Green Open Access added to TU Delft Institutional Repository

‘You share, we take care!’ – Taverne project

https://www.openaccess.nl/en/you-share-we-take-care

Otherwise as indicated in the copyright section: the publisher is the copyright holder of this work and the author uses the Dutch legislation to make this work public.
An Empirical Analysis to Assess the Operational Design Domain of Lane Keeping System Equipped Vehicles Combining Objective and Subjective Risk Measures

Haneen Farah\textsuperscript{a}, Shubham Bhusari, Paul van Gent, Freddy Antony Mullakkal Babu\textsuperscript{b}, Peter Morsink, Riender Happee\textsuperscript{a}, and Bart van Arem\textsuperscript{c}

Abstract—Lower levels of automation are designed to work in specific conditions referred to as the Operational Design Domain (ODD). Beyond these conditions, the human driver is expected to take control. A mismatch between a driver’s understanding and expectations of the automated vehicle capabilities and its actual capabilities as prescribed in the Original Equipment Manufacturers (OEMs) manual, could affect their safety and trust in automation. The main aim of this study is to develop a method for assessing the ODD of lane keeping system equipped vehicles. The analysis method is composed of an objective driving risk measure based on the Probabilistic Driving Risk Field (PDRF), and a subjective risk measure based on driver behavior, trust and situation awareness. We demonstrate the method applicability and the potential to aid OEMs and road operators in defining more clearly the ODD while taking into account the driver’s safety and awareness of the system capabilities.

Index Terms—Automated vehicles, field test, lane keeping assistance systems, operational design domain, risk analysis.

I. INTRODUCTION

THE Society of Automotive Engineers (SAE) categorizes vehicle automation into six levels ranging from no-automation (level 0), where the human driver has full control of the vehicle, to full automation (level 5) where the vehicle controls all the driving tasks [1]. Each level of automation is designed to work in specific conditions referred to as the Operational Design Domain (ODD). The ODD is defined as ‘the specific conditions under which a given driving automation system or feature thereof is designed to function, including, but not limited to, driving modes. An ODD may include geographic, roadway, environmental, traffic, speed, and/or temporal limitations’ [1]. The ODD for all levels of automation, except for full automation, is limited. Currently, SAE level 2 of vehicle automation (i.e. partial automation) is available in several commercial vehicles. Partial automation, refers to ‘sustained and ODD specific execution by a driving automation system of both the lateral and longitudinal vehicle motion control subtasks of the Dynamic Driving Task (DDT) with the expectation that the driver completes the Object and Event Detection and Ranging (OEDR) subtask and supervises the driving automation system’. This means that with partial automation, the driver may be requested to take over control of the vehicle in certain conditions that are not part of the ODD. As each Original Equipment Manufacturer (OEM) specifies their own specific ODD, this could result in uncertainties about the capabilities of different vehicles within the same level of automation, and to mode confusion. A mismatch between a driver’s understanding and expectations of the capabilities of the automated vehicle and its actual capabilities as prescribed in the OEM’s manual, could affect drivers’ trust in automation and their perceived risks.

There is ample research, at level 2 automation, on the impact of the longitudinal automation function (such as Adaptive Cruise Control (ACC) and Cooperative ACC) on driver behavior, traffic safety and traffic flow [2]–[4], but limited research on the impacts of Lane Keeping Assistance System (LKAS) on driver behavior and traffic flow [5].
A. Lane Keeping Assistance System (LKAS)

There are three, commonly known, types of LKAS [6]: Warning, Intervention and Control based systems. Warning based systems, do not directly alter the vehicle trajectory and require that the driver acts on the warning to have any effect. Drivers are warned by the system if they are swaying away from their current lane and the lane change indicator is not turned on by the driver. The principle of an Intervention based system is to provide a steering wheel torque to avoid unintended lane departures. This torque is related to the vehicle’s lateral position and speed. The system has limited authority and is meant to only augment driver commands but not to replace them continuously. Finally, the Control based system, which keeps track of potential unintended lane departures, and continuously steers the vehicle to ensure that it is ideally positioned at the center of the lane (with a small allowable buffer). These systems, not requiring active human steering inputs, are the focus of this paper, combined with ACC and a collision avoidance function to provide SAE level 2 automation.

B. Factors Influencing the LKAS

There are several factors that can affect the performance of the LKAS as described in different vehicle owners’ manuals and can be classified into the following categories [7]:

Road Infrastructure and Traffic: This includes the type of road (i.e. highway, city road or rural road), and whether the road section includes discontinuities such as weaving sections, on-ramps and off-ramps. The road surface quality, in terms of smoothness, wetness or dust/slush are also important factors. The LKAS performance is also affected by the presence of lane-markings and their clarity, consistency and quality [8]. Finally, the traffic state of surrounding vehicles on the road has impact on the vehicle dynamics in both the longitudinal and lateral direction. In high traffic intensities vehicle cut-in has impact on the vehicle dynamics in both the longitudinal and lateral direction. In high traffic intensities vehicle cut-in situations are more probable, which could impact the LKAS performance.

Road Sensing: The first step for an automated vehicle is to sense, process and interpret its surroundings. This step may employ several different types of sensors (such as camera or LIDAR based), either individually or in combination (i.e. sensor fusion). Irrespective of the method there are a few common factors determining their functioning, and as a result the LKAS performance. These factors include: (1) nearby vehicles creating severe occlusions; (2) shadows from nearby trees, noise barriers and buildings creating misleading edges and texture on the road; and (3) abrupt changes in the illumination level, such as when exiting a tunnel, which impact the clarity of images gathered by the sensors and thereby affecting the accuracy and precision. In addition to this, for vision based systems, glare, bright sunlight, oncoming headlights and improper illumination hampers the detection capabilities of the cameras (such as detection of lane-marking).

Vehicle and LKAS: LKAS are designed to function only under certain speed ranges and this varies between different OEMs and depends on the type of LKAS. For all speeds outside the specified ranges, the LKAS either stops functioning or its performance reduces.

Driver: The driver’s behavior and interaction with the vehicle also affects the LKAS performance. Factors like, whether the driver has his/her hands on the steering wheel and the driving style could affect decisions such as taking back control and disconnecting the system.

C. Assessment of the LKAS Performance

Research regarding lateral driving performance dates back to 1982, when most studies in this domain focused on the effect of pharmaceutical drugs on driving performance by real-road driving tests [9]. Since then primary parameters were used to assess the lateral driving performance including the Mean Lane Position (MLP), the Standard Deviation of Lateral Position (SDLP), and the Steering Reversal Rates (SRR). For example, recently Das et al. [10] examined the effect of fog conditions on the lane keeping performance of manual driving, using ‘lane offset’ as an indicator. Chu, et al. [11] applied this measure as an indicator to evaluate the lane keeping performance of the designed controller responsible for the lateral movement of the vehicle. The SDLP was as well applied in several research studies for evaluating manual driving behavior [12], [13]. It is important to indicate that the SDLP, even though used in several studies frequently, by itself cannot adequately describe the lane keeping performance. This is because a low SDLP could also mean that the vehicle is travelling to the left or right of the lane center without much variations. The SRR is generally defined as ‘the number, per minute, of steering wheel reversals larger than a certain minimum angular value, referred to here as the gap size’, [14]. Other measures used to assess the risk of lane-keeping performance include the Time to Lane Departure (TLD)/ Time to Lane Crossing (TLC), see for example Tarko [15] and Li et al. [12]. Another recently newly developed and applied risk measure is the Probabilistic Driving Risk Field (PDRF), [16], [17]. This is based on the concept that safety, or rather, ‘unsafety’ can be regarded as a combination of risk and its consequence. The PDRF aims to model the objective risk on the road, identifying the ‘potential risk’ from static objects such as lane markings and curb edge, and the ‘kinetic risk’ from moving objects, mainly other vehicles.

In the literature, there are limited studies on the interaction of drivers with LKAS. Most studies focused on drivers’ interaction with the longitudinal assistance system such as Adaptive Cruise Control (ACC) [2], [18], [19]. Nevertheless, there are few studies like Navarro et al. [20], Rudin-Brown and Ian Noy [21], Pohl and Ekmark [22] investigating the influence of LKAS on driver’s behavior, behavioral adaptation, and the interaction between the driver and the system. However, these studies, only consider the ‘warning’ type LKAS and not the more sophisticated ‘control’ type LKAS. Therefore, there is a knowledge gap regarding the factors determining drivers’ interaction with more sophisticated ‘control’ type LKAS. Furthermore, most studies are either questionnaire-based or simulator-based. In other words, the participant does not experience real direct risk in case of automation failure. In questionnaire-based studies, the participants may say that they would trust an automated system, yet act in a way that demonstrates that they do not trust it [23]. From a design
standpoint, it is important to design systems that individuals will trust [24]. The assessment of driving behavior in automation includes understanding the factors that affect drivers’ trust and awareness of the capabilities and functionality (ODD) of the system. Therefore, this raises the need for a method that combines a real-road test with a questionnaire-based approach. Developing such a method could be useful for vehicle manufacturers to assess the ODD of their vehicles equipped with LKAS and for road operators to assess the needed changes in the infrastructure.

D. Research Objective, Research Questions, and Main Contributions

The main objective of this study is to develop an analysis method that combines objective and subjective risk measures for the assessment of the ODD of vehicles equipped with LKAS. To do that the following research questions were defined:

- How does the LKAS perform when it is within and when it is exceeding its pre-defined ODD?
- To what extent can the proposed risk measurement metric be used to determine the objective driving risk across different test situations?
- Is there a mismatch between the ODD specified by the OEM and the one specified by the drivers? and which factors contribute to this mismatch?

The main contributions of this paper are as following:

- Development of a method for assessing the ODD of LKAS equipped vehicles considering objective and subjective risk measures;
- Data collection in real driving environment overcoming the limitations of previous studies based on data from a simulated environment or surveys;
- Focusing on ‘control’ type LKAS while most previous studies focused on studying drivers’ interactions with the ‘warning’ type LKAS.

II. Research Approach & Experimental Setup

To answer the research questions in this study, a research approach composed of a field test (including a pilot test) and a set of questionnaires was developed. As illustrated in Figure 1, we first conducted a literature review on LKAS, the factors that influence their performance, methods to assess their performance and as well the defined ODD in the vehicle owner’s manual. This was followed by the experiment setup which consisted of the selection of test situations of interest, vehicle instrumentation and the preparation of the questionnaires. The test situations were selected so that some are classified within the ODD, some out of the ODD, and the remaining as ODD not in or out. Based on that the test route was selected. After the research plan was approved by the human research ethical committee of the TU Delft, we started with the participants recruitment procedure and test scheduling. This was followed by the pilot test, field test, and data processing. The final step included the data analysis which consisted of the lane keeping performance assessment and the objective and subjective risk assessment. These components of the approach are further detailed in the following sub-sections.

A. Field Test Including a Pilot Test

For the field test several routes were considered taking into consideration several factors for determining the final test route. These factors included: the variation in road curvature, road width, presence and quality of lane marking, number of lanes, presence of on-ramps, off-ramps, tunnels, and bridges, and non-highway route sections. Based on these considered factors the final chosen test route consisted of the A5, Zwanenburg to A4 (E19) Schiphol as illustrated in Figure 2.

A pilot test was conducted prior to the main experiment to examine the appropriateness of the proposed test route with respect to the different ODD classified situations. A test map indicating possible locations that fall within and outside the ODD was made prior to the pilot test. During the pilot test the vehicle instrumentation, such as the sensors and cameras, were tested to make sure they function appropriately. The pilot test was conducted on the 9th of May, 2018 with three participants who were part of the research team. Furthermore, a safety driver from the Tesla car rental organization was present during all the test drives to ensure the safety of the participants and the other road users. The safety driver also explained to the participants how to use the automated system of the car.
**B. Vehicle Instrumentation**

A Tesla Model S was used as the test vehicle. The LKAS function of the Tesla referred to as the Autosteer function v8.1 (218.18.2.301aeee), was therefore the system whose performance across different test situations, was assessed. The vehicle was instrumented with a LIDAR (LIDAR Lite v3), Go Pro’s video cameras, a combined Inertial Measurement Unit (IMU), and a GPS unit. The location of the different sensors and equipment installed on the vehicle are illustrated in Figure 3 together with an explanation of the type of data collected. The synchronization of the data was done using beep sounds coming from the IMU at a specific frequency for 5 s at the beginning of every drive.

**C. Questionnaires**

The participants completed two sets of questionnaires: pre-drive and during the field test.

*Pre-Drive Questionnaire*: included questions regarding the participants’ personal demographics, their initial trust, confidence, attitudes, and prior experience (such as: usage frequency, satisfaction, perceived risk) with semi-automated vehicles and LKAS systems in general. The participants were also asked whether they are familiar with the term ‘Operational Design Domain’ and they were asked to explain it in their own words. Prior to this, explanations regarding what semi-automated vehicles are, and the different types of LKAS (warning versus control) were given.

*During the Field Test*: This part captures the real-time trust of the participants during their drive, and more specifically just after they drove through each of the selected testing situations on a scale ranging from 1 to 5 (with score 5 meaning a very high trust). Following this, they were asked to report if they thought the situation was inside, outside or maybe in/out of the ODD of the LKAS.

**D. Participants**

The participants for the test route were recruited through various channels depending on the targeted driver group.

<table>
<thead>
<tr>
<th>Test situation</th>
<th>Description</th>
<th>Relation to ODD</th>
</tr>
</thead>
<tbody>
<tr>
<td>S1-No-LM</td>
<td>Road inside the city, two – lanes per direction, with no lane marking on its boundary; Speed limit of 50 km/h;</td>
<td>ODD-Out</td>
</tr>
<tr>
<td>S2-tunnel</td>
<td>Single lane tunnel of 3.25m and concrete walls at 4.5m and 2m on the right and left sides, respectively; Well illuminated tunnel and well-marked lane markings; Speed limit of 50 km/h;</td>
<td>ODD-In</td>
</tr>
<tr>
<td>S3-off-ramp</td>
<td>Driving close to an off-ramp on the highway; Speed limit 120 km/h;</td>
<td>ODD-Not In Or Out</td>
</tr>
<tr>
<td>S4-curve</td>
<td>Right-turning curve on the highway (radius=295.5m); Speed limit 120 km/h;</td>
<td>ODD-In</td>
</tr>
</tbody>
</table>

The Tesla experienced drivers were recruited via online advertisements on the Tesla Motors Club forum, while the non-Tesla drivers were recruited via social media and digital mailing lists. All participants were compensated for their time with either a gift voucher of €50 or a dinner invitation.

A total of 19 participants (16 males; 3 females) were recruited with age ranging from 24 to 59 years old (mean=41.32; Std.=12.24), and with a driving license ranging from 3 to 41 years (mean=21.05; Std.=12.77). On average, the vehicle kilometers driven by the participants in the past 6 months was between 1,000 and 60,000 kilometers (mean=15,657; Std.=13,268), and vehicle kilometers in a semi-automated vehicle between 100 and 100,000 kilometers (mean=24,74; Std.=30,74).

**E. Selected Test Situations**

To test the ODD of the Autosteer function the selected situations were classified into three main categories:

1. ODD-In: i.e. situations where the Autosteer is designed to work properly according to the manufacturer;
2. ODD-Out: i.e. situations where the Autosteer is not intended to work;
3. ODD-Not In Or Out: i.e. situations where the Autosteer may or may not function adequately.

The selected situations and their categorization based on the OEM manual of the Autosteer are described in TABLE I:

**F. Procedure**

The test procedure involved the following steps:

1. signing the informed consent form and completing the pre-drive online questionnaire: In order to conduct the field tests on public roads within the Netherlands, an approval from the Delft University of Technology ethical committee was obtained. In addition to this, all participants had to sign an informed consent form prior to their participation in the field test. The form provided the participants with information about the general purpose of the research, its procedure, and about the presence of the on-board safety driver during the field test. The participants were informed about all possible risks and discomforts that they could face during the field...
test, and that the collected data is confidential, and also about their rights to refuse or withdraw from the field test at any time and with no positive or negative consequences.

(2) introduction by the Tesla safety driver about the Autosteer system and familiarization with the Tesla Model S vehicle: A participant instructions protocol was developed and used by the on-board safety driver to provide instructions to the participants before the commencement of the test drives. The protocol was developed based on the driver manual of the Tesla Model S and in consultation with the safety driver.

(3) participation in the field test and completion of the real-time questionnaire: drivers were asked about their trust levels just after they drove through each of the selected testing situations and also if they thought the situation was inside, outside or not in or out of the ODD of the LKAS.

III. ANALYSIS METHOD

The analysis method was composed of an objective and subjective driving risk assessment. The objective driving risk assessment was based on the Probabilistic Driving Risk Field (PDRF) method [16], [17], while the subjective risk assessment was based on driver behavior, trust and situation awareness when driving with the LKAS.

A. Objective Risk Assessment (Potential Risk Field)

In the PDRF method, risk is defined by the magnitude of the consequences of a collision and the chance of its occurrence. The Potential Risk Field (PRF) first proposed and defined by Wang, et al. [25] and later by Mullakkal-Babu et al. [17] is composed of the risks experienced due to the non-moving road entities such as lane marking, guard rails, road medians. Lane-marking, in this case is a virtual obstacle that does not involve in a collision but impose a behavioral constraints on driving. The crash severity is then based on the crash energy transferred during a possible collision between the subject vehicle and a road entity. Accordingly, the risk taken by vehicle $s$ due to fixed road boundary object is formulated as:

$$ R_{b,s} = 0.5kM \left( V_{s,b} \right)^2 \cdot \max \left( e^{\left( \frac{r_{s,b}}{k} \right)}, 0.001 \right) $$

where $R_{b,s}$ denotes the risk to subject vehicle $s$ due to road boundary $b$, and $r_{s,b}$ is a vector that denotes the shortest distance between $s$ and $b$. The first part of the equation $0.5kM \left( V_{s,b} \right)^2$ describes the physical crash energy in case of an inelastic collision between $s$ and $b$. $M$ denotes the mass of $s$, and $V_{s,b}$ denotes the velocity of $s$ along $r_{s,b}$, $k$ is a parameter to scale the expected crash energy with respect to the type of road boundary object. This parameter has a range [0-1], representing the level of rigidity of the road boundary object. For further information on how to scale this parameter please refer to [17].

The second term in the formulation, $e^{\left( \frac{r_{s,b}}{k} \right)}$, constitutes the probability of the crash, and receives values between [0-1]. This term equals 1 when $r_{s,b} = 0$. As the distance between $s$ and $b$ increases the probability of the crash decreases. Intuitively, a road object further away offers more possibility for the driver to avoid the collision. $D$ is the coefficient that determines the steepness of descent of the potential risk field. The risk due to a road boundary decreases as one moves away from the road boundary object. In this study we have chosen a value of $D = W/14$, where $W$ is the lane width, indicating that the collision probability term attains a marginal value (0.00091) at the lane center. However, crashes are not always perfectly inelastic and roadside objects allow finite deformation thereby absorbing some amount of crash energy and decreasing the inflicted crash severity. This assumption is consistent with the empirical studies by Zou, et al. [26] where it was shown that the odds of injury due to collisions with a guard rail is lower than that with a concrete median barrier and a concrete wall.

Being based on artificial field theory, the PDRF method possess several benefits in comparison to other surrogate safety measures ([16]). First, the PDRF accounts for the combined risk in both the longitudinal and the lateral directions, while other risk measures, such as the Time to Collision and the Time to Lane Crossing account for the longitudinal or the lateral risk, respectively. Secondly, the PDRF method is applicable to a wider range of evaluation scenarios and driving conditions. This could result in a more informative and realistic representation of driving risk.

The objective risk measure could be very useful for OEMs and road authorities to assess the implications of different mismatches on safety and based on that prioritize their actions and interventions to reduce the mismatches that lead to the highest objective risks.

B. Subjective Risk Assessment

It is important to examine if the objective risk significantly affects drivers’ perceived risk and trust in vehicles equipped with LKAS. This is important especially for the situations in which it is safe for the LKAS to function [27], [28].

Several questionnaire-based studies have identified various factors that affect driver’s trust in automation. Jian et al. [23] identified factors such as predictability, reliability and dependability which impact trust in automation. Drivers’ trust varies dynamically, changing over time [29]. Past experiences and knowledge about the system influence the initial trust while drivers’ interaction with the system influences their dynamic trust. Therefore, in the underlying research, drivers’ trust and awareness about system’s capabilities were measured/reported before and during driving in the Tesla Model S on the test route. The trust questionnaire used in this study was designed based on the three layers of variability in human-automation trust (dispositional trust, situational trust, and learned trust) defined by Hoff and Bashir [30].

IV. RESULTS

We first present the characteristics of the recruited participants for this experiment, followed by the results of the LKAS performance, the objective and subjective risk measurements and finally the ODD assessment.

A. Participants

Out of the 19 participants, 17 participants had previously used a control/intervention type LKAS and all of them had
prior experience in the warning type of LKAS as this was a recruitment criterion. Furthermore, 9 participants reported that they always use the LKAS while they drive, 4 usually use it, 2 sometimes use it, and the remaining 4 reported that this question was not applicable to them. 5 participants took part in prior on-road test experiments, and 7 participants were aware of the term ‘Operational Design Domain’.

B. LKAS Performance

To assess the performance of the LKAS the lane position data, extracted from the processed images, were analyzed across the different selected situations (S1-No-LM (no lane-marking on road boundaries in the city), S2-tunnel, S3-off-ramp, and S4-curve). The objective was to understand how the system performs inside and/or outside the OEM specified ODD, and where the vehicle aligns itself (with respect to the lane center). Data corresponding to the first 15 seconds in each situation of each drive was selected for the assessment. The comparison of the LKAS performance between different test situations was first done using boxplots (see Figure 4). This was followed by pairwise comparisons.

C. Objective Risk Measurement

As expected, for the situations ODD-In, i.e. situations S2-Tunnel and S4-Curve, the vehicle was closer to the lane center as compared to the other situations (Figure 4(a)). Between these two situations, the relative positions from lane center were more skewed to the right in S2-Tunnel than in S4-Curve. This could be explained by the fact that in S2-Tunnel there were walls on both sides of the tunnel, with the wall on the left side being closer to the lane compared to the wall on the right side. This could explain why the LKAS positioned the vehicle away from the wall on the left side. The slight left skew in the curve situation could be explained by the presence of the guard rail on the right side of the lane.

Driving close to the off-ramp (S3-Off-ramp, ODD-Not In Or Out) resulted in the highest variance of the relative positions from lane center compared to the other situations (Figure 4(b)). Furthermore, the relative positions are slightly skewed to the left which can be explained by the unrecognized changes in the lane marking type close to the off-ramp.

While driving inside the city with no lane marking on the road boundaries (S1-No-LM, ODD-Out), it was observed that the LKAS positioned the vehicle closer to the road center. This can be explained by the fact that when the LKAS recognises that there is no lane marking on the right side of the lane, it positions the vehicle closer to the road center.

The Probabilistic Driving Risk Field (PDRF) metric was used to determine the objective driving risks across all the different test situations. In this study we have only considered the Potential Risk Field (i.e. the risks due to non-moving fixed road entities) and a duration of 15 seconds of each test drive. The front left wheel of the subject vehicle was considered as the reference point to determine distances between the subject vehicle and the road entity. The features of the road boundary determine the level of risk it poses to road users [26]. The parameter, $k$ in Equation 1 is used to differentiate the risk posed by each road entity. Four types of road boundaries (non-moving road entities) were considered in the calculation of the risk, with the following values of ‘$k$’: $k=0.1$ for a lane marking strip; $k=0.2$ for a curb stone in the city; $k=0.5$ for a concrete median on the highway; $k=0.7$ for a concrete wall inside the tunnel. Accordingly, lane marking strips pose the least and concrete walls pose the highest risk to driving.

The potential driving risk field for any time instance of a situation, is the sum of the risks due to all the boundary types surrounding the subject vehicle at that time instance. The maximum risk during the 15s in each situation was used as the representative value. The average risk values were not used because in general, for the test duration of 15s, these were equal to zero in most situations. The results are shown in Figure 5.

The objective risk is expressed in Joules and a higher value represents a higher risk. As shown in Figure 5 the mean objective risk is the highest in S1-No-LM and decreases in the
order S1-No-LM > S2-Tunnel > S3-Off-ramp > S4-Curve. The corresponding median values also show a similar trend but with lower magnitudes. This also means that there is deviation in the median and mean values in all the situations and thereby an asymmetry in the objective risk values within each situation. The asymmetry is also observed in the unequal whisker lengths in all situations except for S4-Curve. This means that the risk values are more asymmetric towards values larger than the median. The non-parametric Friedman’s test showed a statistically significant difference in the means of the Objective Risk values, \( \chi^2(3)=17.925, p < 0.001 \). Post-hoc analysis with Wilcoxon signed-rank test with a Bonferroni correction showed a statistically significant difference between S1-No-LM and S4-Curve (\( Z = -3.206, p = 0.001 \)), between S2-Tunnel and S4-Curve (\( Z = -2.534, p = 0.011 \)), and between S3-Off-ramp and S4-Curve (\( Z = -2.482, p = 0.013 \)).

D. Subjective Safety Evaluation

The main aim of this analysis was to investigate if there is a mismatch between the awareness of drivers about the ODD of the LKAS (i.e., inside ODD, outside ODD, or not sure) and the ODD functionality as defined by the OEM.

1) Pre-Drive Subjective Safety: 10 of the 19 participants reported that they had prior negative experiences while driving with their LKAS ON. A moderate positive correlation was found between ‘Awareness of the conditions in which LKAS can function’ with ‘Having prior knowledge about ODD of LKAS equipped vehicles’, \( r_s(16)=0.59, p=0.016 \); and with prior negative experiences, \( r_s(16)=0.60, p=0.014 \). These correlations should be treated with caution because of the relatively small sample size.

2) Real-Time Trust and ODD Awareness: The average trust across all the four situations was found to be relatively high, 4 on a scale from 1 (very low trust) to 5 (very high trust), with a std.=0.88. A Friedman test showed that there is no significant influence of the time spent in the vehicle on the real-time trust of drivers across the different situations (\( \chi^2(3)=3.418, p=0.332 \)). No statistically significant correlation was found between the real-time trust ratings of drivers and their ODD state awareness. Furthermore, no statistically significant correlation was found between the real-time trust across all test situations and the participant’s experience of driving in a Tesla.

The analysis of the mismatch between ODD state awareness of driver and ODD specified by OEM revealed that the highest percentage of mismatch was in S3-Off-ramp (81.2% mismatch), followed by S1-No-LM (68.7%), S2-Tunnel (12.5%), and S4-Curve (6.25%). Interestingly, none of the drivers had a correct ODD state awareness in all of the four situations. A Cochran’s Q test [31] with Bonferroni correction revealed that the ODD mismatch of drivers is significantly different between the four different test situations (Q(3)=24.6, \( p<0.0001 \)). It is important to avoid such mismatches as this could lead to risky situations.

E. ODD Assessment

The ODD of the LKAS was assessed in the different selected situations based on the objective risk measurement (i.e. potential risk field), and based on the subjective risk measurement (i.e., drivers’ questionnaires). The results are summarized in TABLE II.

1) Inside the ODD Situations: The lane keeping performance was best in S4-Curve followed by S2-Tunnel. S4-Curve also had the smallest variation in the lateral position. The relative distances of the concrete tunnel walls from the lane boundaries affected the vehicle location within the lane. The vehicle drifts away from the closest wall, in our situation that was the left wall. Drivers’ real-time trust in the system in this situation was negatively correlated with their perceived risk of the situation. This might explain the relatively low ODD mismatches (12.5%).

2) Neither Inside Nor Outside the ODD Situation: The vehicle lane positions had the highest variation and deviations relative to the lane center with a slight bias to the left (away from the off-ramp). The lane keeping performance was therefore considered the poorest in this situation. The mismatch was also the highest (82.8%) and the majority of drivers (77%) believed that the vehicle was inside its ODD.

3) Outside the ODD Situation: This situation (S1-No-LM), as TABLE II shows, had the highest bias in the lane position towards the left of the lane center, and with considerable variation. Therefore, the lateral objective risk of driving in this situation was the highest. The majority of drivers believed that this situation was inside the ODD of the vehicle, and therefore, the second highest mismatch between drivers’ perception of ODD and OEM specified ODD (68.7%). One possible explanation for this result is the fact that the system could always be switched ON. Therefore, to ensure that drivers understand the capabilities of the system better, the OEMs must either not allow for the system to be activated in these situations, or have a better form of communication with drivers regarding the system possible decrease in performance. 62.5% of the drivers reported that they would have trusted and used the system more if timely information about its capabilities was provided to them.
V. Conclusion and Future Research

We first present a summary and main conclusions from this study followed by the limitations and future research directions.

A. Summary and Conclusions

In this research, an analysis method to assess the Operational Design Domain (ODD) of Lane Keeping Assistance System (LKAS) equipped vehicles was developed. The performance of the LKAS across different ODD-classification types (ODD-In, ODD-Out, ODD-Not In Or Out) was first assessed based on the mean and standard deviation of the vehicle position from the lane center, followed by the calculation of the lateral driving risk based on the Probabilistic Driver Risk Field (PDRF) metric. The results are based on data collected from a field test on a public road under real traffic conditions of which 19 drivers took part, each driving 55 km (about 75 minutes) covering different road types and situations. The main conclusions from this study are as following:

• For ODD-In situations the deviation from the lane center was the smallest, while for ODD-Out, it was the biggest among the three categories. For ODD-Not In Or Out (in this situation off-ramp), even though the mean deviation from the lane center was relatively small, the variation of the deviation was the highest. This can be attributed to the variation in lane marking types on the right side of the lane center.
• Based on the PDRF metric clear differences in the risk of driving between the different situations could be identified and possible explanations for these were given. For example, the maximum risk was measured to be the highest in the situation that was ODD-Out (driving in the city with no lane marking on the road boundaries). On the other hand, even though driving inside the tunnel was considered ODD-In, the lateral risk in this situation was the second highest. This can be attributed to the type of road barrier, i.e. the concrete tunnel wall. In comparison to guardrails, the consequence of a crash with the concrete wall is relatively higher. However, it should be noted that these differences in risk might change if the conflicts with adjacent vehicles are considered.
• Across all test situations, a mismatch was observed between the ODD specified by the OEM and by the drivers. This mismatch can negatively affect the establishment of a correct mental model which is crucial for effective and safe interaction of the drivers with automated systems, such as the LKAS [32]. A higher mismatch was observed in both situations ODD-Out (i.e. no-lane markings) and ODD-Not In Or Out (i.e. off-ramp). Drivers mostly reported these situations to be inside the ODD. This type of mismatch (false-positive ODD) is more dangerous intuitively and also as observed by the PDRF. Therefore OEMs may be recommended to minimize such mismatches by proactively informing/warning the driver. On the other hand, the mismatch in situations that were ODD-In (i.e. tunnel and curve) were found to be minimal. In such situations, the OEMs may focus on improving the LKAS performance.
• Driving experience in a Tesla did not have significant impact on the mismatch. This indicates that regardless of having experience of driving in semi-automated vehicles, increasing drivers’ awareness of the automated system capabilities is still very important. Therefore, training and increasing drivers’ knowledge and understanding of the LKAS capabilities and in which conditions it works appropriately could support drivers in building correct mental models and well-calibrated trust [33], [34]. Using this assessment method, it was only possible to compare the test situations with each other, and not make decisions regarding the inclusion/exclusion of situations from the LKAS’s ODD. This is because, acceptable threshold values for each assessment component (i.e. maximum acceptable risk) vary between vehicle manufacturers and is confidential information. Therefore, this method has the potential to aid OEMs in deciding if a situation should remain inside or moved outside the lane keeping system’s ODD while keeping the drivers’ safety and awareness of the system capabilities at the core of the decision-making process.

B. Limitations & Future Research Directions

This study has several limitations:

• The first limitation concerns the sample size, with 19 participating drivers of which only 3 were female. Therefore, we could not analyze differences in trust and ODD perceptions as a function of gender. Due to the relatively small sample size also the impact of age was not examined.
• A second limitation is the consideration of only static road entities in the determination of the risk, i.e. the potential risk field. The kinetic field risk due to other moving objects (such as adjacent moving vehicles) also plays a very important role in the determination of total lateral risk assessment. Especially for the case of driving close to off-ramps and on-ramps, adding the kinetic field risk could result in different ranking of the risk level for the different situations. It is recommended in future studies to account for the kinetic field risk to increase the realism of the risk measurement.
• A third limitation is the potential errors in lane position stemmed from image processing. Average errors of 3.5% (highway) and 4% (city) were identified in the values of the lane positions. This largely depends on the camera’s image resolution, their calibrations and the applied algorithms. Future research could explore ways to reduce these errors.
• A fourth limitation relates to other factors that might affect the definition of the ODD such as weather and time of day which are important to be considered in future studies.

This assessment method could also be used to study variations among different vehicles’ LKAS performance within the same SAE level 2. The results could accelerate harmonization...
among different vehicles manufacturers with respect to a more detailed definition of the ODD of each automated system.

ACKNOWLEDGMENT

The authors would like to thank Mr. Tom Alkim, at that time senior advisor C-ITS & Automated Driving Rijkswaterstaat, and currently Policy Officer (SNE) Connected & Automated Driving at European Commission for financially supporting this study and partially funding the field test.

REFERENCES


Shuhham Bhusari received the M.Sc. degree in civil engineering with a specialization in transport & planning from the Delft University of Technology. He works as an Innovator and a Product Manager of the Department of Transport & Planning, Royal HaskoningDHV. Given his interests in the field, he works on projects for several provinces, municipalities in The Netherlands and Interregional projects such as the Interregional Automated Transport funded partially by the European Union. These projects, amongst others, involve the assessment of road infrastructure and its surroundings and the behavior of drivers in the car for the efficient functioning of advanced driver assistance systems and their testing. In addition to this, he also works on the development of in-house products. These products include the collection of data from several sources such as in-car, satellites and image data, as well as using it for road inspections and driver behavior monitoring.

Paul van Gent is currently a Ph.D. Researcher with the Transport & Planning Department, Delft University of Technology. He is interested in how the position and role of the human component changes in an ever more automated world. Accommodating these changes, so that the increased automation actually benefits humans and is safe, brings its own sets of challenges that he loves working on. His recent work focuses on applying deep learning and machine learning methods to improve the interaction between drivers and smart in-car systems. He has worked on driver persuasion, developing driver monitoring software, and driver state classification. Before joining TU Delft, he worked at the Dutch Road Safety Research Institute (SWOV), where he developed and implemented sensing equipment for use in naturalistic (instrumented) studies on both cars and bikes.

Freddy Antony Mullakkal Babu received the bachelor’s degree in civil engineering from the National Institute of Technology, Calicut, in 2011, and the master’s degree in transportation systems engineering from the Indian Institute of Technology, Mumbai, in 2014. He is currently pursuing the Ph.D. degree with the Department of Transport and Planning, Delft University of Technology. He is also an Application Engineer at Siemens PLM Software, since 2019. His main research interests include driver behavior modelling, traffic safety and design, and assessment of driving automation systems.

Peter Morsink is currently the Project Manager and a leading professional in road safety and smart mobility with the Royal HaskoningDHV, since 2008. Before joining Royal HaskoningDHV, he worked at the Dutch research institute TNO Automotive for six years, and at the SWOV Institute for Road Safety Research for five years. He is a mechanical engineer, graduated in biomechanics and computational mechanics. Since 1998, he has developed into all-around road safety and smart mobility consultant through a wide range of assignments on different types of traffic and transport-related research projects. He has gained a lot of international experience in executing and managing research and consultancy projects in the field of vehicle and infrastructure safety, (inter)national road safety policy, and intelligent transport systems (ITS). In these areas, he has been active as a lecturer in training programs and he has written many articles. He is currently involved in many projects on road safety and traffic management from different perspectives (infrastructure, technology, autonomous vehicles, education), on both a strategic and operational level, in The Netherlands and abroad.

Riender Happee received the M.Sc. degree in mechanical engineering and the Ph.D. degree from the Delft University of Technology, in 1986 and 1992, respectively. He investigates the human interaction with automated vehicles focusing on safety, comfort, and acceptance at the Delft University of Technology, The Netherlands, where he is currently an Associate Professor at the Faculties of Mechanical, Maritime and Materials Engineering, and Civil Engineering and Geosciences. His key projects include HF/Auto (coordinator), MOTORIST, SHAPE-IT, and HADRIAN. He investigated road safety and introduced biomechanical human models for impact and comfort at TNO Automotive (1992–2007).

Bart van Arem received the M.S. and Ph.D. degrees in applied mathematics from the University of Twente, Enschede, The Netherlands, in 1986 and 1990, respectively. From 1992 and 2009, he was a Researcher and Program Manager at TNO, working on intelligent transport systems, in which he has been active in various national and international projects. Since 2009, he has been a Full Professor of intelligent transport modelling with the Department of Transport and Planning, Faculty of Civil Engineering and Geosciences, Delft University of Technology, Delft, The Netherlands. His research focuses on modelling the impact of intelligent transport systems on mobility. He serves as the Editor-in-Chief of the IEEE OPEN JOURNAL OF INTELLIGENT TRANSPORTATION SYSTEMS.