A Methodology for the assessment of Operational Design Domain for lane keeping system equipped vehicles: The case of Tesla Model S

MSc Thesis: Shubham Bhusari
Transport & Planning, Faculty of CEG, TU Delft

TU Delft

Royal HaskoningDHV
Enhancing Society Together

A Methodology for the assessment of Operational Design Domain for lane keeping system equipped vehicles: The case of Tesla Model S

by

Shubham Bhusari

Defended on

25 October 2018

Thesis submitted in partial fulfilment of the regulations for the degree of Master of Science at the Delft University of Technology, Faculty of Civil Engineering and Geosciences, Department of Transport and Planning.

Student Number
4592964

Graduation committee

Chairperson	Prof.Dr.ir. Bart van Arem	TU Delft, Faculty of Civil Engineering & Geosciences, Transport & Planning
Daily Supervisor	Dr.ir. Haneen Farah	TU Delft, Faculty of Civil Engineering & Geosciences, Transport & Planning
External Supervisor	Dr.ir. Riender Happee	TU Delft, Faculty of Mechanical, Maritime & Materials Engineering
Company supervisor	Ir. Peter Morsink	Royal HaskoningDHV

Preface

The basis for this research originally stemmed from my passion towards development and the need for automated vehicles on the road. The advent of such vehicles has several barriers but most importantly it must not compromise the safety of its users and the other road users. It is my passion to not only find out, but to develop tools which could provide important steps towards further development in this research domain.

Successful completion of the underlying thesis project would not have been possible without the support and assistance of a group of people. I would like to thank Professor dr.ir. Bart van Arem for his guidance and help during all stages of this project. Especially at the earlier stages, his support in finding a company to graduate at, was monumental. Furthermore, I am grateful to my daily supervisor dr.ir. Haneen Farah, for her prompt and informative feedback during the whole thesis project. She was always available for my untimely short visits to her office and I heartily appreciate her contribution to this project even during her vacation breaks. A very special thanks goes out to my external supervisor from Royal HaskoningDHV, ir. Peter Morsink, who left no stone unturned in making sure that the thesis project went according to schedule and for trusting in my capabilities to carry out this project given its financial requirements. His practical knowledge and down to earth attitude is something that kept me motivated while conducting this project. A special gratitude goes out to dr.ir. Riender Happee for his feedback and support during the thesis. His concise and vital suggestions helped me in making important decisions during the project. I would like to thank the whole thesis committee for devoting their time to review my report.

Furthermore, I would also like to thank ir. Jork Stapel (PhD candidate at TU Delft) for his insights and great ideas that helped me while setting up the real-road tests for this project. Data collection and processing would not have been possible without the dedicated and continuous support of two other PhD candidates MSc. Paul van Gent and Freddy Antony Mullakkal Babu. I am grateful to Paul for taking out time from his busy schedule to help me with the instrumentation of the Tesla and post processing of the collected data and, to Freddy for allowing me to implement the driving risk measure currently under development by him and for always being there to explain the nitty gritties of his method. During the road test, Daud Pechler (safety driver from Tesla rental company) also helped with last minute changes to the instruments and helped in ensuring that all the test drives went smoothly and all the participants were safe.

I am indebted to my family, Dhananjay (my dad), Sadhana (my mom) and my brother Karan for supporting me in every possible way throughout the thesis and my life as a student in Delft and even before that in India. They were very patient and provided me continuous love and support which gave me strength and confidence to overcome problems I faced during my studies in Delft. Without them my entire quest for knowledge would not have been a reality.

Last, but not least I would like to sincerely thank my friends and roommates here in Delft and from back in India. They were always there to listen to my (interesting) stories and complains during the thesis and during my stay in Delft, in general. Especially, during the last months of my work my roommates (specially Ivo Hindriks) had to bear the direct consequences of my thesis stress and listen to my rather unnecessary technical explanations and ambitions, I thank them for this.

At this point, I will let my research unfold to you, dear reader.

Executive Summary

Background and Research Objective

Vehicle automation is introduced at different levels, and categorized following the Society for Automotive Engineers to six levels, ranging from no-automation (level 0) to full automation (level 5). Each level of automation is designed to work in specific conditions referred to as the Operational Design Domain (ODD); this includes geographic, environmental, traffic, road geometry, speed and/or temporal dimensions. This research focusses on SAE level 2 automated vehicles (i.e. partial automation) and specifically, on its Lane Keeping Assistance Systems (LKAS - assistance in lateral vehicle control). At this level, the execution of both the lateral and longitudinal vehicle motion control tasks are performed by an automation system with the expectation that the driver supervises the driving automation system. A well-researched problem at this level of automation, is the task of transition of control from the system to the driver when the automated system cannot handle the driving situation. Drivers of level 2 vehicles are responsible to monitor the driving environment at all times. If the ODD is not accurately defined and made clear to the drivers, this could make the drivers more vulnerable to accidents, as they could have misconceptions about capabilities of the system.

Situations where the driver needs to take-over control from the system are defined by the Original Equipment Manufacturers (OEM's)/vehicle manufacturers in their owner's manuals. As each manufacturer specifies their own operational conditions and constraints, this could result in uncertainties about the capabilities of different vehicles within the same level of automation. This is a problem, because generally very few drivers read owner manuals and are therefore, unaware of their own vehicle's functional constraints, let alone that of other vehicles. In addition to this, another problem could be a mismatch between the drivers' understanding of the capabilities of the automated vehicle and its actual capabilities. This mismatch could also lead to serious situations and road accidents.

Furthermore, an important dimension of the ODD is road type, including its geometry and design. Closeness to non-moving road entities (barrier, lane markings) and other vehicles has an impact on accident risks while driving. These risks can be quantified and their magnitude reflect on the performance of the driver assistance systems.

Therefore, it is believed that there is a need for a methodology to assess the ODD of semi-automated vehicles that are currently already available in the market. Development of such a methodology forms the *aim of this research*. This methodology could help vehicle manufacturers in making drivers more aware about the situations in which their semi-automated vehicles can/cannot be used, thereby, increasing their acceptance and trust in driver assistance equipped vehicles over time. To achieve this objective the following main research question is proposed:

To what extent can the Operational Design Domain of vehicles equipped with lane keeping systems be assessed by understanding the subjective and objective risk of driving in pre-specified test situations?

To answer this question, first the *objective vehicle-related aspects* of ODD assessment are measured, followed by the *subjective driver-related aspects*. Assessing performance of the LKAS inside and outside its ODD, and measurement of lateral driving risks in selected situations, constitute the objective aspects. Understanding divers' attitude and response towards LKAS at different stages of testing, constitutes the subjective aspect. These two aspects are assessed using a real-road case study of a Tesla Model S.

Research Methodology

The assessment methodology is implemented using an approach that combines literature study, real-road tests with instrumented Tesla Model S, and driver behavior surveys at three stages of testing (before, during and after the road tests). This methodology assesses the Autosteer system of the Tesla in specific situations. These situations are identified by analysing the Tesla Model S owner's manual and classifying situations where the system is intended to work (*Inside the ODD)*, not intended to work (*Outside the ODD)* and where it may or not work as intended (*Neither inside nor outside the ODD)*.

Next, one out of three candidate routes from within the Netherlands, was selected as the test route for the real-road tests. Participants for the tests are recruited based on specific criteria such as their age (between 25 and 60years) and having prior experience of driving in vehicle equipped lane keeping assistance systems, the Tesla Model S is instrumented to gather the research specific data, survey questionnaires are developed and final ethical requirements are fulfilled before conducting the tests. The final selected situations are: (S1) Inside the city with no lane marking strips on the road boundaries (*Outside the ODD);* (S2) Inside a tunnel within city *(Inside the ODD);* (S3) Close to an off-ramp on the highway *(Neither inside nor outside the ODD);* (S4) On a curve (right turning) on the highway *(Inside the ODD)*.

These situations are tested on a route which starts and ends at the parking lot of Den Ruygen Hoek-Oost, Rijsenhout. The test route first, includes a part in which the drivers are familiarized with the ADAS system of the Tesla by the on-board safety drive. After familiarization, the drivers then drive first on the highway and then in a short city road section and then back on the highway, towards the end of the test route. Each driver is requested to fill in a questionnaire before their drive about their initial attitude towards LKAS and automated vehicle in general, respond to specific questions during their drive on their trust and awareness of the capabilities of the LKAS of the Tesla, and fill in a questionnaire after their drive which is focussed at understanding drivers' behaviour in specific test situations. The gathered data is then processed using image processing techniques and a data visualization tool is developed to help during the analysis phase of the research.

The lane keeping performance is assessed using the Mean Lane Position and the Standard Deviation of Lane Positions (SDLP) metrics, and the lateral risks are measured using a novel metric, referred as the Probabilistic Driver Risk Field metric. These measurements form the objective part of the methodology. On the other hand, drivers' behaviour and responses towards the LKAS system, are assessed using a statistical hypothesis based approach with an aim of identifying and investigating reasons for mismatch between drivers' perceptions of the system's capabilities and its actual capabilities. This forms the subjective part of the assessment methodology.

Results: Lane keeping performance of LKAS

The mean lane positions in the two 'Inside ODD situations' showed that in the tunnel (S2), the Tesla is slightly closer to the lane centre than on the curve. There is a significant right bias in its mean positions confirming that the Tesla is attempting to move away from the left lane marking strip as it is closer to the left side concrete tunnel wall. Similarly, in the situation with no lane marking on the road boundaries (S1), the Tesla aligns itself with a left bias closer to the left lane marking strip, away from the road edge. Close to the off-ramp (S3), the mean position of the Tesla is closer to the lane centre, than in (S1). This is attributed to the lack of a lane marking strip on the road boundary in S1, forcing the Tesla to bias its position closer to the road centre. Moreover, as expected, the standard deviation of lane positions is lower in the situations 'inside the ODD' compared to the two other situation types. Between the other two types of situations ('Outside the ODD' and 'Neither inside nor outside the ODD'), standard deviation in lane position is higher close to the off-ramp compared to the no lane marking situation. Based on this, it is concluded that the lane keeping performance of the Autosteer 'Inside the ODD' is better than the performance when it is 'outside the ODD' and 'Neither inside nor outside the ODD'. The performance is slightly better in the situation with no lane marking (S1) compared to driving close to an off-ramp (S3), as there is larger range of variation in standard deviation of lane positions in (S3). It is also concluded that the Autosteer predominantly attempts to move away from road edges, guard rails and tunnel walls, but at the same time moves closer to other road traffic which could lead to unsafe situations.

Results: Lateral driving risk measurements

Using the field theory based approach referred as Probabilistic Driving Risk field method, it is found that the lateral driving risk across the different situations is in the order (S1)>(S2)>(S3)>(S4). This was different from what was expected based on the lane keeping performance across the different test situations. There is a large difference between risks in S1 (outside the ODD) and S4 (Inside the ODD). It is observed that the maximum-minimum range of the objective risk values also follow the same order (S1 through to S4) across the situations and there is considerable skew in higher risk values in S1, S2 and S3 but not in S4. Relatively, S4 (right turning curve on the highway) is measured to be the safest situation to drive in. In this research, to interpret and give reasons for the measured risks, a relationship between lane keeping performance indicators (mean and standard deviation of lane positions) and the components of the risk measurement metric (severity and probability of collisions), is established.

Moreover, the risk field approach used in this research is also compared to risk measurements using time to lane crossing metric. The results indicate that both metrics depict the same trends in risk as they are both dependent on the lateral distance to road barriers, but differ in the realism of the magnitude of the risk they represent. The risk field approach shows advantages in terms of sensitivity to different road barrier types and shows additive properties (risks due to different road entities in both lateral and longitudinal direction can be represented as one risk measurement).

Results: Driver behavior in LKAS equipped vehicles

Statistical results show that there are factors such as perceived risk of driving, frequency of using LKAS and perceived ease of driving which correlate with the drivers' real-time trust and ODD perception whilst driving in the vehicle. Most importantly, there are mismatches between driver's perception and vehicle manufacturer's specification of the capabilities of the Autosteer.

In terms of number/percentage of mismatches, maximum mismatches (81.2% of drivers) were observed in the ***'Neither inside nor outside the ODD'*** situation of driving close to an off-ramp. In this situation, most drivers (77%) believed that the Tesla is inside its ODD. Next highest mismatch (68.7%) was seen in the '*Outside the ODD'* situation of driving in the city with no lane marking (S1). In this situation, most drivers (82%) believe that the Tesla is inside its ODD and very few drivers are not sure about it. This mismatch could lead to very dangerous situations as drivers might not be completely ready to take over control from the LKAS system. In both situations *'Inside the ODD'*, there were very less mismatches. It is also found that across the different test situations, factors such as drivers' real-time trust, their perceived risk of driving in a situation, their initial trust in AV's in general, are possible factors that contribute to such mismatches and therefore, provides a first indication of which factors to focus on to avoid these mismatches.

ODD assessment for Tesla Model S

After analyzing both the objective and subjective aspects of the ODD assessment methodology, each of the test situations are assessed. While driving close to an off-ramp, which is neither inside nor outside the Autosteer's ODD, lane keeping system performs the poorest, but most drivers think this situation is inside the ODD. This shows the need for a better form of communication between the vehicle and the driver and/or changes in road design (like having a single lane marking strip instead of multiple strips leading to the off-ramp). On city roads with no lane marking on its boundaries (outside the ODD of Autosteer), the lane keeping system also performs poorly (but better than while driving close to an off ramp) and most drivers believe that the vehicle is inside its ODD. Given these results, the Autosteer is still allowed be switched ON in this situation, making the driver more vulnerable to road accidents. Therefore, it is concluded that either the Autosteer must not be allowed to be turned ON in this situation, or that drivers must be given more information about the constraints of the Autosteer's functionality in this situation or make them aware of the risks of driving in this situation (on the on-board LCD screens).

While driving inside the Tunnel (inside the ODD), lane keeping performance is good but there is a skew away from left lane marking strip, this results in higher lateral driving risks as the lateral velocity of the vehicle is high thereby increasing the severity of collisions with the barriers. Therefore, it is advised not to attempt at swaying away from the left strip, inform the driver of the risk and the vehicle to maintain its mean position. Furthermore, driving on the curve has the best lane keeping performance, the least lateral driving risk and minimum mismatch between driver and OEM specified ODD. It is not possible to decide if any of these situations must be included or excluded from the ODD, as the thresholds for lane keeping performance and driving risks are unavailable due to OEM confidentiality.

Finally, it is important to keep in mind that there are cross correlations between drivers' perception of risk and trust on the system, while making infrastructural changes to reduce risks in either of the situations that are not inside the ODD.

Main contributions

This research makes scientific and practical contributions. It fills existing gaps in scientific research as, it involves a successful implementation of a novel potential driving risk field method for risk measurement for real-road experiments. This is useful for the calibration of the components of this method which is still under development at TU Delft. It also identifies relationships between drivers' attitude and response towards LKAS across different testing stages using a combined real-road test and questionnaire based approach, which is also a relatively lesser used approach in the scientific community (most researches are either simulation or survey based). Furthermore, there is limited or no literature that attempts at development of a methodology for assessing ODD for vehicles at any level of automation, this research mainly fills this scientific gap.

Practically, this research provides a road environment sensing tool using which, vehicle dynamics data (acceleration, speed in both lateral and longitudinal direction), vehicle co-ordinates and geometry, distances to surrounding moving and non-moving road entities, can be determined in continuous motion. This could be of great value to Royal HaskoningDHV in future mobility related projects.

Limitations and next steps

This research has scientific and technical limitations. 1) Main technical limitations: Autosteer function of the Tesla cannot be turned on without the Adaptive Cruise Control (ACC) function of the vehicle also active. This can have an impact on the drivers' perception of risk and their trust on the vehicle in general. Furthermore, between the first two and the next two test days, two different versions of Tesla Model S (60D and 90D) were used (but they had the same software version), this could also have an impact on the vehicle's performance across the different situations.

2) Main scientific limitations: There is a maximum error of 16% and an average error of 3.5% (highway) and 4% (city), in the image processed lane positions, which is another limitation of the research. Moreover, the sample size for the statistical tests is limited, leading to an increase in chances of type 1 errors (showing significant results when they are not) in the statistical tests conducted in this research.

Next steps from this research include, extension of the risk field approach so it also includes risks due to other moving vehicles on the road. This could further increase the realism of the risk measurement. Furthermore, driver facial videos collected in this research, can be used for research into understanding drivers' psychological mindset across different road situations within semi- automated vehicles. This method can also be extended to other ADAS and eventually be used by vehicle manufacturers for a complete assessment of situations before deciding ODD. Finally, Infrastructure developers can use the driving risk measurements to identify hotspots (highly risky road sections) and test infrastructure changes to reduce these risks.

Table of Contents

List of Figures

List of Tables

Chapter 1. Introduction

In this chapter, background information and motivation for the conduction of this research is provided. It first introduces the main problems that require attention and then possess research questions, answers to which, fulfil the objectives of this research. It then briefly describes the research methodology adopted to answer the posed research questions followed by, a detailed overview of the scope and scientific relevance of this research. Finally, it describes the structure/ outline of this thesis report.

1.1. Background and Motivation

There are uncertainties about the advent of automated vehicles in the eyes of the public as well as the experts within the automotive industry. Recent fatal accidents such as [1] in which an Uber test vehicle collided with a pedestrian crossing a city street with a bicycle and [2] in which, just seconds before a fatal crash, a Tesla sped up and steered into a concrete barrier as it stopped following the path of a vehicle in front of it; have only added to these uncertainties. In the former case, the cause of the accident was investigated to be due to the driver's inattention. In the latter case, Tesla noted that the driver had received multiple warnings to keep their hand on the wheel, and did not intervene in the seconds before the accident. Both these accidents indicate that there is a need for driver's to be aware of the capabilities and limits of the vehicle automation they drive in. On the other hand, it also creates a need for the evaluation of situations in which the vehicle automation can or cannot be used.

A general consensus is that having fully autonomous vehicles on the road would certainly take a considerable number of years and therefore, vehicle automation is introduced in stages termed levels [3]. Each level of automation is designed to work in specific conditions; this includes geographic, environmental, traffic, road geometry, speed and/or temporal dimensions. This is referred to as Operational Design Domain (ODD) that differs for every level of automation and is one of several measures that is used to define the levels of autonomy [4].

As the level of autonomy increases, the role of drivers diminishes from a controlling function to mostly a supervisory function. Defined in [4], SAE level 2 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'.

From [1], [2] and [5] it can be seen that ensuring driver's awareness and knowledge about the systems, is vital for the development of such systems. Annual road safety statistics published by the European Commission in the year 2015, reported that over 90% of the traffic accidents were caused by human errors [6]. In addition to this, drivers in contrast to pilots, vary widely in their abilities, seldom have extensive training, and cannot be expected to understand complex literature (technical information about systems in the owners' manual) about the automated driving features in their own vehicle [7, 8].

Within SAE level 2 vehicles, a well-researched problem is the task of authority transition. In this level of autonomy, the driver is always responsible for the Object and Event Detection and Ranging (OEDR). If the ODD is not accurately defined, this could make the driver more vulnerable to accidents. Situations where the driver needs to take-over control from the assistance system/vehicle are defined by its ODD, which is in-turn defined by the Original Equipment Manufacturers (OEM's)/vehicle manufacturers in their owner's manuals. As each OEM specifies their own ODD, this could result in uncertainties about the capabilities of different vehicles within the same level of automation. This is a problem, because most of the drivers do not read the owners' manual and are therefore, unaware of their own vehicle's functional constraints, let alone that of other vehicles. This could lead to either under or over-utilisation of the ADAS and potential safety related problems associated with it [9, 10, 11].

When analysing the manuals of several SAE level 2 vehicles, such as Tesla Model S [12], Mercedes E class [13] and Volvo XC90 [14], it was found that the ODD definitions are not completely clear and leave room for the interpretation of the driver. For example, in the owners' manual of a Tesla Model S, it is mentioned that the LKAS/Autosteer function (a system that actively steers the vehicle to the centre of its lane) may not work in sharp curves but they do not specify when a curve can be classified as being sharp. This ambiguity could make it difficult for drivers to distinguish between situations that are within or outside the ODD of the Advanced Driver Assistance System (ADAS) (Lane Keeping assistance or adaptive cruise control) installed in their vehicles. To avoid this, some manufactures do provide visual aid (in form of in-vehicle pictorial representation of the vehicle and its surroundings) to drivers regarding its ODD, but it is uncertain if drivers understand this, as not all drivers read the owners' manual (based on post drive control group discussion and [15, 16]. In addition to this, performance of the ADAS inside and outside its prescribed ODD could also have an impact on drivers' understanding of the vehicle's capabilities, which in-turn could have an impact on the drivers' trust and acceptance of the system in the long run [17].

In addition, a mismatch between the driver's understanding of the capabilities of the automated vehicle and its actual capabilities, as prescribed in the owners' manual, could also be because of the intrinsic attitude of drivers towards AV's. This includes factors such as their initial trust, ease of use and perceived risks based on prior experience, learned trust during driving and prior experiences within these systems [18, 19]. This a topic that requires more scientific attention.

Finally, an important dimension of the ODD is road type, including its geometry and influence of its components on driving risk. Closeness to other vehicles and non-moving objects (such as: road components and barriers) has an impact on accident risks while driving. These risks can be quantified and their magnitude reflect on the performance of the ADAS [20]. Moreover, drivers might modify or change their behavior when they are aware of the risk in a situation. [5].

Therefore, it is believed that to solve the above problem, a methodology needs to be developed to assess the Operational Design Domain (ODD) of semi-automated vehicles that are currently already available in the market. The proposed methodology in this study, as will be further explained in the coming chapters, assumes that development of semi-automated vehicles should happen with a central focus of improving the drivers experience within the system, both subjectively and objectively. The development of such a methodology could help Original Equipment Manufacturers (OEM's) in making drivers more aware about the situations in which their semi-automated vehicles can/cannot be used. Thereby, increasing their acceptance and trust of ADAS equipped vehicles over time.

1.2. Research objective and questions

Following the problem description, the main objective of this research is to develop a methodology for the assessment of the Operational Design Domain for ADAS equipped vehicles, specifically for vehicles equipped with Lane keeping Assistance systems (LKAS). This approach may be used by OEM's while defining ODD for their ADAS equipped vehicles and/or for assessing their already defined ODD. Furthermore, infrastructure developers may also use this method while planning changes in road design to accommodate for the advent of semi-automated vehicle on the road. This research aims at making practical and scientific contributions in the following four aspects:

1) Understanding performance variation of Lane Keeping Assistance Systems inside and outside the ODD.

2) Implementation and assessment of a novel metric for objective driving risk measurements and its comparison with existing Surrogate Measures of Safety (SMoS). This constitutes the <u>objective aspects of ODD assessment</u>.
3) Understanding drivers' attitude and response towards LKAS and AV's in general, at different stages of testing, within and between various road situations. This also includes, identification of reasons for mismatch between the driver's awareness about automation capabilities and those prescribed by OEM's. This constitutes the <u>subjective aspects of ODD assessment</u>.
4) A combined subjective and objective assessment of ODD for LKAS equipped vehicles.

To achieve the above objective the following main and sub-research questions were proposed:

To what extent can the Operational Design Domain of vehicles equipped with lane keeping systems be assessed by understanding the subjective and objective risk of driving in pre-specified test situations?

The research Sub-Questions are:

1) What are the components of a LKAS and their potential reasons for failure?
2) Which criteria can be used to identify the pre-specified real-road test situations?
3) How does the Lane keeping assistance system perform when it is within and when it is exceeding its pre-defined ODD?
4) To what extent can the proposed risk measurement metric be used to determine the objective driving risk across different test situations?
5) Which factors contribute to the mismatch in the ODD, in the selected situations, between the one specified by the OEM's and that which is specified by the drivers?

1.3. Research Methodology

To answer the above posted research questions, this study employs a literature research, a survey, and a real-road driving test. The survey involves questionnaire development for three stages of the real-road driving test conducted on a selected route within the Netherlands.

The literature research was used to identify the research gaps and motivate the various decisions made within the research. The focus of the literature research was on first understanding how lane keeping systems work and potential reasons for their failure. Based on that, research methods and research gaps were identified for the use and implantation of SMoS's for driving risk assessment. Finally, for assessing and understanding existing literature on trust, acceptance and behaviour/response of drivers specifically towards LKAS equipped vehicles and ADAS in general.

The next step was the development of a pre-and post-drive behaviour assessment questionnaires based on the reviewed scientific literature. After this, the field experiment was setup and this involved selection of a test vehicle, selection of test situations, recruiting participants and other miscellaneous steps that will be discussed in more detail in Chapter 4 of this report. The next step was to collect, store and process the data to extract required inputs for further analysis. This included image processing of continuous pictures taken during the test drives and filtration of vehicle related data collected during the test drives. These inputs were then used to answer the research sub-questions and subsequently for the development of the ODD assessment methodology. The research approach and associated analyses results will be presented in detail in Chapter 4 of this report.

1.4. Research Scope

This research involves a real-road experiment in which drivers were asked to drive in a SAE level 2 ADAS equipped vehicle. Due to time constraints, the focus of the research was to assess the ODD only for the LKAS function (lane keeping assistance function) and not the Adaptive Cruise Control function (longitudinal assistance function). The vehicle chosen for this research was Tesla Model S. It is important to indicate that this research focusses on the LKAS in a SAE level 2 vehicle and not a level 3 vehicle because, there are very few commercial vehicles that fall under SAE level 3 category and in these vehicles the driver is more outside the loop than in level 2 vehicles [21], making the drivers more vulnerable while real-road testing. Moreover, a SAE level 1 vehicle was not selected because, there are very few or no vehicles in this level which are equipped with control type LKAS (vehicle is steered by the system to remain lane cantered).

The participants for this research included only drivers with prior experience in active LKAS or Lane departure system (LDW) equipped vehicles. This include drivers with prior experience of driving in a Tesla or any other vehicle with similar SAE level 2 features.

Moreover, for this research the latest edition (June 2018) of the SAE document [4] was used to understand the specifications and definitions within the AV domain.

Furthermore, out of the various dimensions of the ODD as mentioned in [4] only the 'Infrastructure/ road geometry' aspects which fall under the road environment dimension is the focus of this research, other dimensions such as surrounding geography, time of day, other road way characteristics are not considered in this research.

The test route was selected such that a few selected situations mentioned in the OEM's owner's manual, could be tested. The detail description of the situations will be provided in Chapter 4 of this report, but it is important to note that the research is focussed only at these few situations and not on a complete exhaustive list of situations mentioned in the vehicle manuals. The driving route is chosen such that there is sufficient situations that can be seen on both highway and city road sections of the route.

Corresponding to the objective aspects of the ODD assessment method, in this research, only the lateral driving risks due to non-moving (fixed) road boundaries and barriers (lane marking, road median, guard rails etc.). The potential impacts of other moving road users were ignored due to time constraints for the completion of this research.

Corresponding to the subjective aspects of the ODD assessment method, in this research. the drivers' physical and physiological states, such as the level of drowsiness or fatigue and aggression, were not recognised or classified. Recognizing drivers' intentions, such as lane changing, was also beyond the scope of this research.

Furthermore, in this research the drivers were asked to report their trust on the LKAS in different pre-defined situations in real-time during their test drive. It is important to note that they were asked not to base their trust ratings on the performance of the ACC function of the vehicle, but purely on the LKAS function of the vehicle. This was done to mitigate the effect of the performance of the ACC function of the vehicle, on drivers' self-reported trust and ODD state awareness ratings.

1.5. Scientific Relevance

Within scientific literature, there is abundant research on the impact of ACC function of semi-automated vehicles on driver behavior, traffic safety, traffic flow etc**.**, [22, 23-28]. But, relatively there is lesser research on the impacts of LKAS on traffic entities and on the driver [29]. Current literature within the lateral vehicle control domain focusses mainly on development of the control algorithms and limited literature on the identification of driver state and relevant traffic related impacts [23, 30-34].

According to [35, 36], there is a need for a constructive dialogue between the infrastructure developers and the automotive industry. The objective of this research is to develop an assessment methodology for the ODD of LKAS equipped vehicles, by taking the road environment dimension of ODD into focus, and therefore filling this gap in literature. In addition to this, there is no scientific research on design requirements, defining or assessment of ODD that includes subjective and objective aspects of driving in the ADAS equipped vehicle, this is another gap this research attempts to fill. The research attempts at the integration of the three main stakeholders for AV development; the driver, infrastructure designers and operators and technology providers [3], [37].

Amongst these stakeholders, in this research the driver is assumed to play a central role. Creating a joint cognitive system between the driver and the system can greatly enhance overall safety and performance [38]**,**[7]**.** This is because humans non-consciously treat computers and robotic systems as humanlike entities, developing a relationship through interaction. As situations requiring trust in automation may occur in matter of seconds, this relationship between the driver and system needs to be considered in the efficient design of AV's [39, 40].

Existing studies on automation have used interviews, online questionnaires [41, 42, 43], or different contexts of driving simulation [44, 45]. Whereas, in this research a real-road driving testing approach is adopted. This is done to observe and measure drivers' true response within ADAS equipped vehicles. Such responses are difficult to capture when the driver is in a third-person frame (doesn't face consequences of accidents) and has no perception of the actual driving risks [46]. In a first-person framing (faces direct consequence of accidents), the test of trust and behavior in general, has more gravity and realism. In addition to this, in this research behavioural adaptation in drivers when they are in semi-automated vehicles, is also taken into account to some extent, while drawing conclusions. This can have a considerable influence on the drivers perceived risk [34].

In this research, a novel Surrogate Measure of Safety (SMoS) is implemented for a real-road case study using an instrumented vehicle. This driving risk measurement metric is based on risk field theory concept [47]. There is very limited literature in which such a method applied for a real-road field experiment and is therefore, a potential major contribution of this research. The probabilistic deterministic risk field approach used in this research will also be compared to existing SMoS's such as the Time to Lane Crossing (TLC) metric.

Finally, the research involves measurement of, and analyzing the relation between; real-road objective risk of driving in LKAS systems and self-reported subjective responses of drivers while driving in specific pre-selected situations. By analyzing the relationship between the two types of risks, reasons for mismatch between the drivers' awareness of the systems capabilities and its actual capabilities can be determined. This could lead to increased driver awareness about LKS and thereby lead to an increased acceptance of these assistance systems and increased safety. This would also mean that over the long run, there would be more semi-automated vehicles on the road thereby resulting in their associated benefits within the traffic efficiency, safety domain.

1.6. Thesis Outline

In this thesis report as described Figure 1, Chapter 2 gives an overview of the literature that was studied to identify the research gaps, understand relevant concepts and methodology to answer research sub-questions. This also forms an input for Chapter 3, that describes the research approach that was used to answer questions within the research. The ODD assessment methodology is developed and demonstrated using a case study. This case study is introduced in Chapter 4. Within this chapter, the experiment related pre-requisites, decisions, specifications and steps involved in its setup are presented. Following which in Chapter 5, the steps involved in the processing of the gathered experiment data is presented. Chapter 6, then elaborates on the data analysis, development and implementation of the proposed ODD assessment methodology with the help of literature overview and based on the Research approach proposed in this thesis. It also describes the variables tested and methods used for their statistical testing. Finally, Chapter 7 presents the drawn conclusions, followed by a discussion of the research results and its limitations, and finally future recommendations resulting from this research.

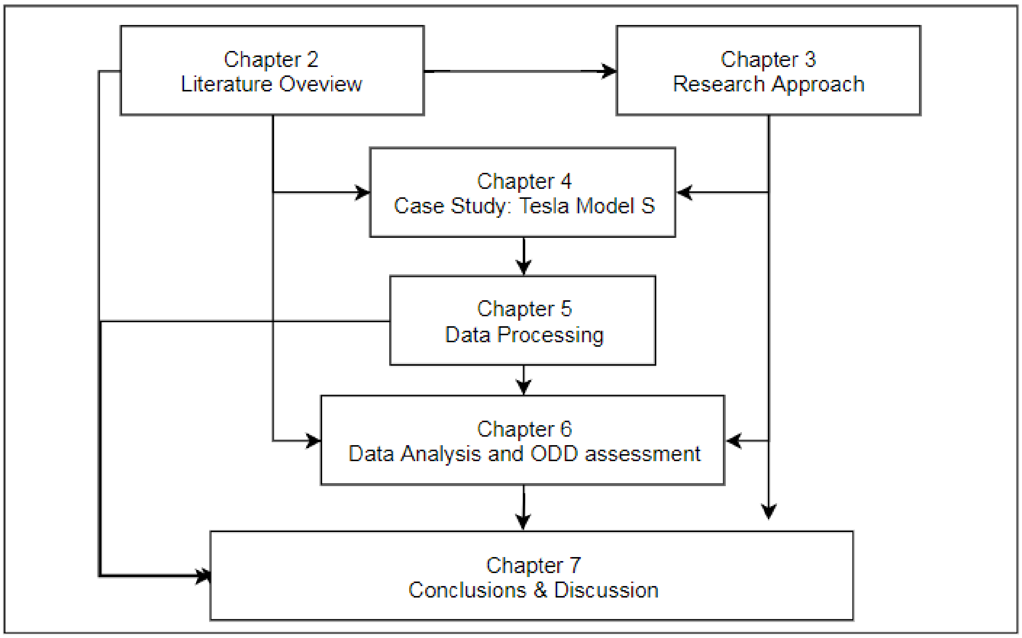


Figure 1. Thesis Outline

Chapter 2. Overview of Literature and Concepts

This chapter provides an overview of the literature which was reviewed to understand the research related concepts and gaps in existing literature relevant to the objective of this research. It first describes the components, performance indicators and methods to assess lane keeping systems in general. Followed by this it describes, state-of-the-art with respect to measurement of driving risks and also describes the novel risk measurement metric which is implemented in this research along with its potential advantages over conventional risk measurement metrics. Next, existing scientific literature pertaining to driver behaviour, situation awareness and their response in general to AV's and LKAS, is discussed and this is followed by an overview existing work with respect to relationships between driver behaviour, driving risk and ODD. Finally, it concludes with the identified research gaps and concepts, which is useful during the course of this research.

2.1. Lane Keeping Assistant systems

In this sub-chapter, first the general components of a Lane Keeping Assistance System (LKAS) will be described. Following this, factors affecting performance of LKAS will be discussed, followed by an overview of how the LKAS performance has been assessed in existing literature.

2.1.1 Components

Before describing a general LKAS, it is important to mention that there are three types of LKAS commonly known in literature [31, 48]: Warning, Intervention and Control based systems.

Warning based systems, do not directly alter the vehicle trajectory and require that the driver must choose to act on the warning for the warning to have any effect. The driver is warned by the system if she is swaying away from her current lane without indicating a lane change. 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 not really replace them. Finally, the most sophisticated of the three is the *Control based system*, in which the system not only keeps a track of potential unintended lane departures, it also continuously steers the vehicle to ensure that it is ideally positioned at the centre of the lane (of course with a small allowable buffer). In these systems, since the system has automatic control of the steering wheel, it effectively removes the driver from the loop (but the driver must be always ready to take over control)

The components of a general LKS described in this report is based on [31, 49, 50, 51, 52]. There are basically three major components of a LKAS: The Driver, The Vehicle and its surroundings and The Lane keeping module. A continuous interaction between the three, is the basis of the LKAS, with an ultimate objective of ensuring unintended lane departures. A detailed description of these components and steps for lane keeping assistance, is provided with the help of Figure 2.

1) *Road Sensing*:

The first step is for the vehicle to sense, process and realise its surroundings. This involves lane marking and other road geometry characteristics identification. These observations can be done by several different types of sensors, either individually or in combination. Once the images are collected they need to be processed for lane and object detection, followed by image to world correspondence, which is generally done with the help of in-built map systems. Road sensing can be camera based, LIDAR based, Stereo imaging based (dual camera approach), combined GPS & Inertial Measuring Unit (IMU) based, Radar based and by several other combined methods. Each of these methods have their advantages and disadvantage, but this is not the scope of this research. This information is then fed to the Lane Keeping Module, which consist of several sub-modules.

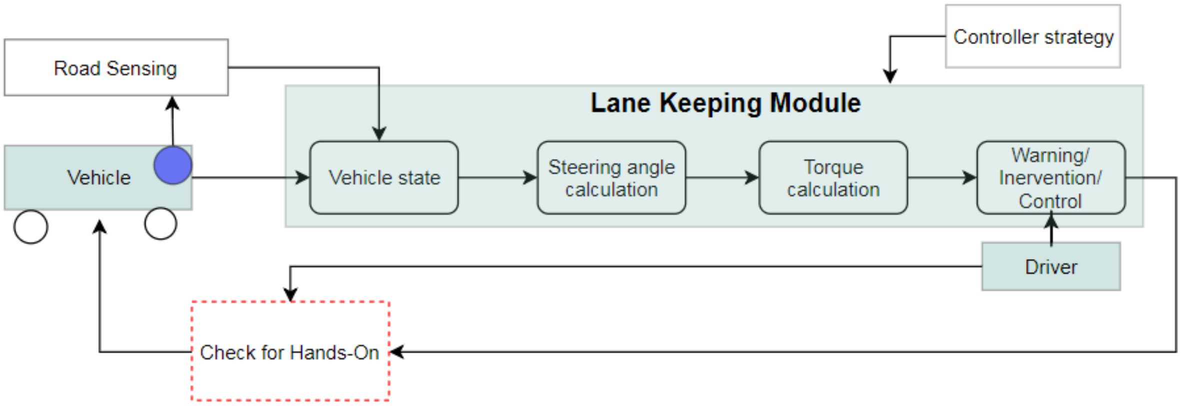


Figure 2. General components of an LKAS (blue dot refers to in-built camera/sensing hardware of the vehicle)

2) *Vehicle state observer*:

This step involves measurement of the dynamics of the vehicle and understanding the road geometry, using existing vehicle equations of motion. Two typically used models to understand vehicle dynamics, are the ‘bicycle model’ and ‘point mass model’ [53]. The difference between the two being that in the former model the vehicle’s rotation around the x, y and z-axis are neglected and involves calculation of road curvature and information about the rotation of the vehicle about the three axes. This is something that is not required with point mass models. In general, with this step the lateral motion of the vehicle relative to the lane boundary is measured.

3) *Steering angle and torque calculation:*

The steering angle sub-module calculates the optimal steering angle required to guide the vehicle back into the lane using observed states from the camera and sensors. Factors such as road curvature, heading angle, lateral displacement are generally used to determine the optimal steering angle. On the other hand, required torque calculation is based on the deviation from the optimal steering angle using a combined feed forward and feed-back controller and is dependent on the vehicle’s speed.

4) *Intervention/Control/Warning:*

Once the vehicle dynamics is determined, the next step is to make the decision whether to intervene, control or warn the driver. This depends on the intended purposed of the on-board Lane Keeping Assistant. The intervention, control or warning is made based on a *Controller Strategy* that varies for different OEM’s and is generally rendered confidential by manufacturers. The driver assistance should occur when a set of conditions are fulfilled. The conditions are generally classified into three clusters, the driver, vehicle and road cluster.

- The driver cluster contains information about driver state in terms drowsiness, distraction and whether the driver has his/her hands on the steering wheel or not. This generally determined by additional steering torque information obtained using systems such EPAS for Renault [49]**.** By doing this the system identifies how attentive or inside the loop, the driver is as al existing LKAS are *Hands-On* features.
- The vehicle cluster contains conditions about the vehicle state, like the vehicle speed, acceleration in all directions, lane change indication etc.

- The road cluster includes information about road environment such as curvature, lateral distance to lane marking and heading angle.

5) *Controller Strategy*

The intervention, continuous steering feedback control or lane departure warning is provided when the vehicle enters a 'threshold zone'. This threshold zone depends on the internal controller, sensor accuracy, pre-selected departure buffers and various other OEM specific and controller purpose specific criteria. Therefore, this is where LKAS between OEMs differ from each other and is something which will be reflected upon in further stages of this report. For vehicles that have a control type LKAS, the additional required torque will essentially bring back or help the vehicle maintain its intended position with the lane centre as reference.

It is important to note that the above description of the components of LKAS, is not specific to any OEM and therefore, differ considerably between different OEMs. But it gives an idea of how these systems work. In addition to this, some OEMs [12, 54] also rely on the vehicle state of the leading or adjacent vehicle in determining the appropriate intervention or control strategy.

2.1.2 Factors affecting performance of LKS

To determine factors affecting LKAS performance, vehicle owner manuals of different vehicles such as Mercedes E350 [13], Volvo XC90 [14] and Tesla Model S [12], were analysed in addition to referring to a comprehensive study in [52]. The factors were identified and classified into the following four categories.

Road and infrastructure related

- The type of road, i.e. whether it is a well-marked highway, city road or rural road has a huge impact on the LKAS performance.
- Roads branching off, merging, winding or narrowing, also has an impact on the performance of the LKAS.
- The performance also depends on presence of lane marking on either side of the road and the type and quality of the lane marking. If the lane markings are present, well painted, not changing or crossing over often, this is ideal for a good LKAS performance.
- The road surface quality, in terms of smoothness, wetness or dust/slush.
- LKAS have hampered performance, or does not work at all at intersections, roundabouts and close to toll booths.
- Traffic states of surrounding vehicles on the road affects the LKAS performance as this has impact on the vehicle dynamics in both longitudinal and lateral direction. Moreover, more traffic leads to more vehicle cut-in situations, this has a major impact on the LKAS performance.

Road sensing

As mentioned earlier, road sensing can be done by several methods and these methods have a few common factors affecting their functioning:

- Nearby vehicles can create severe occlusions. Shadows from nearby trees and buildings may create misleading edges and texture on the road. In some cases, like when the host vehicle comes out of a tunnel there are abrupt changes of several orders of magnitude in the illumination level, leading to over exposed image. This has impact on the clarity of images gathered by the sensors and thereby effecting its accuracy/precision, which is very important in such systems.
- Factors such as weather, presence of obstacles in front of the sensors, high proximity of leading vehicles blocking view of cameras etc. have a big impact on the visibility which again affect the accuracy and/or range of the on-board sensors. The system should operate, or at least identify the condition and lower its confidence level, under rain, fog, haze and night conditions. In addition to this, for vision based

systems, glare, bright sunlight, oncoming headlights and improper illumination hampers the detection capabilities of the cameras.

Vehicle

- LKAS are designed to function only under certain speed ranges and this varies between different OEMs and depends on the type of LKAS. For a Mercedes E350, the Active Lane Assist (Control type LKAS) function works only between 60km/h and 150km/h and this range is between 70km/h and 140km/h for a Tesla Model S. For all speeds outside this range, the LKAS either stops functioning or its performance reduces.

Driver

The drivers' behaviour within the vehicle also affects the LKAS performance. Factors like, whether the driver is wearing the seat belt, if they have their hands on the steering wheel, and their driving style (mainly sensation seeking, as the drivers then tend to experiment with the functionality of the system), are a few driver related reasons for the dysfunction or poor performance of LKAS.

2.1.3. Assessment of the LKAS performance

After understanding how a LKAS works and the potential factors that affect its performance, the next step is to understand how its performance can be measured or quantified. This helps in the comparison of LKAS between vehicles from different OEMs or between different road situations, which is essential to answer the proposed research questions.

Research regarding lateral driving performance dates back to 1982, when most studies in this domain focussed on the effect of pharmaceutical drugs on driving performance by real-road driving tests [55]. Since then a primary parameter used to assess lateral driving performance has been the Standard Deviation of Lateral Position (SDLP), the Mean Lane Position (MLP) and the Steering Reversal Rates (SRR).

These methods are still used in research and has proved to be a useful measure of driving performance [20, 56, 57, 58]. In the current state-of-the-art, Standard Deviation of Lane Position (SDLP), Mean Lane Position (MLP) and Steering Reversal Rate (SRR) used to assess "driver's" lateral performance, the focus of this research is to assess the performance of the LKAS and variation in LKAS and driver combined lane keeping performance across different aspects. This is something that still needs more investigation as there is very limited or no literature pertaining to this. This switch from a driver centric to a machine/system centric performance assessment should be possible as the main aim of these methods is to determine either by how much the vehicle strays away from the lane centre (MLP), how much variation exists in the lane position of the vehicle over time across different road segments (not necessarily about the lane centre) or)—and the effort dedicated to lateral control [59]—and the frequency with which subjects departed their lanes. Each of these measurements are still possible with the LKAS as focus.

In general, any hypothesis that posits improved lateral control, predicts a decrease in the frequency of lane departures which is marked by a decrease in Standard Deviation of Lane Position (SDLP) and Steering Reversal Rate (SRR). SDLP, in a basic sense predicts actual traffic safety, i.e. the likelihood of becoming involved in a traffic accident. This was proved by an experiment by [60]**.** A degraded lane keeping performance (driver or system), may lead to run-off-road crashes or collisions with other vehicles [61]. Specifically, the increase in SDLP can dramatically increase the probability of lane departures that lead to a crash [61]. Steering reversal rate, is defined as the number of changes in steering wheel direction per unit time [62, 63] and a higher SRR implies a poor lane keeping performance and vice versa. Given, the usability of these methods for LKAS performance, the next step is to understand the procedure for their measurement for real-road experiments.

First, an instrumented vehicle is needed which captures video images of the road and its lane markings. Using image processing techniques, the vehicle's position relative to the lane markings, can then be determined. The raw lateral distance measurements have high distortions in measurement and unwanted information, this is because the driving tests are performed in normal traffic, events may occur where the vehicle/driver makes unintended manoeuvres, data corresponding to such instances need to be filtered out [64]. Once, the data is filtered, the SDLP, MLP and SRR measurements are ready to be used for several types of comparisons.

Standard deviation of lane positions, even though considered reliable in literature and used in several studies frequently, by itself, cannot adequately describe 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 centre without much variations. This means that it is still closer to the lane boundaries and at high speeds this is very hazardous. Therefore, to quantify the imminence of lane crossing, a time based metric (Time to Lane Crossing (TLC)) that measures time left until the outer edge of a moving vehicle crosses either side of the lane boundaries, is recently employed in literature to enhance the assessment of lane keeping performance systems [64], [65]. The essence of TLC is that it incorporates relevant longitudinal and lateral motions simultaneously, and provides an assessment of the lateral control safety margin. Unlike Standard Deviation of Lane position (SDLP), TLC is a synthetic variable and is dependent on several measurements, which will be explained in Chapter 2.2 of this report. In addition to this, a novel real-time surrogate measure of driving safety, 'Probabilistic Driving Risk Field (PDRF) method', which has the potential of adding value to LKAS/ driver lane keeping performance assessment, will be described in Chapter 2.2.

2.2. Objective driving risk measurement

As mentioned in the previous sub-chapter, for a comprehensive LKAS performance assessment, there is a need for including measurements of real-road driving risk in addition to metrics such as SDLP, MLP and SRR. Real-road driving risk/safety has gained a lot of attention in the research community. This approach makes use non-crash vehicle interactions as a substitute of actual crashes and is referred to as Surrogate Measures of Safety (SMoS) in the research community.

There are two types of SMoS, 1) based on observed trajectories, 2) based on motion prediction. The first type, involves measures such as Deceleration to Avoid a Crash, and Post Encroachment Time (PET). In these measures, the movement of the interacting vehicles is tracked without considering any probability of their paths intersecting (crash). On the other hand, measures such as Time to Collison (TTC) and TLC, describe a chance of the paths of interacting vehicles to conflict based on motion prediction techniques. The latter type of SMoS is time continuous by definition and is more relevant for the proposed real-road research as it results in more information about pre-crash events giving the approach more reality but at the same time being less simplistic.

The focus of this research is on lateral safety, previous research [66, 67], suggests that driver lateral safety models are mostly based on vehicle kinematics, dynamics and are based on information regarding the vehicle's state. Vehicle's state includes its position, velocity (lateral and longitudinal), acceleration, yaw velocity and its relative motion, velocity and distance. However, these models do not capture the effects of all types of traffic factors on driving safety such as describing the interaction among driver behaviour characteristics, vehicle states, and road environments, which are very important for accurately describing driving risk.

Several advanced safety algorithms have been studied that are based on artificial intelligence, risk homeostasis theory or other advanced modern mathematical techniques to account for driver's behaviour [68, 69]. These studies were focussed on making safety algorithms better adapted to driver's behaviour using self-learning risk (impact of historical perceived risk and acceptable risk of driver on their acceleration or deceleration) and stimulus response (response of the driver to the changes in dynamics relative to the leading vehicle, such as differences in speeds or headways) concepts, but for car-following models.

Furthermore, the concept of field has been used to described risk that a driver faces in ADAS equipped vehicles. [70], extending the concepts of the elastic band theory to autonomous vehicle motion prediction and based it on the potential field theory. [71], proposed a field theory in traffic flow where lanes and vehicles form a potential field active within the minds of drivers, and that drivers always drove along low points of the potential field. A novel gravitational field concept was developed in [72], in which once again car-following behaviour was described using a series of attractive and repulsive forces related to the vehicle and the space in front of it. More recently, [47] proposed a field-based model that included risk factors such as personality, psychological characteristics of the driver, complex road conditions and driver-vehicle-road interactions, for rear-end crash avoidance [73]. However, the model parameters are not directly related to the crash dynamics and motion uncertainty and their calibration and validation seems to be very tedious.

Moreover, these methods and models do not successfully include a very important factor [74], the different levels of the 'severity' of a crash. Crash severity underpins the physics of dynamic collision mechanism and is independent of interacting objects. This parameter is calculated based on energy absorbed by the colliding bodies during a collision, independent of the road user type be it a motor vehicle or a pedestrian, with a possibility of resulting in different levels of injury for the same crash energy exchange, based on the type of collision. This when combined with probabilistic prediction of the occurrence of collision between to interacting objects, could serve to be an apt and comprehensive risk measure. In this research, a first attempt will be made to implement an SSoM currently under development at TU Delft by PhD researcher F.A. Mullakkal Babu [75] that accounts for both, crash 'Severity' and 'Probability', for the real-road assessment of the objective risk of driving in a LKAS equipped vehicle. This SSoM hereafter, would be referred to as the Probabilistic Driving Risk Field (PDRF) method.

In the PDRF method, risk is defined by the magnitude of the consequences of a collision and the chance of its occurrence. It is based on the field theory concept as discussed earlier and assumes that risks are experienced by a vehicle that is driving alongside other moving and non-moving road entities. The risks experienced due to the non-moving road entities such as lane marking, guard rails, road medians etc., contribute to the Potential Risk Field (PRF) experienced by the vehicle. The risks experienced due to the moving road entities such as the other road users, contribute to the Kinetic Risk Field experienced by the vehicle. The sum of these risk fields determines the total risks experienced by a vehicle. The crash severity in both these fields is based on the crash energy transferred during a possible collision between the subject vehicle and a road entity. For this research, due to time constraints only the potential risk fields were calculated and the formula used for this is described below Equation 1.

The potential risk taken by s due to fixed road boundary object is formulated as follows

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

Equation 1. Potential Risk Field taken by subject vehicle S and fixed road entity b

where $R_{b,s}$ denotes the potential risk 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 formulation $0.5M\left(V_{s,b}\right)^2$ 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 $r_{s,b}$. However, all such crashes are not perfectly inelastic and roadside object allow finite deformation thereby absorbing some amount of crash energy thereby decreasing the inflicted crash severity. This assumption is consistent with the empirical studies [76]. In [76], 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.

The second term in the formulation is $e^{\left(\frac{-|r_{s,b}|}{D}\right)}$ constitutes the probability of the crash, and has a range [0-1]. This term tends to a maximum of 1 at $r_{s,b} = 0$, and depicts a decrease in crash probability with an increase in $r_{s,b}$. Intuitively, a road object further away offers more possibility for the driver to evade the collision.

The gradient $\frac{dR_b}{dr_{s,b}} = -0.5kM\left(V_{s,b}\right)^2 \cdot \frac{e^{\left(\frac{-|r_{s,b}|}{D}\right)}}{D}$ is a continuous and decreasing function of $r_{s,b}$, and the risk reaches a finite maximum $R_b = 0.5Mk\left(V_{s,b}\right)^2$ solely at the position of b when $r_{s,b} = 0$ and is the theoretical crash energy.

The potential benefits of using the PDRF method compared to other SMoS's are as follows:

- It could offer more diversity in terms of scenarios and driving conditions that is can be used for. This means that it would not only be a measure of longitudinal or lateral driving risk alone, but a combined risk measure in both these directions.
- It could offer more sensitivity in risk trend plots than using other SMoS's. This refers to the fact that this method involves a collision type - crash severity based risk determination. This means that it will generate different driving risk measurements for potential collisions with different type of stationary and moving objects on the road (including lane markings). This could result in a more informative and real representation of driving risk.
- Additivity: Another benefit of using the PDRF approach is that it gives a risk measure that considers moving and non-moving objects of the road in both the longitudinal and lateral directions, using just one risk magnitude value. On the other hand, when using existing SMoS's separately to determine longitudinal risk (TTC) and lateral risk (TLC), it is not possible to generate one value for total risk by simply adding these two, time-based metrics as this value (in secs) cannot represent total risk.

This research aims at successfully implementing this novel SMoS and take a step forward towards justifying its potential benefits over exiting SMoS's.

In addition to this, it is important to assess if these objectively measured real-road risks that the vehicle experiences due to other real-road moving and on-moving objects, even have an impact on the drivers' behaviour whilst driving inside vehicles equipped with ADAS. Current state-of-the art with respect to this aspect of driver behaviour and perception of risk and its impact on their trust, will be discussed in the following chapter.

2.3. Driver behaviour, trust and situation awareness in ADAS

As mentioned in the previous sub-chapter, it is important to understand the drivers' behaviour and their perception of the risk of driving in an ADAS equipped vehicle. It is important especially when deciding the situations in which it is safe for the ADAS to function [77, 78]. Therefore, for this the driver is assumed to play a central role in the assessment of Operational Design Domain for LKAS based ADAS.

In the current state-of-the-art, there is limited literary work on the behavioural responses to, and interaction of drivers with, Lane Keeping Assistance systems. Most of the literature focusses on the driver's interaction with the longitudinal assistance systems such as Adaptive Cruise Control (ACC) ([25, 79, 80, 81, 82]). For the literature that does focus on lateral assistance systems, they generally give attention to the development of lateral control system considering drivers' behaviour [83-86], Nevertheless, there are few studies like [33], [34] and [48] in which the focus is on the influence of LKAS on driver behaviour, investigating behaviour adaptation in drivers whilst in LKAS and understanding the interaction between driver and system, respectively. These studies, only consider the 'warning' type of LKAS and not the more sophisticated 'control' type of LKAS. Therefore, this creates a need for driver behaviour assessment in more sophisticated LKAS, which is one of the objectives of the underlying research.

Most studies on driver behaviour analysis in ADAS are either questionnaire [87, 88, 89], or simulator based [45, 57, 61, 90], i.e. it has been performed in a third-person frame, where the participant has no perception of direct risk in case of automation failure. However, few studies also use real-road testing approach to understand driver trust and behaviour in ADAS equipped vehicles [5, 91, 92]. In questionnaire studies, people may say they would trust an automated system, yet act in a way that demonstrates that they do not trust it. Instruments such as the questionnaire by [93] inquire as to one's beliefs in the system's capabilities and trustworthiness, but one's beliefs may not translate to behaviours. Therefore, this strikes the need for an approach that combines a real-road test where the driver in first-person experiences the risk of driving, with a questionnaire based approach at different stages of testing.

Assessment of driving behaviour in automation includes understanding the factors that affect their trust, awareness of capabilities and functionality (ODD) and other physiological, psychological factors that have an impact on their factors. From a design standpoint, it is important to design systems that individuals will trust appropriately, granting the system authority when appropriate, and taking control when necessary [8].

Several questionnaire based studies have identified various factors that affect driver's trust in automation. [93] identified factors such as predictability, reliability and dependability having an impact on the trust in automation. This list was expanded by [46, 94] with factors such as faith, responsibility, robustness, familiarity, understandability, usefulness and dependence. [89], highlights certain issues faced in studies of trust and human intervention. Drivers' trust and awareness are multi-dimensional entities and require calibration in different stages of its measurement. Moreover, trust varies dynamically, changing over time as relationships develop. For instance, [95] established a hierarchical model of trust, and believed that certain factors of trust may change with time and increasing emotional investment. Keeping this in mind, in the underlying research, drivers' trust and awareness about system's capabilities were measured/reported before, during and post driving in the test vehicle on the test route.

The pre-drive and post-drive questionnaires were designed using concepts and relationships developed in [96]. In this research, a three-layered trust was described. 1) *Dispositional trust*, referring to an individual's overall tendency to trust in automation that include factors like culture, age, gender, personality traits. 2) *Situational trust*, referring to situational specific trust of the driver on automation and consists of both external and internal factors. External factors include type of system, complexity of the system, task difficulty, perceived risk etc. and internal factors include drivers' self-confidence, attentional capacity. 3) *Learned trust,* which refers to drivers' evaluation of a system drawn from past experiences or the real-time interaction with the system and is divide into initial trust and dynamic trust. Past experiences, knowledge about the system influence the initial trust and the interaction between driver and system influences her dynamic trust. Keeping these factors and a 3-layered approach in mind the pre-, Real-Time and post- drive questionnaires, were designed.

2.4. Driver behaviour, driving risk & ODD

The aim of the underlying research is to assess ODD for LKAS equipped vehicle by understanding drivers' behaviour whilst in the vehicle and at the same time also accounting for the actual objective risk of driving in these vehicles. It is also important to consider that the two factors, subjectively measured driver behaviour such as trust and perceived risk could also be associated with the actual risk that the driver experiences while driving [5, 97].

Moreover, while defining the ODD for their vehicles, OEMs such as GM and Waymo use either test track experiments or real-road driving tests, but with experienced tests drivers who have an inherent high trust and anticipation level as compared to an average driver of such vehicles [54, 98]. This again stresses the need for an average driver based approach while testing ADAS. Moreover, it is unclear what threshold values and factors (both subjective and objective) the OEM's use to decide if a driving condition/situation is safe (included in the ODD) or unsafe (excluded from the ODD). This information is difficult to obtain from OEMs as it is very confidential and therefore requires investigation.

Another important factor is the perception error or mismatch between driver expectation and reality. In [7], it was observed that comparing drivers' expectations of the automated driving system's behaviour to their own inputs did not yield significant correspondences between expectations and actions across the different conditions tested in this research. This indicated that the participants behaved by trusting the car, even though they stated they did not expect the computer to act and thus would need to act themselves The results of this research [7] are real because, if a driver cannot predict accurately what an automated system will do a few seconds into the future, or is unaware of its capabilities in a specific situation and does not respond in a way that is appropriate, disaster can result.

Finally, as indicated earlier some of the research in this domain consider only subjective judgments of the driver and the others consider evaluation and quantification of objective risks experienced by drivers, there is limited research that combines these two methods. To increase the reliability of ADAS technologies and customise the users' driving experience, it is important to use an approach that aims at investigating relationships between these two subjective and objective safety measures and identify potential reasons for mismatch between the expectations and the reality of driving in LKAS (ADAS) equipped vehicles. Such an approach is followed to achieve the objectives of this research.

2.5. Conclusion

In this chapter, the three major components and their sub-components of a general LKAS, were described. Next, road and infrastructure related, road sensing related, vehicle related and driver related factors affecting the performance of LKAS were listed. This was based on a detailed analysis of the owner manuals of three LKAS equipped vehicles. This was followed by an overview of indicators for lane keeping performance of drivers and their usability for assessment of the performance of LKAS. The main indicators identified were mean and standard deviation of lane positions, surrogate metrics such as time to lane crossing. Furthermore, existing literature regarding the two types of surrogate measures of safety (based on observed trajectories and based on motion prediction) were discussed along with the shortcomings of widely used metrics such as time to collision, time to lane crossing etc. A novel risk measurement metric 'Probabilistic driving risk field' was introduced and its potential advantages over existing risk measures were listed. It was also identified that current literature related to driver behaviour within ADAS in general, is more focussed towards to ACC rather than LKAS equipped vehicles and mainly use a survey or a simulation based approach for drivers' subjective risk assessment. Finally, the need for a combined subjective and objective risk based approach for ODD assessment was highlighted based on existing shortcomings of the practises of a few vehicle manufacturers for such an assessment of their ADAS systems.

Chapter 3. Research Approach

In this chapter, a complete overview of the approach used in this research, is provided. It first gives a step by step description of the various components of this research and then focusses on a detailed description of the preliminary requirements for the conduction of the road tests. It first presents a detailed analysis of the vehicle owner's manual of the Tesla Model S (selected test vehicle) along with its conditions of operation and limitations. Resulting from this analysis, driving situations are classified into three categories of 'Inside the ODD', 'Outside the ODD' and 'Neither inside nor outside ODD'. This chapter then discusses the procedure and criteria used for recruitment of participants for the road tests and finally, gives an overview of the external instrumentation of the Tesla Model S for gathering required research data.

Given the objective of this research, the first step was to thoroughly read through the *owners' manual* of several vehicles that are equipped with LKAS (Mercedes E350, Tesla Model S, Volvo XC90). This helped in understanding the capabilities and specification of the LKAS in SAE level 2 vehicles, in general. It also helped in understanding the situations/conditions where these systems can/cannot/maybe function (i.e. Operational Design Domain). Using this, the next step was to *classify the ODD* of these vehicles into ODD-in, ODD-Out and ODD-Not Sure categories. This served as a key input while setting up this research.

The *Experimental Setup,* includes several sub-steps. First, from the classified ODD situations a list of possible *test situations* was made such that there was good representation of situations that fall under the three ODD classifications (ODD-in, ODD-Out and ODD-Not Sure). The next step, was to recruit participants for the research and the procedure used for this is described in (Chapter 3.3). Following which, the *Test vehicl*e was selected, rented and instrumented (Chapter 3.4).

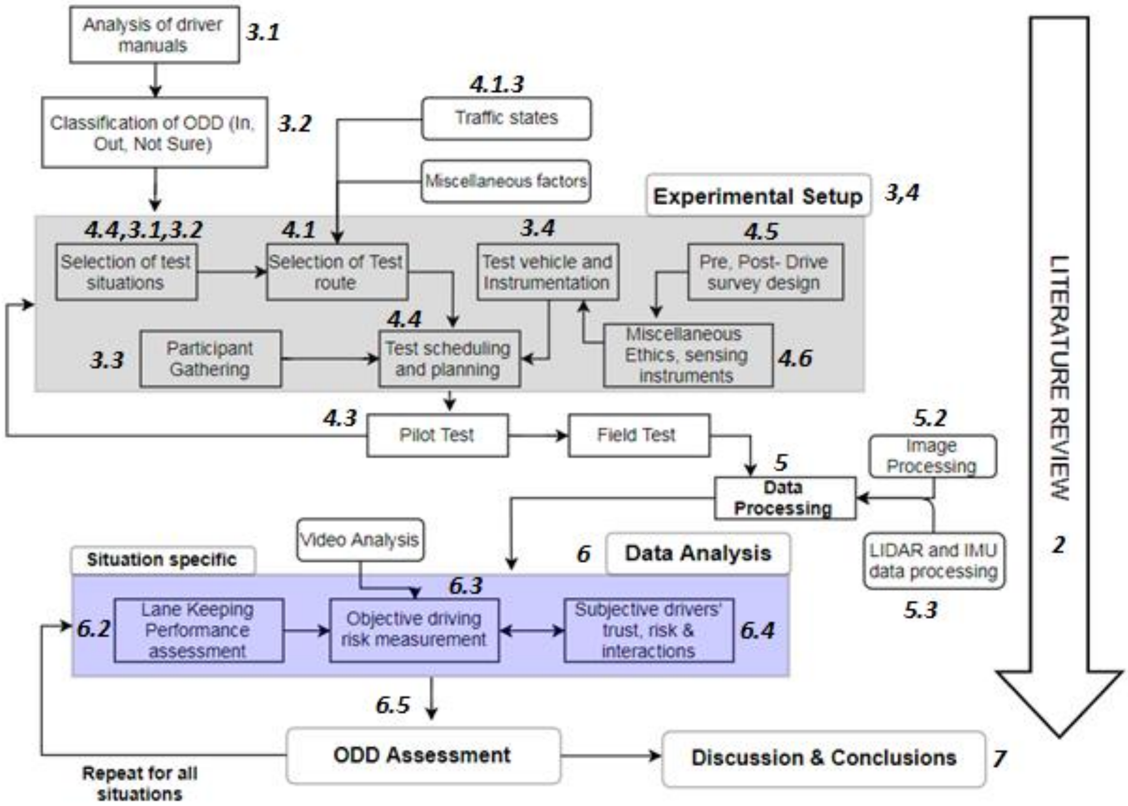


Figure 3. Research Approach with indication of relevant chapters

The next steps were more focussed towards deciding the specifics of the real-road experiment (detailed description in Chapter 4). Out of which the first was step was to *Select the test route* (Chapter 4.1) from a list of candidate routes that were determined by visually examining the map of the Netherlands and referring to prior routes where similar research was conducted in the past [99]. The selection was based on the list of test situations, analysis of *Traffic states* (to ensure off-peak testing since the focus is on the road geometry) and other *Miscellaneous factors* such as accessibility of test route, closeness to vehicle charging stations etc.

Using the details about gathered participants, availability of instruments and other necessary time related inputs a testing environment (testing route and test schedule) was first proposed (Chapter 4.2). After which, to finalize this testing environment, a real-road *Pilot test* was conducted (Chapter 4.3). This test helped to identify practical, logistical constraints of the proposed testing environment. Using this information, the final testing environment was developed (Chapter 4.4).

The next, important step was to *Design the Pre, Post-drive and real-time behaviour assessment* questionnaires (Chapter 4.5) based on the literature overview and objective of this research. The final step before commencing with the road testing was, to obtain *Ethical permission and decide participant instructions* (Chapter 4.6) for the conduction of these tests. Following which, the *Field tests* were conducted. The section filled in grey in Figure 3, depicts the steps involved in setting up this experiment

Upon the conduction of the tests, the gathered raw data needed to be processed, to enrich for further data analysis. The *Data Processing* (Chapter 5) involved *Image processing* (Chapter 5.3) for lane position determination and data filtration (Chapter 5.2) of the data gathered by the on-board sensors and GPS.

The next step was *Data Analysis* (Chapter 6), this included *Lane Keeping performance assessment* (Chapter 6.1), *Objective driving risk determination* (Chapter 6.2) and a subjective *Assessment of drivers' trust, risks and their interactions* while driving in the test vehicle (Chapter 6.3). These analyses were performed for each of the test situations and the determination of objective driving risk also involved *Video Analysis* of test situation. The OEM specified (in driver manuals) *ODD for each situation was then assessed* (Chapter 6.5) by considering lane keeping performance in the situation, objective risk of driving in the situation and the drivers' behaviour and interaction with the system in the situation. This assessment procedure was then repeated for all the other selected test situations.

The following sub-chapters give insight into the analysed owner's manuals (Chapter 3.1), classification of ODD for the test vehicle and development of a list of possible situations that could be assessed in the research (Chapter 3.2), participant recruitment procedure (Chapter 3.3) and finally, test vehicle's instrumentation (Chapter 3.4). The subsequent, road-test specific aspects will be discussed in Chapter 4 of this report.

3.1. Analysis of vehicle owner's manual

This research involved setting up a real-road driving test on a specific route inside the Netherlands using an LKAS equipped SAE level 2 vehicle. A Tesla Model S, was selected as the test vehicle. This was because from prior projects at Royal HaskoningDHV [100], it was seen that amongst the other SAE level 2 vehicles like the Volvo XC90, Mercedes E350, which also have LKAS equipped in them, the range of situations in which the Tesla could perform wider than the other vehicles. For example, a Tesla's Autopilot function could be turned on (in the presence of the required road and lane markings, speeds) even in city roads around intersections (even though this does not completely fall into the ODD of vehicles of this category). At the same time drivers are also subjected to higher risks of driving. Since, the aim of this research is to assess situations not only inside the said ODD, but also situations where the system may or may not work, choosing a Tesla was the best option, by of course giving the safety of driving in these situations the highest priority (making required arrangements to ensure participant safety).

The first step in identification of the test situations that should be assessed in this research, was to thoroughly go through the owner's manual of a few SAE level 2 vehicles such as Mercedes E350, Volvo XC90 and Tesla Model S. Within these manuals, the focus was on the content regarding the Lane Keeping Assistance functions and specification. It was seen that in terms of the content, each of these vehicles had similar specification, limitations and driver instructions corresponding to the ODD/functionality of their LKAS.

For this research, the test vehicle is a Tesla Model S. Its control type LKAS is referred to as 'Autosteer' and comes in a combined package with Adaptive Cruise Control (ACC) function. This package is referred to as its Autopilot (AP) option. A few highlights from its specifications and limitations are presented below, these were used as an input while classifying the Operational Design Domain for its Autosteer function in the next sub-chapter.

- Autosteer is a hands-on feature. Driver must keep their hands on the steering wheel always.
- Autosteer is intended for use on freeways and highways where access is limited by entry and exit ramps. When using Autosteer on residential roads, a road without a centre divider, or a road where access is not limited (city with pedestrian and vehicle crossings), Autosteer limits the driving speed.
- Autosteer should not be used on city streets, in construction zones, or in areas where bicyclists or pedestrians may be present. Drivers must never depend on Autosteer to determine an appropriate driving path and always be prepared to take immediate action.
- To initiate Autosteer, the vehicle must be driving at least 5 mph (8 km/h) on a roadway with visible lane markings. If a vehicle is detected in front, Autosteer can be initiated at any speed, even when stationary.
- To indicate that Autosteer is now actively steering Model S, the instrument panel displays the Autosteer icon in blue. When Autosteer can detect lane markings, it also displays the driving lane in blue (Figure 4).

Figure 4. Conditions to initiate Autosteer

- Autosteer attempts to centre the Model S in the driving lane. However, if the sensors detect the presence of an obstacle (such as a vehicle or guard rail), Autosteer may steer Model S in a driving path that is offset from the centre of the lane.
- Autosteer is dependent on the real-road speed limit. In situations where the speed limits are absent, its functioning speed is limited to 40mph (70km/h).
- In situations where the driver attempts to engage Autosteer, but is not driving within the required driving speed for Autosteer to operate, or Autosteer is not receiving adequate data from the camera or sensors, a display on the instrument panel will indicate that Autosteer is temporarily unavailable.

Limitations of Auto Steer

Autosteer is particularly unlikely to operate as intended in the following situations

- Autosteer is unable to accurately determine lane markings due to poor visibility (heavy rain, snow, fog, etc.), or an obstructed, covered, or damaged camera or sensor.
- When driving on hills.
- When approaching a toll booth.
- The road has sharp curves or is excessively rough.
- Bright light (such as direct sunlight) is interfering with the camera's view.
- The sensors are affected by other electrical equipment or devices that generate ultrasonic waves.
- Lane markings are excessively worn, have been adjusted due to road construction, or are changing quickly (for example, lanes branching off, crossing over, or merging).
- The road is narrow or winding
- If there are strong shadows casted on the lane markings.
- You are drifting into another lane but an object such as a vehicle is not present
- A vehicle in another lane cuts in front of you or drifts into your driving lane

An understanding of its functionality, specifications and limitations was useful while deciding test situations and while interpreting the results for this research.

3.2. Classification of the ODD

Understanding the functional requirements and limitations of the Autosteer function of the Tesla from its owners' manual gave a clear idea of its Operational Design Domain as specified by its OEM, Tesla. This was then classified into three categories of situations. The first category, ODD-In, referred to those situations where the Autosteer is designed to work for sure. The second category ODD-Out, referred to those situations where the Autosteer is not intended to work. Third, ODD-Not Sure, referred to those situations where the Autosteer may or may not function adequately.

Table 1, shows different driving situations (not exhaustive) as classified into three different ODD categories. The final test situations for this research were then selected from each of these categories. This was an important input also while selecting the test route as the candidate route on which most of these situations are visible/possible, had more likelihood of being selected.

Table 1. Classification of ODD for Tesla Model S

ODD-In	*ODD-Out*	*ODD-Not Sure*
Straight stretch of highway road with good quality lane markings, no intersecting traffic and ideal driving conditions.	When there are no lane markings on either side of the road and speed is greater than 70kmph.	changes in lane marking type (close off-ramps, on-ramps and beginning and ending of emergency lane)
At long curves (not very sharp) on the highway	In construction zones and close to toll booths.	Poor visibility due to adverse weather conditions
	City roads, with intersecting traffic and traffic lights.	Bright light interfering with camera view (existing a tunnel)
		Sharp curves on the road
		Under bridges and in tunnels
		Narrowing or winding road
		Slow moving traffic, with a lot of cut-in situations on the highway

3.3. Participant recruitment

The aim of this research is to assess the ODD of LKAS equipped vehicles, an important aspect of this assessment is the drivers' trust, perceived risk and his/her behaviour in general, while driving in specific situations. A driver without any experience of driving in ADAS equipped vehicles, has been proved to have a higher cognitive workload, hazard response time than those with experience in ADAS [101, 102]. These drivers also lack the knowledge of the functionality of the systems and therefore less aware of its limits and capabilities then drivers who use these systems regularly. Therefore, it was important that drivers with some experience of driving in ADAS (LKAS specifically) equipped participate in this research.

Having drivers with prior LKAS experience, will ensure that there is consistency in the level of driver responses in the different test situations. This will make comparing driver responses across situations easier. This is because, there could be a considerable variation in responses of ADAS inexperienced drivers in different test situations making it difficult to conclude if this variation in response is due to the characteristics of the test situation or due to their inexperience.

Furthermore, first time users are also much more vulnerable to potential unsafe driving situations and their inability to know how to react in such situations could be life threatening. Therefore, keeping driver safety and ethics in mind, it is better to have LKAS experienced drivers participating in this research. Finally, it was important to avoid young/ old drivers as their responses to such ADAS is also at opposite ends of the spectrum as compared to an average driver with ADAS experience [45, 103] , making it again difficult for between situation response comparisons.

Therefore, for this research the following criteria were proposed to recruit its participants. If a driver fulfilled these criteria, he or she would be permitted to participate in this research.

- Driver has driven in or drives in Lane Keeping Assistance equipped vehicles.
- Driver is between the age 25 and 60years.

However, as the test vehicle for this research is a Tesla Model S, the participant set was divided into two groups 1) Tesla experienced 2) Non-Tesla experienced drivers. There could also potentially be a difference between driver responses between Tesla and Non-Tesla experienced drivers, this will also be investigated in this research.

In drivers that have experience in such vehicles, this is also a problem that the OEM's must identify and try to solve. This is something which will also be investigated in this research.

The participants for this research were recruited using the following methods:

- Through online advertisements on the Tesla Motors Club forum [104] (focussed at Tesla drivers).
- Through distribution of paper advertisements and display of digital advertisement (on Televisions) at the department of Civil Engineering & Geosciences at TU Delft.
- Though distribution of paper advertisements and broadcast emails at the Amersfoort office of Royal HaskoningDHV.
- By visiting the vehicle charging station at Den Ruygen Hoek Oost, Amsterdam (also the test start location).
- By word of mouth and contacts of university and company supervisors and colleagues.

Each participant as a compensation for their time, was also given an option of choosing either a €50 gift voucher or a night of dinner and drinks at the offices of Royal HaskoningDHV. The recruitment advertisements differed based on the targeted driver group. The respective, advertisements for Tesla and Non-Tesla experienced drivers are show in Figure 27 of Appendix B: Experimental Setup.

3.4. Vehicle Instrumentation

As mentioned earlier, for this research a Tesla Model S was used as the test vehicle. The LKAS function of the Tesla referred to as the Autosteer function, was therefore the system whose performance across different test situation, is assessed. The software version of the Autosteer v8.1 (218.18.2.301aeee) was the LKAS system under consideration. It is also important to mention that for the first 2 test days (22nd and 23rd May), a Tesla Model S 60D was used and for the next two (29th and 30th May) a Tesla Model S 90D was used. This may have an influence on this research but given its objective, as the difference between these two versions of Tesla's is mainly the slight difference in power and range (on one full battery charge), but the functionality, software and hardware of the Autosteer function is the same for both these versions as the software versions are the same.

Furthermore, for this research the Tesla Model S, was also instrumented by LIDAR's, Go Pro's and a combined IMU and GPD unit. The LIDAR's (Light Detection and Ranging) was installed with an intention to obtain distances to adjacent vehicles (on both sides) and to leading vehicles. During the pilot test the LIDAR was mounted on the side of the car inside the window (Figure 28), on close inspection of the corresponding data it was noticed in some cases that the distance data had a lot of distortion because the light beams from the LIDAR were probably reflecting into the LIDAR from the test vehicles window, rather than from objects on the outside. This meant that the that was not the right position for the LIDAR to placed and they had to be placed outside the vehicle. After trying out several option, the LIDAR's were places on the top of the vehicle facing the three directions as shown in Figure 5.

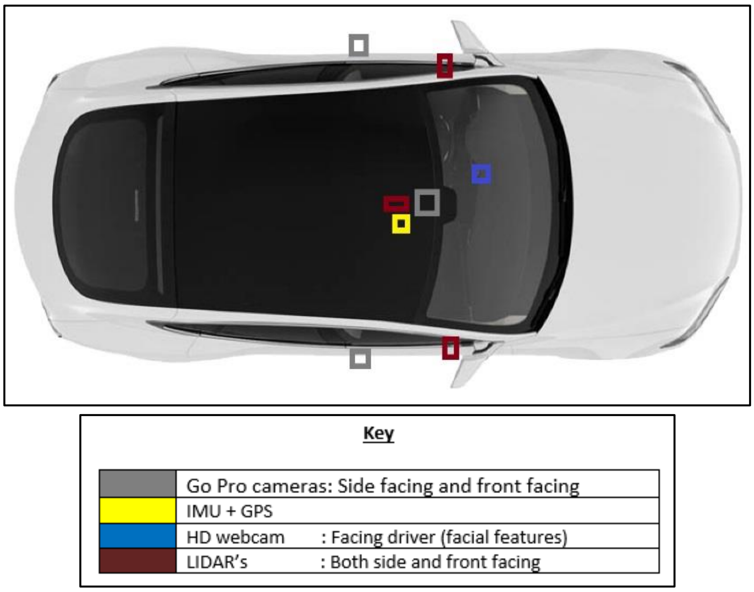


Figure 5. Vehicle Instrumentation

The Go Pro's, on the other hand were installed on the bottom of the front door on both sides of the vehicle with an objective of determining the position of the vehicle in its lane. In addition to this, there was a road facing camera installed on the windshield from the inside of the vehicle along with a HD webcam that faced the driver, which was intended to record the drivers' reactions while driving on the test route. Finally, a GPS was installed onto the dashboard and along with the LIDAR's, it was connected to the Inertial Mearing Unit (IMU), that along with the GPS measured vehicle dynamics in all directions (i.e. velocity, acceleration and position in all directions). Moreover, the IMU was responsible for the synchronization of the internal cameras and LIDAR data as it made planned beep sounds (at known interval and frequency) at the start of every test drive. All the data, i.e. the vehicle dynamics and LIDAR data, were then stored into the SD card that was mounted on the IMU. This entire setup (internal and external) is shown through Figure 5 and Figure 6.

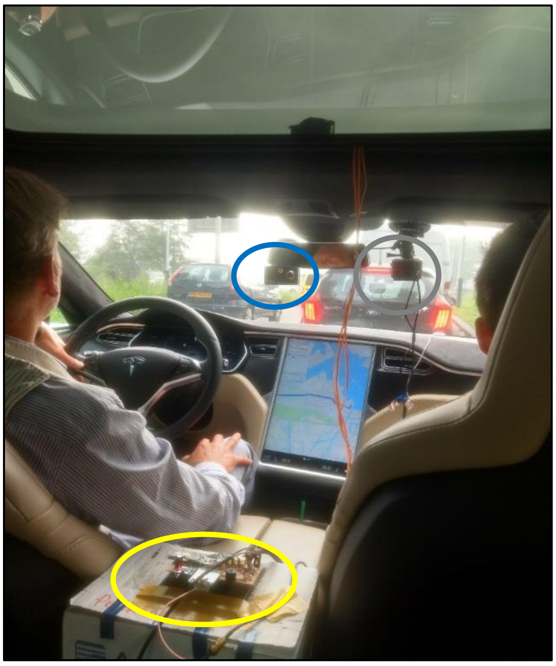


Figure 6. Left image: internal instrumentation view; Right image: Exteral instrumentation view (color coding same as Figure 5.)

3.5. Conclusion

In this chapter, the research approach and the preliminary steps involved for setting up the road tests, were described. This analysis was also used to identify, operational requirements and constraints for the intended functionality of the Autosteer function. This understanding of the functionality serves as a vital input for the next steps of this research. This chapter also discussed the specific criteria regarding the age and driving experience, which were used to select participants for the road tests. Finally, it also gives a, pictorial representation and detailed information about the various instruments that were externally installed onto the Tesla for data recording. For this research, distance measuring devices such as LIDAR's, vehicle dynamics measuring devices such as inertial measuring units, vehicle position measuring devices such as GPS and, road sensing devices such Go Pros, were used.

Chapter 4. Experimental Setup- Case of Tesla Model S

The previous chapter, provided an overview of the research approach along with a description of the initial steps taken to setup the real-road for this research. In this chapter, the next steps (for the given test vehicle, possible test situations and type of participants) that were taken to decide the testing environment, are described.

First, the chapter describes steps involved in the selection of the test route from a set of candidate routes based on various factors. This was then used to propose a testing environment along with a test schedule. The chapter than describes the pilot test, which was conducted to understand the practical and logistical constraints of the road tests. Using the information from the pilot test the final testing environment is developed and described. As mentioned earlier, in this research, driver behaviour was assessed in three different stages, before the test (Pre-Drive Questionnaire), during the test (Real-time self-reported behaviour) and after the driving test. The design and contents of each of these questionnaires is also provided in this chapter. Finally, a description of ethical requirements before commencing the road testing is given, before describing the participant wise testing procedure.

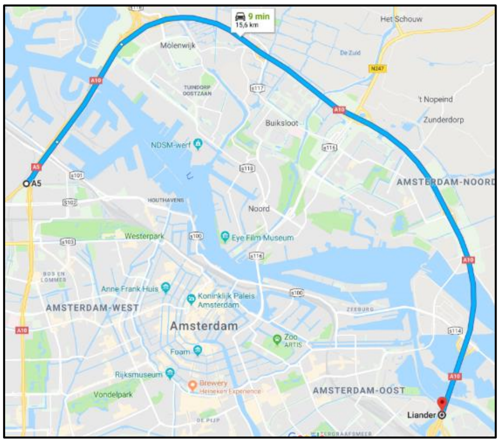
4.1. Test route selection

Three candidate routes were selected based on prior research at Royal HaskoningDHV and visual inspection of the routes, keeping in mind the ODD classification of the Tesla Model S. Out of these candidate routes, based on several factors (research and external), list of possible testing situations and verification of traffic states; the testing route for this research was proposed. This sub-chapter gives a walk-through to each of steps.

4.1.1. Candidate routes

The three candidate routes were as follows, one out of which, was proposed as the test route:

- Candidate route 1: A10 Amsterdam, starting from AS near Meininger hotel and ending at A1 towards Diemen (Figure 7).
- Candidate route 2: Starting at Einsteinweg A4, going towards exit at A5 (Westrandweg) via A9 (Figure 8).
- Candidate route 3: A5, Zwanenburg to E19 (A4) Schiphol (Figure 9).



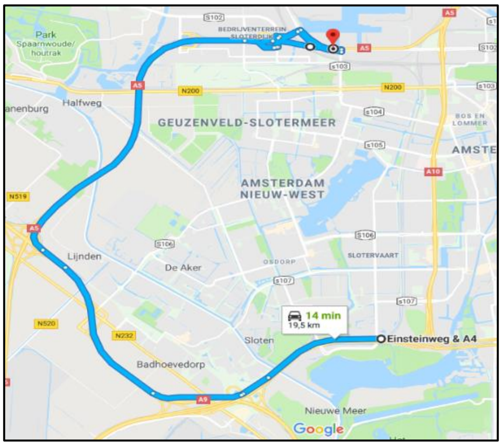


Figure 7. Candidate route 1 (Left)

Figure 8. Candidate route 2 (Right)

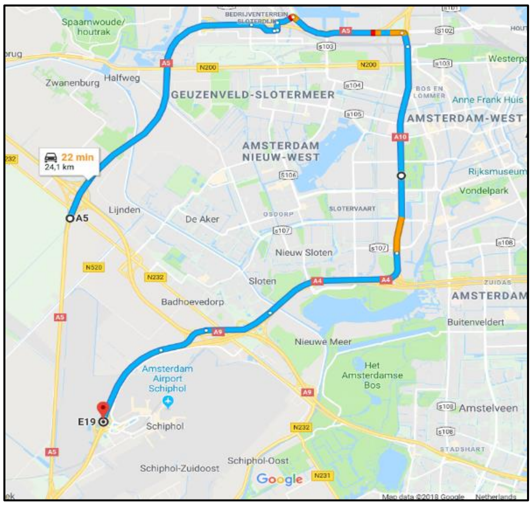


Figure 9. Candidate route 3

4.1.2. Factors affecting route choice

There were several factors that were considered while deciding the test route. These were divided into two broad types: Research related and General.

Research related factors

- Sufficient curves on the route
- Changes in lane width along the route.
- At least 2-3 off-ramps and on-ramps along the route.
- Route has varying lane marking type.
- Presence of usable emergency lane on the route.
- Proneness to roadworks or traffic jams.
- Presence of tunnel and/or bridges on the route.
- Length of the route and the cycle time for completion of the route.
- Presence of narrowing or winding route sections.
- Presence of non-highway route sections.

General factors

- Sufficient traffic data available for the entire route.
- Cyclic nature of the route.
- Closeness to supercharger/charger.
- Closeness to parking space.
- Closeness to a place to a place to freshen up.
- Closeness to train station or bus stop.

4.1.3. Verification of traffic states

As understood from the vehicle manuals, most LKAS's do not turn on during traffic jams and slow-moving traffic. Therefore, increased traffic intensity may result in more situations where the driver rather than the LKAS controls the vehicle, which does serve the purpose of this research as it aims at assessing the performance of the system and not the driver. Therefore, to ensure that the selected route has moderate traffic intensity levels and is not prone to traffic jams, traffic states for all the candidate routes were verified using a tool from the DiTTLab of TU Delft [105].

As it was planned to have the test route in the 3rd and 4th of weeks of May 2018, it was only possible to verify the traffic states for the candidate route for the 3rd and 4th weeks of April 2018. Moreover, traffic states were only verified for Tuesdays and Wednesdays between 9AM and 4PM (off-peak hours). This is because in the Netherlands, the traffic on Mondays, Thursday and Fridays are generally higher as people either travel to and from work or decide to take extended weekends. It was also decided to conduct the tests in traffic off-peak hours once again to avoid driving in heavy traffic and ensuring that all the test drives are conducted
The following traffic characteristics per candidate route, per day were verified:

1) Average travel time
2) Standard deviation travel time
3) Average flow rate
4) Traffic jam start time
5) Traffic jam end time
6) Average speed

Table 14, summarises the traffic states for each candidate route, for the off-peak hours on Tuesdays and Wednesdays of the 3rd and 4th Week of April. This then serves as input while proposing the test route.

From Table 14, on all the considered test days (Tuesday, Wednesday and Thursday) the travel times and the traffic flows for candidate route 1 and 2 were comparable and were significantly lower than that of candidate route 3 (as it was a longer route). On all the candidate route, traffic congestions generally begin close to 16:00, which is close to the proposed test end time.

4.2. Proposed testing environment

For this research, first, based on the analysis of the owner's manual for the Tesla, list of possible testing situations and the several factors affecting test route choice, a testing environment was proposed. This sub-chapter gives a description of the proposed test route, testing situations and test schedule.

4.2.1. Proposed test route

The test route for this research was proposed based on a visual inspection of each of the candidate routes on google street view. As shown in Table 15, (Appendix B: Experimental Setup) for each candidate route a score was generated. This score was based on all the factors that affect test route choice, i.e. the research related factors, general factors and visibility of test situations in all three ODD classifications. A score of 1 was given when a factor was either visible on the candidate route or if it was relevant for that route, and a score of 0 was given otherwise. Traffic state/intensity at all the candidate routes was included as a factor under the 'general factors' category. It must be noted that from Table 14**,** each of the candidate routes generally show potential traffic jams towards the end of the off-peak period (at around 1600 hrs), this was kept in mind while making the test schedule.

From Table 15 (Appendix B: Experimental Setup), candidate route 2 and 3 have almost the same score and therefore the proposed test route would have to be one of the two. For this research, candidate route 3 (A5, Zwanenburg to E19(A4) Schiphol) was selected as the proposed test route. This is because, it is much easier to a have a closed cyclic route for candidate route 3 (Figure 9) than for candidate route 2 (Figure 8), as in the later, the cycle would have to be completed on the A10, which is much more prone to traffic jams than the A5 that splits away from A4 close to Bastion Hotel. Moreover, this route has its starting point around parking lot P1 of Schiphol Airport, which was believed to be a better access point for all participants than the other candidate routes. Moreover, it is also the longest route out of the three candidate routes and therefore, allowing more situations to be tested.

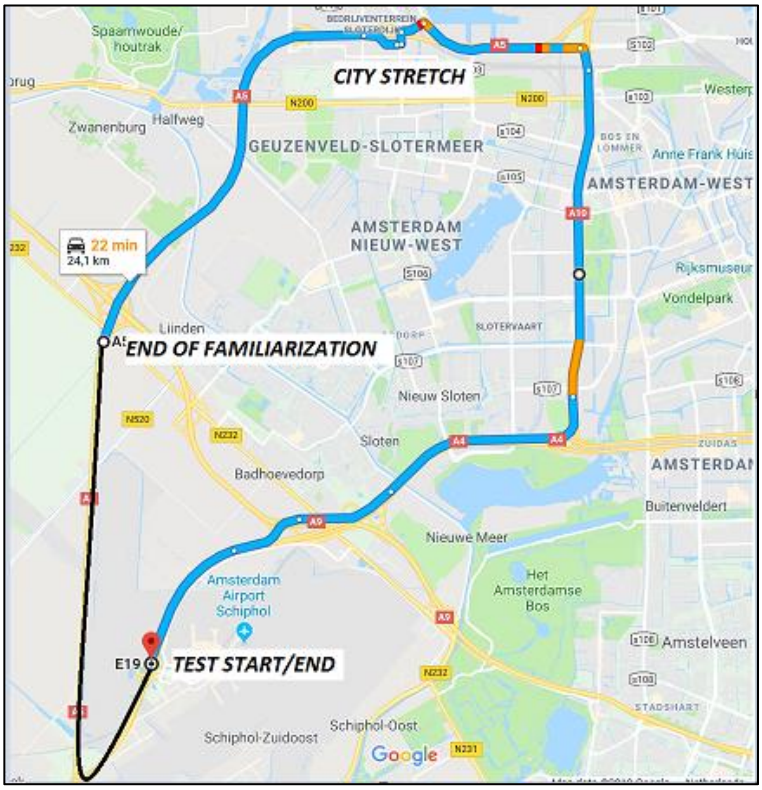


Figure 10. Proposed test route

Therefore, the proposed testing environment was follows:

- Test route = A5, Zwanenburg to E19(A4) Schiphol. As shown in Figure 10, the test was proposed to start and end close to Schiphol airport P1. After which the driver will be given some time to familiarize with the Tesla (marked as the black curved line) and receive the test instructions at the same time (especially important for the Non-Tesla experienced drivers). After the familiarization phase ends the driver then carries on driving until he/she returns to the point where the test began, this then marked the end of one test drive and this process was to be repeated for all the other participants in the same chronological order. This route also includes driving through a small city road stretch as indicated in Figure 10.
- All test drives were to take place during off-peak hours.
- Length of test route = 24.2kms
- Length of familiarization phase = 10kms
- Total distance travelled per participant = 34.2kms.

A test schedule was also prepared for the proposed test route and this will be described in the next sub-chapter. Moreover, it must be noted that this proposed was not yet the final test route and was subject to change post the pilot tests.

4.2.2. Proposed test schedule

While proposing the test schedule, the following factors were included referring to Figure 10.

1) Familiarization time - This refers to the travel time taken to complete the familiarization phase of the test.
2) Actual test travel time - This refers to the travel time taken to complete test route (i.e. from End of familiarization to the end of the test)
3) Route completion time - The route completion time is equal to the sum of the familiarization travel time and the actual testing travel time along with a buffer of 10 mins added to this time, to account for variability in the travel times.
4) Time taken for post drive questionnaire - The driver, after the completion of the test was supposed to fill in the post drive questionnaire.
5) Time between participants & participant reachability – This together referred to the total time before the next participant.
6) Total cycle completion time – This refers to the sum of the time components 3), 4) and 5).
7) Test start time – This referred to the time at which it was proposed to start the test drives on each test day. 9:30 every day.
8) Test end time – This referred to the time at which it was proposed to end a test day. 16:00 every day.
9) Total available testing time – This referred to difference between test start time and test end time.
10) Other factors – Such as range of the vehicle, Distance to charger were included to determine if the vehicle needed to be charged during a test day, or one complete charge would be sufficient.

From Table 16, (Appendix B: Experimental Setup) the Total cycle completion time per driver is 75mins and the total time available per test day is 375mins, therefore, the number of possible participant drives per day was calculated to be 5.

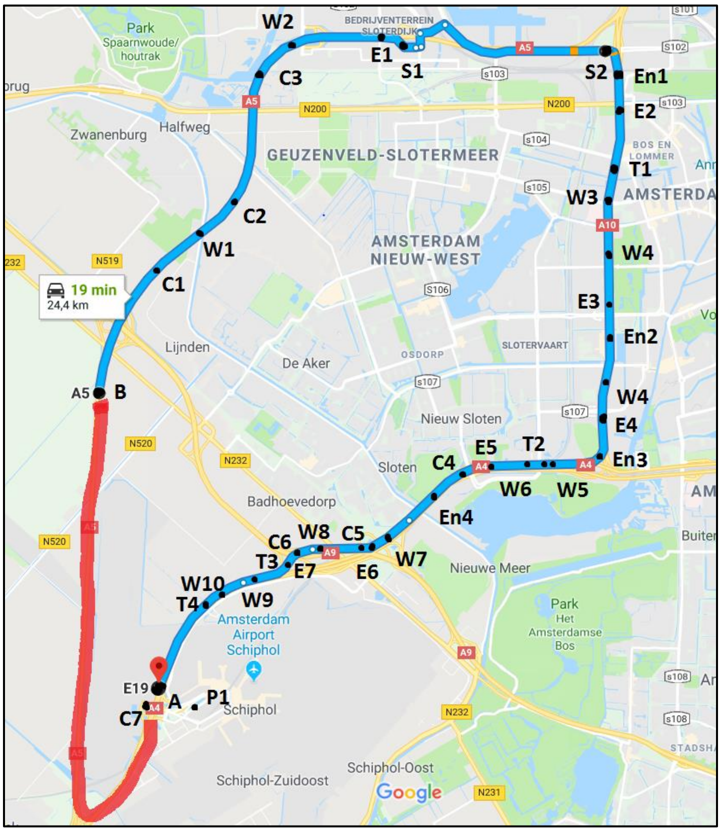
Furthermore, from Table 17 (Appendix B: Experimental Setup), the total route distance per driver was calculated to be 37.2kms and the range of the Tesla Model S is 310kms. This shows that for the proposed time schedule 5 participant drives would be possible per day and this will be possible with full charge of the vehicle. Of course, this was only the proposed test schedule and this was subject to change after the pilot test was conducted and this discussed in the following sub-chapter.

4.3. Pilot test

In this research, a pilot test was conducted to determine if the proposed test route is as it was observed while visually inspecting it on google street view and to determine if the time and distance components considered while making the scheduled tests are representation of their actual values. Moreover, during the pilot test possible locations of the sensors and cameras, were also tested.

To verify if different ODD classified situations are visible on the test route, a test map indicating possible locations of these situations was made (Figure 11). During the pilot drive, each of these situations were verified with the help of a checklist.

The pilot test was conducted on the 9th of May and the pilot participants were Dr. Haneen Farah (University Supervisor for this research), Ir. Peter Morsink (Company Supervisor for this research), Paul Van Gent (PhD researcher at TU Delft). Paul, was also present to provide a helping hand with the placement and locations of the instruments on the test vehicle, which was his area of expertise. Furthermore, Daud Pletcher the safety driver from the Tesla car rental organization, was also present to help with understanding how the ADAS on the Tesla work.



Notation	Road element/ situation
A	Start of Familiarization phase / End of test
B	Start of actual test route
C1 to C7	Curves on the road
W1 to W10	Change in number of lanes / change in Lane marking type
E1 to E7	Presence of exit / off-ramps
En1 to En4	Presence of entry/ on-ramps
S1	Start of city road stretch
S2	End of city road stretch
T1 TO T4	Tunnels or bridges
P1	Parking lot at Schiphol
	Familiarization phase

Figure 11. Route map for Pilot Test and key

During the Pilot test, possible locations for the sensors and cameras were also tested. The sensor used for this research was a LIDAR and it was going to be used determine the distance between the test vehicle and other moving and non-moving objects on the road, to is front, left and right side. In addition to this, a combined Inertial Measuring Unit (IMU) and Global Positioning System (GPS) were also going to be installed in the vehicle. Finally, Go Pro cameras were also going to be installed to determine the position in the lane, of the test vehicle. The purpose of these instruments will further be discussed in further chapters of this report. During the pilot, only one of the LIDAR's was tested, and it was coupled to the IMU and GPS to check if the systems work as intended. The LIDAR was placed inside the car and facing the right through the window. The Go Pros were located on the both sides of the car at the bottom of the front doors, respectively. These locations are shown in Figure 28,Figure 29 of (Appendix B: Experimental Setup)**.**

In addition to this, another important factor to practice was the interaction between the participant and safety driver and myself, as it was very important to ensure that these interactions result in the least possible workload on the driver. Finally, the total route times were also verified to compare with the proposed times.

The key takeaways from the Pilot test were as follows:

- With the help of practice interactions with the pilot participants and after a thorough discussion with the safety driver, a driver instruction and interaction protocol was decided.
- Some important nuances of the test route, regarding which off-ramps and on-ramps to take or avoid, were also decided and saved onto the Tesla's in-built map.
- It was noticed that the route could be extended to include more city road sections, thereby increasing the number of possible situations that could be tested. A time compromise would have to be made by conducting the post-drive questionnaire online instead of right after the test drive.

- It was also noticed that accessibility of the test route could be a problem for participants arriving by personal vehicles (as the short time parking in Schiphol was expensive and could lead to an increase in participant reachability time. The route could be extended to start and end at Den Ruygen Hoek Oost, which has a charging station and a place for the participants to relax and freshen up and is well connected to the bus stop as well.

These key takeaways, lead to small alterations to the testing environment and will be discussed in the following sub-chapter.

4.4. Final testing environment

By adapting to the key takeaways from the Pilot test, the proposed test environment and test schedule were slightly altered. The resulting final test environment is described in this sub-chapter.

4.4.1. Test situations and Route

The final test route now has an extended city road stretch and does not start and end close to Parking lot P1 of Schiphol Airport. The test starting point was the parking lot at Den Ruygen Hoek-Oost, Rijsenhout. This was where the test familiarization was to begin. As seen in (Figure 12), the actual test starting point was the same as the proposed route but the test end was extended further South of the test end point in the proposed test route. Furthermore, the drivers now had to drive from the Test point to the test starting point and this part was also not included for data analysis. In addition to this, the city roads stretch was extended considerably from that in the proposed route, this was done to include more variation (in terms of ODD classifications) of the possible test situations.

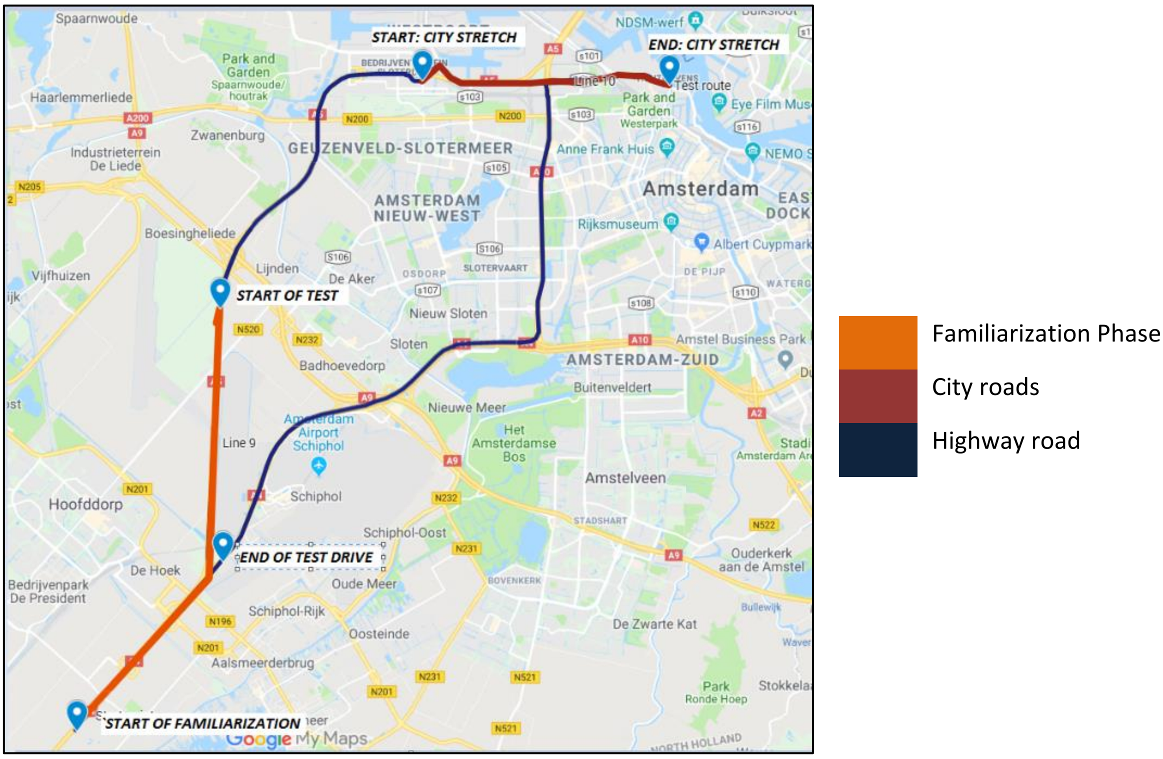


Figure 12. Final Test Route

Out of the various possible situations described in Table 1, the final situations that would be assessed (per ODD classification), is described in Table 2. These were decided after visual inspection during the pilot tests.

Table 2. Final list of test situations

ODD-In	ODD-Out	ODD-Not Sure
Straight stretch of road with no problems in lane marking, no reasons for the LKS to not work properly.	Approaching traffic light around slow-moving traffic inside the city	changes in lane marking type (at exits, entry and beginning and ending of emergency lane)
At standard curves (not very sharp, but long) on the road.	No lane marking on road boundary inside the city	Bright light interfering with camera view under bridges and in tunnels
Inside a well illuminated tunnel	Blind curve inside the city	Driving close to on-ramps or off-ramps or emergency lanes (changing lane markings)
		Driving around slow-moving traffic on highways

4.4.2. Final test plan

Based on the takeaways from the pilot study, the proposed test plan was altered. The test route was extended ad this meant that the total distance travelled per driver and the subsequent travel time, also increased. Furthermore, the drivers also had to drive a short distance from the 'End of test drive' to the 'Start of Familiarization' / starting point of the test. To compensate for this increased travel time, it was decided to conduct the post-drive questionnaire online and therefore still have the same number of participants per day as was in the proposed schedule.

The time and distance components of the final test route are presented in Table 18, Table 19 (Appendix B: Experimental Setup), respectively. These tables confirm that the total number of participants per day is still 5. For the road tests, there were a total of 19 participants recruited. Since, one day only 5 participants tests were possible, the road testing had to span over 4 days, with the last test day having 4 participants instead of 5. The final test days were Tuesday 22nd, Wednesday 23rd, Tuesday 29th and Wednesday 30th of May. For each day, the tests began at 9:30 and ended at approximately 16:00 and the participant slots were allocated on a first-come-first-serve basis and based on the preference of the participants. The final participant slots across the four test days is shown in Table 20 (Appendix B: Experimental Setup), and this was based on a detailed schedule as shown in Table 21 of (Appendix B: Experimental Setup).

From Table 19 (Appendix B: Experimental Setup), the total distance travelled per participant was approximately 55kms and the estimated total time taken per participant drive was 75mins. This again meant that the total mileage (220kms) per day (was still less than 280kms (range of the Tesla) and therefore, no additional charging was required. Before each test day, the vehicle had to be instrumented with the necessary sensors and cameras, this is discussed in detail in the next sub-chapter.

4.5. Pre-drive, Real-time & Post drive questionnaires

Before the conduction of the field tests, an important step was to design the pre-drive, post drive and real-time driver behaviour assessment questionnaires. This sub-chapter will provide a walkthrough to each of these questionnaires.

4.5.1. Pre-drive Questionnaire

The purpose of the pre-drive questionnaire was to assess the drivers initial attitude towards semi-automated vehicles and LKAS systems in general. The questionnaire is presented in (Appendix C: Pre-drive questionnaire). At the beginning of this questionnaire, it was made clear to the participants what semi-automated vehicles are, and they were given clear comparisons between 'control' type, 'warning' type LKAS (which is commonly misunderstood by people) and the Adaptive Cruise Control (ACC) type of ADAS. This was followed by questions regarding participants' demographics, mileage and experience with LKAS and semi-automated vehicles in general. In addition to this, participants were also asked if they have participated in such a research in the past, as this could have an impact on their driving behaviour as compared to the other participants of the research.

Furthermore, the participants were also inquired about their own judgement of how safe they are as a driver, how confident they are with using ADAS in vehicles, how much they trust semi-automated vehicles, and whether they are aware of the term Operational Design Domain (ODD) and its meaning. The participants were also given an opportunity to verify their knowledge about Operational Design Domain. This question was asked especially because in other questionnaires of this research (Real-time and post-drive), the participants were going to encounter this term quite often.

The final part of the questionnaire was more focused at the participants initial attitude towards LKAS specifically. They were inquired about their perceived risk of driving in LKAS equipped vehicles, their awareness of its capabilities, frequency with which they use or drive in LKAS equipped vehicles and how easy, useful and satisfactory they consider driving in these vehicles. Finally, participants were also asked questions about their prior negative experiences of driving in such vehicles, if any. On average, this questionnaire was expected to take 10mins to complete.

4.5.2. Real time behaviour assessment

This research also captures the real-time behaviour and response of the participants during their drive. Each participant was asked two specific questions during their drive. These questions were asked after the participant drove through any of the testing situations as specified in Table 2. In each situation, the driver was first asked to report their trust on the Autosteer (LKAS) on a scale from 1 to 5, with 5 meaning they had very high trust on the Autosteer in the situation they just drove through and alternatively, a rating of 1 meant that they had very low trust on the Autosteer in that situation. After which, for the same situation the participant was also asked to report If they thought the situation was inside, outside or maybe in/out of the ODD of the LKAS they were driving in. Each driver was asked these two specific questions a maximum of 18 times during their drive of approximately 35minutes (some drivers did not encounter all the possible test situations and therefore, responded less than 18 times).

It is important to note that the participants were already aware what the term ODD refers to and that they were asked these two questions only after the situation was passed and not when they were in the test situation. This was done to ensure that these questions had the least influence on their workload and therefore not affect their response [106]. As mentioned earlier, the driver already knew what the term ODD referred to and this also helped in keeping the interaction between the driver and myself short.

4.5.3. Post-drive Questionnaire

Post their test drive, each participant was sent an online questionnaire to fill in. This questionnaire is presented in Appendix D: Post-drive questionnaire. The aim of the questionnaire was to get an impression of the driver's experience in the LKAS equipped Tesla Model S, on the test route. The questionnaire began with general questions about how easy found driving on the test route and how reliable they considered the performance of the Autosteer during the whole drive. They were also asked if they could have beforehand predicted the situations in which the Autosteer would not work, and to also rate how much they trusted the Autosteer after their test experience.

Following which, the participants were asked situation specific questions regarding their trust, awareness of the usability of the Autosteer function, the need for Autosteer, difficulty of using it and their dependability on Autosteer, in four selected test situations/situation types. The four specific situations/ situation types were based on test situations specified in Table 2. They were, 1) Around off-ramps/ on-ramps when changes in lane marking type was prominent. 2) While driving at a curve on the highway. 3) While driving inside a tunnel. 4) While driving through the city road stretch.

For each of these situations, the participants were asked how they would rate the ease of driving in that situation and whether they had kept the Autosteer ON whilst in that situation. If the Autosteer was ON, the participants were asked to rate how risky they perceived driving in that situations with Autosteer ON, how satisfied they were with the performance of the Autosteer and whether they had experienced an error in the automation (Autosteer) whilst in that situation. If they did experience an error in the automation, they were also asked how they coped with it and how safe they perceived that error in automation to be. On the other hand, participants that had the Autosteer OFF in that situation were asked to specify a reason for not turning ON the Autosteer. In addition to this, for all situations, regardless of the Autosteer being ON or OFF, all participants were also asked to indicate whether they were satisfied with the information provided to them by the vehicle (regarding its functionality) and if they thought that Autosteer was needed in that situation or not.

In addition to these questions, specifically for the curve situation the participants were also inquired if they were driving in the vicinity of a large vehicle(s) such as trucks and how much of an influence did these vehicles have on their perceived risk of driving in that situation. For the tunnel situation (3), the participants were also asked additional questions about the influence the features of the tunnel (illumination and closeness to tunnel walls) had on their trust on the Autosteer. Finally, for the city road stretch (4), participants were also asked to report their trust on the Autosteer whilst driving on the city roads when lane markings were missing on the boundaries of the road.

The final part of the questionnaire involved asking participants to report their trust on the Autosteer around slow-moving traffic and whether they had any other negative experiences, discomfort or general comments about the performance of the Autosteer from their test experience.

It is important to note that, there may have been instances when participants experienced certain situations more than once, in such an instance they were asked to respond to those specific questions by considering all instances of that situation, that they experienced.

4.6. Ethical requirements

Before the road-tests could commence, there were three additional steps that were necessary to be carried out. The first step was to develop a *Participant Instructions* protocol, that was to be used by the on-board safety driver as a reference while providing instructions to the participants before the commencement of their test drive. This was developed based on the driver manual of the Tesla Model S and though interactions with the safety driver during the pilot test. The instruction protocol is shown in Appendix E: Pre-drive driver instructions**.**

The second step was to prepare an *Informed Consent Form* that all participants were requested to sign before the commencement of their test drive. This consent form gave the participants information about the researchers involved, the purpose of the research and its general procedure. They were informed of the role they played in the research and were also made aware of the possible risk and discomforts that they could face during the test. Most importantly, it was made clear to all participants that the data collected in this research is confidential and that it would not be shared with any other parties' other than the respective representatives from TU Delft and Royal HaskoningDHV. They were also informed about the presence of the on-board safety driver and that they had the right to refuse or withdraw from the test at any time during the testing procedure. The informed consent form was as shown in Appendix F.

Finally, to conduct road tests with human drivers on public roads within the Netherlands, it was necessary to have permission from the ethics committee at TU Delft. This involved a standard procedure of submitting an *Ethics Review* application form (Appendix G) to the committee along with the informed consent form, the driver instructions protocol and a copy of the advertisement that was used to recruit participants for the research. The ethics review application required a complete explanation of the research, its purpose, possible danger to its participants and the steps taken to ensure safety of its participants. The road tests could be conducted only after receiving an approval from the committee. This procedure took almost two weeks to go through after which the road tests were conducted on the fixed test dates.

4.7. Test procedure

After setting up the test and fulfilling the ethical requirements for the road testing, a participant-wise test procedure was developed. A participant test involves the following steps:

- Before their test day, each driver had to submit a signed copy of the informed consent form and fill in the pre-drive online questionnaire.
- On the test day, they were requested to arrive 5 mins before scheduled time slot (Table 20).
- Before the test drive, they were provided with an introduction and familiarization to the functions of a Tesla by the test safety driver.
- While driving, they were supposed to respond to two very specific questions about their real-time behaviour in certain situations.
- Finally, after their test drive, they were sent an online post-drive questionnaire and they were requested to fill this questionnaire as soon as possible, to make sure that their driving test experience is still fresh in their minds.

4.8. Conclusion

In this chapter, the complete setup for the experiment was described. It included a description of the three candidate routes out of which the final test route was selected. The various research related and general factors affecting the choice of the test route were listed. Furthermore, traffic states of the different candidate routes were verified to be moderate and the possible location and time for traffic jams were identified. Following which, using a score based approach and the affecting factors, a test route was selected. The main takeaways of an on-road pilot test were also used to then finalise the testing route and eventually a list of testing situations as classified in Chapter 3, constituted the final testing environment. These situations are tested on a route which starts and ends at the parking lot of Den Ruygen Hoek-Oost, Rijsenhout and the test dates were fixed and a testing plan was generated. Before the road tests, pre-drive, real-time and post-drive driver behaviour questionnaires were designed and the pre-drive questionnaires were also sent to the selected participants. Finally, before conduction of the roads tests, ethical requirements in the form of the informed consent and permission from the ethical board of TU Delft was also got and the road-tests were conducted based on the described test procedure.

Chapter 5. Data Processing

This chapter gives a description of the data collected during the road tests and questionnaires. It also describes the methods that are used to process the raw data, such that it could be used to answer the research questions defined in this study.

5.1. Gathered data

Upon successful conduction of the road tests, the following research data were collected:

- Side Go Pro data – The Go Pros were placed on either side of the front doors of the test vehicle as shown in Figure 29. The Go Pros on either side of the car took pictures of the respective front wheel and the closest lane marking strip (Figure 31). These images were taken at a frequency of one image per second at 720p quality.
- Front facing Go Pro data – The front facing Go Pro was placed on the windshield of the vehicle as shown in Figure 6. The cameras took continuous video footage of the road surroundings facing forward, for every participant drive video footage was taken at 1080p quality.
- Driver facing HD webcam –continuous video footage at a 720p HD quality of all participants for their entire drive, was collected. It was located also on the windshield, facing the participants.
- LIDAR data – The three lidars placed as shown in Figure 6, measured distances to other moving and non-moving objects on the road in three directions (front, left and right). These distances were supposed to be measured during all the drives but there were some unexpected problems that were encountered affecting a complete collection of this data.
- IMU + GPS data – The onboard GPS and Inertial measuring unit, measured vehicle dynamics i.e. its velocity, acceleration and position along the x, y and z axis. These measurements were collected for all participating drivers.
- Data from the pre-drive, post-drive and real-time driver assessment questionnaires.

5.2. Processing LIDAR, IMU Data & video synchronization

As mentioned earlier, all the LIDARs were coupled with the IMU unit. The raw distances were measured at a very high frequency of 100 data points per second and this raw data (Figure 33, Figure 34) in itself was not useful for analysis. To acquire the data a sensing device was developed. It is based on an ARM Cortex-M4 chipset (MK66FX1M0VMD18), on a Teensy 3.6 board (shown in Figure 35). Acceleration data was collected using an MPU-6050 accelerometer/gyroscope sensor combination. Furthermore, GPS data and speed were collected using a G.top 0.13 (PA6H) GPS sensor with an external antenna attached. Three Garmin Lidar Lite V3 sensors (Figure 36) were used to acquire the ranging data. All the collected data were logged to an SD card at a rate of 100Hz.

The Video data was collected using 3 GoPro cameras directed towards the left, right and front of the test vehicle. Synchronisation of collected sensor data and camera data was done through the sensing board. At power-on a series of 5 one-second beeps sounded at a frequency of 2.7KHz. An algorithm was developed to detect this series of beeps in the video and cut them all at the same point. Doing so ensured that there was a common t=0 moment for all data streams.

In terms of data filtering, acceleration data was cleaned from high frequency vibrations using a Butterworth lowpass filter (2nd order, cut-off at 2Hz). The filter used the forwards and backwards 'filtfilt' function, leading to no delay in the filtration. LIDAR data was cleaned from artefacts using a variation of a Hampel Filter. For each given data point, the filter took 15 data points on each side, giving a moving window size of 31. The median absolute deviation (MAD) was calculated for each window. If the data point differed from the window median by more than three MAD, it was replaced by the window median. This suppressed artefacts commonly encountered in the signal. The filtered data is represented by the bottom image in Figure 33 and the orange line in Figure 34 (Appendix H).

The synchronized road facing video data and the LIDAR + IMU data were visualized on a python based notebook Figure 30 (Appendix H). With this visualization, it was possible to obtain ranging data, vehicle dynamics for every second during the tests for when the LIDAR was ON. An example of this visualization is presented in Figure 38 (Appendix H). This figure represents distances to moving (other road entities) and non-moving objects (road boundaries and barriers) in lateral and longitudinal directions. All distances are measured in cm, with the blue lines corresponding to distances to the right, grey line corresponds to distances in front, and the red line corresponds to distances to the left, of the subject vehicle. Figure 38, depicts all these distances for 17.5 secs of a corresponding road facing camera video footage.

It can be concluded with confidence that the detection accuracy of the LIDAR was high, this is proved by Figure 37 (Appendix H), that shows the LIDAR detecting even a very small cavity on the wall of the tunnel. Finally, it is important to mention that the development of the IMU, its coupling to the LIDARs, data filtration procedures and python based coding were conducted by MSc. Paul Van Gent, a PhD researcher at the department of Transport & Planning of TU Delft.

5.3. Image processing

The images taken by the Go Pros had to be processed to obtain the vehicles position with its lane. This image processing procedure is described using Figure 13, the known entities here are the width of the vehicle and the width of the lane. Using the Go Pro images, the objective of the image processing was to determine D1 and D2, that represent the distance of the front wheel of the test vehicle to the lane marking strips on either side of the vehicle.

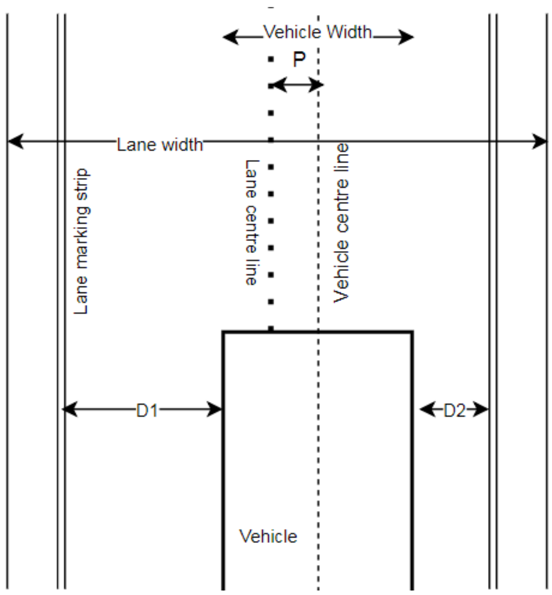


Figure 13. Reference for image processing

A python based algorithm was developed to carry out the image processing Appendix J (confidential)**.** The adopted procedure is summarised in Appendix J (confidential).

This procedure results in distances D1 and D2. The processed images were converted into a continuous video of easier visualization and this helped in determining distances D1 and D2 for every second of the test duration for which the Go Pros were switched ON. A snapshot of this video is shown in Figure 39. This snapshot combines processed images for both sides of the vehicle. On each side, the number on the top left represents the number of pixels between the wheel and the lane marking and the number on the top right represents the distance between the two, in meters. Subsequently, on the bottom left of each image the timestamp of the image is displayed and on the right bottom the corresponding frame number is displayed. This comprehensive visualization of distances was very useful during the analysis phase of the research. On each side, if looked closely the 128x128 pixel images are visible. As explained earlier, these were used to locate lines on the Go Pro images. It is important to mention that the python-neural network based algorithm described here, was developed by Max Maton, a freelance software expert.

The next step was to determine 'P', which represents the position of the vehicle center line relative to the lane center line. This distance was dependent on the width of the lane and the width of the vehicle. It was calculated using the following formula, where W_l is the Width of the lane, W_v is the width of the vehicle and *D2* is the distance between right front wheel and the closest lane marking strip. This formula also be formulated using *D1* and it will result in the same value for P. Using this formula, if the measured P was positive, this meant that the vehicles' center is P units towards to the right of the lane center. Alternatively, a negative P meant that the vehicle was P units to the left of the lane center.

$$P = \frac{W_l}{2} - \frac{W_v}{2} - D2$$

Equation 2. Vehicle position relative to lane centre

The image processing was performed with fixed lane widths for highways at 3.5m and city roads at 3.25m. Therefore, the validation of the lane positions had to be computed with these lane widths as reference. To do so, percentage errors in these measurements were computed for each drive and for all the measurements for which both D1 and D2 were available (there were a few measurements for which only one of them was available). Across all the drives the (absolute) average and maximum percentage errors were computed, for the highway and city sections of the test route, respectively (Table 52). From Table 52, on the highway on average the percentage error was 3.5 % and the maximum error was 16%. In the city road section, on average the percentage error was 4% and the maximum error was also 16%. It must be noted that these errors only represent the accuracy of the line detection algorithm for the reference lane widths of 3.5m and 3.25m, for the highway and city, respectively.

The lane positions about a result of this image processing, before being used for analysing lane positions of the test vehicle and thereby, assessing its lane keeping performance, had to undergo one last filtration. This was because the raw position values across all participants also consisted of measurements that corresponded to lane change manoeuvres. Moreover, there could also be a few seconds of positions that have unusually high position values. Data corresponding to all these instances needed to be filtered out. This data omission process was based solely on the specific python code used for the image processing procedure. Therefore, all lane position values that were greater than -1m and 1m from the lane centre, were omitted as this would represent either a lane change manoeuvre or that the Autosteer of the Tesla was not ON. Each lane position value corresponds to the distance between the centre of the lane and the vehicle centre, as described in Figure 13, where a positive value of lane position means that the vehicle is to the right of the lane centre line and alternatively, on the left side of the lane centre, when the position value was negative.

5.4. Summary of collected data

Before commencing with the data analysis phase of the research, it was important to have a clear overview of the available data across all the test days. This is provided in Table 3, where a 'Complete' refers to the availability of complete data set for that data component i.e. no data was missing. The table also highlights what type of data and at what time during the testing period, was data missing for any component.

On 22nd May, due to unexpected rain during the testing period, the LIDARs had to be tuned off. This was because, a rain protecting cover was not put over the LIDAR. For all the following test days, the LIDARs were covered (Figure 40), to avoid any more loss of data. There were also some test days on which Go Pro image data was missing on specific times, this was due to insufficient battery back-up due to human error (forgetting to switch cameras). It is important to note that there was no time instance during the entire testing period where Go Pro image data was missing for both sides of the vehicle, at the same time. This meant that, the lane position measurements were still possible for all participant drives during the testing period without any missing positions.

On the 30th May, LIDAR data corresponding to the final participant was not available. This was because the IMU turned off automatically. A possible reason for this is could be that the power source to which the IMU board was connected to (a portable charger) would have stopped detecting that the IMU was drawing power from it, as the IMU requires very low power to function.

Finally, the data collected from the questionnaires also had to undergo one final filtration before the conduction of sub-question specific hypothesis testing. This filtration was based on missing data values corresponding to a few of the participant drives and was done to ensure that there was consistency in the data used across the analysis. The data corresponding to the drivers/ drive 1,2 and 4 was excluded as the trust ratings of these drivers were not available for all the test situations mostly because they did not experience the test relevant situations or the questions were accidently framed a little different for these drivers. Therefore, the total sample size was reduced to 16 drivers, from 19 drivers.

Keeping the available and missing data in mind, the next step was to analyse this data to achieve the objectives of this research.

Table 3. Summary of available data from the instrumented vehicle

Data component	**22nd May**	**23rd May**	**29th May**	**30th May**
Driver facing	Only until participant 4 i.e. until 14:45	Complete	Complete	Complete
Road facing	Complete	Complete	Complete	Complete
Left Go Pro	not available from 14:36 to 15:20 for D4	Complete	Left, off between 14:19 and 14:34 for D14, the rest is available.	Complete
Right Go Pro	Complete	Complete	right turned off at 15:48, but rest is available.	Complete
LIDAR	Available only for D3	Complete	Complete	Not complete for D19
Valid data sets	2 (1 Tesla exp. Driver)	5 (2 Tesla exp. Drivers)	5 (1 Tesla exp. Driver)	5 (2 Tesla exp. Drivers)

Chapter 6. Data Analysis and Assessment of ODD

Upon conduction of the field test and the processing of the gathered data. The next step was to analysis the data to answer and attempt to achieve the objective of this research. This chapter presents the analysis of the research data for the objective and subjective risk measurements and finally, combine the results to assess the operational design domain of the Autosteer for a few selected test situations.

Frist, Chapter 6.1 describes the final situations that were selected for analysis in this research. It then presents the results of the lane keeping performance of the Autosteer over four different aspects, in Chapter 6.2. This is followed by results of the lateral objective risks measurements across different test situations, Chapter 6.3 and then, by the presentation of the results of several statistical analyses and the relevant hypotheses to measure the driver behaviour/subjective risk in the Tesla, in Chapter 6.4. This is followed by, finally assessing the operational design domain of the selected test situations in Chapter 6.5.

6.1. Final selected situations

During the road tests, the real-time driver behaviour assessment questions were asked to each participant in all the situations mentioned in Table 2. Four out of these situations were eventually selected for analysis. The real-time ratings for some of the situations in Table 2 were not encountered or recorded for all participants, therefore for the final analysis only situations that had responses from all participants were selected. Moreover, the four situations were also consistent with the situations that were specifically questioned in the post-drive questionnaire. Finally, while selecting the final situations for analysis, it was also ensured that there was at least one situation per ODD classification type ('Inside the ODD', 'Outside the ODD' and 'Neither inside nor outside the ODD'). Subsequently, the situations that were analysed in this research were as follows, and are depicted in (Figure 14).

1) Driving on city roads with no lane marking on road boundaries.
2) Driving inside a tunnel.
3) Close to an off-ramp with changes in lane marking type.
4) Driving in a curve on the highway.

Given below is a description of each of these selected situations:

Situation 1: Inside the city with no lane marking strips on the road boundaries (refer Figure 42, for top view of the situation)

- Direct road edge on the either sides of the road boundaries.
- Speed limit is 50km/h.
- It is a two-lane road with traffic moving in the same direction. There is no lane marking on road boundaries, only the centre lane marking is present.
- Un-signalised interactions at crossings.

Situation 2: Inside a Tunnel within the city

- Single lane of 3.25m and concrete walls at approximately 4.5m and 2m on the right and left, respectively.
- Speed limit is 50Km/h.
- Good lighting conditions and well-marked lane markings. The visibility is poor only while entering and exiting the tunnel.
- There is also a curve present inside the tunnel.

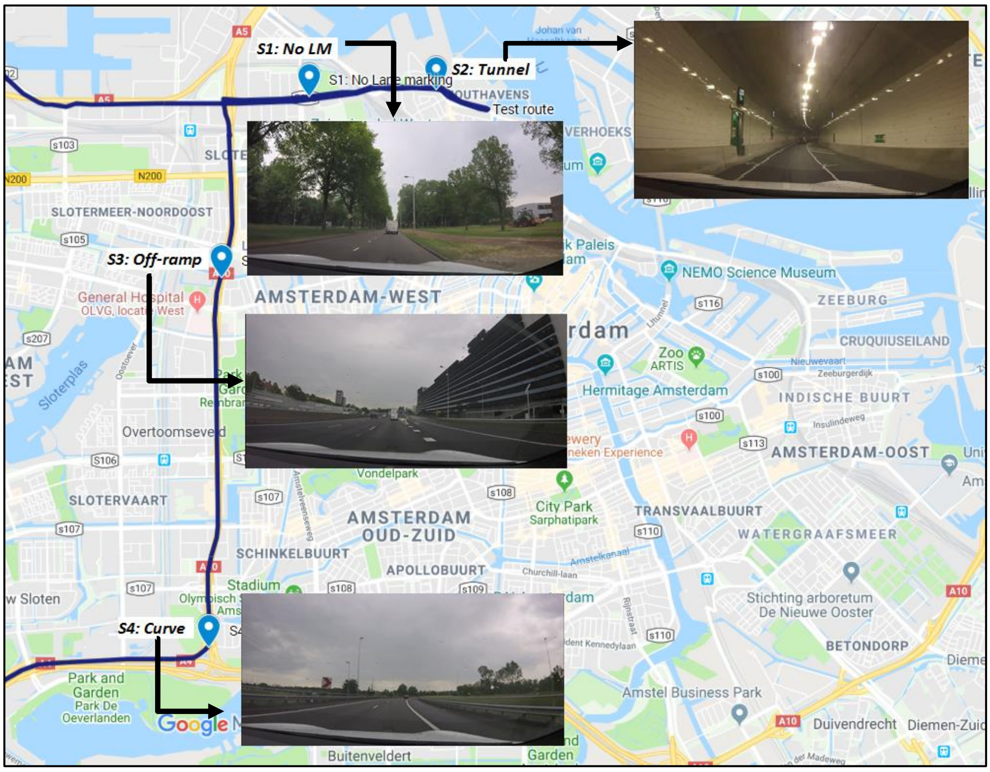


Figure 14. Section of test route with final test situations

Situation 3: Close to an Off-ramp on the highway (refer Figure 43 for top view of situation)

- Hard shoulder on the right.
- Vehicles driving between 80 and 120km/hr.
- The beginning of this situation is defined as when the Lane Marking begin to change from a solid line to a dotted thicker line and remains like this until the off-ramp is passed.

Situation 4: Curve (refer Figure 41 for top view of situation)

- Constant but relatively long right turning curve (radius of curvature = 295.51m and curvature = $0.0038m^{-1}$)
- Vehicle speeds between 80 and 120 km/hr.
- Guard rail on both sides and soft shoulder of 0.5m.

6.2. Lane positions: LKAS performance

With the help of the image processed lane position data. The performance of the LKAS (Autosteer) of the Tesla Model S was assessed over the following four aspects, with the objective of understanding how the system performs inside and/or outside the OEM specified ODD:

1) Between drives variation in lane position from the lane centre.
2) Between days' variation in lane position from the lane centre.
3) Between road sections variation in lane position from the lane centre.
4) Between situations variation in LKAS (Autosteer) performance.

Analyses of these aspects were assessed based on either boxplot, Standard deviation of lane position(SDLP), mean lane position (MLP) calculations, or a combination of these metrics.

It was expected that for aspect 1), since the majority of the testing time the Autosteer function was ON, the lane positions of the Tesla for all the drives across all the days must be close to the lane centre with slight deviations on both sides of the centre (allowable buffer which is different for different vehicle manufacturers). This is because the Autosteer is designed to steer the vehicle such that it is lane centred (with an exception maybe around road obstacles). But during the tests drives there were also few common instances when the Autosteer had to be OFF/ could not be turned ON (at off and on-ramps, at intersections, around slow-moving traffic or based on the driver's preference). Therefore, it was expected that there would be some variations in lane positions about the lane centre, on both sides of it. The resulting lane positions therefore, correspond to driver and LKAS's combined lane keeping performance.

For aspect 2), it was expected that since the same LKAS system (of a Tesla Model S) is under considerations at all test days, there should not be significant variation in its performance across test days. Of course, at the same time, there would be factors such as traffic intensity and weather that could have an impact on this. This is something that will be investigated in the coming sub-chapters.

For aspect 3), the test route was divided into three sections. The first section involved driving on highway roads until the point 'Start of city stretch' in Figure 12 (length of section 7.2kms), the second section involved driving in the city stretch of the route indicated by a brown line in Figure 12 (length of section 7.7kms) and, the third section involved driving again on the highway after the city stretch until the 'end of the test drive' point (length of section 8.5kms) in the same figure. It was expected that, the performance of the Autosteer would be better on the highway roads than on city roads as according to the owner's manual the system is not supposed to be turned on inside the city and this trend would not vary drastically between test days.

Finally, in aspect 4) which is of key focus, Autosteer's performance is assessed alone (Autosteer is always ON) across situations; S1 (No LM), S2 (Tunnel), S3 (Off-ramp) and S4 (Curve). Based on the vehicle owners' manual of the Tesla, it was expected that the order of the Autosteer's performance across the situations will be S4≈S2>S3>S1. Situations S4 (curve) and S2 (Tunnel) are both inside the ODD as specified in the manuals, therefore it is expected that lane keeping is performed in these situations. On the other hand, S3 (close to off-ramp) is neither inside or outside the ODD and S1 (No Lane marking on road boundaries in the city) is strictly outside the ODD. Therefore, lane keeping performance in S3 is expected to be better than in S1. More, it is important to understand where the vehicle aligns itself (with respect to lane centre) in these ODD classified situations.

It is important to note that the LKAS performance assessment was based on the description of how a general LKAS works and the specific explanation of how the Autosteer function works in a Tesla. Moreover, the final assessment of the LKAS in different ODD classified situations, will also include measurements of driving risks in these situations which is dependent on lane keeping performance of the system.

6.2.1 Between drives variation

The frequency plots of the lane positions across all the drivers are shown in Appendix A. Keeping this in mind along with Figure 15**,** the between drives variation of lane positions are assessed. Figure 15**,** shows the position of the vehicle from the lane centre across different drives on the same test route. This plot excludes the outliers which are the extreme high and low values (lane change manoeuvres). Each boxplot shows the mean, median, data range of 50% of the position values, maximum and minimum values for each drive. A negative value of position means that vehicle was on the left of the lane centre and consequently, a positive value means that the vehicle was to the right of the lane centre.

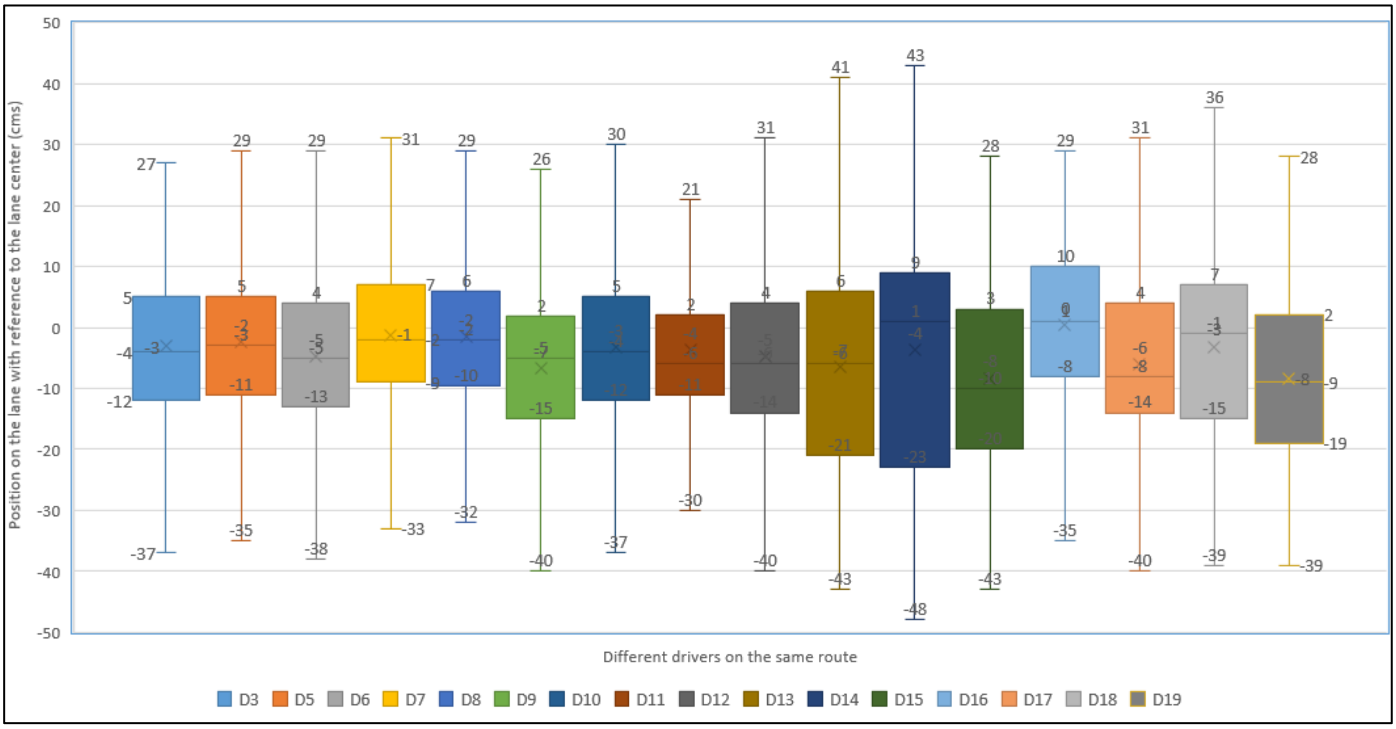


Figure 15. Between drives variation in the position of the vehicle from the lane centre (driver and LKAS performance)

As seen in the plot, mean lane positions for all drives other than D13, D14, D15, D17 and D19 are somewhat similar and between the values -5cm and +1cm from the lane centre. This means that in terms of overall lane keeping performance (driver and LKAS) for most drives over the same route, the Autosteer and drivers in combination did a good job at ensuring the vehicle is somewhat lane centred.

On the other hand, for drives D13,14,15,17,19 the mean position values are away from the lane centre and range from -4cm to -10cm to the left of it. This could imply that on average during these drives the Lane keeping performance was worse than the other drives on the same route in the same LKAS system. This is an indication of the inconsistent performance of the Autosteer on the same route. Overall across all drives, it can also be seen that the vehicle tends to be positioned more towards the left of the lane centre rather than to its right.

Another important observation is the skew of the data. For most drives except for D14, D17 and D18, the lane positions are evenly distributed about their median values, the mean and median values are almost the same and the length of the whiskers is the same above and below the median. This shows that the positions are evenly distributed between its highest and lowest value and about its median value. Thereby, reflecting on the symmetry in the performance (driver + Autosteer) even though for most of these drives are not lane centred. On the other hand, for D14, D17, D18 the difference between median and mean values are large, with the median values being more towards the left of the lane centre than the mean, signifying a higher variation in vehicle position towards the left of the lane centre, than on the right. Especially for D14, find which for a majority part of the drive the vehicle was to the left of the lane centre.

The maximum and minimum lane position values for all drives other than D13, D14 are between the range -40cm to +40 cm from the lane centre and 50% of the lane positions on these drives (expect D15 and D19) is between the range -15 and +10cms from the lane centre. For D13 and D14, the maximum and minimum values range between -48cm and +43cm which is higher than the other days.

This however, does not reflect on its performance in certain specific situations that were selected for this research and it also consists of instances when the driver was in control of the vehicle rather than the Autosteer (in control majority of the times).

6.2.2 Between days' variation

The next step was to compare the combined Autosteer and driver performance between the 4 test days. This was done using two widely used metrics, Mean Lane Position (MLP) and Standard Deviation of Lane Position (SDLP). Figure 16 and Figure 17 respectively, depict the variation in the SDLP and MLP values across different drives and between different test days. In general, a smaller SDLP value refers to a better lane keeping performance and the intended MLP for a control type LKAS must be around 0 (along with a small buffer).

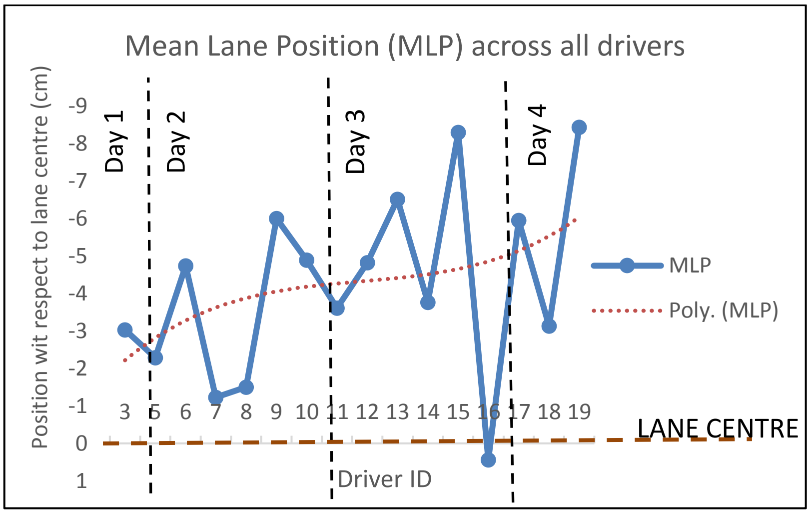


Figure 16. Variation in MLP across all drives and between different test days

In Figure 16, it is not easy to identify a trend in the MLP between the different test days. Therefore, a polynomial order 3 trendline ($y = -0{,}0033x^3 + 0{,}0864x^2 - 0{,}8317x - 1{,}4673$) was plotted to make it easier to spot a trend. Based on this, it was observed that the MLP is predominately to the left of the lane center and shows a small increasing (negative, more towards the left) trend between test days, with exceptions on drives D7, D8 and D16. Moreover, the vehicle lane position was most left oriented on Day 3. This also confirms the lane position trends as seen in the boxplots of Figure 15.

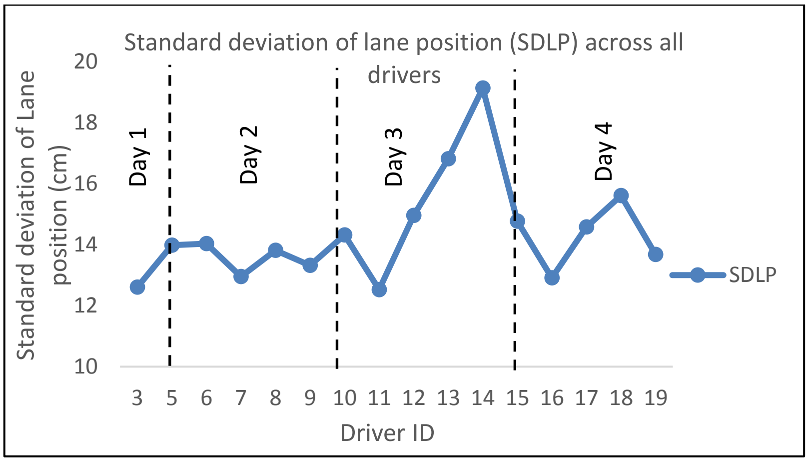


Figure 17. Variation in SDLP across all drives and between different test days

Based on the Standard Deviation of Lane Position (SDLP) values in **Figure 17,** between days' LKAS (and driver) lane keeping performance can be assessed. The values for Days 1,2 and 4 tend to be similar and that for Day 3 is considerably higher than the other days. Therefore, this could mean that the combined driver and LKAS lane keeping performance was the worst on Day 3, relative to the other test days. But to confirm this, a distinction between driver and LKAS must be made. This is out of the scope of this research.

6.2.3. Between road sections variation

To assess how the combined driver and LKAS performance varies across different road sections (Highway (7.2kms), City (7.7kms) and Highway (8.5kms) again), a plot with the mean values of lane positions per section per day was made. The standard errors for these mean values were also calculated as show in Figure 18. From this figure, it was once again observed that in general, across these road sections the vehicle was biased mostly towards the left of the lane centre (a negative lane position means that vehicle was on the left of the lane centre). Furthermore, the mean lean positions were observed to be further away (to the left) from the lane centre in the City road section as compared to the Highway sections. But for both, the city and highway sections the vehicle was predominantly positioned to the left of the lane.

On comparing section performances across different test days, it was seen that apart from Day 3, on every other test day the mean lane positions first increased (more away from the lane centre) from Section1 to Section 2 and then decreased (closer to the lane centre) after going from the city road section to the highway again. But on Day 3, the mean lane position values only increased (went further away from the lane centre) from Section 1 to Section 3, not showing a similar trend as the other test days.

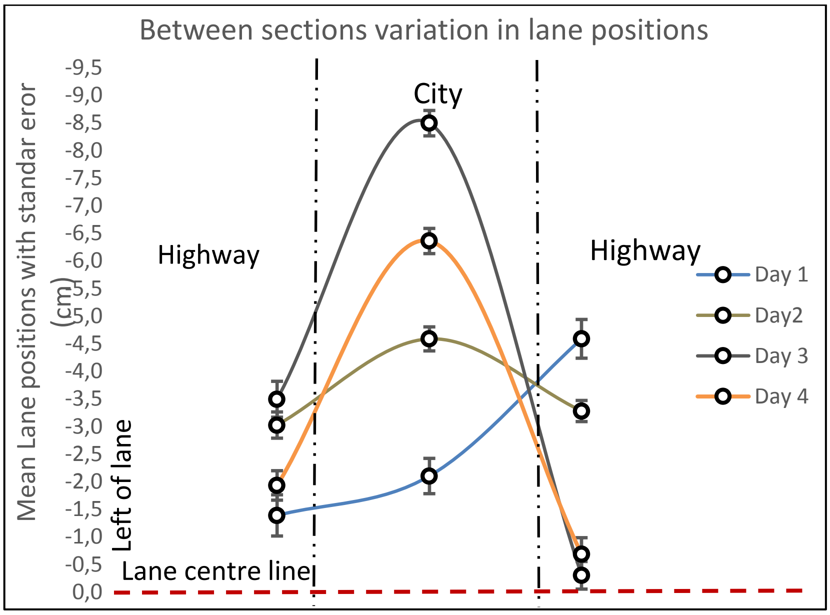


Figure 18. Between sections variation of lane positions

Finally, it was also observed that the standard error to the mean lane position in the city road section, across all the test days, was always lower than that in the highway sections even though in the city the lane positions were further to the left of the lane centre than on the highway.

6.2.4. Between situations variation in Autosteer performance

While comparing Autosteer performance between the different test situations (S1 to S4), certain adjustments to the data were made. For each drive, only data corresponding to the first 15 seconds of driving in each situation, was selected for lane keeping performance comparison. This was done to ensure that the Autosteer was ON in these situations and to have consistency with the objective driving risk measurements (also calculated for a duration of 15secs).

This meant that a maximum of 15 lane position values were available per drive per situation. To increase the sample size for better box plot representation, the lane position values for each situation across different drives were combined into one data set. This was possible, because in these situations (in the duration of 15secs) the Autosteer was ON and the system was responsible for lane keeping and not the driver.

Comparison of Autosteer performance between different test situations was first done using boxplots of the lane positions in these situations, Standard Deviation of Lane Positions (SDLP) and the Mean Lane Positions (MLP). This was then followed up by pairwise comparisons of lane positions in ODD-in, ODD-out and OUT-Not sure situations (as classified earlier). The differences/similarities are also verified statistically (S1 – No LM, S2 – Tunnel, S3 – Off-ramp and S4 – right turning curve).

From Figure 19, first, it can be observed that the length of the whiskers on either side of the median is the highest in Situation S3, this means that the range of lane positions is the highest in this situation as compared to the other situations. The next observation is regarding the skew in the lane position values about its median. In all the situations, slight asymmetry is visible. In S2, S3 and S4, the skew (small) is towards the top (right of lane) and in S1, the skew is towards the bottom (left of lane). Suggesting that the lane positions in S1 are more spread out towards the left of the lane centre and in situations S2, S3 and S4 the lane positions are more spread out towards the right of the lane centre.

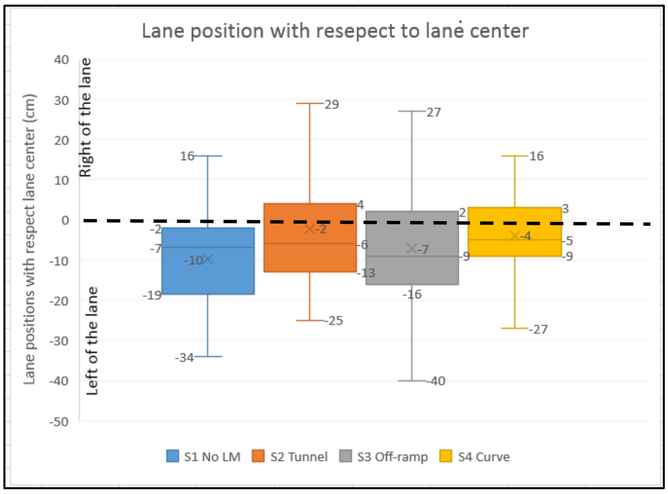


Figure 19. Situation specific position of the vehicle from the centre of the lane

Furthermore, S1 and S3 have more lane positions to the left of the lane centre compared to S2 and S4, this is confirmed by the fact that more than 50% of the lane position values in these situations are negative (data between minimum value and upper quartile). In fact, 75% of the lane position values in S1 are negative meaning that the vehicle is predominantly aligning itself towards the left of the lane centre in this situation. A similar trend is also seen in S3.

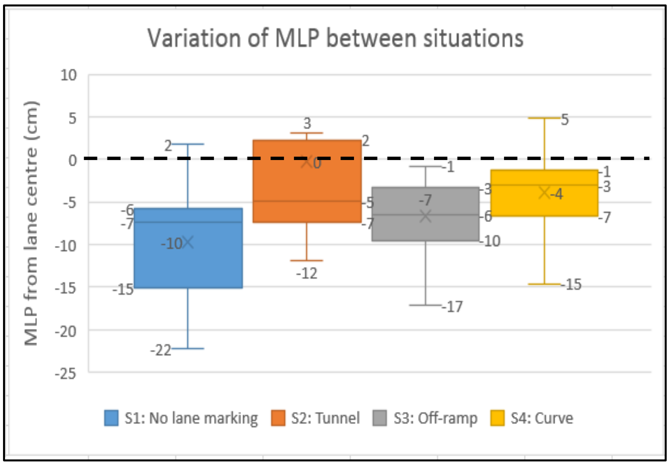


Figure 20. Variation of MLP between various test situations

To obtain further insights into the LKAS performance across the test situations, the boxplots of the SDLP and MLP were also observed. Figure 20 and Figure 21**,** depict the variation of MLP and SDLP measurements across different test situations. The corresponding values for these metrics are shown in Table 12. In this data set, the mean and standard deviation for the lane position values for each driver in each situation (over a duration of 15sec) is listed.

From Figure 20, it is visible that MLP for S1, S3 and S4 are 75% of the time towards the left of the lane centre, again confirming that in this situation the vehicle is predominantly positioned closer to the left lane marking. Out of all the situations, in S2, 50% of the data (between 1st and 3rd quartile) are the closest to the lane centre. But, there is also skew in length of whiskers and data about the median in S2, this means that Mean Lane Positions in this situation is more spread out close to the lane centre and condensed to the left of the lane centre. In general, this figure also confirms that in S1, S3 and S4 the vehicle is more biased towards the left of the lane centre than to the right.

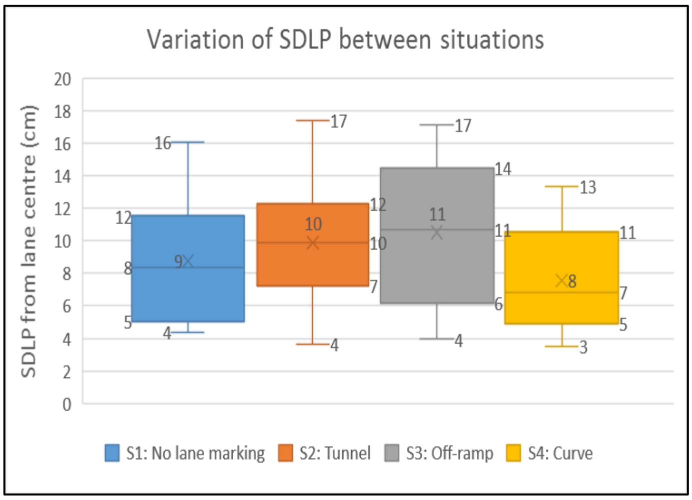


Figure 21. Variation between SDLP between different test situations

In Figure 21**,** it can be observed that the range of the maximum and minimum values of the SDLP is almost similar for all the situations with a slightly different range in S4 (Curve). The off-ramp situation S3, has 50 % of it values between 6 and 14, which is greater range of values than the 50% data points of the other situations. This gives an indication that compared to the other situations, the Autosteer performance in S3 has more variation (more spread out) than the other situations. This is an indication that the Autosteer performance in S3 is poorer than the other situations, but this needs statistical confirmation.

On the other hand; the range, length of whiskers, median/mean values of SDLP is the lowest in S4, thereby visually implying that the Autosteer performed the best in this situation, relative to the other situations. There is also asymmetry in the lengths of the whiskers in S1 suggesting that in general for this situation, there was a larger difference between the higher and lower lane position deviation values relative to the other situations. In the other situations (especially S2, S3) the length of the whiskers is almost the same on either side of the median (symmetric), this means that there were as many drives in these situations with low deviation in lane positions, as the drives with high deviation in the lane positions. It also important to note that the median value of the standard deviation is slightly lesser than in S2.

The above findings were purely based on visualization of the lane position data, but these findings were also verified using statistical tests. Since this was a case of K-dependent repeated non-parametric measurements of SDLP and MLP (normality was checked using procedure described in [107]), between different situations each driver was subjected to. The best statistical method to use according to Table 4, was the Friedman test.

The first comparison was done for the ***MLP*** values between different situations. The null hypothesis for the test was: *'There is no difference between the mean values of MLP between different test situations each driver was subjected to'.*

Table 23**,** summarises the results for the Friedman test. As a typical requirement to avoid type 1 errors in the results, the p-value for these tests need to be corrected for number of repetitions, using the Bonferroni correction. Therefore, the new α value is 0.05/4 =0.0125. For a situation where the Bonferonni correction is not applied, from Table 23**,** it can be concluded that there is statistical evidence of a difference in the MLP values for the different situations [χ (3) = 9.213, p=0.027]. This is because the p-value is lesser than α- value 0.05 and therefore, this leads to the rejection of the null hypothesis. On the other hand, when the Bonferonni correction would be applied the p-value is higher than the α (0.0125), which means that there is no evidence to reject the null hypothesis.

For the case of when the Bonferroni correction is not applied, due to a significant difference between MLP of situations, a post hoc test to the Friedman test was also conducted to get more insights into the results. For this pairwise Wilcoxon singed rank tests, were conducted. This test helps in comparing each situation with the others and gives a comprehensive understanding of the MLP variation. Just like for the Friedman test, the null hypothesis for each of the situation comparisons was as follows: *"There is no difference between the MLP values between the two considered situations that each of the drivers were subjected to".* Table 24*,* summarizes the results of the Wilcoxon tests for all situation pairs.

From Table 24**,** it is seen that there is a significant difference between MLP pairs of S2 and S1 (hereafter referred as MLP1), MLP pairs of S4 and S1 (hereafter referred as MLP2), MLP pairs of S4 and S3 (hereafter referred as MLP3). The statistical results for the respective pairs is as follows MLP1 [z=-2.147, p=0.032, r=-0.38)], MLP2 [z=-2.076, p=0.038, r=-0.37] and MLP3 [z=-2.171, p=0.030, r=-0.38]. This gives statistical confirmation that the LKS performance in these situation pairs, is different and the effect sizes (r values) also reflect a large difference between the MLP values between these situation pairs. By comparing the number of positive and negative ranks between the pairs that show a significant difference, the relative right or left alignment of the vehicle in these situations can be determined.

For the pair MLP1, negative ranks for No LM situation are higher this means that in this situation, the vehicle lane position is relatively more away from the lane center towards the left than, while driving in the Tunnel situation (because the sign of the lane position refers to direction of the vehicle's position with reference to the lane center). Similarly, the vehicle lane positions while driving on the Curve situation is more towards the right of lane positions than in the No LM situation. Finally, vehicle positions on the curve are relatively more towards the lane center than while driving close to an off-ramp.

The next step was to perform the Friedman test to compare the *SDLP* values between the different test situations. The null hypothesis for this test was "There is no difference between the SDLP values of the different test situations. "

From Table 25**,** it can be concluded that there are no statistically significant differences between the SDLP values between the situations [χ (2) =4.082, p=0.253]. Therefore, there were no further post-hoc tests conducted and the comparisons were done based on the boxplot information. To visually compare the lane position frequencies, pairwise plots were also made [2.) Appendix A]. The plots confirmed the differences as seen using situation-wise boxplots and statistical testing.

6.2.4. Discussion & Summary

The lane keeping performance ***between the drives*** was expected to be similar for the drives as it was the same route and the same LKAS system and all the test drives happened during off-peak hours. It was seen that for most of the drives the combined lane keeping performance of the driver and Autosteer was consistent but biased to the left of the lane centre. But, for drives D13, D14, D15, D17 and D19 the lane keeping performance was poorer as compared to the other drives. This showed that there is some inconsistency in the lane keeping performance based on the spread in the positions from the lane centre. But in general, the maximum and minimum range of the lane positions was similar across most of the drives, this suggested that overall the vehicle never goes more than roughly 45cms away from the lane centre on either side of it.

The inconsistent performance in D13, D14, D15, D17 and D19 could be attributed to higher traffic on the road during those drives forcing the driver to take control of the vehicle more often than on the other drives. Moreover, even when the Autosteer is ON but the vehicle is close to slower moving traffic, Autosteer's performance could be heavily influenced by the car-following function (ACC) and this could also lead to reduced lane keeping performance when surrounding traffic states are higher.

In fact, the route travel times for drives D13, D15 and D17 were higher than the other drives expect for D5 (verified by test times on the LIDAR), and on visually going through the road facing camera data for these drives, it was verified that indeed on these drives the traffic especially towards the end of the city road section and the end of entire drive was higher than the rest of the drives. This, to some extent justifies the relatively poor lane keeping performance on these drives. But it is important to note that there where the travel time was higher (more traffic) but the Lane keeping performance was biased on an average closer to the lane centre (Drive 5).

During the drives, the general alignment of the vehicle was to the left could be because, for most of each drive the vehicle is one the right most lane of the road and there is generally a guard rail / road barrier towards its right. As mentioned in the owners' manual the Autosteer may attempt to steer away and keep an offset distance to the road obstacles and road barriers. This could be a reason or the left skew in its lane position. It is important to understand, to go way from the guard rails, the vehicle is moving closer to road traffic on its left, and thereby increasing its chances of collisions with these road users.

Like expectations regarding the lane keeping performance between drives, it was also expected that the performance of the Autosteer (and the driver), should not vary considerably ***between days*** unless there was varying weather conditions or drastic changes in traffic intensity. It was observed that lane positions were moving very slightly away from the lane centre towards the left from Day 2 to Day 3 and Day 4. There is no statistically significant proof for this and it could once again be attributed to the higher average travel times (higher traffic) on days 3 and 4 compared day 2. Unusually, on Day 1 the travel time was higher and the weather conditions were worse than the other test days, yet the lane keeping performance over the two test drives on this day were better than the other test days. A plausible reason for this could be that the control algorithms on the Autosteer become stricter during adverse weather conditions (If it is ON) or during rain, drivers control more often and are more cautious and drive slower than usual. This requires further research.

On visual inspection of all the drives on Day 3, it was also seen that especially for the last three drives on this day the traffic volume on the road was higher than the other days. This explains the high standard deviation in lane positions (SDLP) on Day 3.

By comparing combined, lane keeping performance of the driver and Autosteer ***between 3 road sections*** (highway then city and then back to highway), provided results mostly as were expected. The combined performance was better (closer to the lane centre) on highway sections than the city road section for all days except day 1. The performance was better on the highway as these systems function better when there are no intersections, lesser closely driving slow moving vehicles and better marked roads. Moreover, the Autosteer does not turn ON most of the time inside the city and therefore, the vehicle is more controlled by the driver in the city than on the highway. As the performance of the driver in lane keeping can be assumed to be worse than the system (it is designed to be better), this explains why combined lane keeping performance in the city was worse than on the highway section. Moreover, it is difficult to explain the unusual trend on Day 1, because it corresponded to lane positions only for two drives. An important observation was the lane positions were predominantly left biased in all sections, this could be because the vehicles mostly drove on the right most lanes of the road which is closer to the right guard rail on the highway and close to road edges in the city.

On the other hand, standard errors to the average lane positions were smaller in the city road stretch than on the highway. This could be because, drivers are more often in charge of the driving task in the city and as they drive at slower speeds and stop quite often at intersections, their lane positions may not change much from their average lane position (this does not mean that they are closer to the lane centre, they can be further away from the lane centre but not have deviations in their average positions).

In addition to a general overview of the LKAS and driver performance across all drives, between test days and between road sections, the focus of this research was on assessment of the Autosteer's lane keeping performance between the different situations classified based on whether they were inside, outside or neither inside/outside the ODD, as specified by the vehicle manufacturers.

For the situations deemed ***inside the ODD*** based on manuals provided by the OEM, i.e. situations S2 and S4, it was seen that maximum percentage of lane positions were closer the lane centre as compared to the S1 (No Lane marking), which was outside the ODD and S3 (for which the ODD was not sure). Between these situations (S2 and S4), the lane positions had more skew to the right in S2 than in S4 and lane positions in S4 were more concentrated to the lane centre than in S2. This could be because, in S2 there were tunnel walls on either side of the lane, but the wall was closer to the left than to the right. This could explain why the Autosteer attempted to align the vehicle away from the wall that it is closer on the left. Moreover, there was also a slight left skew in the curve situation and this could again be because the vehicle was predominantly on the right lane and close to the guard rail on the right.

Driving close to the off-ramp (S3) was deemed as ***neither inside/ outside the ODD*** situation. In this situation, the range of deviation of the lane positions (between maximum and minimum) was the highest compared to the other situations. Moreover, the vehicle positions had more frequency towards the left of the lane but a slight skew towards the right. This could be because of the changing lane marking types in this situation (there are two lane markings strips on the right of the vehicle) and may be Autosteer does not know for sure which lane marking strips to be in the middle of.

On the other hand, while driving inside the city with no lane marking on the road boundaries (S1), which was classified as being ***outside the ODD***, it was seen that the Autosteer was attempting to stay close to the lane centre but majority of its lane positions are away to the left of the lane centre along with a significant skew towards the left. This could once again, be because the Autosteer recognises that there is no lane marking strip on the right and therefore, it attempts to stay close to the road centre. This could also be because there was slow moving traffic it encounters in these situations. Given the car-following behaviour of the Autosteer, it follows the trajectory of the leading vehicle (when both lane markings are absent). Since, manually driven vehicles would also ideally prefer to stray away from the road edge, this could mean that the test vehicle was just following the trajectory of a manually driven leading vehicle.

The Mean Lane Positions (MLP) and Standard deviation of Lane Position (SDLP) were also used to compare Autosteer's performance between different situations types. In terms of lane position trends, similar trends as described earlier were observed with the only difference being the mean lane positions in between the two inside ODD situations (S2 and S4) showed that more percentage of mean lane positions in the tunnel were slightly closer the lane centre than in S4 (On the curve), but there was a significant right skew in the mean lane positions. Confirming that the vehicle was indeed attempting to move towards the right of the lane centre to avoid the concrete tunnel on its left. On the other hand, the standard deviation of lane positions between the different situation types (ODD-in, ODD-out and ODD-In/Out) suggested that, in the two situations that were not inside the specified ODD, S1 and S3, the deviation in lane positions were lesser in S1 (outside the ODD) than in S3 (may be inside or outside the ODD). But the range of the deviations in these situations were more than the other two situations which were inside the ODD (S2 and S4), this was expected given their ODD classification.

This indicates that even though mean lane positions of the vehicle in S3 was closer to the lane centre compared to S1, its performance was poorer. Implying that a mean lane position closer to zero is not always an indication of the performance of the lane keeping system, but it must be combined with standard deviation observations as the Autosteer is designed to have a slightly altered lane alignment in certain situations. These differences between the mean lane position values were statistically significant but on the other hand, differences in standard deviation in lane positions, were not.

Finally, it can be concluded that the lane keeping performance of the Autosteer in the ODD-In situations (S4: Curve and S2: Tunnel situations) were better than the other situations types. Between S1: No lane marking on the road boundaries (ODD-out situation) and S3: close to off-ramp (ODD-in/out situation), the performance was slightly better in S1 than in S3 as there was larger range of deviation in lane positions in the off-ramp situation. Even though in both cases the vehicle was attempting to sway towards the left of the lane centre. It is also important to bear in mind that the variation in lane positions was in the range of a few 10's of centimetres (away from the lane centre), which when looked at macroscopically is still better than the general lane keeping performance of drivers.

6.3. Objective risk measurement

in this research, the Probabilistic Driving Risk Field (PDRF) metric is used to determine the driving risks experienced by the test vehicle, across all the different test situations. It was also mentioned that, risks only due to the non-moving fixed road entities (Potential Risk Field) were measured in this research using Equation 1 in sub-chapter 2.2 of this report.

Since this was the first attempt to use such a metric for data obtained from a specific control field test, there were a few assumptions that had to be made. These assumptions are described in this sub-chapter. For this research, the Objective Risks (OR) using the Probabilistic Driving Risk Field method, were measured for a duration of 15secs for all the test situations (S1 – No Lane markings on the road boundaries, S2 –

Driving inside a Tunnel, S3 – Around an off-ramp and S4 – At a curve) across all the test drives. These objective risks were compared between the situations, both visually using boxplots and statistically.

Furthermore, given the novelty of this Surrogate Measure of Safety (SMoS), for verification its results were also compared to an existing widely used SMoS, the Time to Lane Crossing (TLC) metric. This comparison is presented within this sub-chapter. Finally, the sub-chapter concludes by providing a discussion about all the results obtained through this analysis.

6.3.1 Assumptions and steps taken to determine Objective Risks

1) Only the first 15 seconds of each situation was considered for objective risk assessment. This method employed the lane positions determined in the previous section of this report. If for any instance during the test duration of 15secs, the lane position was not available (due to filtering or missing data), Objective risk was not calculated for that instance. This meant that there were situations where less than 15 seconds of data was available, to ensure that no situation had more than 15seconds of data (for consistency), no extra data was included to make up for missing data seconds.
2) All speeds (both lateral and longitudinal direction) and accelerations (both lateral and longitudinal direction) were calculated for every second of every situation for every drive. This was obtained by manually going through the videos of each testing situation across all drives and simultaneously using the python based visualization tool (Figure 38). In instances, where due to certain technical issues such information was not available, the subject vehicle's longitudinal motion was restricted to the speed limits of the road.
3) The reference co-ordinates for the implementation for this method, were the front left wheel of the subject vehicle, i.e. a local co-ordinate system was used to determine distances between the subject vehicle and the road entity.
4) All the filtrations performed (to account for lane change maneuvers and extreme values) while determining Standard Deviation of Lane Positions (SDLP) and the Mean Lane Position (MLP) values, were also done while objective risk determination.
5) Because of simplification for image processing, the lane width on all roads other than inside the tunnel in the city, was 3.5m per lane (excluding lane marking width). Inside the city tunnel the lane width was 3.25m (excluding lane marking width). Based on the [108] the width of the paved shoulder on the highway was 2.75m (for the off-ramp situation) and the distance to the road medians was 0.5m.
6) When there is a vehicle leading the subject vehicle, the latter speed is restricted to that of the leading vehicle because of the ACC function included in the AP for Tesla. This was used to determine the speed of the subject vehicle when GPS information about the subject vehicle was missing.
7) There were 4 type of road boundaries (non-moving road entities) considered for this research, the corresponding 'K' factors values used in Equation 1 based on the results of [76], were as follows
    - K=0.1 for a lane marking strip.
    - K=0.2 for a curb stone of the rod (when there is no lane marking strips on the road boundaries in the city).
    - K=0.5 for a concrete median on the highway.
    - K=0.7 for a concrete wall (inside the tunnel).
8) Potential field risk for an instance of a situation, is the sum of risks due to all the barrier/boundary types surrounding the subject vehicle, at that time instance.
9) Only the maximum potential risk field measurement of the 15secs duration of each situation was used for further analysis. The average values (for the 15secs duration) were not used because, in general the average risk values for the test duration were zero in most situations. The potential risk

field risk metric is dependent on the square of the lateral velocity of the test/subject vehicle. But the test vehicle is an LKAS equipped vehicle and as the Autosteer was ON in all these test situations the lateral velocity of the vehicle was quite often close to zero, thereby resulting in very small risk measurements.

10) The effect of the fork at the end of the off-ramp situation is neglected for this analysis as for this case the longitudinal velocity will also be included for the analysis.

6.3.2. Situation specific Objective Risk of driving (lateral objective risk)

In this sub-chapter, the maximum potential risk in the lateral direction (maximum of both, left and right of subject-vehicle risk measurements) experienced by vehicles in a 15sec duration within each of the test situations, will be compared using boxplots.

Based on the lane keeping performance of the Autosteer, it was expected that the lateral driving risk will be the highest in S3 (close to an off-ramp), this is because the lane keeping performance shows a largest range in this situation and in addition to this the vehicles are moving at high speeds as well close to changing lane markings and guard rail. Followed by this, in S1 (No Lane Marking on road boundaries) the lane keeping performance was poorer than S2 and S4. Therefore, lateral risk in this situation was second highest. This was followed by S2 and S4 in the same order of their lane keeping performance and the presence of concrete walls on both sides in S2 (Inside the tunnel). Therefore, the expected order of lateral risks was: S3>S1>S2>S4.

Figure 22, shows a box plot of the maximum 15sec potential field based objective risk measurement for each of the test situations and for each driver. The objective risk is expressed in Joules and a higher value represents a higher risk due to fixed barriers/boundaries on the road.

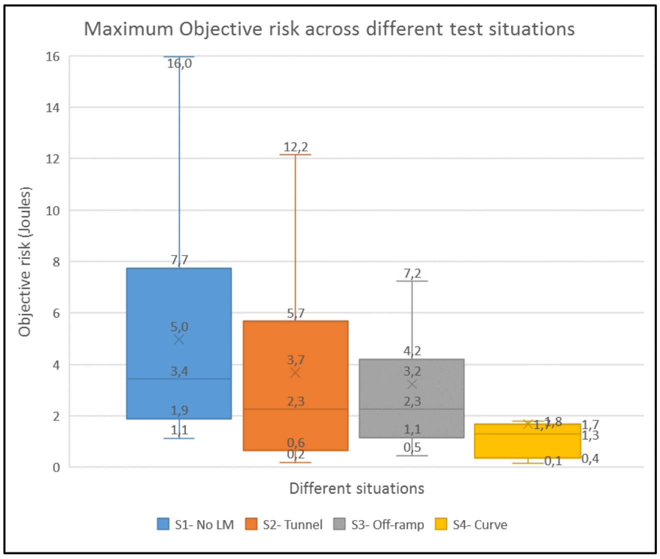


Figure 22. Variation in Maximum Objective Risk between different test situations

The mean of the Objective Risk is the highest in S1 and decreases in the order S1>S2>S3>S4. 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. An asymmetry is also observed by the unequal whisker lengths in all situations other than S4. This means that the risk values are more asymmetric towards values larger than the median, i.e. there is more spread in the risk in situations S1, S2 and S3 towards the large values. The range of the risk values is again largest in S1 followed by S2, S3 and then S4, again implying highest variation in risk while driving in S1 as compared to the other situations (this is also visible by the decreasing Interquartile range from S1 to S4). The mean, median and spread of the OR values in S4 is the smallest and most symmetric, respectively. This implies that the risk of driving on the curve (S4) was the least risky and the risk values were less varying.

6.3.3. Statistically comparing the OR between different situations

The results provided in the previous sub-chapter are visual based. There was also a need for these results to be verified statistically. Therefore, in this sub-chapter results of a statistical comparison between OR's of different situations, are described.

The objective risk measurements were computed for each test drive across the same 4 situations, making this a repeated measures experiment. Furthermore, the objective of the test was to compare the risk measurements between different test situations. Given this objective and type of experiment, from Table 4**,** either a repeated measures ANOVA or Friedman test could have been used. To decide which method to use, it was important to check for normality (to check if the data is parametric or not) of the OR measures. To do so a procedure described in [109] was used. The results of this normality check are provided in Table 26 (Appendix I: Statistical tests) and Figure 44 (Appendix I: Statistical tests). The results of the tests suggested that the objective risk measurement is not normally distributed and therefore non-parametric.

Therefore, Friedman test was used to compare risks between different situations. The null hypothesis for this Friedman test was *"There is no difference in the means of the Objective Risk values between different test situations.*"

Once again, the Bonferroni correction could be applied for a Friedman test to have control over type 1 error while reporting statistical significance. If this correction is applied, the α value would change from 0.05 to 0.0125 (again four comparisons are done). From Table 27 (Appendix I: Statistical tests)**,** the p-value is 0.000 and this means that the null hypothesis must be rejected and that there is a significant difference between the Objective risk measurements between the situations [χ (3) = 17.925, p=0.00], even after Bonferroni correction.

Furthermore, a post hoc to the Friedman test is the Wilcoxon Singed rank test. Using this test, pairwise comparisons between risk measurements between situations can be done. Table 28 **,** Table 29 (Appendix I: Statistical tests) summarize the results of these pairwise tests. The null hypothesis for the pairwise Wilcoxon tests was: "There *is no difference between the mean Objective Risk values between the two selected situations. "*

Hereafter, for the comparisons, the objective risk pair of 'S2:Tunnel' and 'S1:No Lane Marking' situations is referred to as 'OR1', the pair of 'S3:Off-ramp' and 'S1:No LM' situations is referred to as 'OR2', 'S4: Curve' and 'S1:No LM' situation pair as 'OR3', the 'S3:Off-ramp' and 'S2:Tunnel' situations pairs as 'OR4', the 'S4:Curve' and 'S2:Tunnel' situations pairs as 'OR5', and the 'S4:Curve' and "S3:off-ramp' situations pair is referred to as 'OR6'.

If the Bonferroni correction was done for these pairwise tests, the α value would have to be changed from 0.05 to 0.017 (as each situation is being compared to 3 other situations). Based on the p-values for the comparisons in Table 29**,** this would mean that only OR3 and OR5, would have a significant difference.

But, when the Bonferroni correction is not applied (α value = 0.05), then there is clearly a significant difference between Objective risk values between situation pairs OR2, OR3, 0R5 and OR6. This is because the p-values for these pairs is less than 0.05, [OR2 (z=-2.068, p=0.039, r=0.036); OR3 (z=-3.206, p=0.001, r=0.57); OR5 (z=-2.534, p=0.011, r=0.44); OR6 (z=-2.482, p=0.013, r=0.44)] and this means that there is enough evidence to reject the null hypotheses along with high effect sizes (r values greater than 0.3 [110]. The risk values for the OR1 and OR4 pairs are not significantly different from each other as the p-value is greater than 0,05 [OR1 (z=-0.362, p=0,717); 0R4 (z=-0.724, p=0.469)] and the null hypothesis cannot be rejected.

Finally, from Table 28, the signed ranks between the pairs that show a significant difference (when Bonferroni correction is not applied), can be compared. For pair OR2, the negative rank is higher meaning that the risk in the 'No Lane Marking on road boundaries' situations are higher than that in the off-ramp situation. Next, for pair OR3 the negative ranks are again more than the positive ranks, meaning that the driving risk in the No LM situation is also higher than that in the Tunnel situation. For pair OR5, the negative ranks are again higher than the positive ranks meaning that the driving risk is more in the Tunnel than while driving in the 'Curve'. Finally, for pair OR6, the negative ranks are again higher meaning that the driving risk is higher in the off-ramp situation than while driving in the 'Curve' situation. These results also confirm the visual observation that the OR values between the situations have the order S1>S2>S3>S4.

6.3.4. Comparing TLC and PRF values

An important step in this research was to verify if the novel risk field is comparable to the TLC which has been used predominantly as a surrogate safety measure while assessing risks and road safety. In this sub-section, the trends in maximum risk measurements computed by using the PDRF method and the corresponding TLC values, will be compared across the (15sec duration) different test situations (Table 13). The magnitudes of these measures cannot be compared as they different units of measurement. Therefore, the measurement trends will be compared.

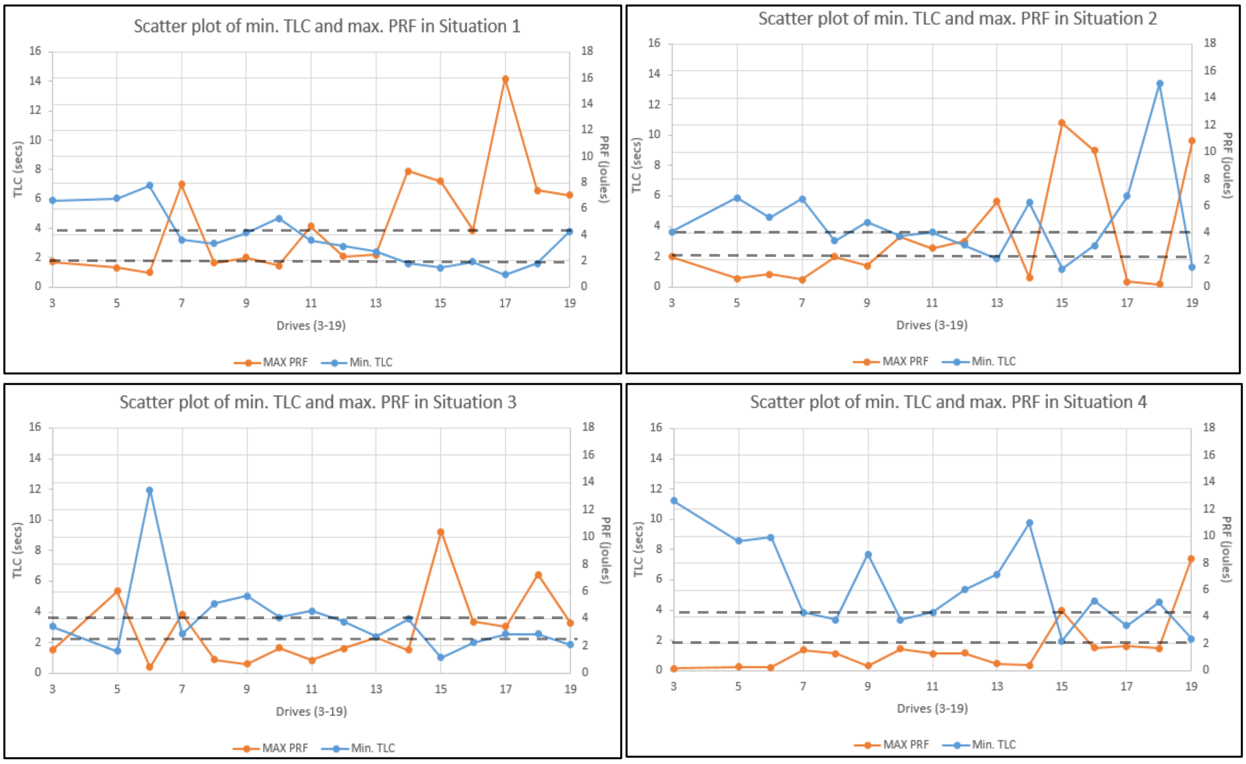


Figure 23. Comparison between PRF and TLC trends for different test situations

Figure 23**,** compares the trends in objective risk measured using the PRF (Potential Risk Field) and the TLC method, for each situation**.** Each TLC value corresponds to the absolute minimum TLC value between the left and right TLC values for each driver for 15sec duration of each situation. On the other hand, each PRF value corresponds to the maximum value of PRF between left and right of the lane centre for each driver for a duration of 15sec of each situation.

From Figure 23**,** For each of the situations, as the TLC values increase the corresponding PRF values decrease and vice versa. A large value of TLC corresponds to lesser risk of crossing the lane boundaries and thereby a safer driving situation, and vice versa. On the other hand, a low value of PRF corresponds to a lower severity and/or probability of a collision to occur and thereby, a safer driving situation, and vice versa. Keeping this in mind, the observed inverse relation between TLC and PRF means that both these measures are similar in representing the driving risk/road safety trend and the use of PRF in this research is justified. The difference between the two mainly lies in the realism in the magnitude of the risk they represent, which will be discussed in sub-chapter 6.3.5 of this report.

Additionally, the risks measured using TLC can also be classified into three categories; low risk, medium risk and high risk. From [111], TLC values lesser than (or equal to) 2 seconds are classified as 'high risk', between 2 and 4 seconds as 'medium risk' and TLC values above 4 seconds as 'low risk'. Based on this, in Situation 1 (No Lane marking), there are 5 drives with TLC less than or equal to 2 seconds and majority of the risks in the medium risk range, in situation 2 (Tunnel) there are 3 drives with TLC less than or equal to 2 seconds and an equal representation of risk values in the medium and low range. In Situation 3 (close to an off-ramp) there are 4 drives with TLC less than 2 seconds and majority of the TLC values spread over the medium and low range, and finally, in situation 4 (Curve) there are 2 drives with TLC values less than 2 seconds and majority of the TLC values in the low range.

A similar classification of the risk values measured using the risk field approach implemented in this research is not possible at this moment. This is because, the units of measurement of the two metrics are different and their magnitude represent different entities, energy transfer (PDRF) and time (TLC). Therefore, this indicates a need for further research into this novel PDRF risk metric to be able to classify the measured risks.

6.3.5. Relation between objective risk and lane keeping performance

As mentioned earlier, the PDRF metric consists of a severity term and a probability term. Looking at Equation 1 and Figure 16, Figure 17 plots, a relation between lateral objective risk between situations can be made with the lane keeping performance in these situations. The Mean Lane Position (MLP) values plot the mean positions of the vehicle within its lane therefore, its median/middle value relates to the probability of a collision to occur. The closer a vehicle is to the lane marking strip on either side of the lane, more is the probability of a collision on the respective side. On the other hand, Standard Deviation of Lane Position values and the skew in the MLP values represent the deviation in lane positions, thereby relating to the lateral velocity and hence, the severity of a collision.

Furthermore, the probability term in the Potential Risk Field formulation corresponds to an exponential function that increases as the vehicle moves closer to a lane marking and is relatively lower, otherwise (Figure 24). On the other hand, the severity is polynomial function of degree two and strictly depends on the lateral velocity of the vehicle, which is determined by the variation in its lateral position within the lane.

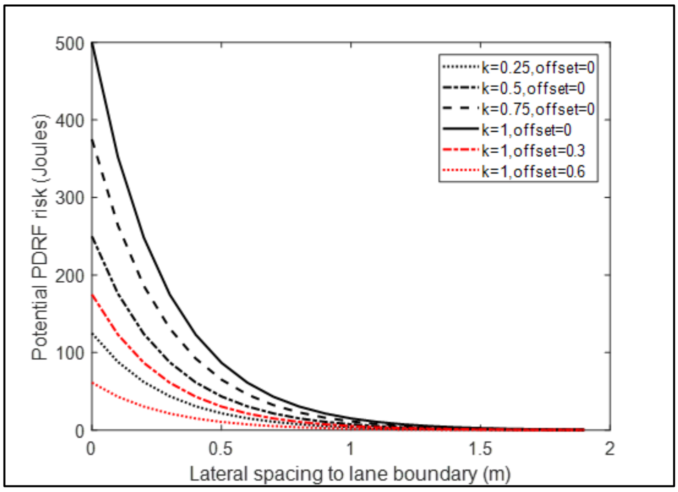


Figure 24. Variation in PDRF risk based on the spacing to lane boundary (source [75])

Therefore, this relationship between the lane keeping performances indicators and lateral objective risk, will be used while interpreting the situation specific lateral driving risks.

6.3.5. Discussion & Summary

For this research, it was important to compare the lateral objective risks of driving in the three types of ODD classified situations. That is, in situations deemed inside the ODD (S2 and S4), outside the ODD (S2: Tunnel) and situations which is neither inside nor outside the ODD S3 (Off-ramp).

The results of the lateral risk calculations were different from what were expected based on the observed lane keeping performance and characteristics of the situations (S3>S1>S2>S4). It was observed that, the lateral driving risks across the different situations were in the order S1>S2>S3>S4, with a large difference between risks in Situation 1 and Situation 4. It was observed that the maximum-minimum range of the objective risk values also follows the same order (S1 through to S4) across the situations and there is considerable top skew in the risk values in S1, S2 and S3 and not in Situation 4 was measured to be the safest situation to drive in relative to the other situations.

The Probabilistic Driving risk field (PDRF) method used for this research is based on the calculations of the severity and probability of a lateral collision. The severity was dependent the lateral velocity of the vehicle, the type of road barrier (k) and the probability was determined by the lateral distance to the road barrier and lane width.

For Situation 4, the maximum lateral risk was the lowest, a reason for this could be because its position in the lane was the most condensed close to the lane centre and the variation in its lane positions was the least. Therefore, severity and probability of collision was the least.

Furthermore, it can be seen from Equation 1, the magnitude of the lateral velocity has a significant contribution to the risk measurement. The plausible reason for the highest lateral velocity in Situation1 (severity), could be because of the condensed lane positions closer to the centre and significantly large skew towards the left of the centre and at the same time having a top skew in the Standard deviation in lane position. Simultaneously, it is seen that in this situation the range of the lane position and its mean is the most negative (close to the left lane marking strip), leading to the highest probability of collision. Therefore, as both the probability and severity of the collision is high, its maximum risk was the highest relative to all the other situations.

On the other hand, while driving inside the Tunnel (S2), the lane keeping performance was the second best but the vehicle was trying to move away to the right of the lane centre (right bias) as the tunnel wall was further away towards the right. Moreover, there was a considerable variation in lane positions as well, therefore, this could lead to a considerable lateral velocity value. Most importantly, the median of the mean lane positions was condensed to the left of the lane centre and the concrete tunnel wall closer on this side 'k' for this barrier type was the highest (0.7). Moreover, the lane width inside the tunnel was also slightly lower than in the other situations. Due to a combined somewhat high probability of collision and a fairly high crash severity, the maximum lateral driving risk in this situation was the second highest. The bias in lane positions in this situation was lower than in S1 and that could be a reason for a lower risk in S2 (Inside a Tunnel) as compared to that in S1 (Inside the city with no lane marking on the road boundaries).

While driving close to an off-ramp (S3), the severity is quite high as the possible lateral velocity is high, but the probability is low as Mean Lane Positions has no skew and the vehicle positions are more condensed towards a position close to the centre of the lane. Due to the nature of the PDRF curve (Figure 24) and the exponential probability function, the probability of collision could be low and therefore lead to a somewhat low lateral driving risk.

On statistically comparing the lateral objective risks across different situations it was seen that even after correcting for type 1 errors, there was a statistically significant difference between lateral driving risk (even between the two situations that were inside the ODD a specified by the OEM). Moreover, the post-hoc tests also resulted in significant differences between most of the situation pairs for the case when the Bonferroni corrections was not applied and showed the same trend in risk S1>S2>S3>S4. In addition to this, even when Bonferroni correction was applied, the differences between S2 and S4 (both inside the ODD) were statistically significant.

While looking at the three types of situations classified based on ODD specified by the OEM. For the ODD-in situations (S2 and S4) there was a statistically significant difference between the lateral risks of driving in these situations (even after statistical corrections), with risk of driving in the Tunnel being significantly higher than that on the Curve. Furthermore, the risk of driving in the city without lane markings on the road boundaries which was deemed outside the ODD, was the highest. The risks of driving close to an off-ramp, for which it was unsure if the situation was inside or outside the ODD were the second highest.

Finally, the risks measured by the probabilistic driving risk field (PDRF) method used in this research were also compared with risks calculated by Time to Lane Crossing (TLC), an existing widely used counterpart. Maximum risks calculated for the same instances of each situations using both the metrics, were compared. The results showed that both the situations depict the same trends in risk as they are both dependent on the lateral distance to the road barrier, but they differ in the realism in the magnitude of the risk they represent. The risk magnitudes produced by the PDRF method, also include the sensitivity to the type of road barrier the vehicle encounters. Furthermore, the PDRF method also shows an advantage of being additive as it can represent risk due to all the different types of road barriers it encounters, using only one single risk value (as performed in this research). This shows its advantages over the conventional TLC, for which the risks due to different road entities (measured in units of time) cannot simply be summed up to represent a total lateral risk.

6.4. Statistical Analyses of driver behaviour in LKAS

As mentioned earlier, for this research it was considered that enhancing driver experience within LKAS equipped vehicles, is key for its development. Therefore, it is very important to understand and address the interaction between the driver and the LKAS system. For each driver, this interaction was assessed in three stages; before the test drive, during the test drive and post the test drive. The main aim of this analysis was to investigate if there is a mismatch between, awareness of drivers about the operational design domain of the LKAS (ODD state, i.e. whether a situation is inside, outside the ODD or they are not sure)) and the operational design domain/ functionality as defined by Original Equipment Manufacturers (OEM), using a case study of Tesla Model S. The former, for this analysis, will be referred to as ODD state awareness of the driver and later as the ODD specified by the OEM. To go one step further, this analysis also investigates driver related factors that could be potential reasons for this mismatch (referring to research sub-questions 4. Of this research).

This sub-chapter first provides the descriptive statistics for the pre-drive, post-drive and real-time behavioural assessment questionnaires. Following which, further sub-questions and research hypotheses are proposed to understand behaviour of drivers in LKAS equipped vehicle. Following which the results of the statistical analyses will be presented. The sub-chapter will also include a description of the motivation behind the choice, data preparations and conduction of each statistical test before, finally summarizing and discussing the results.

6.4.1. Descriptive statistics

6.4.1.1. Pre-drive questionnaire

As mentioned earlier, the objective of this questionnaire was to understand the initial attitude of the driver towards LKAS equipped vehicles and semi-automated vehicle in general. In addition to this it was also used to generate a driver demographic data base. For this test a total of 19 participants were recruited including 3 women and 16 men with an age range from 24 to 59 ($M=41.32$, $SD=12.24$) and on average with 21years with a license ($M=21.05$, $SD=12.77$). On average, the vehicle kilometres driven by the participants in the past 6 months was in the range 1000 to 60000kms ($M=15657$, $SD = 13268$kms) and vehicle kilometres in a semi-automated vehicle in the range 100 to 100000kms ($M= 24742.1$, $SD= 30742.88$).

Certain situational factors were also inquired in the questionnaire. Out of the 19 participants, 89.47% had used a control/intervention type LKAS earlier, all participants had prior experience in at least warning type LKAS as it was a criterion for participant recruitment, and 73.68% also had experience of driving adaptive cruise control equipped vehicles. 26.32% of the participants had participated in prior road-tests and 36.84% of the drivers were aware of the term 'Operational Design Domain'. Furthermore, 57.89% drivers reported that they drove in LKAS equipped vehicles every day, 10.53% drove in LKAS equipped vehicles a few times a week, 10.53% a few times a month and the remaining 21.05% reported that they do not drive in such vehicles anymore. Furthermore, 47.37% participants reported that they always use the LKAS while they drive, 21.05% usually use LKAS while they drive, 10.53% of participants sometimes use LKAS when they drive and the remaining 21.05% reported that this question was not applicable to them.

Furthermore, 52.63% of the participants reported that they had prior negative experiences while driving with their LKAS ON. Out of these participants, 50% of them have negative experiences a few times a week, 20% once a month, 20% have such experiences less than once a month and the remaining 10% reported that they do not drive in LKAS equipped vehicles anymore.

6.4.1.2. Real-time trust and ODD awareness

All participants were asked two specific questions at specific situations during their test drive. These questions corresponded to their trust on the LKAS in the situation they drove in, and whether they believe the situation they drove in was inside, outside or neither inside/outside the operational design domain of the Autosteer (hereafter referred as 'ODD state awareness'). A snapshot of these real-time driver reported measurements is presented in Table 22 (Appendix I: Statistical tests). This table includes several questions based on the list of situations presented in Table 2, were asked to the participants. It is important to note that the participants were not asked to answer test questions while they were negotiating a situation, only once they passed a situation they were questioned about their trust and ODD state awareness in that situation. In Table 22, the situations marked in grey correspond to highway situations and the ones marked in orange correspond to city road situations.

6.4.1.3. Post-drive questionnaire

After their test drives, participants were asked to fill in an online questionnaire about their driving experience during the road test. In addition to the situation specific questions, a few general questions were also asked. The data gathered using the questionnaire was used for testing several statistical hypotheses discussed in the following chapter.

6.4.2. Research hypotheses and questions

Before investigating for a mismatch between the driver's awareness, and OEM's specification about the ODD, first step was to understand in general how drivers behave in LKAS equipped vehicles. This included investigating for relationships between, drivers attitude and responses before (initial), during (real-time) and after (post) their drives, and other factors that affect their behaviour.

This research includes dispositional, situational and learned factors affecting driver trust. But due to shortage of time and a small sample size of 19 participants the dispositional (demographics related factors) were excluded from the analysis.

Factors such as situation specific perceived risk, perceived ease of using LKAS systems, frequency of using these systems, prior negative experiences in the system and other factors (from the questionnaires) were still included for the analysis.

Therefore, this analysis required the following *sub-questions* to be answered to finally answer research sub-question 4.) of this research. The analysis (whenever possible) considered both, between drives and between situations, aspects.

1) Is there an influence of time spent by a driver in the vehicle on his/her real-time trust ratings?
2) What is the influence of pre-drive driver behaviour variables on their awareness about Operational Design Domain (ODD) of the Autosteer and on their real-time trust ratings?
3) Is there a relationship between the real-time trust ratings of the drivers and their awareness about the ODD of the Autosteer function (ODD state awareness)?
4) Is there a relationship between the drivers' perceived risk of driving in a situation and their awareness of the ODD state, and on real-time trust ratings in that situation?
5) Does the ease of driving in a situation have an influence on the drivers' real-time trust ratings and the ODD state awareness in that situation?
6) Are the real-time trust ratings different for Tesla and Non-tesla experienced drivers, across the different test situations?
7) Does the real-road objective risk of driving measured across different situations, have an influence on the drivers' awareness of the ODD state, and on their real-time trust ratings?
8) Is there a mismatch between the drivers' awareness of ODD state (In/out/ maybe in or out) and the ODD as specified by the OEM (Tesla)? Are these mismatches statistically different in the different test situations?
9) What are potential factors that could lead to a mismatch between the ODD state awareness of drivers' and that determined by the OEM's (Tesla)?

Each of these questions have a research hypothesis associated with it and will be described and tested individually using the methods described in Table 5**.** While analysing each hypothesis, the motivation behind choosing a method for the analysis will be described individually in the subsequent sub-chapters. This chapter will also include a description of the steps taken to enrich or filter the data for the various statistical analyses that were performed.

6.4.3. Selection of analysis method

Depending on the type of data, type of outcome and independent variables, the and objective of the hypothesis, the chosen methods of analysis differ considerably. A lot of research ([107, 112, 113]) have been focusing on the advantages and disadvantages of several statistical techniques, concluding that it is crucial that the appropriate method is chosen to yield reliable and trustful results. With the help of these researches and Table 4, the methods used to answer each of these proposed hypotheses were selected and depicted in Table 5.

Table 4. Standard statistical tests based on research objective (Source [114])

	Parametric	*Non-parametric*
Assumed distribution	Normal	No assumption
Assumed variance	Homogenous	no assumption
Typical data	ratio or interval	ordinal or nominal
Observations	independent	any
Usual central measure	mean	median/mode
Benefits	more solid conclusions	simplicity, unaffected by outliers
Correlation test	pearson	spearman
Relation between categorical variables	chi-sqaure tests	
Independent measures, 2 groups	independent measures t-test	Mann-whitney test
independent measures > 2 groups	one way ANOVA	kruskal-wallis H test
dependent measures, 2 measures	Dependent measrues t-test	Wilcoxon signed rank test, McNemar test
dependent measures > 2 measures	Repeated measures ANOVA	Friedman test. Cochran Q
one categorical independent measure and >2 dependent measures	one-way MANOVA	

Table 5. Research hypotheses, tested variables and method for hypothesis testing

Subject of Analysis	***Tested variables***	***Type of Variables***	***Step 1 of the analysis***	***Use of variables in the analysis***	***Step 2 of the analysis***	***Use of variables in the analysis***
$H_0{}^1$ **(SQ1)**	Driving situations (S1 – S4)	Categorical	Friedman test	No lane marking Tunnel Off-ramp Curve	Wilcoxon signed rank test (Post hoc)	6 tested pairs No lane marking with tunnel/off-ramp/curve. Tunnel with Off-ramp/curve. Off-ramp with Curve
	Real-time trust on LKAS (Dependent)	Ordinal		Likert scale (1= Very Less Trust, 5= Very High Trust)		

$\boldsymbol{H_0^2}$ **(SQ2)**	"Initial trust" "prior knowledge ODD" "frequency LKS" "Initial perceived risk" "Ease of use" "prior negative exp." "Awareness of capabilities"	Ordinal Categorical Ordinal Ordinal Ordinal Categorical Ordinal	Spearman correlation	Ordinal Likert scale (1=Very less xx, 5= Very high xx) Where XX are different variables Categorical Yes, No	-	-
	Real time Trust	Ordinal	Chi-square test for ODD_driver and the categorical variables	Likert scale (1= Very Less Trust, 5= Very High Trust		-
	ODD_driver	Categorical		Categorical In, Out, Not sure		-
$\boldsymbol{H_0^3}$ **(SQ3)**	Real time Trust	Ordinal	Spearman correlation	Likert scale (1= Very Less Trust, 5= Very High Trust	-	-
	ODD_driver	Categorical		Categorical In, Out, Not sure		
$\boldsymbol{H_0^4}$ **(SQ4)**	Post drive perceived risk	Ordinal		Likert scale (1= Very Less risk, 5= Very High risk)	Phi Cramer V	Paired comparisons
	ODD_driver	Categorical	Chi-square test	In, Out, Not sure		
	Real time trust	Ordinal	Spearman correlation	Likert scale (1= Very Less trust, 5= Very High trust)		
$\boldsymbol{H_0^5}$ **(SQ5)**	Ease of driving (post drive)	Ordinal		Likert scale (1=Very Easy, 5= Very Difficult)	-	-
	Real time trust	Ordinal	Spearman correlation	Likert scale (1= Very Less trust, 5= Very High trust)	Cramer V	Paired comparisons
	ODD_driver	Categorical	Chi-square test	In, out, Not sure	Cramer V	Paired comparisons
$\boldsymbol{H_0^6}$ **(SQ6)**	Category of driver (Tesla and Non-Tesla drivers)	Categorical	Mann-Whitney test		-	-
	Real time trust	Ordinal		(1= Very Less trust, 5= Very High trust)	-	-
$\boldsymbol{H_0^7}$ **(SQ7)**	Objective risk	Interval	Chi-square test		Eta squared coefficient	-
	ODD_driver	Categorical		In, Out, Not sure		-
	ODD_oem	Categorical	Chi=square test	In, Out, Not sure		-
	Real time trust	Ordinal	Spearman correlation	Likert scale (1= Very Less Trust, 5= Very High Trust		-
$\boldsymbol{H_0^8}$ **(SQ8)**	Different test situations (S1-S4)	Categorical	Cochran's Q test	In, Out, Not sure	McNemar's test	Pairwise comparisons between all situations
	Mismatch between ODD driver and ODD by OEM	Categorical dichotomous		In, Out, Not sure		-
$\boldsymbol{H_0^9}$ **(SQ9)**	"Objective driving risk" "Real-time trust"				Eta, Cramer's V, phi	-

	"Frequency of using LKAS" "Initial trust on AV's" "Perceived risk" "Awareness of LKAS capabilities" "Prior negative experience"		Pairwise chi-square tests			
	Mismatch in different test situations (S1-S4)	Categorical dichotomous		(1=Mismatch, 0= match)		-

6.4.5. Results of statistical tests

This sub-chapter, presents the results of all the statistical test that were conducted to answer the questions in Chapter 6.4.2. For all the statistical tests, the α-value (threshold value for statistical significance) was chosen as 0.05. For any of the proposed null hypotheses to be rejected, the corresponding p-value had to greater than this α-value. If this was not the case, the null hypothesis could not be rejected.

SQ1: Trends in real-time trust ratings

Drivers were asked to report their trust on the Autosteer at the specific test situations, during their drive. Prior research indicated that drivers' trust in automation could vary with the time spent in the automation system. Following are the results of the visual and statistical checks for these trust trends.

In Figure 25, the test situations (S1-S4) were arranged in the chronological order they were experienced by drivers. The real-time trust ratings were recorded in a Likert scale from (1 – Very low trust to 5 – Very High trust). Therefore, in this figure, the maximum trust rating, over the four situations was 20. Visually, there is no clear evidence for an increase or decrease in the trust ratings between situations. It can also be observed that there is quite a lot of variation in the total trust ratings (over all situations), between the drivers and on average all driver generally has high trust (mean of approximately 4 in all situations) across the different situations.

The next step was to check statistically, if there was an influence of time spent in the system on drivers' real-time trust. The following null hypothesis was set for this test:

H_0^1: There is no influence of time spent in the vehicle on the real- time trust of drivers across different situations.

The real-time trust ratings of the drivers were measured at an ordinal scale and therefore, considered as a non-parametric variable. Moreover, this is a repeated measures test as the dependent measures here are the real-time trust ratings of different drivers for the same situations. Based on this and from Table 4, the Friedman test was used for this analysis.

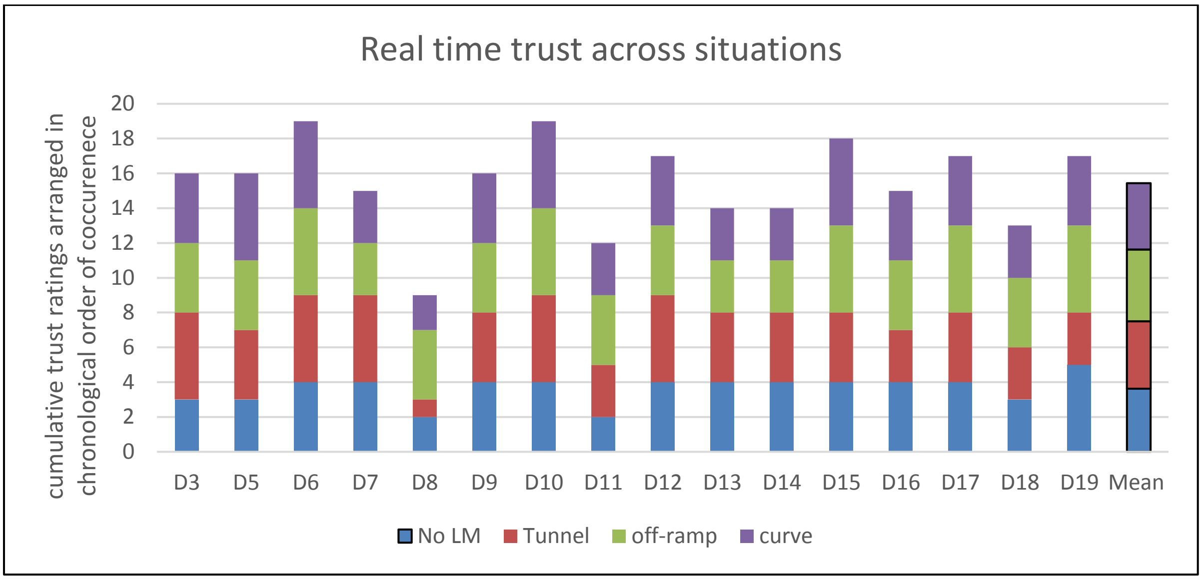


Figure 25. Real time trust ratings between and within drivers across situations

On performing a Friedman test, the $\chi^2(3) = 3.418$, $p=0.332$ (Table 30). Since the p-value is greater than the α value of 0,05 this means that the null hypothesis cannot be rejected and therefore, ***there is no statistical evidence that the trust ratings of drivers was influenced by the time spent in the vehicle during the testing.*** This meant that further post-hoc tests were not needed.

SQ2: Relationship between pre-drive attitude of drivers towards AV's and LKAS, on their real-time trust ratings and ODD state awareness

Based on the questions asked in the pre-drive questionnaire, a list of factors that could have an influence on the drivers' behavior with LKAS were selected. This was based on the overview of relevant literature presented in this report. The list therefore, does not include all the factors that could affect driver behavior in automation, which were included in the questionnaire. The factors considered were:

1) Drivers' Initial trust on AV's in general. (Likert scale from 1-Very less trust to 5- Very high trust)
2) Frequency of using LKAS (Likert 0- this question not applicable to me to 5 – Always)
3) Initial perceived risk of driving in LKAS equipped vehicles (Likert scale from 1- Not at all risky to 5- very risky)
4) Perceived ease of driving in a test situation. (Likert scale from 1-very easy to 5- very difficult)
5) Awareness of the conditions in which LKS can function (1- not at all aware t0 5- completely aware)
6) Having prior knowledge about ODD of LKS equipped vehicles (1- yes, 0 – no)
7) Having prior negative experiences while driving in LKS equipped vehicles. (1- yes, 0- no)

The relationships were tested pairwise for each factor, with the drivers' real-time trust rating and their ODD state awareness, for each test situation. Since, each of the factors are of either categorical or ordinal type, the correlations were tested either using spearman's correlation or the chi-square test for independence, depending upon whether the independent-dependent variable pair was categorical-categorical (chi-square test), categorical-ordinal (spearman's correlation) or of ordinal-ordinal (Spearman correlation) type. For each of the tests the null hypothesis was the same and as follows:

H_0^2: There is no relationship between the drivers' initial attitude factor(s) (listed above) and their real-time trust on the LKAS, and their ODD state awareness, they are independent of each other.

In the above null hypothesis, the independent variables were all the factors affecting initial attitude of the driver and the dependent variable was either the real-time trust rating or the drivers' ODD state awareness, in each situation.

First, spearman rank correlation tests ***real-time trust ratings*** as the ***dependent variable***, were conducted. The factors were included into a correlation matrix and therefore, in addition to analysing the intended relationships, other internal correlations are also reported (between factors relationships). For each of the tests the threshold α-value was 0.05, that determined if the null hypothesis was rejected or not.

From Table 31, Table 32 the p-values for all correlation tests with ***Trust in S1*** *(No Lane marking on road boundaries)* and ***Trust in S2*** *(while driving inside the city tunnel)* as the dependent variables, are greater than 0.05. This means that there is no evidence to reject the null hypothesis for each of the independent-dependent pairs, for both these situations. ***Therefore, there is no influence of initial attitude of driver towards LKAS and AV'S on their trust on the system while driving in both these situations.***

However, for the case of ***Trust in S3*** *(Close to an off-ramp)* as the dependent variable, from Table 33, there is a strong and significant relationship between real-time trust of drivers while driving close an off-ramp and their 'initial perceived risk of driving in LKAS equipped vehicles', and their 'frequency of using LKAS'. ***There is a statistically significant negative association between the real-time trust and initial perceived risk [$r_s(16) = -0.644$, $p = 0.007$] and a positive association between real-time trust and frequency of using LKAS [$r_s(16) = 0.741$, $p = 0.001$].*** Both these associations are strong as the r_s values are greater than 0.5 [110].

Finally, for the case of ***Trust in S4*** *(Driving on a curve)* as the dependent variable, from Table 34, ***there is a strong positive correlation between the real-time trust rating of drivers while driving on a curve on the highway and their 'frequency of using LKAS' [$r_s(16) = 0.869$, $p = 0.000$].*** This correlation is strong as the r_s values is again greater than 0.5.

In addition to this, there were also a few strong between-factors correlations observed, such as:

1) 'Perceived Ease of use' and 'Initial perceived risk of driving' are positively correlated. As the initial perceived risk increases the difficulty level of using LKAS systems as reported by drivers also increases [$r_s(16) = 0.587$, $p = 0.017$].
2) 'Awareness of the conditions in which LKS can function' is positively correlated with 'Having prior knowledge about ODD of LKS equipped vehicles' and with prior negative experience [$r_s(16) = 0.589$, $p = 0.016$], [$r_s(16) = 0.599$, $p = 0.014$], respectively.
3) Prior knowledge about the capabilities of LKAS and prior negative experience are correlated positively [$r_s(16) = 0.63$, $p = 0.009$]. This implies that as a result of prior negative experiences in LKAS equipped vehicles, drivers have more knowledge about its functional capabilities.

All these correlations are strong as the r_s values are greater than 0.5.

The *next* set of analyses, included the ***Drivers' ODD state awareness*** as the ***dependent variables*** and the same independent variables as were for the first set of analyses with real-time trust ratings.

For this analysis, depending upon whether it was a categorical-categorical type or a categorical-ordinal type of correlation either the chi-squared test for independence or the spearman rho correlations test, was used. For correlations between ODD state awareness of the driver (categorical variable) and Categorical dichotomous variable 'prior knowledge about ODD functions' and 'prior negative experiences in LKAS equipped vehicles' the chi-square test of independence is preferred over spearman's rank correlation. On the other hand, for correlations between ODD state awareness of the driver and the other pre-drive initial attitude variables, spearman rho correlation was used. The null hypothesis was the same as $\boldsymbol{H}_0{}^{\mathbf{2}}$, only that the dependent variable was the 'ODD state awareness'.

Table 35**,** presents the results of spearman correlation tests. From this table, it is observed that ***there were no statistically significant correlations between ODD state awareness of drivers' and any of the (ordinal type) factors constituting the initial attitude of drivers towards AV's and LKAS***, for any of the test situations. This was because all the p-values are greater than the threshold α-value of 0.05.

Furthermore, the results of chi-square test for independence between ODD state awareness and the remaining two categorical dichotomous variables 'prior knowledge about ODD functions' and 'prior negative experience with LKAS', across all the test situations, are presented in Table 36. From the table, ***there is no sufficient evidence to reject the null hypothesis.*** This is because both the likelihood ration and linear-by-linear association statistics are greater than 0.05. The reason why these test statistics were chosen to decide statistical significance instead of the chi-square statistic, is because for both these the general assumptions for chi-square tests were violated (percentage cells with insufficient count must be less than 20%).

SQ3: Relationship between real-time trust ratings of drivers and their ODD state awareness

This was an ordinal-nominal type of correlation and therefore, spearman's rho test was used. The null hypothesis for the test was as follows:

H_0^3: There is no statistically significant relationship between the drivers' ODD state awareness and their real-time situation specific trust ratings.

From Table 37, it can be stated that ***there are no correlations between the real-time trust ratings of drivers and their ODD state awareness, in any of the selected test situations***. This is because the p-values for all the tests are greater than α-value 0.05, giving no evidence to reject the null hypothesis of independence.

SQ4: Relationship between post-drive perceived risk and real-time driver behavior measurements

The real-time measurements include drivers' real-time trust ratings and their ODD state awareness about LKAS, across all situations. Therefore, for this analysis, there were two pairs of variables for which statistical testing needed to be conducted. The first test pair, was 'ODD state awareness' of the driver and 'post-drive perceived risk' (across all test situations), and the second test pair was 'real-time trust ratings' of drivers and their 'post drive perceived risk' (across the 4 situations). For each pair, the null hypothesis was as follows:

H_0^4: There is no statistically significant relationship between the drivers' post-drive perceived risk of driving in a situation and their real-time situation specific trust ratings/ ODD state awareness across all test situation.

For the ***first pair of tests***, Table 38, presents the results of the chi-square tests that was used to test for independence across the 4 different situations. For all correlations across the situations, assumptions for chi-square tests were violated (cells with no counts must not be less than 20%), therefore as mentioned in [113] the p-value of the likelihood ratio test statistic and/or the linear-by-linear test statistic was used, instead of the chi-square test statistic.

Keeping this in mind, it was observed that for ***Situation S1 (No Lane markings on road boundaries) and Situation S2 (Driving inside a tunnel), there is no statistically significant correlations***

between the ODD state awareness of drivers and their post-drive perceived risk of driving in those situations, as the corresponding p-values are greater than 0.05. On the other hand***, in Situations S3 [LL (3) =7.951 p=0.047] and S4 [Linear-linear (1) =7.335 p=0.007], it was observed that there is a statistically significant relationship between the two variables*** in consideration as in both the cases the null hypothesis was rejected as the p-values greater than 0.05.

For the ***second pair of tests,*** Table 39***,*** presents the results of the spearman correlation tests (as both are ordinal variables) across all the test situations. From the table, it was observed that for ***S1, there were no statistically significant correlations*** as the p-value was greater than 0.05. ***In S2 (Driving inside a tunnel), there was a statistically significant negative correlation between real-time trust ratings and post-drive perceived risk*** as the p-value is lesser than 0.05 and the null hypothesis was rejected [r_S(*16*) = *-0.583*, *p = 0.018*]. This is a strong correlation as Ir_SI > 0.5. Furthermore, in both ***situations S3 (close to an off-ramp) and S4 (On the curve), there was no statistically significant correlations between their respective, real-time trust ratings and post-drive perceived risk of driving*** in both these situations as the p-values were greater than 0.05 and the null hypothesis could not be rejected.

However, once again there were certain notable internal correlations that were also observed. ***The real-time trust in Situation S4 (on the curve) is negatively correlated with post-drive perceived risk of driving inside the tunnel,*** as the p-value is greater than α-value 0.05 [r_S (*16*) = *-0.621*, *p = 0.01*]. Furthermore***, the post-drive perceived risk of driving close to the off-ramp is strongly positively correlated to the perceived risk of driving on the road with no lane marking*** on its boundaries [r_S(*16*) = 0.845, *p = 0.001*]. These correlations were strong as Ir_SI > 0.5.

SQ5: Relationship between perceived ease of driving in a situation and the real-time behavior measurements

Like the previous statistical test. For this analysis, there were two test pairs for which correlation tests were conducted. The first test pair, was 'ODD state awareness' of the driver and 'perceived ease of driving' (across all test situations), and the second test pair was 'real-time trust ratings' of drivers and their 'perceived ease of driving' (across the four situations). For each pair, the null hypothesis was as follows:

$H_0{}^5$: There is no statistically significant relationship between the drivers' perceived ease of driving in a situation and their real-time situation specific trust ratings/ ODD state awareness across all test situation.

For the ***both test pairs,*** spearman's rank correlation tests were used to investigate their respective relationships. Table 40 and Table 41***,*** presents the correlation matrix resulting from tests for both test pairs, respectively. It was observed that ***there is no statistically significant relationship between both, ODD state awareness and the drivers real-time trust ratings, with perceived risk of driving in a situation, across all the test situations***.

This is because for all correlations in both test pairs p-values are greater than α-value 0.05, and the null hypothesis could not be rejected. However, there was an internal factor ***strong positive correlation observed between perceived ease of driving in S4 (Curve) and ease of driving in S2 (Inside a tunnel***) [r_S(*16*) = 0.631, *p = 0.009*].

SQ6: Comparison of real-time trust ratings between drivers with, and without experience in a Tesla

Differences/similarities between real-time situation specific trust ratings between drivers with and without experience of driving in a Tesla, were investigated both visually and statistically. Figure 26, depicts the trust ratings of drivers (Drivers 3,6,7,11,16 and 18 were Tesla experienced and the rest were not experienced of driving in a Tesla).

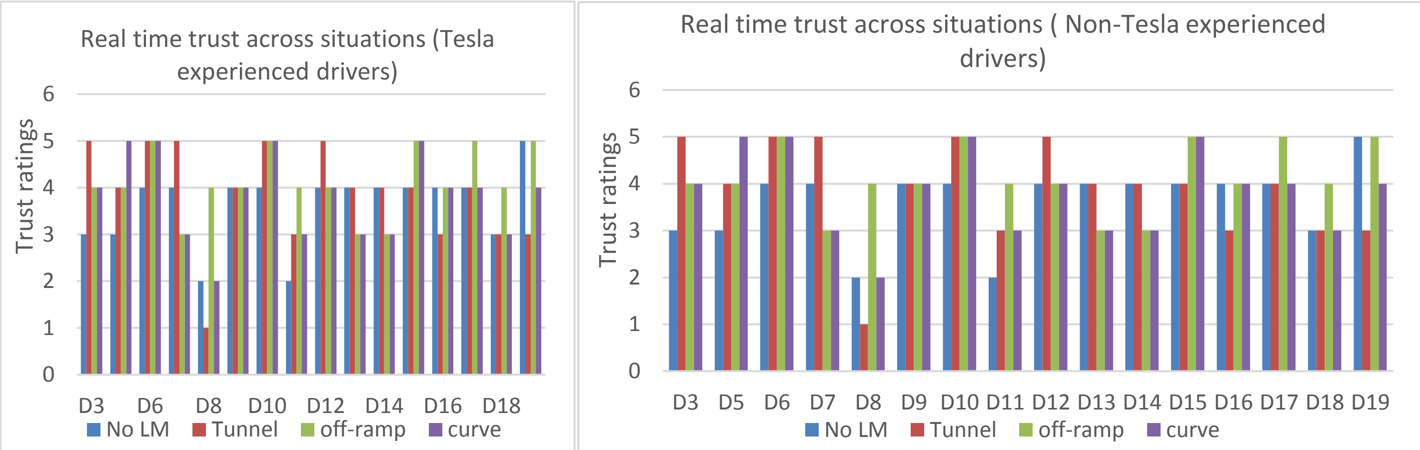


Figure 26. Real-time trust ratings Tesla VS Non-Tesla experienced drivers

From this Figure, there is no clear difference between the two categories of drivers observed visually. Therefore, the next step was to investigate the same statistically.

Since the dependent variable in consideration for this test (real-time trust ratings) is of ordinal type and the driver categories are independent, the Mann-Whitney U test was used. Median comparisons were done between the two groups for each of the situations. Having only 6 drivers with prior experience in Tesla as opposed to 13 who weren't, does not affect the choice of the statistical test [113]. The null hypothesis for the test was:

H_0^6: There is no statistically significant difference between the real-time trust ratings of drivers with, and without experience of driving in a Tesla, across all test situations.

The results of this test are presented in Table 42, it was observed that ***across all the test situations, there was no statistically significant different between trust ratings of the two categories of drivers***. The null hypothesis could not be rejected as the p-value for each situation was greater than α-value 0.05.

SQ7: Relationship between objective risk and real-time trust, ODD state awareness

In this research, potential relationships between objective measured driving risk (interval type variable) and real-time trust ratings of drivers, and their awareness of the ODD of the Tesla. This analysis was split into two parts. The first part investigating the relationship between Objective Risk and real-time trust across all test situations and the second, investigating relationship between Objective Risk and ODD state awareness of drivers across all situations. The null hypothesis for both test parts was:

$H_0{}^7$: There is no statistically significant relationship between objective risk of driving in a situation and the drivers' real-time trust ratings/ and their ODD state awareness, across all test situations.

For the first part, as these variables are of interval (Objective risk) and ordinal type (real-time trust) respectively, using the spearman's rank correlation test was the ideal choice. From Table 43**,** since the p-value for all variable pairs is greater than α-value 0.05, there is no evidence for the null hypothesis to be rejected, implying that ***Objective Risk has no effect on the real-time trust ratings of the driver in any of the situations.***

For the second part, as these variables are in the interval (Objective risk) and categorical type (ODD state awareness), respectively. The best way to measure the associations between them is by using chi-square tests and following up by eta squared coefficients to determine effect size (if any). From Table 44, it was observed that ***there is no statistically significant relationship between objective risk of driving in a situation and the drivers ODD state awareness in that situation, across all situations***. This was because the p-values for all situations were greater than 0.05 and the null hypothesis could not be rejected. Therefore, follow up eta-squared tests were also not conducted.

SQ8: Mismatch between ODD state awareness of driver and ODD specified by OEM

During their drive after each test situation, each driver was asked to report whether they think that the situation they drove on was inside, outside or they were not sure about, the Operational Design Domain of the test vehicles LKAS (Autosteer). This was referred as the ODD state awareness of the driver. However, for each of the situations based on the owners' manual provided by the OEM (Tesla in this case) the situations were also classified into three ODD categories (In, Out and Not sure) in Table 1. This research investigated if there was a mismatch between these above-mentioned variables.Table 6, depicts mismatches between ODD states specified by driver and the OEM across all the four test situations (S1 – No Lane marking, S2- Inside a Tunnel, S3- Close to an off-ram and S4 – On a curve). In the table, a mismatch corresponds to when either of the ODD stated by driver is different from that of the OEM. If There is a mismatch this was denoted by the number '1' in the table and '0', otherwise. The ODD stated by drivers in real-time at each of the test situations are as shown in Table 51, and were used an input to identify ODD mismatches.

In this table, the total mismatches and its corresponding percentages, were computed for each driver and for each situation. The drivers with prior experience of driving in a Tesla are highlighted in grey. From Table 6, it was observed that ***there was no driver without at least one mismatch between their ODD state awareness and the ODD as specified by the OEM, across the test situations***. Looking at the mismatches, it was seen that the highest number of mismatches was in Situation 3 (close to an off-ramp) and majority of the drivers believed that the situation was inside the ODD, a few of the drivers were right in reporting that the situation was neither inside nor outside the ODD and there were no drivers who thought that the situations was outside. In Situation 1 (No lane marking on road boundaries), there were no drivers who believed that this situation is neither inside nor outside the ODD (they either thought it was inside or Outside the ODD). In Situations 2 and Situations 4, the number of mismatches were the lowest, with mismatches in the Tunnel situation being slightly higher.

To check statistically if the ODD mismatches were significant, logistic regression had to be performed between the categorical variables using dummy coding. But, given the sample size of the data set, the results were not significant. Nevertheless, the ODD mismatch between different situations could be compared.

Table 6. Mismatch between ODD state awareness of driver and ODD specified by Tesla (OEM)

ID	*Mismatch S1*	*Mismatch S2*	*Mismatch S3*	*Mismatch S4*	*Mismatches per driver*	*%mismatch per driver*
3	0	0	1	0	1	25
5	0	0	1	0	1	25
6	1	0	1	0	2	50
7	1	1	1	0	3	75
8	1	0	1	1	3	75
9	0	0	1	0	1	25
10	1	0	1	0	2	50
11	0	1	0	0	1	25
12	1	0	0	0	1	25
13	1	0	1	0	2	50
14	1	0	0	0	1	25
15	1	0	1	0	2	50
16	1	0	1	0	2	50
17	1	0	1	0	2	50
18	0	0	1	0	1	25
19	1	0	1	0	2	50
Total	11	2	13	1		
% mismatch	68,75	12,5	81,25	6,25		

To verify statistically, if the differences between ODD mismatch between situations were significant a Cochran's Q test was used. This test has been used widely in pharmaceutical and psychological research [114], the test involved comparing categorical dichotomous mismatch in ODD between different situations as a repeated measure. The α-value 0.05 was again changed to 0.0125 (4 comparisons) complying with the Bonferroni correction for type 1 error. The null hypothesis for the test was:
H_0^8*: There is difference between the ODD mismatch of drivers between the different test situation they drove on*.

Table 49, presents the results of the Cochran's Q test. ***It was seen that even after applying Bonferroni correction the ODD mismatch in the test situations were statistically significant*** [*Q* (3) = 24.6, *p = 0.00*], as the p-value was lower than the corrected α-value of 0.0125.

Furthermore, to compare the mismatch between the situations pairwise, the McNemar test was conducted post-hoc. The null hypothesis of this test was similar to the Cochran-Q test but now only mismatch between two situations were compared. Once again, the Bonferroni correction was conducted and the α-value was change to 0.0167 (each situation compared to three other situations) Table 50, presents the results of these tests and it was seen that for all the test pairs other than Situation 1/Situation 3 and Situation 2/Situation 4, ***the differences in ODD mismatch between the situations were statistically significant even after Bonferroni corrections,*** as the p-values for the pairs with significant difference was less than α-value 0.0125.

Finally, it was also observed from Table 6 that the six drivers with prior experience of driving in a Tesla had either two or more than two mismatches (out of the 4 situations). ***The odds of a mismatch by a Tesla experienced driver was 0.42 and that by a driver not experienced of driving in a Tesla was 0.425, resulting in an odds ratio of 0.99.*** This could imply that having prior experience of driving in Tesla may not mean that the drivers are more aware of the ODD of the vehicle. Finally, while comparing mismatches between situations, it was observed that order of ODD mismatch between the situations was as follows: S3> S1> S2>S4.

SQ9: Potential reasons for Mismatch between ODD state awareness of driver and ODD specified by OEM

Once it was verified that there are mismatches between the ODD state awareness of the drive and that specified by the OEM. The next step further, was to identify potential reasons for these mismatches.

To assess reasons for a mismatch (and/or a match) between ODD state awareness of the driver and ODD specified by the OEM, situation specific pairwise chi-square tests were conducted. The data was divided based on whether there was a mismatch or not. For each situation, the dependent variable was the dichotomous variables that suggested if there was a mismatch (represented by the number '1')/match (represented by the number '0'), and the independent variables included all those factors that showed significant relationships/associations with the drivers' real-time behaviour measurements, because of all the prior hypothesis testing in this chapter.

The factors chosen as independent variables were: objective risk of driving (Interval type data), real-time trust ratings of drivers (ordinal data type), perceived ease of driving in a situation (ordinal type), frequency of using LKAS equipped vehicles (ordinal scale), drivers' initial pre-drive trust on AV's (ordinal type), post-drive perceived risk of driving in a situation (ordinal type), Awareness about capabilities of LKAS (categorical dichotomous) and prior negative experience in LKAS equipped vehicles (categorical dichotomous type data). The null hypothesis for all the tests was as follows:

$H_0{}^9$: There is no statistically significant relationship between mismatch of ODD state awareness of drivers' and the ODD as specified by the OEM (Tesla), and the various (independent variables), across all test situations.

It is important to note that, for all the chi-square tests performed in the following analyses, the assumptions for chi-square tests as discussed earlier, were violated. Therefore, instead of the p-value for the chi-square test statistic, the corresponding values for Likelihood ration test statics and/or linear-by-linear test statistic were used for the categorical-ordinal type interactions and for the categorical dichotomous-categorical dichotomous interactions the fisher's exact test statistic was used.

In Situation 1: No Lane Marking on road boundaries, from Table 45, it was observed that there was almost ***statistically significant relationship between ODD mismatch and real-time trust ratings of drivers in this situation using likelihood ratio statistic*** [$\chi^2(15) = 7.84$, $p=0.05$], at α-value 0.05. For test with the ***rest of the independent variables, there were no statistically significant relationships observed*** as the p-values for all these pair-wise tests were greater than 0.05.

In Situation 2*: While driving inside a tunnel*, from Table 46, it was observed that ***there were no statistically significant relationships between ODD mismatch in this situation and any of the selected independent variables***. This is because p-values in each of the test pairs was greater than 0.05.

In Situation 3: *Driving close to an off-ramp*, from Table 47, it was observed that ***there is a statistically significant relationship between ODD mismatch in this situation and the drivers' perceived risk of driving*** in this situation [$\chi^2(15) = 7.9$, $p=0.047$], at α-value 0.05 (as the null hypothesis could not be rejected). To determine the effect size of this relationship, the Cramer's V statistic was used. The observed effect size could not be deemed statistically significant as its p-value was 0.083, greater than α-value 0.05. Furthermore, it was observed that there is ***almost statistically significant relationship between drivers' initial trust on AV's and ODD mismatch*** [$\chi^2(15) = 5.61$, $p=0.052$] with α-value 0.05.

Finally, ***in Situation 4***: *Driving at a curve*, from Table 48, it was observed that ***there is a statistically significant relationship between ODD mismatch in this situation and the drivers' perceived risk of driving*** in this situation [$\chi^2(15) = 7.9$, $p=0.047$], at α-value 0.05 (as the null hypothesis could not be rejected). Furthermore, it was also observed that ***there is a statistically significant relationship between the drivers´ real-time trust in this situation and mismatch in ODD state awareness and ODD specified by OEM*** [$\chi^2(15)$ =4.3, p=0.04] for linear-by-linear test statistic, at α-value 0.05. ***There was no statistically significant relationship observed between the other independent variables and ODD mismatch,*** as their corresponding p-values were greater than 0.05.

6.4.6. Discussion & Summary

Exploratory statistical tests were conducted to investigate relationships between the pre-drive, post-drive driver behaviour, and their real-time responses. Before doing so, it was also important to verify if there was an increase or decrease in drivers' trust ratings with more time they spent in the system. It was seen that, there was no statistically significant effect of time on drivers' trust. This could maybe because the four situations at which drivers reported their real-time trusts, were selected from a list of situations on which the drivers were asked the same question. In fact, all of the selected test situations were closer to the end of the test and maybe time already had its effect on drivers' trust. But this requires further investigation and is beyond the scope of this research.

The key results from these exploratory analyses of the questionnaire data obtained from 16 participants (after filtering) in three stages are provided situation-wise below. In each of the tables, Table 7,Table 8,Table 9 and Table 10, the dependent variable corresponding to that test situation, the variable it was statistically related with and the type of relationship, is summarised.

Table 7. Results for exploratory tests in Situation 1 (Outside the ODD as specified by Tesla)

Dependent variable	***Significant relationship with***	***Type of relationship***
Post-drive perceived risk	Perceived risk in Situation 3(Close to off-ramp)	Strong, positive

Table 8. Results for exploratory tests in Situation 2 (Inside the ODD as specified by Tesla)

Dependent variable	*Significant relationship with*	*Type of relationship*
Real-time trust	Perceived risk in this situation	Negative, strong

Table 9. Results for exploratory tests in Situation 3 (Maybe inside/outside the ODD as specified by Tesla)

Dependent variable	*Significant relationship with*	*Type of relationship*
Real-time trust	Initial perceived risk of driving in LKAS	Negative, strong
Real-time trust	Frequency of using LKAS	Positive, strong
ODD state awareness	perceived risk of driving in this situation	

Table 10. Results for exploratory test in Situation 4 (Inside the ODD as specified by Tesla)

Dependent variable	*Significant relationship with*	*Type of relationship*
Real-time trust	Frequency of using LKAS	Positive, strong
ODD state awareness	Perceived risk of driving in this situation	
Real-time trust	Perceived risk inside the tunnel (Situation 2)	Negative, strong
Ease of driving	Ease of driving inside the tunnel (Situation 2)	Strong, positive

From the above tables, it is seen that in situations ***inside the ODD*** as specified by Tesla (driving inside the tunnel and on the curve), perceived risk of driving in these situations had a negative and strong relationship with the real-time trust reported by drivers. Furthermore, while driving in the curve situation (Situation 4), there is a relationship between drivers' perceived risk of driving in this situation and their ODD state awareness. In fact, through internal correlations, it was observed that the perceived risk of driving inside the tunnel also had a negative and strong relationship with real-time trust ratings while driving on the curve and, the perceived ease of driving inside the tunnel also had a strong and positive relationship with the perceived ease of driving on the curve. The reason for these observations could be that there was a small curve inside the tunnel and maybe this also led to correlations between variables in these two situations.

In the situation, ***may be inside/outside the ODD*** i.e. while driving close to an off-ramp, drivers that used LKAS more frequently had a higher trust on the LKAS in real-time while driving close to an off-ramp. Furthermore, drivers that perceived driving in LKAS equipped vehicles less risky had more trust in real-time on the LKAS while driving in this situation and on the other hand, there is a relationship between perceived risk of driving close to the off-ramp and a driver's ODD state awareness.

While driving inside the city on a road with no lane marking on its boundaries (Situation 1), which was ***outside the ODD***, it was seen that there were no direct relationships between any of the pre-drive/post-drive measurements and the real-tie measurements. However, perceived risk of driving in this situation is strongly (positive) correlated with that, of driving close to an off-ramp, this could be an indication that in the situations that are not inside the ODD, the perceived risks of drivers are correlated. This correlation could be because in both these situations there is either no lane marking on one side (City) or there is changing lane marking in the other (off-ramp).

In addition to these results, a few internal statistically significant correlations between initial attitude variables (from pre-drive questionnaire) were also observed. It was seen that drivers with higher perceived ease of using LKAS system in vehicles, perceive less risk of driving in these vehicles. It was also proved that having prior negative experiences in LKAS equipped vehicles, does increase the driver's knowledge about the ODD of semi-automated vehicles. But it is not the safest method to learn more about the ODD of semi-automated vehicles. Therefore, this implies that there is a need for a better method of increasing driver's knowledge about the functionality of semi-automated vehicles.

It is important to mention that there was relationship between lateral objective risk of driving in each situation and, the driver's real-time trust and their ODD state awareness.

Furthermore, there were no statistically significant (and visual) difference between real- time trust ratings between drivers with prior experience in a Tesla and drivers without. This could be because of the unequal representation of the two categories and the rather small sample size.

From the exploratory results, it was concluded that lack of sample size, led to many variables not showing statistical significance relationships. This raises the need for more of such studies to understand drivers' behaviour in three stages while driving in LKAS equipped vehicles. It is important to note that the significant relationships determined in the analysis could also be due to multiple significance effect and therefore, require more sample size for verification.

After conduction of the exploratory statistical test, given the aim of this part of the research, presence of mismatches between ODD state awareness of drivers and the ODD as specified by the OEM's (Tesla), were investigated. From the statistical tests, it was observed that there is a statistically significant difference between the ODD mismatch across all the situation (even after correcting for type 1 error). Moreover, the difference between most of the situation pairs (other than S1/S3 and S2/S4), were statistically different from each other.

In terms of number/percentage of mismatches, maximum mismatches were observed in the ***'Neither inside nor outside the ODD'*** situation of driving close to an off-ramp. In this situation, most of the drivers believed that vehicle was inside its ODD. Next highest mismatches were seen in the '***Outside the ODD'*** situation of driving in the city with no lane marking on road boundaries. In this situation, most drivers believed that the vehicle was inside its ODD and a very few were not sure. This mismatch could lead to very dangerous situations as the drivers might not be completely fall-back ready (ready to take over from the LKAS system).

In both the situations ***'Inside the ODD'***, there were very less mismatches. While driving inside the tunnel only two drivers were not sure if the vehicle was inside/outside Its ODD and while driving on the curve only one driver was not sure about the vehicle being inside its ODD.

To go one step further, an attempt was also made to identify possible reasons for these ODD mismatches using driver behaviour variables from within the research. Even though the sample size of the data set was very small it was possible to identify a few possible factors that could have an impact on ODD mismatch. For one of the ***'ODD-In situations'*** of driving on a curve, ODD mismatch was related to drivers' real-time trust and perceive risk of driving in that situation. In the ***'ODD neither in nor out'*** situation of driving close to an off-ramp, it was seen that ODD mismatch was related once again the driver's perceived risk in that situation and to their initial trust on AV's in general (almost significant at 95% confidence interval). Finally, in the ***'Outside the ODD'*** situation of driving the city with no lane marking on road boundaries, ODD mismatch was almost related (p-value = 0.05) to the driver's real-time trust on the LKAS in that situation.

This is an indication that there could be several other factors that lead to the ODD mismatch in drivers across the different type of ODD classifications. It is important to avoid such a mismatch between the driver's understanding and the OEM's specification of the functionality of the LKAS equipped vehicles as this could lead to fatal road accidents within these systems. To avoid such mismatches, the OEM's and road developers must focus on some of the driver behaviour aspects presented in this research as it provides an indication of potential factors.

6.5. General Observations

During the road tests, in addition to the observations regarding the four selected situations, there were also certain general and rather potentially dangerous situations that were observed. This sub-chapter describes two such situations and provides plausible reasons for these observations.

6.5.1. Off-ramp and rush-hour lane dilemma

Close to the end of the test route, on the first exit after getting on to the A4, there was a situation that was encountered in few of the test drives. Shown in Figure 45, the Tesla is approaching an off-ramp at 100km/h, intending to go straight onto the rush-hour lane rather than taking the off-ramp. In Figure 46, the Tesla gets closer to the off-ramp fork and slowly begins to move to the right of the lane centre. After which, in Figure 47, the Tesla begins to go into the off-ramp rather than maintaining the intended straight trajectory. Followed by this, in Figure 48, the Tesla is heading straight towards the fork and if the driver doesn't take control here, this could lead to a collision with the fork. At this point, the driver then takes control of the steering wheel and steers to the left to ensure that the vehicle goes back to the rush-hour lane (Figure 49).

In this situation, if the driver failed to take over from the vehicle at the right time, it could have resulted in a possible collision with the off-ramp fork. A potential reason for the occurrence of such a situation could be type and number of lane markings at the fork as seen in Figure 45. There are solid line lane markings at both the off-ramp and at the beginning of the rush-hour lane, this could mislead the LKAS into taking the off-ramp when it was not intended to. A possible solution for this on the infrastructure side, could be to change the lane marking type at such road situations such that there is a clear distinction between the off-ramp and the rush-hour lane. On the other hand, the LKAS could also be trained to avoid such situations by ensuring that the controller strategy of the LKAS gives more weightage to the in-built map and intended direction, in such situations.

6.5.2. Car-following inside the city

Another unusual situation was experienced by a few drivers in the city section of the test route. In Figure 50, the Tesla is driving on a city road with no lane markings on its boundaries and following a vehicle at the same time. The leading vehicle then starts taking a right turn (Figure 51), after which the Tesla gets closer to the leading vehicle (which has turned) and aligns itself to continue following the leading vehicle into the turn (Figure 52). Before the Tesla takes the (unintended) turn, the driver takes over and ensures that the vehicle is heading straight again (Figure 53).

In this situation, if the driver failed to take over from the LKAS, the Tesla would have taken the right turn without giving an indication of this turn to its following vehicles. This could be lead to a dangerous situation and possible collision with its following vehicle. The occurrence of this situation could be because, in situations when a lane marking on either side of the road is missing (in the presence of a leading vehicle), the Autosteer is designed to follow the trajectory of the leading vehicle and thereby, taking the unintended right turn. Such situations could be avoided either by ensuring that the Autosteer cannot be turned ON in such city roads, or the on-board cameras detect the turning indicator of the leading vehicle or, by ensuring that the controller strategy of the LKAS gives more weightage to the in-built map and intended direction, in such situations.

6.6. ODD assessment for Tesla Model S

The ODD of the Autosteer function of Tesla Model S is assessed using the lane keeping performance of the Autosteer, lateral driving risks and the driver behaviour in the vehicle, in different pre-selected situations. Through the method proposed in this research, the different ODD-classified situations can only be compared with each other (Table 11) and decisions regarding including or excluding situations from the ODD cannot be made with certainty. This is because, the thresholds for the acceptable lane keeping performance and objective risk values are defined by the vehicle manufacturers and therefore, confidential.

Table 11. Assessment of selected test situations

Test Situation	*Autosteer's Lane Keeping performance*	*Risk of driving*	*ODD mismatch*	*Main subjective relationships*
Situation 1: No LM	High bias towards left of lane centre with considerable deviations in lane positions.	Highest risk of driving	Second highest mismatches (68.7%); 82% of drivers believe vehicle is inside ODD	ODD mismatch dependent on real-time trust
Situation 2: Tunnel	On average aligned close to lane centre along with a bias away from left lane marking strip avoiding tunnel wall.	Second highest risk of driving, mainly due to sway within the lane	Less mismatches (12.5%)	Real-time trust negatively correlated to perceived risk
Situation 3: Off-ramp	Highest range of lane position variation with a slight bias to left of lane centre	Relatively low driving risks because probability of collision is low	Highest mismatch (81.2%); 77% of drivers believe vehicle is inside ODD	ODD mismatch dependent on perceived risk and initial trust in AV's
Situation 4: Curve	Lane position condensed close to lane centre and has least deviations in its values	Lowest risks of driving	Least mismatches (6.2%)	Awareness about vehicle's ODD is dependent on perceived risk

Outside the ODD | Inside the ODD | Neither inside nor outside the ODD

Inside the ODD situations:

Situation 2: Driving inside a tunnel inside the city

Lane keeping performance was good with the vehicle mostly close to the lane centre but there was bias in lane positions away from the left lane marking strip. This was due to the concrete tunnel wall being closer on the left of the vehicle compared to its right. This is good, but at the same time this bias leads to vehicle swaying a lot within its lane, in turn leading to higher lateral velocity increasing its lateral driving risk. Therefore, risks were the second highest in this situation. Moreover, drivers real-time trust on the system in this situation specifically, is negatively related to their perceived risk of the situation. The number of ODD mismatches is low (12.5%) and even if there is, drivers do not think it is outside the ODD. In this situation, it would be advisable to inform the driver about risks and the vehicle should not sway much within the lane, maintain position even if it is to the right of the lane centre.

Situation 4: Driving on a curve on the highway

Overall best lane keeping performance (relative to the other vehicles) as the vehicle has condensed lane centre alignment and the deviation in its positions was the least. Once again, slight left skew in mean lane positions to avoid guard rail but the lateral objective risk of driving was still the least. Perceived risk of drivers in this situation is related to their awareness of the ODD in this situation. There is very low mismatch (6.2%) between driver perceived and OEM stated ODD. The ODD mismatch is related to the real-time trust and the perceived risk of driving in this situation.

In general, within the situations inside the ODD, there is a correlation between driver behaviour in the situations. This is important to keep in mind while altering specifications and vehicle behaviour in any of these situations, as it could have an indirect impact on the driver's behaviour in other 'Inside ODD' situations.

Outside the ODD situation:

Situation 1: Driving inside the city with no lane markings on road boundary

Lane keeping performance was poor as there was very high bias in the lane positions towards the left of the lane centre and there was considerable deviation in the lane positions. Vehicle was attempting to move away from the road edge as is a design specification of the Autosteer function. Because of a very high left skew (largest of all situations), the (maximum for 15 seconds) lateral velocity was the highest in this situation (relative to other situations). Therefore, lateral risk of driving in this situation was the highest. Majority of drivers believe that this situation is inside the ODD of the vehicle (81.2%) and there is second highest mismatch (68.7%). This mismatch is related to the drivers trust on the system in this situation.

The lane keeping performance was poor, lateral driving risk was the highest and there is very mismatch in driver's perception of ODD and OEM specified ODD in this situation. The situation must either be removed from the ODD or necessary steps must be taken to improve driver trust in this situation as it influences mismatch ODD state.

Neither inside/outside the ODD

Situation 3: Driving close to an off-ramp on the highway

Vehicle lane positions had the highest range between maximum and minimum values (away from the lane centre) and the deviations in the lane positions were the highest along with a slight bias to the left. The lane keeping performance was therefore, considered the poorest. The objective risk on the other hand is fairly low even though lane keeping performance is poor, this means that the probability of collisions is low. On the other hand, the mismatch in driver believed and OEM specified ODD, was the highest (82.8%) and majority (77%) of drivers believe that the vehicle is inside its ODD. The mismatch in driver's ODD awareness correlates with their perceived risk of driving in this situation and their initial trust in AV's.

Perceived driving risks in Situation 1 and Situation 3 are positively correlated and must be kept in mind while making vehicle functionality alterations or infrastructural changes directed towards either of these types of situations. This is important as the objective risks while driving inside the city with no lane marking on the road boundaries, was higher and if any alterations are made to the situation environment or vehicle functionality, directed towards this situation, it could indirectly effect driver behaviour in situations while they drive close to off-ramps.

Especially for the situation 1, which was specified to be outside the ODD of the Autosteer by Tesla, since the fact the system could be switched ON, might be a possible reason for majority of drivers thinking that the situation was inside the ODD of the vehicle. Moreover, the lateral driving risk in this situation was shown to be the highest. Therefore, to ensure that the drivers understand the capabilities of the system better, the OEM's must either not allow for the system to be turned ON in these situations or have a better form of communication with the driver regarding the systems possible decrease in performance. This will help in a proper calibration in trust of the drivers as they generally correlate their responses and awareness of the capabilities in one situation with another, as shown in this research. Furthermore, from a questioned asked in the post-drive questionnaire it was seen that 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.

The proposed assessment methodology helps in comparing each of the ODD-classified situations with each other but a final decision for the inclusion or exclusion of a situation to or from the operational design domain of an LKAS system depends on the acceptable threshold values for each of the components of this method, which varies between OEM's and is confidential information. Therefore, it is advised that vehicle manufacturers could use this method to generate the required input while deciding ODD for their vehicle. In addition to this while deciding ODD, Vehicle manufacturers must keep drivers' behaviour and response in such LKAS systems as a very important input, and calibrate their systems to promote trust when justified, and encourage intervention when necessary.

Chapter 7. Conclusion

7.1. Summary

In this research, a methodology to assess the Operational Design Domain of Lane Keeping Assistance System equipped vehicles is developed. The method is mainly intended for use by manufacturers of semi-automated vehicles to assess the situations and conditions in which their lane keeping systems that are already available in the market, can and cannot perform. It involves assessment of the situations based on three main components: lane keeping performance of the system in these situations; the lateral objective risks of driving in these situations; and the driver's attitude/behaviour using questionnaires at three stages of testing. The methodology is implemented to a real-road case study using an instrumented Tesla Model S and assesses its Autosteer system at certain pre-specified situations.

It first involves identifying the situations that should be tested and this was done by analysing the Tesla Model S owner's manual and classifying situations where the system is intended to work, not intended to work and where it may or not work as intended. One out of three candidate routes from within the Netherlands, was then selected for the real-road test. Participants for the tests were recruited based on specific criteria, the Tesla Model S was instrumented to gather the research specific data, survey questionnaires were developed and the road test were conducted.

Raw data collected from the tests were processed and a data visualization tool was developed to help with data look-up during the analysis phase of the research. Data analysis included assessment of lane keeping performance of the Autosteer in different pre-classified test situations, measurement of the lateral risk of driving in these situations using a novel risk metric, assessment of drivers' behaviour before-during-after the tests to identify mismatches in their perception, and these results were then combined to assess each of the ODD-classified situations and thereby, the Autosteers Operational Design Domain.

This chapter concludes the thesis by answering the research questions in sub-chapter 7.2, discussing the main scientific and practical contributions of this research in sub-chapter 7.3, listing the limitations of the research in sub-chapter 7.4, and finally recommending possible next steps to take this research forward, in sub-chapter 7.5.

7.2. Answers to research questions

Chapter 1.2 presented several research sub-questions, corresponding to the steps involved in the development of the ODD assessment methodology. Here, the five sub-questions are answered first followed by the main research question.

1) What are the components of a LKAS and their potential reasons for failure?

There are three different types of lane keeping assistance systems based on the level of involvement of the vehicle in the steering task. If the system merely warns the driver of departure from lane, it is referred as 'warning type'; if the system provides intervention in the form of opposing torques to the steering wheel, it is referred as 'intervention type'; and when the system actively steers the vehicle to ensure lane centring, it is referred as 'control type' LKAS. For this research, the control type LKAS of a Tesla Model S was the focus.

A LKAS involves continuous interaction between its three components; the driver, the vehicle and its surroundings, and the lane keeping module. The vehicle is responsible for sensing of the road and its environment using a combination of cameras and/or sensors and it identifies the lane marking and other vehicle surroundings. This information is then fed to the lane keeping module which is responsible for identification of the dynamics of the vehicle, calculate the optimal steering angle and torque for either intervening, or actively steering the vehicle to its desired position in the lane (including a vehicle manufacturer dependent buffer). The information regarding the need and extent of required lane centring is fed to it by the controller strategy. An important input for control type LKAS is the check for 'Hands on', where it is verified that the driver is inside the loop or not.

Factors affecting its performance were identified as potential reasons for the system's failure. It was classified into three main types of failure factors; road and infrastructure related, road sensing related and driver related. These were based on a detailed analysis of the owner's manuals of a Mercedes E-350, Volvo XC90 and Tesla Model S. Road and infrastructure factors in general, include changes in road geometry, change in type of lane marking, type of road and traffic states of surrounding vehicles. The road sending type failures includes factors that block or restrict the sensing capabilities of the on-board cameras and sensors. Factors such as glare, bright sunlight, oncoming headlights and improper illumination could lead to such types of failure. Driver related factors include waring of seat-belt, hand on steering wheel and their driving style in terms of their risk taking and sensation seeking. An elaborate list of factors is provided in Chapter 2.1.2.

This understanding of the functional specifications and constrains of LKAS in general, served as very important input while setting reference standard and interpretation of keeping performance of the Tesla.

1) *Which criteria can be used to identify the pre-specified real-road test situations?*

Given the aim of this research, selection of test situations was vital for its outcome. To first understand the operational design domain of the Tesla Model S, a detailed analysis of its owner's manual was performed. Using this, three classifications of its operational design domain (focused on its infrastructural dimension) were made. 'Inside the ODD' situations corresponded to those road situations where the Autosteer of the Tesla is intended to work, 'Outside the ODD' situations correspond to those situations when the Autosteer is not intended to be used, and 'Neither inside/outside the ODD' situations correspond to those situations where the Autosteer may or may not work as intended. A list of possible situations for each of these ODD classifications was made. These situations were fed as an input for the test route selection and were observed during the road tests. A final filtration of these situations was during the analysis phase of the research due to missing data corresponding to some situations, ensuring that at least one situation per ODD-classification is analysed.

2) *How does the Lane keeping assistance system perform when it is within and when it is exceeding its pre-defined ODD?*

This question can be answered using two aspects from this research. First, performance of the Autosteer in combination with the driver in different road sections (inside and outside the ODD). Second, performance of the Autosteer alone across different ODD-classified situations.

The test route was divided into three sections, out of which two were highway sections and one was a city road section. The combined driver and Autosteer performance was better on the highway sections on majority of the test days. This was because, on the highway there are no intersections, lesser closely driving slow moving traffic, clear and well recognizable lane marking strips and lack of obstacles and other occlusion that could hamper the sensing capabilities and functionality of the LKAS. On the other hand, inside the city, which is predominantly specified to be outside the ODD of Autosteer, the Autosteer can be turned ON only in a few specific situations. Therefore, since the driver contributes more to the overall lane keeping performance in the city, the performance was naturally poorer than on the city. However, lane keeping performance was different between the two highway sections. An important observation was that the mean lane position in all the road sections across all test days were predominantly biased towards the left of the centre, this was attributed to the fact that vehicles were mostly at the right most lane closer to the guard rails and road edges, and always attempting to stay away from them. This aspect gives an overview of the Autosteer and driver performance over two broad ODD-classified sections.

In the second aspect, solely the performance of the LKAS across different ODD-classification types, is assessed. This assessment is based on the mean lane positions and standard deviations of lane positions across these situations. For the situations deemed 'inside the ODD', majority of the lane positions were concentrated close to the lane centre with slight skews towards the direction away from the closest road barrier. For the situation outside the ODD of Autosteer, the range of the vehicle's alignment within the lane was the highest with again a skew towards the left, away from the road curb stone. Finally, for the situation neither inside nor outside the ODD, even though the mean lane positions were closer to the lane centre, the range of standard deviation in its positions was the highest. This was attributed to the changing and multiple lane marking types on the right of the lane centre, in this situation. Since the performance of the system is tested, more importance was given to the standard deviation in the lane positions over mean lane positions, and based on this the performance in the situations were ranked (from high performance to low performance): Driving on the curve (Inside the ODD) > Driving inside the Tunnel (Inside the ODD) > Driving on city roads with no lane marking on its boundaries (Outside the ODD) > Driving close to an off-ramp on the highway (Neither inside nor outside the ODD).

3) *To what extent can the proposed risk measurement metric be used to determine the objective driving risk across different test situations?*

In this research, a novel risk measurement metric based on the field theory concept, was used. This metric, Probabilistic Driver Risk Field (PDRF), was implemented to determine the lateral driving risk in each of the ODD- classified situations. Several assumptions were made to aid its implementation. The maximum risk was measured to be the highest in the situation that was outside the ODD (driving in the city with no lane marking on the road boundaries). This was attributed to its large skew and predominant alignment closer to the left lane marking strip. On the other hand, even though driving inside the tunnel was considered inside the ODD, the lateral risks in this situation was the second highest. This was attributed to the road barrier type, significant skew in mean lane positions away from the closer concrete tunnel wall. In the situation of driving on the curve, which was also inside the ODD, the driving risks were measured to be the least. On the other hand, for the situation that was neither inside or outside the ODD (close to off-ramp) the risks were not as high as inside the tunnel and in the city, but higher than on the curve. This meant that clear differences between the risk of driving in different situations were visible and possible explanations for these were given. Furthermore, to justify its usability, the risks were also compared to corresponding time to lane crossing measurement. The results of which, implied that the metric does represent appropriate risk trends and may even be relatively more comprehensive in its approach as it is sensitive to the type of interaction between the vehicle and he road entity.

Using the metric, it was also possible to depict that Autosteer's road obstacle evasive property could also be a dangerous as the lateral objective risks in such instances were computed to be high. However, the proposed metric also had few limitations. It involved several practical assumptions, measurements corresponding to only 15 secs duration of each situation and that it only considered risks due to non -moving road entities. If the risks due to other moving road entities were also included, the lateral driving risk measurements could be different, thereby indicating that there is room for further improving the accuracy of risk measurements.

4) Is there a mismatch? and which factors contribute to this mismatch in the ODD, in the selected situations, between the one specified by the OEM's and that which is specified by the drivers?

Across all the test situations, mismatch was observed between ODD specified by the OEM and by the driver. Higher mismatch was observed in both situations that were outside the ODD (i.e. no-lane markings and on/off-ramp). Drivers mostly reported these situations to be inside the ODD. On the other hand, the mismatch in situations that were inside the ODD (i.e. tunnel and curve) were minimal. Moreover, the odds of a mismatch by drivers with prior experience in a Tesla was almost the same as that, by drivers without. This indicates that in addition to having experience of driving in semi-automated vehicles, increasing drivers' awareness of its capabilities is also important.

Several pre-drive and post-drive factors had a statistically significant relationship with the real-time behaviour of the driver whilst inside the vehicle. From amongst these factors there a few which were found to contribute also to the mismatch between driver and OEM specified ODD. For the situation of driving on the curve, which is inside the ODD, this mismatch was found to be related directly to the driver's real-time trust in the system and their perceived risk of driving in this situation. For the situation outside the ODD (driving in a city with no lane marking on its lane boundary), the ODD was found to be related also to the drivers' real-time trust while driving in this situation (almost at 95% confidence interval). Finally, in the situation neither inside nor outside the ODD (driving close to an off-ramp), ODD mismatch is found to be directly related to the driver's perceived risk of driving in this situation and (almost at 95% confidence interval) to their initial trust on semi-automated vehicles in general. There were signs of indirect contribution to ODD mismatch by a few other factors that showed internal correlations with factors related to the mismatch, but are subject to future research.

Having answered the research sub-questions, it is now possible to answer the main research question:

To what extent can the Operational Design Domain of vehicles equipped with lane keeping systems be assessed by understanding the subjective and objective risk of driving in pre-specified test situations?

The answer to the main research question is the implementation of the proposed methodology to assess the operational design domain of vehicles equipped with lane keeping systems, using a real-road case study for a Tesla Model S. The design domain was assessed using a combined objective and subjective method, in specific test situations. In this method, the lane keeping performance of the Tesla model S and its lateral driving risks constitute the objective measurements; and conduction of a three-staged (pre-drive, real-time, post-drive) driver behaviour assessment constitutes the subjective measurements.

This method required a thorough understanding of the functional components, limitations and constraints of the Autosteer function of the Tesla Model S. Based on its owner's manual the Operational design domain of the Tesla was classified into situations where the Autosteer is intended to be used, not intended to be used and where it may or may not work as intended. The objective and subjective measurements were obtained for four specific situations that corresponded to these three classifications. An assessment of each of these situations was then made, by combining lane keeping performance, driving risk and driver responses, in these situations. Using this assessment methodology, it was only possible to compare the test situations with each other, and not make decisions regarding the inclusion or exclusion of situations from the Autosteer's ODD. This was because, acceptable threshold values for each assessment component (i.e. maximum acceptable risk) varies between vehicle manufacturers and is confidential information. Therefore, this method has the potential to aid vehicle manufacturers while deciding if a situation should remain inside or moved outside their lane keeping system's operational design domain. For situations that exceed/may exceed the design domain, this assessment provides the necessary steps to take and the driver, infrastructural and system related factors to keep in mind if these situations were to be included into the design domain, keeping the drivers safety and awareness of the system capabilities as the crux of the decision-making process.

This is a first step towards such an assessment and implements methodologies that can be extended to other situations and other types of ADAS systems, then those tested in this research.

7.3. Main contributions

Performance variation of Lane Keeping Assistance Systems inside and outside the ODD

There is limited practical evidence and literature about the lane keeping performance of lane keeping assistance systems, and this research fills this gap as it compares the performance across different days, different road sections/road types, different drives on the same route and between different test situations classified based on the operational design domain of the Tesla. Moreover, strong evidence for the obstacle avoidance lateral behaviour of the Autosteer was also provided in this research. This could also serve as a useful information for infrastructure developers, to aid in designing roads to increase the operational design domain of semi-automated vehicles, so that the systems can be used in more situations. The methodology used to generate vehicle trajectories (image processing technique) used for lane keeping assessment, is robust and can also be extended to other track trajectories of other vehicles (two wheelers and four wheelers). This could be of practical use for Royal HaskoningDHV as they conduct several projects in this domain.

Implementation and assessment of a novel metric for objective driving risk measurements and its comparison with existing Surrogate Measures of Safety (SMoS)

The field theory concept is niche and in fact very few researches have used this method for real-road tests. The probabilistic driving risk field theory method used in this research for driving risk measurement is still under development and this real-road research could be useful for the calibration of its parameters and its further development. Specifically, in the severity component of the risk metric the road barrier-type sensitivity factor ('k') and, in the probability component of the metric, the degree of the exponential function need to be calibrated such that they represent the road environment and chances of collision better. This research also contributes to the limited existing literature about the relationship between lane keeping performance of ADAS and the lateral risk of driving. Furthermore, this research also gives an indication that the implemented risk measurement metric shows similar trends as existing metrics such as TLC.

Drivers' attitude and response towards LKAS and AV's in general

There is limited literature about how drivers behave in lane keeping assistance systems in real-road experiments (when they face the direct consequences of driving in these systems). Most researches use a questionnaire based or a simulation based approach for this, in doing so they fail to capture the real and practical response from the drivers. This research used a combined questionnaire and real-road test based approach to comprehensively assess driver behaviour in LKAS. Moreover, real-time responses from drivers were collected and related to their pre- and post-drive attitudes and to make a comprehensive assessment of driver behaviour within these systems and therefore, contribute to existing literature. Furthermore, in this research, mismatch between the driver perceived and vehicle manufacturer specified ODD is researched and the obtained results confirm the need for a better communication between the vehicle and driver and make first indications towards factors that could have an impact on such mismatches, therefore paving way for future research in this domain.

A combined subjective and objective assessment of ODD for LKAS equipped vehicles.

There is very little or no scientific literature that attempts at assessing operational design domain for vehicles at any level of automation. The ODD assessment method proposed and implemented in this research fills this gap in the scientific community. In the commercial community, by using thresholds (i.e. maximum allowable driving risk values and maximum allowable deviation from lane centre) for each component of this assessment method, different vehicle manufacturers can assess situations in which their LKAS system can, cannot perform. In situations where the performance is undetermined, it helps in identifying vehicle, infrastructural and driver related factors that could be altered to include these situations into the systems ODD.

Road surrounding sensing tool

With the help of Paul van Gent, a PhD candidate at Delft university of technology, for this research a road sensing tool was developed by coupling, camera captured road surrounding videos, LIDAR data, GPS and Inertial measuring unit data. With the help of this tool, vehicle dynamics data (acceleration, speed in both lateral and longitudinal direction), vehicle co-ordinates and geometry, distances to surrounding moving and non-moving road entities can be determined in continuous motion. This could be of great value to Royal HaskoningDHV in future mobility related projects as it is robust and extendable to different vehicles and vehicle types.

7.4. Research limitations

Given the nature of the study, novelty of some of the methods used and other decision made over the course of the research. There were a few limitations of this research. In this sub-chapter, these limitations are listed for each phase of the research and possible ways to overcome them are provided.

Data collection/test conduction

- The Autosteer function of the Tesla cannot be turned on without the Adaptive Cruise Control (ACC) function of the vehicle also active. This could have an impact on the drivers' perception of risk and their trust on the vehicle in general, and is one of the main limitations of the research. To account for this, during the road-tests, it was ensured to remind the driver to base their responses only on the Autosteer's performance and their preferred ACC settings (desired vehicle lengths from leading vehicle).
- Real-time (during the drive) measurements might affect the drivers' attention, workload and thereby, affect their natural driving behaviour and perceptions within the automation. To avoid this as much possible, drivers were asked to respond verbally rather than performing a secondary task, but its effects might still be present.
- Data from a couple of test drives is not complete or not available, due to unexpected rain and instrument board turning off automatically because of insufficient power drawn from the power source. To avoid this for similar future studies, it is best to be prepared for unexpected weather conditions on days even if weather is predicted to be favourable for testing. To solve the problem of

automatic instrument turning off, it is best to keep another device connected to the power source at all times to make sure it does not turn off automatically.

- LIDAR data was missing in some situations due to the orientation of the LIDAR changing during the drive. To avoid this for future studies, proper calculations for orientation based on the required range of the LIDARS, need to be made and the orientation must be checked before every test drive.

Data processing

- There is a maximum of 16% error and average errors of 3.5% (highway) and 4% (city) in the values of the image processed lane positions. This was mainly due to assumptions that had to be made while developing the algorithms because of inexperience with conduction of such tests and lack of man-power available during the testing days. The steps that must be taken in future such vehicle trajectory tracking experiments are provide in Appendix J (confidential).

Data analysis

- Lane keeping performance assessment: While filtering out possible lane change manoeuvres and unusual lane position values it was assumed that the wheel positions could not be greater than 1m away from the closest lane marking strip. This assumption might have also resulted in certain lane position values that were border line lane change manoeuvres to be excluded. To avoid this uncertainty, more research into lane change behaviour could be done to determine better threshold lane positions corresponding to a lane change.
- Objective risk:
    - Due to missing LIDAR data for some parts of a couple of test drives, certain assumptions regarding the dynamics of the subject vehicle, were made (their speed were limited to either leading vehicle speed or road speed limit). This can be avoided by the LIDAR data related steps mentioned earlier.
    - The Probabilistic Driver Risk Field method used in this research, only included potential field risk (risk due to non-moving risk entities), the kinetic field risk due to other moving objects also plays a very important role in total risk determination, even for Lateral risk assessment. Especially for the case of driving close to the fork close to the off-ramp (Situation 3), as here if the kinetic field risk would be included, there are high chances of it being the riskiest situation relative to the other test situations.
    - Only maximum risks for a 15seconds duration were considered.
- Statistical analyses:
    - Number of test participants is less (only 16 after filtrations), this negatively affects having concrete evidence in statistical testing. Moreover, the statistical significance between tested relationships could also be due to ‘multiple significance effect’ (that arises to multiple correlation test conductions for the same data). The only way to avoid not using such statistical corrections, is to have a larger data set, i.e. more participant drives.
    - Bonferonni corrections: For some of the statistical tests, the results were reported without correcting for type 1 errors (quite possible in this research due to several repeated measure analyses). Once again, this could be avoided by having a large data set.
- ODD assessment:
    - It is not possible to make ODD related decisions (to include or exclude situations from the ODD) based on this research, due to varying and unknown threshold values for individual components of this method. To make such decisions, a constructive dialogue with OEM’s must be developed.

7.5. Next steps

In this sub-chapter, possible future directions and steps to take this research forward are discussed. There are two perspectives for which possible next steps are described.

Scientific perspective

- The immediate next step for this research is to include kinetic field risk component (risks due to other moving road entities) to the risk assessment method used in this research. This would result in an even more realistic representation of driving risks. In fact, the required data set for including this aspect to the risk measurement is *already made available* through the underlying research.
- During this research, driver facial video data corresponding to all participant drives was collected. This can be used for further research into driver psychology, work load and stress analysis whilst driving in semi-automated vehicle. Using the video footage, facial movement analysis, blink rate analysis, field of view analysis (gaze direction analysis) can be performed. Additionally, using machine learning and existing facial action coding systems [115], it is possible to assess driver behaviour across different situations within such automated vehicles.
- This underlying research also stems further research into the concept of mismatch between driver's awareness about LKAS capabilities and its actual capabilities. The indications for possible causes for such mismatches as provided in this research, could therefore, serve as first step for research within this domain.
- This research was focussed only towards LKAS, but a similar methodology is applicable to the other ADAS systems such as Adaptive cruise control. This would require alteration to system's performance assessment method and necessary alterations to the probabilistic driver risk field metric. This required further investigation.
- This assessment methodology could also be used to research variation of different vehicles' performance within the same SAE level 2 over different ODD classifications (ODD-inside situation, ODD-outside situations, ODD- maybe in or out situations). This could result in useful information about the operational design domain of ADAS of different vehicle manufacturers (which is general confidential information).

Practical perspective

- Original Equipment Manufacturers (OEM's):
  - Using pre-set threshold values for the objective components of this methodology, OEM's can test different road situations before including them or excluding them from ODD of their vehicles.
  - Using this methodology, OEM's could also assess the LKAS (or any other ADAS) performance variations between software updates.
  - To increase the situations awareness of the driver, vehicle manufacturers could also use the objective risk value computation in real time, to show to the drivers on the internal LCD screens by adding at as a layer to the existing on-board map systems. This would help the OEM's warn drivers that they are approaching a risky situation and ned to be ready to take over from the system.
- Infrastructure developers: The potential risk field component of the objective risk measurement method implemented in this research can also be used by infrastructure developers for the identification of hotspots (road situations or road sections with higher risk than the other sections). Once these hotspots are identified various options for infrastructural changes can also be tested. Moreover, these tests can be conducted for both autonomous and non-autonomous vehicles.

Chapter 8. References

1. Gartenberg, C. *Safety driver of fatal self-driving Uber crash was reportedly watching Hulu at time of accident*. 2018 2018]; Available from: https://www.theverge.com/2018/6/22/17492320/safety-driver-self-driving-uber-crash-hulu-police-report.
2. Levin, S. *Tesla fatal crash: 'autopilot' mode sped up car before driver killed, report finds*. 2018 [cited 2018; Available from: https://www.theguardian.com/technology/2018/jun/07/tesla-fatal-crash-silicon-valley-autopilot-mode-report.
3. Milakis, D., M. Snelder, B. Van Arem, G. Van Wee, and G. Homem de Almeida Correia, *Development of automated vehicles in the Netherlands: scenarios for 2030 and 2050.* 2015.
4. International, S., *Taxonomy and Definitions for Terms Related to Driving Automation Systems for On-Road Motor Vehicles*. 2016.
5. Eboli, L., G. Mazzulla, and G. Pungillo, *Measuring the driver's perception error in the traffic accident risk evaluation.* IET Intelligent Transport Systems, 2017. 11(10): p. 659-666.
6. Commission, E., *Road safety in the European Unio: Trends, statistics and main challenges*. 2015.
7. Miller, D., M. Johns, B. Mok, N. Gowda, D. Sirkin, K. Lee, and W. Ju. *Behavioral Measurement of Trust in Automation: The Trust Fall*. in *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*. 2016. SAGE Publications Sage CA: Los Angeles, CA.
8. Lee, J.D. and K.A. See, *Trust in automation: Designing for appropriate reliance.* Human factors, 2004. 46(1): p. 50-80.
9. Sagberg, F., Selpi, G.F. Bianchi Piccinini, and J. Engström, *A review of research on driving styles and road safety.* Human factors, 2015. 57(7): p. 1248-1275.
10. Recarte, M.A. and L.M. Nunes, *Mental workload while driving: effects on visual search, discrimination, and decision making.* Journal of experimental psychology: Applied, 2003. 9(2): p. 119.
11. Lindgren, A. and F. Chen, *State of the art analysis: An overview of advanced driver assistance systems (adas) and possible human factors issues.* Human factors and economics aspects on safety, 2006: p. 38-50.
12. Tesla, *Model S Owner's Manual*. 2017.
13. Mercedes, *Owner's Manual Mercedes E-350.* 2017.
14. Volvo, *Volvo XC90 Driver's Manual.* 2017.
15. Dickie, D.A. and L.N. Boyle. *Drivers' understanding of adaptive cruise control limitations*. in *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*. 2009. Sage Publications Sage CA: Los Angeles, CA.
16. Larsson, A.F., *Driver usage and understanding of adaptive cruise control.* Applied ergonomics, 2012. 43(3): p. 501-506.
17. Parasuraman, R. and V. Riley, *Humans and automation: Use, misuse, disuse, abuse.* Human factors, 1997. 39(2): p. 230-253.
18. Mattsson, D., *ADAS: A simulation study comparing two safety improving Advanced Driver Assistance Systems*. 2012.
19. Elyasi-Pour, R., *Simulation Based Evaluation of Advanced Driver Assistance Systems*. 2015, Linköping University Electronic Press.
20. He, J., J.S. McCarley, and A.F. Kramer, *Lane keeping under cognitive load: performance changes and mechanisms.* Human factors, 2014. 56(2): p. 414-426.
21. Kyriakidis, M., J.C. de Winter, N. Stanton, T. Bellet, B. van Arem, K. Brookhuis, M.H. Martens, K. Bengler, J. Andersson, and N. Merat, *A human factors perspective on automated driving.* Theoretical Issues in Ergonomics Science, 2017: p. 1-27.
22. Donikian, S., S. Espie, M. Parent, and G. Rousseau. *Simulation studies on the impact of acc*. in *5th World Congress on Intelligent Transport Systems*. 1998.
23. Amditis, A., M. Bimpas, G. Thomaidis, M. Tsogas, M. Netto, S. Mammar, A. Beutner, N. Mohler, T. Wirthgen, and S. Zipser, *A situation-adaptive lane-keeping support system: Overview of the safelane approach.* IEEE Transactions on Intelligent Transportation Systems, 2010. 11(3): p. 617-629.
24. Gorter, C., *Adaptive Cruise Control in Practice: A Field Study and Questionnaire into its influence on Driver, Traffic Flows and Safety.* 2015.
25. Hoedemaeker, D.M., *Driving with intelligent vehicles: driving behaviour with adaptive cruise control and the acceptance by individual drivers*. 1999, TU Delft, Delft University of Technology.

26. Huisman, M., *Impacts of (Cooperative) Adaptive Cruise Control Systems on Traffic Flow A simulation case study on the effects of (cooperative) adaptive cruise control on the A15 highway.*, in *Transport & Planning*. 2016, Technical University of Delft.
27. Nilsson, L., *Safety effects of adaptive cruise controls in critical traffic situations*. 1996: Statens väg-och transportforskningsinstitut., VTI särtryck 265.
28. Piccinini, G.F.B., C.M. Rodrigues, M. Leitão, and A. Simões, *Reaction to a critical situation during driving with Adaptive Cruise Control for users and non-users of the system.* Safety science, 2015. 72: p. 116-126.
29. Wang, W. and D. Zhao, *Evaluation of Lane Departure Correction Systems Using a Stochastic Driver Model.* arXiv preprint arXiv:1702.05779, 2017.
30. Brandt, T., T. Sattel, and M. Bohm. *Combining haptic human-machine interaction with predictive path planning for lane-keeping and collision avoidance systems*. in *Intelligent Vehicles Symposium, 2007 IEEE*. 2007. IEEE.
31. Mammeri, A., G. Lu, and A. Boukerche. *Design of lane keeping assist system for autonomous vehicles*. in *New Technologies, Mobility and Security (NTMS), 2015 7th International Conference on*. 2015. IEEE.
32. Soualmi, B., C. Sentouh, J.-C. Popieul, and S. Debernard, *A shared control driving assistance system: interest of using a driver model in both lane keeping and obstacle avoidance situations.* IFAC Proceedings Volumes, 2013. 46(15): p. 173-178.
33. Navarro, J., J. Deniel, E. Yousfi, C. Jallais, M. Bueno, and A. Fort, *Influence of lane departure warnings onset and reliability on car drivers' behaviors.* Applied ergonomics, 2017. 59: p. 123-131.
34. Rudin-Brown, C. and Y. Ian Noy, *Investigation of behavioral adaptation to lane departure warnings.* Transportation Research Record: Journal of the Transportation Research Board, 2002(1803): p. 30-37.
35. program, E.I. 2018; Available from: https://eip.its-platform.eu/activities/sa-42-facilitating-automated-driving.
36. Farah, H., S.M. Erkens, T. Alkim, and B. van Arem, *Infrastructure for Automated and Connected Driving: State of the Art and Future Research Directions*, in *Road Vehicle Automation 4*. 2018, Springer. p. 187-197.
37. Milakis, D., B. Van Arem, and B. Van Wee, *Policy and society related implications of automated driving: A review of literature and directions for future research.* Journal of Intelligent Transportation Systems, 2017. 21(4): p. 324-348.
38. Mok, B.K.-J., D. Sirkin, S. Sibi, D.B. Miller, and W. Ju, *Understanding driver-automated vehicle interactions through Wizard of Oz design improvisation.* 2015.
39. Reeves, B. and C.I. Nass, *The media equation: How people treat computers, television, and new media like real people and places*. 1996: Cambridge university press.
40. Nass, C. and Y. Moon, *Machines and mindlessness: Social responses to computers.* Journal of social issues, 2000. 56(1): p. 81-103.
41. KPMG, C. and G. Silberg, *Self-driving cars: Are we Ready.* 2013.
42. Payre, W., J. Cestac, and P. Delhomme, *Fully automated driving: Impact of trust and practice on manual control recovery.* Human factors, 2016. 58(2): p. 229-241.
43. Rödel, C., S. Stadler, A. Meschtscherjakov, and M. Tscheligi. *Towards autonomous cars: the effect of autonomy levels on acceptance and user experience*. in *Proceedings of the 6th International Conference on Automotive User Interfaces and Interactive Vehicular Applications*. 2014. ACM.
44. Strand, N., J. Nilsson, I.M. Karlsson, and L. Nilsson, *Semi-automated versus highly automated driving in critical situations caused by automation failures.* Transportation research part F: traffic psychology and behaviour, 2014. 27: p. 218-228.
45. Rumschlag, G., T. Palumbo, A. Martin, D. Head, R. George, and R.L. Commissaris, *The effects of texting on driving performance in a driving simulator: The influence of driver age.* Accident Analysis & Prevention, 2015. 74: p. 145-149.
46. Muir, B.M. and N. Moray, *Trust in automation. Part II. Experimental studies of trust and human intervention in a process control simulation.* Ergonomics, 1996. 39(3): p. 429-460.
47. Wang, J., J. Wu, and Y. Li, *The driving safety field based on driver–vehicle–road interactions.* IEEE Transactions on Intelligent Transportation Systems, 2015. 16(4): p. 2203-2214.
48. Pohl, J. and J. Ekmark. *A lane keeping assist system for passenger cars-design aspects of the user interface*. in *Proceedings: International Technical Conference on the Enhanced Safety of Vehicles*. 2003. National Highway Traffic Safety Administration.
49. Renault. *Renaultsafetysite*. 2003 [cited 2018; Available from: www.conceptsecurite.renault.fr. /demain/einxc5.htm.

50. Reichart, G. and K. Naab. *Systems for lateral and longitudinal vehicle guidance*. in *Proc. First World Congress on Applications of Transport Telematics and Intelligent Vehicle-Highway Systems*. 1994.
51. Visvikis, C., T. Smith, M. Pitcher, and R. Smith, *Study on lane departure warning and lane change assistant systems.* Transport Research Laboratory Project Rpt PPR, 2008. 374.
52. Hillel, A.B., R. Lerner, D. Levi, and G. Raz, *Recent progress in road and lane detection: a survey.* Machine vision and applications, 2014. 25(3): p. 727-745.
53. Rajamani, R., *Vehicle dynamics and control*. 2011: Springer Science & Business Media.
54. Motors, G., *Safety Report*. 2018.
55. O'Hanlon, J., *Driving performance under the influence of drugs: rationale for, and application of, a new test.* British journal of clinical pharmacology, 1984. 18(S1): p. 121S-129S.
56. Ghasemzadeh, A. and M.M. Ahmed, *Utilizing naturalistic driving data for in-depth analysis of driver lane-keeping behavior in rain: Non-parametric MARS and parametric logistic regression modeling approaches.* Transportation research part C: emerging technologies, 2018. 90: p. 379-392.
57. Dijksterhuis, C., K.A. Brookhuis, and D. De Waard, *Effects of steering demand on lane keeping behaviour, self-reports, and physiology. A simulator study.* Accident Analysis & Prevention, 2011. 43(3): p. 1074-1081.
58. Peng, Y., L.N. Boyle, and S.L. Hallmark, *Driver's lane keeping ability with eyes off road: Insights from a naturalistic study.* Accident Analysis & Prevention, 2013. 50: p. 628-634.
59. Macdonald, W.A. and E.R. Hoffmann, *Review of relationships between steering wheel reversal rate and driving task demand.* Human Factors, 1980. 22(6): p. 733-739.
60. Owens, K. and J.G. Ramaekers, *Drugs, driving, and models to measure driving impairment*, in *Drugs, driving and traffic safety*. 2009, Springer. p. 43-58.
61. Allen, R.W., Z. Parseghian, and A.C. Stein. *A driving simulator study of the performance effects of low blood alcohol concentration*. in *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*. 1996. SAGE Publications Sage CA: Los Angeles, CA.
62. Van Leeuwen, P., R. Happee, and J. De Winter, *Investigating the effect of a visual search task for simulator-based driver training.* 2013.
63. McLean, J.R. and E.R. Hoffmann, *Steering reversals as a measure of driver performance and steering task difficulty.* Human Factors, 1975. 17(3): p. 248-256.
64. Verster, J.C. and T. Roth, *Standard operation procedures for conducting the on-the-road driving test, and measurement of the standard deviation of lateral position (SDLP).* International journal of general medicine, 2011. 4: p. 359.
65. Zhou, J., H. Peng, and T.J. Gordon, *Characterization of the lateral control performance by human drivers on highways.* SAE International Journal of Passenger Cars-Mechanical Systems, 2008. 1(2008-01-0561): p. 450-458.
66. Mammar, S., S. Glaser, M. Netto, and J.-M. Blosseville. *Time-to-line crossing and vehicle dynamics for lane departure avoidance*. in *Intelligent Transportation Systems, 2004. Proceedings. The 7th International IEEE Conference on*. 2004. IEEE.
67. Pilutti, T. and A.G. Ulsoy, *Fuzzy-logic-based virtual rumble strip for road departure warning systems.* IEEE Transactions on intelligent transportation systems, 2003. 4(1): p. 1-12.
68. Hamdar, S., M. Treiber, H. Mahmassani, and A. Kesting, *Modeling driver behavior as sequential risk-taking task.* Transportation Research Record: Journal of the Transportation Research Board, 2008(2088): p. 208-217.
69. Zhang, J., Y. Wang, and G. Lu, *Extended Desired Safety Margin Car-Following Model That Considers Variation of Historical Perceived Risk and Acceptable Risk.* Transportation Research Record, 2018: p. 0361198118773884.
70. Henkelman, G., B.P. Uberuaga, and H. Jónsson, *A climbing image nudged elastic band method for finding saddle points and minimum energy paths.* The Journal of chemical physics, 2000. 113(22): p. 9901-9904.
71. Wang, J., J. Wu, Y. Li, and K. Li. *The concept and modeling of driving safety field based on driver-vehicle-road interactions*. in *Intelligent Transportation Systems (ITSC), 2014 IEEE 17th International Conference on*. 2014. IEEE.
72. Hsu, T.-P., G.-Y. Weng, and Y.-J. Lin. *Conceptual structure of a novel car-following model upon gravitational field concept*. in *19th ITS World CongressERTICO-ITS EuropeEuropean CommissionITS AmericaITS Asia-Pacific*. 2012.
73. Brown, T.L., J.D. Lee, and D.V. McGehee, *Human performance models and rear-end collision avoidance algorithms.* Human Factors, 2001. 43(3): p. 462-482.

74. Kim, K., L. Nitz, J. Richardson, and L. Li, *Personal and behavioral predictors of automobile crash and injury severity.* Accident Analysis & Prevention, 1995. 27(4): p. 469-481.
75. Anthony-Babu, F., *Probabilistic field approach for highway risk assessment*. 2018, Technical University of Delft. p. 29.
76. Zou, Y., A.P. Tarko, E. Chen, and M.A. Romero, *Effectiveness of cable barriers, guardrails, and concrete barrier walls in reducing the risk of injury.* Accident Analysis & Prevention, 2014. 72: p. 55-65.
77. Brookhuis, K.A., D. De Waard, and W.H. Janssen, *Behavioural impacts of advanced driver assistance systems–an overview.* European Journal of Transport and Infrastructure Research, 2001. 1(3): p. 245-253.
78. Merat, N. and J.D. Lee, *Preface to the special section on human factors and automation in vehicles: Designing highly automated vehicles with the driver in mind.* Human factors, 2012. 54(5): p. 681-686.
79. Filzek, B. and B. Breuer. *Distance behaviour on motorways with regard to active safety: A comparison between adaptive-cruise-control (ACC) and driver*. in *Proceedings of the International Technical Conference on Enhanced Safety of Vehicles*. 2001.
80. Stanton, N.A., M. Young, and B. McCaulder, *Drive-by-wire: the case of driver workload and reclaiming control with adaptive cruise control.* Safety science, 1997. 27(2-3): p. 149-159.
81. Pauwelussen, J. and P.J. Feenstra, *Driver behavior analysis during ACC activation and deactivation in a real traffic environment.* IEEE Transactions on Intelligent Transportation Systems, 2010. 11(2): p. 329-338.
82. Prestl, W., T. Sauer, J. Steinle, and O. Tschernoster, *The BMW active cruise control ACC*. 2000, SAE Technical Paper.
83. McCall, J.C. and M.M. Trivedi, *Video-based lane estimation and tracking for driver assistance: survey, system, and evaluation.* IEEE transactions on intelligent transportation systems, 2006. 7(1): p. 20-37.
84. Reichart, G. and K. Naab. *Heading Control and Active Cruise Support Driver Assistance Systems for Lateral and Longitudinal Vehicle Guidance*. in *TOWARDS AN INTELLIGENT TRANSPORT SYSTEM. PROCEEDINGS OF THE FIRST WORLD CONGRESS ON APPLICATIONS OF TRANSPORT TELEMATICS AND INTELLIGENT VEHICLE-HIGHWAY SYSTEMS, NOVEMBER 30-3RD DECEMBER 1994, PARIS. VOLUME 4*. 1994.
85. Enache, N.M., M. Netto, S. Mammar, and B. Lusetti, *Driver steering assistance for lane departure avoidance.* Control engineering practice, 2009. 17(6): p. 642-651.
86. Naab, K. and G. Reichart, *Driver assistance systems for lateral and longitudinal vehicle guidance-heading control and active cruise support.* JSAE Review, 1995. 2(16): p. 220.
87. Helldin, T., G. Falkman, M. Riveiro, and S. Davidsson. *Presenting system uncertainty in automotive UIs for supporting trust calibration in autonomous driving*. in *Proceedings of the 5th international conference on automotive user interfaces and interactive vehicular applications*. 2013. ACM.
88. Lee, J.D. and N. Moray, *Trust, self-confidence, and operators' adaptation to automation.* International journal of human-computer studies, 1994. 40(1): p. 153-184.
89. Muir, B.M., *Trust in automation: Part I. Theoretical issues in the study of trust and human intervention in automated systems.* Ergonomics, 1994. 37(11): p. 1905-1922.
90. Kokubun, M., H. Konishi, K. Higuchi, T. Kurahashi, Y. Umemura, and H. Nishi, *Assessment of drivers' risk perception using a simulator.* R&D Review of Toyota CRDL, 2004. 39(2).
91. Eboli, L., G. Guido, G. Mazzulla, G. Pungillo, and R. Pungillo, *Investigating Car Users' Driving Behaviour through Speed Analysis.* Promet-Traffic&Transportation, 2017. 29(2): p. 193-202.
92. Eboli, L., G. Mazzulla, and G. Pungillo, *Combining speed and acceleration to define car users' safe or unsafe driving behaviour.* Transportation research part C: emerging technologies, 2016. 68: p. 113-125.
93. Jian, J.-Y., A.M. Bisantz, and C.G. Drury, *Foundations for an empirically determined scale of trust in automated systems.* International Journal of Cognitive Ergonomics, 2000. 4(1): p. 53-71.
94. Sheridan, T., *Trustworthiness of command and control systems*, in *Analysis, Design and Evaluation of Man–Machine Systems 1988*. 1989, Elsevier. p. 427-431.
95. Rempel, J.K., J.G. Holmes, and M.P. Zanna, *Trust in close relationships.* Journal of personality and social psychology, 1985. 49(1): p. 95.
96. Hoff, K.A. and M. Bashir, *Trust in automation: Integrating empirical evidence on factors that influence trust.* Human Factors, 2015. 57(3): p. 407-434.
97. Scialfa, C.T., D. Borkenhagen, J. Lyon, and M. Deschênes, *A comparison of static and dynamic hazard perception tests.* Accident Analysis & Prevention, 2013. 51: p. 268-273.
98. Waymo, *Waymo Safety Report*. 2017. from: https://waymo.com/safety/.
99. Stapel, J., F.A. Mullakkal-Babu, and R. Happee. *Driver behaviour and workload in an on-road automated vehicle*. in *Road Safety & Simulation International Conference*. 2017.

100. Temmink, F., *Trust and experiences of the Dutch Road authority on Lane Keeping systems, accompanied by Adaptive Cruise Control: specific incidents*, in *Transport*. 2016, University of Twente.
101. Hoedemaeker, M. *Driving behaviour with ACC and the acceptance by individual drivers*. in *Intelligent Transportation Systems, 2000. Proceedings. 2000 IEEE*. 2000. IEEE.
102. Rudin-Brown, C.M. and H.A. Parker, *Behavioural adaptation to adaptive cruise control (ACC): implications for preventive strategies.* Transportation Research Part F: Traffic Psychology and Behaviour, 2004. 7(2): p. 59-76.
103. Ezer, N., A.D. Fisk, and W.A. Rogers, *Age-related differences in reliance behavior attributable to costs within a human-decision aid system.* Human factors, 2008. 50(6): p. 853-863.
104. Club, T.M., *Tesla Motors Club*. 2006.
105. Delft, D.T. 2016; Available from: http://dittlab-apps.tudelft.nl:8080/app-regiolab/home.jsp)
106. Naujoks, F., C. Purucker, and A. Neukum, *Secondary task engagement and vehicle automation–Comparing the effects of different automation levels in an on-road experiment.* Transportation research part F: traffic psychology and behaviour, 2016. 38: p. 67-82.
107. Sheskin, D.J., *Handbook of parametric and nonparametric statistical procedures*. 2003: crc Press.
108. CROW, *Handleiding Online Kennisbank CROW Contentmanagement.*
109. Cramer, D., *Fundamental statistics for social research: step-by-step calculations and computer techniques using SPSS for Windows*. 2003: Routledge.
110. Lakens, D., *Calculating and reporting effect sizes to facilitate cumulative science: a practical primer for t-tests and ANOVAs.* Frontiers in psychology, 2013. 4: p. 863.
111. Reymond, G., A. Kemeny, J. Droulez, and A. Berthoz, *Role of lateral acceleration in curve driving: Driver model and experiments on a real vehicle and a driving simulator.* Human factors, 2001. 43(3): p. 483-495.
112. Siegal, S., *Nonparametric statistics for the behavioral sciences*. 1956: McGraw-hill.
113. Field, A., *Discovering statistics using SPSS*. 2009: Sage publications.
114. Seeger, P. and A. Gabrielsson, *Applicability of the Cochran Q test and the F test for statistical analysis of dichotomous data for dependant samples.* Psychological Bulletin, 1968. 69(4): p. 269.
115. Vural, E., M. Cetin, A. Ercil, G. Littlewort, M. Bartlett, and J. Movellan. *Drowsy driver detection through facial movement analysis*. in *International Workshop on Human-Computer Interaction*. 2007. Springer.
116. Ballard, D.H., *Generalizing the Hough transform to detect arbitrary shapes.* Pattern recognition, 1981. 13(2): p. 111-122.