boer (2016). The similarities between the TLC and the $\hat{r}$ arise from the fact that they both have an expanding arc (for TLC) or a field (for DRF). This combined with the threshold, below which these metrics need to be maintained, results in satisficing behaviour (a strategy that aims for a satisfactory result [e.g., stay within lane boundary], rather than the optimal solution [e.g., follow lane centre]). However, there are two key differences.

First, the expansion of the arc for the TLC is fixed, whereas the expansion of the DRF is proportional to the (absolute value of the) steering angle. This means that the DRF expands more (higher $\hat{r}$) for a sharper curve compared to a shallower one, and hence would want the car to drive slower for a sharper curve. The ‘swath TLC’ gets around this problem by having an additional ‘straight-line’ TLC which projects straight ahead and intersects with the outer lane boundary of the curve. The added benefit of having a steering dependent expansion over a straight line TLC is that it is generalisable to scenarios other than interacting with lane boundaries.

The second difference is that, since TLC is calculated using the intersection points of the arc and the lane boundary, it cannot account for the presence of objects outside the lane boundary. For a TLC based controller driving on a road with curbs on both sides is the same as driving on a road with curb on one side and an open oncoming lane on the other. The DRF being a field can expand beyond the lane boundaries and hence can account for objects outside the lane boundaries and the different levels of danger that they pose (curb is more dangerous than an open oncoming lane).

In this study, all the scenarios were static which made it easy to formulate a high definition (HD) cost map of the test-track, offline. However, multi-lane dynamic scenarios need dynamic cost maps that are updated online. Researchers, primarily using computer simulations, have suggested approaches for straight roads and intersections using the ‘Conditional Random Field (CRF)’ (Guofa et al., 2021) and the ‘Potential Field Indicator (PFI)’ (Li et al., 2020) that account for the uncertainties in behaviour of neighbouring vehicles and the noise introduced while sensing their state (e.g., position, velocity, acceleration). Lu et al. (2020) also incorporate the mass of the vehicle and Huang et al. (2020) incorporate the intention of the vehicle in a mixed (manual + automated) traffic condition using long-short term memory (LSTM) networks to determine the potential field generated by external vehicles. In the future, we would like to incorporate these features into the dynamic cost-map and test the validity of $\hat{r}$ in multi-lane dynamic scenarios with test vehicles.

An important aspect of the DRF model is its intent to formulate a unified risk measure over a plethora of driving scenarios. Although further experimentation with other (static and dynamic) scenarios will be needed to test the validity of the model, this unified approach to express driver’s perceived risk has important applications in designing the behaviour of automated systems. It can be used to assess the performance of the automated system in different scenarios and adjusted to meet the preferences of the individual drivers.
5. Conclusions

The aim of this paper was to test if the Driver’s Risk Field (DRF)-based risk estimate ($\hat{r}$) could correlate with the perceived risk of the driver. To ensure that the driver’s perceived a realistic level of risk, the experiment was performed on a test track with a Nissan Leaf test vehicle.

- The DRF risk estimate ($\hat{r}$) correlates with manual driving behaviour expressed in terms of steering angle and speed ($\rho_{\text{steering}} = 0.69$, $\rho_{\text{speed}} = 0.64$, Table 2, Fig. 3)
- In automated driving condition, the $\hat{r}$ could predict the comments and takeovers performed by the drivers in the curve driving section ($r^2 = 0.98$, Fig. 7).
- The $\hat{r}$ could also predict the comments and takeovers by the drivers, in the automated driving condition, while negotiating a car that was parked outside the lane boundary ($r^2 = 0.59$, Fig. 7).

CRediT authorship contribution statement

Sarvesh Kolekar: Conceptualization, Methodology, Software, Formal analysis, Writing - Original Draft, Writing - review & editing, Visualization. Bastiaan Petermeijer: Conceptualization, Methodology, Resources. Erwin Boer: Visualization, Reviewing, Supervision, Project administration, Funding acquisition. Joost de Winter: Conceptualization, Methodology, Writing - review & editing, Supervision, Project administration, Funding acquisition. David Abbink: Conceptualization, Visualization, Methodology, Writing - review & editing, Supervision, Project administration, Funding acquisition.

Acknowledgements

We would like to thank Nissan Research Centre, Japan and in particular Dr. Tsuyoshi Sakuma and Hidetoyo Aoki for their help, guidance, and technical support throughout this research. The Netherlands Organisation for Scientific Research (NWO) funded this project. Sarvesh Kolekar, Bastiaan Petermeijer and David Abbink were supported by the VIDI 14127 project, and Joost de Winter was supported by the VIDI 178047 project.

Appendix A. Supplementary data

Supplementary material related to this article can be found online at https://doi.org/10.1016/j.trc.2021.103428.

References


S. Kolekar et al.


