Operational Design Domain Requirements for Improved Performance of Lane Assistance Systems
A Field Test Study in The Netherlands

Reddy, N.; Farah, H.; Huang, Yilin; Dekker, Thijs; van Arem, B.

DOI
10.1109/OJITS.2020.3040889

Publication date
2020

Document Version
Final published version

Published in
IEEE Open Journal of Intelligent Transportation Systems

IEEE Open Journal of
Intelligent Transportation Systems

Received 29 June 2020; revised 18 November 2020; accepted 24 November 2020. Date of publication 26 November 2020;
date of current version 8 December 2020.

Digital Object Identifier 10.1109/OJITS.2020.3040889

Operational Design Domain Requirements for Improved Performance of Lane Assistance Systems: A Field Test Study in The Netherlands

NAGARJUN REDDY [1], HANEEN FARAH [1], YILIN HUANG [2], THIJS DEKKER [3], AND BART VAN AREM [1] (Senior Member, IEEE)

[1] Department of Transport and Planning, Delft University of Technology, 2628 CN Delft, The Netherlands

[2] Department of Multi-Actor Systems, Delft University of Technology, 2628 CN Delft, The Netherlands

[3] Department of Geo-Data and Innovation, Province of North Holland, 2012 DE Haarlem, The Netherlands

CORRESPONDING AUTHOR: N. REDDY (e-mail: n.reddy@tudelft.nl)

This work was supported in part by the Applied and Technical Sciences (TTW), a subdomain of the Dutch Institute for Scientific Research (NWO) through the Project Safe and Efficient Operation of Automated and Human-Driven Vehicles in Mixed Traffic (SAMEN) under Contract 17187, and in part by the Netherlands Organization for Scientific Research (NWO) through the Project Spatial and Transport Impacts of Automated Driving (STAD) under Contract 438-15-161.

ABSTRACT There is a pressing need for road authorities to take a proactive role in the deployment of automated vehicles on the existing road network. This requires a comprehensive understanding of the driving environment characteristics that affect the performance of automated vehicles. In this context, a field test with Lane Departure Warning (LDW) and Lane Keeping Systems (LKS)-enabled vehicles was conducted in the Netherlands. Empirical data from the experiment was used to estimate the impact of driving environment components such as weather condition and lane width on the performance of the automated vehicles. Driving at night in the presence of streetlights with rain resulted in least detection performance for both the vehicles as compared to other visibility conditions. As for lane-keeping performance, the LKS positioned the vehicle significantly more to the left of the lane on left-curves than on straight sections. The LKS also positioned the vehicle more left on lanes with a width less than 250 cm than on wider lanes. These findings were translated into levels of service of the Operational Design Domain (ODD). Each level of service corresponded to a performance level of the lane assistance systems, classified as “High”, “Medium”, and “Low”, and defined using indicators.

INDEX TERMS Automated vehicles, field test, lane assistance systems, operational design domain, performance evaluation.

I. INTRODUCTION

A. BACKGROUND

While Automated Driving Systems (ADS) are expected to increase traffic safety, instances of crashes involving vehicles with such systems raise questions about their actual safety [1]. It has thus become increasingly relevant, for the involved stakeholders and especially for road authorities, to take action and initiative towards understanding the behavior of these systems on public roads and their implications on the existing road infrastructure. Although there have been a lot of studies that seek to evaluate the performance of ADS, the large focus is on longitudinal driving such as Adaptive Cruise Control (ACC) and Cooperative ACC (CACC) systems [2]–[7]. As of 2022, the new EU regulation makes it mandatory that all vehicles sold in the EU will have a set of ADS to increase road traffic safety [8]. Among these systems is lane assistance systems which support the driver in the lateral control of the vehicle. Lane assistance systems are not as commonly evaluated as the longitudinal driving systems, presumably due to the variations in their modes of operation and relative complexity in evaluating

The review of this article was arranged by Associate Editor Peter Han Joo Chong.

lateral behavior compared to longitudinal behavior. However, the consequences of unsafe lateral behavior can be severe. Therefore, it is important to evaluate the performance of lane assistance systems and to study the effect that driving environment components (such as lane markings, horizontal road curvature, and weather conditions) have on their performance.

The driving environment is considered a key component to be taken into account when evaluating the Operational Design Domain (ODD) of ADS [9]. ODD, for a certain ADS, is the set of driving conditions in which it is designed to function. These driving conditions may include weather conditions, road (and roadside) infrastructure components, and also vehicle-related conditions such as minimum speed. Presently, there is a need to establish standard practices and lexicon towards characterizing and defining the ODD for automated vehicles [10]. The specification of the ODD is paramount for the users (drivers) who need to be aware of the driving conditions in which their automated vehicle may deteriorate its performance or even not be able to function at all. The ODD specification is of equal importance, if not more, for road authorities who design roads and establish driving regulations for vehicles on their roads. If the automated vehicle enters an environment that is outside its ODD, there is a high probability of transfer of control from the automated driving mode to the manual driving mode, which could increase the driving risk. Such transitions may also occur if the vehicle incorrectly detects the environment to be within its ODD, potentially leading to poor performance and also unsafe situations. Roads must therefore be designed to consistently provide a driving environment in which the ADS can operate at a certain level of performance.

To define the ODD, the effect that the components of the driving environment have on the performance of the system must be first determined. The main aim of this research is to evaluate the performance of lane assistance systems by estimating the effect that the driving environment has on their performance. This research also provides a taxonomy for defining levels of performance for lane assistance systems as well as for characterization of the ODD, contributing to the efforts of standardizing this process.

B. PERFORMANCE INDICATORS

Certain indicators are being used in practice to evaluate the performance of lane assistance systems. For instance, the European New Car Assessment Programme [11] mainly uses the Distance to Lane Edge (DTLE), which is the distance between the wheel and the inner edge of the lane marking, to evaluate the safety of lane assistance systems. Several scientific studies study the lane-keeping performance of human drivers. Still, it may be assumed that lane-keeping performance indicators of human driving are also applicable for lane-keeping systems. Reference [12] investigated the effect of fog conditions on the lane-keeping performance of manual driving, using Lane Offset as an indicator. Lane Offset, also referred to as Lateral Offset, is defined as the deviation from the center of the lane. This indicator has been applied in other studies [13]. The Mean Lateral Position (MLP), which is the average Lateral Offset over a stretch of road, has been used in several studies [14], [15].

Another commonly used indicator for manual driving is the Standard Deviation of Lane Position (SDLP) [15]–[17]. Some studies argue that low SDLP does not necessarily mean safe performance as it may be due to certain distracted driving behavior attributed to distracting road infrastructure components (pavement repair patches, shadows) [18], [19]. Still, SDLP certainly can be regarded as a performance measure of the system, especially in combination with MLP.

Another indicator is the Time to Lane Crossing (TLC) which was used by [17] to evaluate the driving performance. TLC is defined by [20] as the time required for the vehicle to cross the road edge, given its trajectory. Reference [21] also proposes using the Time to Lane Departure (TLD), which can be measured at every instant assuming a straight trajectory (straight extension to the vehicle's current direction), or with constant lateral speed.

Other studies on the performance of LKS use indicators that focus on the steering wheel [22], [23] such as Standard Deviation of steering wheel angle (relating to the change in steering wheel angle as an indicator of stability) and Steering wheel reversal rate (SRR) (represents the number of times that the steering wheel is reversed by a magnitude larger than a specific angle, or gap). An even more extensive list of indicators concerning the steering wheel is provided by [24].

C. ROAD INFRASTRUCTURE REQUIREMENTS

Reference [25] provide a review of existing studies on infrastructure for automated vehicles, covering physical and digital infrastructure. It is found that there is a scarcity in knowledge of physical infrastructure as compared to digital infrastructure. Also, there is a need for empirical evidence-based studies on the geometric design of roads for automated vehicles.

Lane assistance systems primarily function by detecting lane markings. References [26] and [27] emphasize the importance of consistent road markings, signage, and pavement by stating that poor quality or unconventional use of lane markings (or cracks in pavement) could cause lane assistance systems to fail in identifying the correct lane boundary, which has direct implications on traffic safety.

Reference [28] conducted a qualitative study on the influencing factors for automated driving based on literature review and an online questionnaire filled by experts and stakeholders. The primary factors are complex urban road environments, quality of lane markings, temporary road work zones, poor visibility due to bad weather, and irregular or damaged road edges or curbs. Other aspects of the road infrastructure, such as low curve radii, slippery road surface, and poor visibility, were determined to be of medium importance.

Reference [29] speculate that the radii of curves could become smaller as automated vehicles are expected to negotiate curves with higher accuracy and at greater speeds. The report also proposed that on multi-lane curves, manually driven vehicles could be allowed to drive on the outer lanes (higher radius), while automated vehicles would be allowed to drive on the inner lanes (lower radius). Other propositions were that the width of lanes could be reduced (expecting more stable steering), improving quality and uniformity of lane markings, reduction in intensities of lights at intersections.

Reference [30] provide a prediction of the ODD for a Level-4 Highway autopilot including highway convoy. For road markings, it was recommended that there must be "Minimum quality of solid or dotted lines painted on the pavement if accurate lateral positioning is based on a camera detecting the location of the lane borders" thus highlighting the importance of visible lane markings. The system was also expected to be able to operate in all weather conditions except for severe conditions such as heavy rain or snow. All the ODD recommendations, including road infrastructure, were theoretical and not based on empirical research.

Attempts to operationalize the infrastructure requirements are mainly in the aspect of lane marking quality. For instance, the European Road Federation (ERF) presents the effect of reduction in quality of lane markings on their readability [31]. Although this focuses on human drivers, it also extends to propose a "good lane marking" that would ensure they are readable also for automated systems such as LKS and LDW. The reflectivity of the lane marking has been the most defined aspect of the quality of lane markings [31]–[33]. A "good lane marking" is defined as that which is visible to human drivers as well as automated vehicles irrespective of lighting conditions, weather conditions, and driver age.

Reference [34] conducted a pilot test with a LKS to study the effect of lane width on its control. It was found that the LKS cannot operate on lanes with a width less than or equal to 2.5 m, and it always can operate on lanes with a width greater than or equal to 2.75 m. It was concluded that widths of current lanes cannot be reduced with the current level of automated vehicles operating on the roads. Reference [35] demonstrated a method for assessing the ODD of lane keeping systems using a field test approach. The maximum risk was found to be in situations outside the ODD, such as roads inside the city with no lane markings. Some driving environments within the ODD were also found to be high risk, such as driving inside tunnels. The tunnel walls even influenced the lane keeping performance, as the vehicle lane position skewed away from the tunnel wall. Therefore, lane keeping systems could be significantly affected by the nature of the driving environment.

D. RESEARCH GAP AND RESEARCH QUESTIONS

Existing literature indicates that although research and studies regarding ADS and driving environment do exist, there is still a major gap in understanding the relationship between them. This is because although there is a lot of anticipation and predictions about the driving environment of the future, hardly any of these studies are based on concrete empirical evidence. This is mainly because the focus has been primarily on the distant future, where the vehicles would be fully automated, and have 100% market penetration rate. This makes it impossible to make any predictions of future driving environments (especially road infrastructure) with a high degree of certainty as the type and performance levels of these automated vehicles are not known. However, the few studies that have aimed to understand the driving environment requirements for some specific functions of automated driving such as Lane Assistance Systems using empirical approaches show the benefit of these studies [34], [35]. There is still a knowledge gap regarding the precise nature of the impact of driving environment components on the performance of the ADS systems. This research aims to define the effects that components of the driving environment have on the performance of Lane Assistance Systems. Therefore, the main aim of this study is to answer the following research question:

How do the different components of the driving environment affect the performance of Level 1 Automated Vehicles with Lane Assistance Systems?

E. CONCEPTUAL MODEL

The operation of the lane-keeping system in relation to its driving environment can be represented as a control structure, with the road authority (Box 1) and the LKS (Box 2) being the two controllers (see Figure 1). For a different system, such as the Lane Departure Warning (LDW) system, the controller would consist of a different algorithm, with the rest of the other components being the same. The LKS is specifically chosen to demonstrate the conceptual model. The LKS's control action is the steering wheel correction that is executed (Box 2.4) based on a defined algorithm (Box 2.3). The road authority's control action consists of modifications to the road infrastructure (Box 1.1) based on the performance evaluation of the LKS (Box 1.2). The LKS performance is affected by the driving environment, which consists of the infrastructure components (Box 3), and the weather conditions (Box 4). Examples of their subcomponents are presented in the conceptual model, such as streetlights and pavement type for infrastructure components, and time of day for weather conditions. The combination of these two components (and their sub-components) form the driving environment, which could be within or outside the ODD of the LKS. The specific algorithm on which the LKS is based and defined by the car manufacturer and is generally not publicly available. The driving environment conditions are known or can be measured or observed explicitly. In the conceptual model, the arrows (represented by

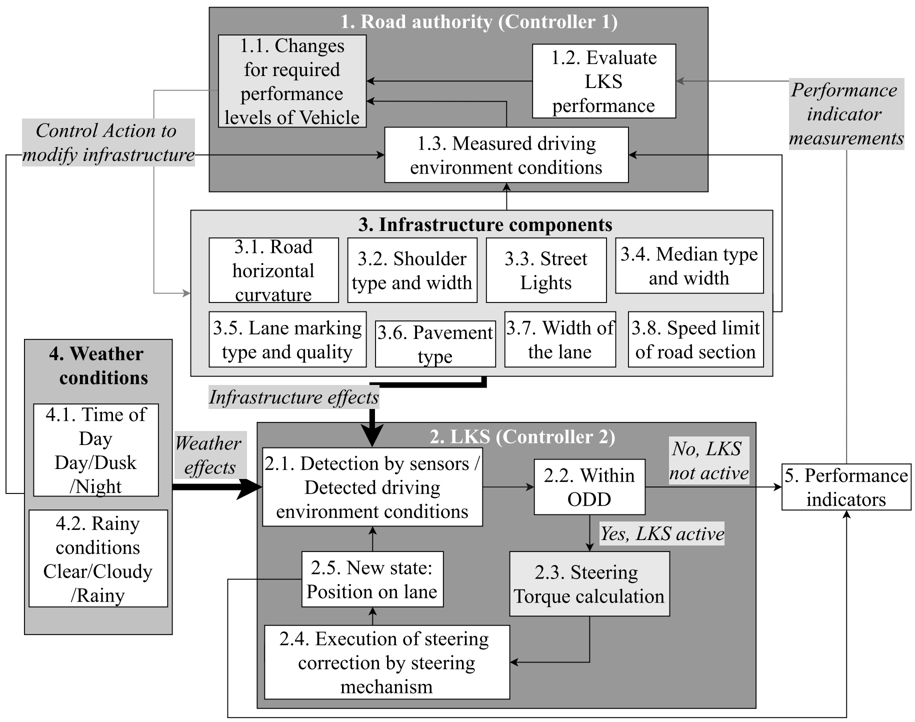


FIGURE 1. Conceptual model of LKS operation in its driving environment.

thick lines) connecting the infrastructure components and the weather conditions to the LKS controller are the focus of this research. These arrows represent the effect that the driving environment has on LKS performance. The conceptual model is not exhaustive. It only focuses on the part of the Dynamic Driving Task (DDT) of which the LKS takes responsibility. A more complete picture would also involve the human driver who does not only continuously monitor the LKS performance but is also responsible for all the other tasks of the DDT. That is, however, out of the scope of this research.

This article is structured as follows: Section II describes the research method including the field test set-up; Section III presents the descriptive analysis with performance evaluation; Section IV involves the modelling of performance; Section V provides the ODD Levels of Service with performance classification; Section VI discusses the results and the limitations; finally, Section VII proposes recommendations.

II. RESEARCH METHOD

This research adopts an empirical approach to answer the research question. A field test was conducted to understand the impact of driving environment components on the performance of lane assistance systems. It involved driving lane assistance systems-enabled vehicles in different driving environments and evaluating their performance. The effect of the driving environment components was estimated using statistical regression models. Finally, the driving environment components were categorized into different levels of service corresponding to the performance of the lane-assistance system in these conditions. The following sub-sections describe the field test setup, followed by the vehicle instrumentation, and data collection and processing.

A. FIELD TEST SETUP

For the field test, two vehicles/systems were selected, namely the Volkswagen e-Golf equipped with a Lane Keeping System (LKS), and the Toyota Auris equipped with a Lane Departure Warning (LDW) system. These two vehicles were chosen based on the popularity of their usage in the Netherlands. The two vehicles were driven at the same time on two routes of approximately 600 kilometers long in the Province of North Holland. The two routes were driven in different test sessions for practical feasibility. Figure 2 depicts the two routes. The intention was to cover as many different types of road environments as possible. Therefore, the two routes covered several parts of the Province of North Holland.

The test drives had an equal number of day and night sessions to account for the potential effect of the time of day on the performance of the two systems. The field test was done in March 2019. Day time test drives started around 10.00 am, and nighttime test drives started between 6.30 pm and 7.00 pm when it was dark. Also, the test sessions were scheduled to cover different weather conditions such as clear weather, cloudy, and rainy. Table 1 presents an overview of the field test sessions that were carried out. Thus, for each route, there were 8 test sessions (2 vehicles x 2 sessions

(a) Route 1 (b) Route 2

FIGURE 2. The two routes selected for the field test.

TABLE 1. Overview of test sessions.

Session	Number of test sessions
Volkswagen Day Route 1	3 (2 + 1 extra day ride)
Toyota Day Route 1	2
Volkswagen Day Route 2	2
Toyota Day Route 2	2
Volkswagen Night Route 1	2
Toyota Night Route 1	2
Volkswagen Night Route 2	2
Toyota Night Route 2	2

x 2 times of day) resulting in 8 sets of data points. The Volkswagen for Day Route 1 had an extra session due to data collection issues that were discovered in one of the previous sessions.

B. VEHICLE INSTRUMENTATION

The field test involved a collection of data from different sources. The Volkswagen was equipped with 5 GoPro cameras: one on the windscreen facing the driving direction, one facing the dashboard display, one capturing the steering wheel, and two cameras on either side of the vehicle facing the front wheel. These cameras were used to evaluate the performance of the lane-marking detection as well as the lane-keeping performance. The Toyota Auris with the lane-departure warning system was equipped with only 2 GoPro cameras, namely the windscreen camera and the dashboard camera to evaluate the lane detection performance. Figure 3 presents the positions of these cameras next to examples of the images captured by each of these cameras.

The camera on the windscreen (Figure 3a) captures the driving environment of the vehicle for later reference and analysis. The camera facing the dashboard (Figure 3b) covered the status of lane detection indicating the number of lines detected by the vehicle (no lines, one line, or both lines). The camera capturing the steering wheel (Figure 3c) was used to monitor transitions of control between the driver and the LKS. The two cameras on either side of the vehicle (Figures 3d and 3e) were used to measure the distance between the edge of the wheel and the edge of the lane marking, which indicated the position of the vehicle within the lane. The data processed from these cameras were used to evaluate the lane detection performance as well as the lane keeping performance. Each of these cameras was synced to local time using a GoPro app. The driver of the Volkswagen was instructed to drive with the LKS activated, and to allow the LKS to take steering control by default. However, the driver was informed to take manual control whenever he desired. The driver of the Toyota drove the vehicle manually at all times. The two drivers of the vehicles were employees of Royal HaskoningDHV and they drove in all the test sessions.

In addition to the cameras, both vehicles had a cellphone that recorded the GPS position for every second of the drive. The time synchronization between the cameras and the cell phone was done using a GoPro app on the cell phone with all the cameras being connected to the app. This was primarily used to measure the speed of the vehicle, but also later during the analysis to know the position where the vehicle was driving. Both vehicles also had co-drivers, who were

(a) A camera facing the windscreen (left) and a still from the windscreen camera (right)

(b) A camera facing the dashboard (left) and a still from dashboard camera (right)

(c) Camera capturing the steering wheel (left) and a still from steering wheel camera (right)

(d) Left camera position for capturing left lane-marking (left) and a still from Left camera (right)

(e) Right camera position for capturing right lane-marking (left) and a still from the right camera (right)

FIGURE 3. **Camera positions and respective still.**

equipped with laptops that had a logbook. They logged the driving conditions, for example, the time of day and weather condition. The co-driver of the Volkswagen also logged the status of the steering wheel control between the driver and the LKS. The next section discusses the data collection and processing.

C. DATA COLLECTION AND PROCESSING

The final dataset resulting from the field test was in csv format. The data collected included weather condition, time of day, driving speed, the status of LKS (explained later), and presence of streetlights. The two systems, that is the LKS and LDW, have certain minimum speed thresholds below which

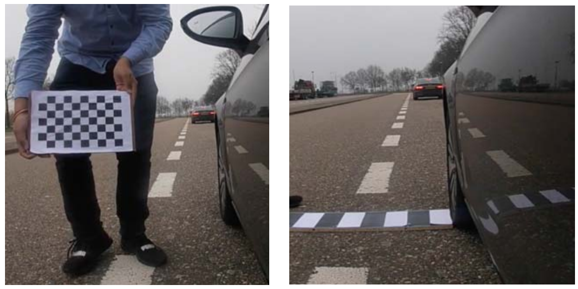
TABLE 2. Overview of dataset duration.

Description	Duration in seconds
Total number of data points	160,451 (about 45 hours)
Number of data points for Toyota Auris (LDW)	72,239 (about 20 hours)
Number of data points for Volkswagen Golf (LKS)	88,213 (about 25 hours)
Number of data points of day rides	82,828 (about 23 hours)
Number of data points of night rides	74,100 (about 21 hours)
Number of data points when both vehicle speeds above 60 kmph	107,605 (about 30 hours)

they cannot be active. The LKS had a minimum speed limit of 65 kmph to be active, and the LDW 50 kmph. While decelerating, the LKS remains active up to 60 kmph. This research therefore only considered the data points where the vehicles were driving above 60 kmph to evaluate their performance. The overview of the dataset duration is presented in Table 2.

The field test covered different weather conditions as well as the times of day. The co-drivers recorded the weather conditions data and time of day as separate variables. However, in reality, these two variables correspond to one variable which is the visibility condition in which the field test was done. Therefore, these two variables were combined into a new variable 'Visibility condition' that consisted of seven categories: 1) Clear (Day + clear weather), 2) Cloudy (Day + cloudy weather), 3) Rainy (Day + rainy weather), 4) Dark (Night + clear/cloudy weather), 5) Dark_Rainy (Night + rainy weather), 6) SL (Night + Streetlights + clear/cloudy weather), 7) SL_Rainy (Night + Streetlights + rainy weather).

For measuring the lane keeping performance, the position of the vehicle in the lane needs to be determined. The side cameras were used to extract the distance from the wheel to the adjacent lane marking using image recognition. However, there was a lack of any 'ground truths' for the measurements and also the fish-eye lens distortion meant that a pixel in one part of the picture did not have the same dimensions as a pixel in another part of the picture. To address these issues, two additional steps were done during the test drives. At the start of activating every camera video, a calibration task and a validation task were done. The calibration task involved holding a black-and-white checkerboard at different positions of the camera view as shown in Figure 4(a). This was used to calibrate the pixel dimensions in different positions of the camera image, given the true dimensions of the checkerboard. The validation task involved placing a black-and-white striped wooden plank perpendicular to the wheel of the car as shown in Figure 4(b). This was used to firstly define the angle of the line of measurement of the distance between the wheel and the lane marking, and secondly, to validate the measurements, given the true dimensions of the wooden plank. The image recognition involved identification



(a) Calibration task (b) Validation task

FIGURE 4. Image recognition tasks.

of edges of objects in the image, recognition of lane markings, and measuring the distance from the lane marking to the wheel.

The output of this dataset was the distance from the left wheel to the left lane marking, and the distance from the right wheel to the right lane marking, at a one-second resolution. This resulted in the lane position measurement at a resolution of one second and as well the lane-width. Due to budget and time constraints, this lane position data was available for only one of the day test sessions. Still, this data allowed for the derivation of some useful insights.

III. PERFORMANCE EVALUATION AND DESCRIPTIVE ANALYSIS

Analysis of the collected data first involved defining the performance indicators for evaluating the performance of the LKS and LDW. Both systems were evaluated using the defined indicators in different driving conditions. Then, regression models were developed and estimated for the performance indicators using the driving conditions as predicting variables. Finally, ODD levels of service were defined for each performance indicator using insights from the literature as well as the findings in this study. This section discusses first the performance indicators selected for the evaluation of these systems, followed by the results of the descriptive analysis.

A. PERFORMANCE INDICATORS

The performance indicators are classified into two categories: Detection Performance indicators, and Lane Keeping Performance indicators. Detection Performance relates to the ability of the systems to detect the lane markings on the road.

The detection performance was evaluated for both the LKS and the LDW system. The Lane Keeping Performance relates to the ability of the system to perform the lane-keeping task. Therefore, the Lane Keeping Performance was evaluated for the LKS only.

To evaluate the Detection Performance, the detection status indicated on the vehicle's dashboard was used. Three possible indicators were used: (1) percentage of both lines detected, (2) percentage of no lines detected, and (3) percentage of one line detected. All three were used for performance evaluation, but the focus was on the percentage of both lines detected and percentage of no lines detected. The percentage of both lines detected for a specific driving condition is defined as the percentage of occurrences of both lines being detected, with respect to all the detection states for that driving condition. The percentage of no lines detected for a specific driving condition was similarly defined as the percentage of occurrences of no lines being detected, with respect to all the detection states for that driving condition.

To evaluate the Lane Keeping Performance, two indicators were defined. The first indicator is the Mean Lateral Position, or MLP. Lateral Position, or Lateral Offset, is defined as the shortest distance between the center of the vehicle front axle and the center line of the lane. For sign convention, the right-hand coordinate system is used, where the distance towards the right is positive, and to the left is negative. The Lateral Lane Position was calculated according to Eq. (1) below. MLP is defined as the average lateral position measured over a defined stretch of distance. It is calculated by dividing the sum of Lateral Positions over that stretch by the number of data points collected in that stretch.

$$\begin{aligned} &\mathit{Lateral\ Lane\ Position} \\ &\quad = 0.5 * (\mathit{Left\_Distance} - \mathit{Right\_Distance}) \end{aligned} \tag{1}$$

The second indicator is the Standard Deviation of Lane Position, or SDLP. SDLP is the most popularly reported indicator to measure lane keeping performance, especially for human driving. SDLP is calculated using Eq. (2) below.

$$SDLP = \sqrt{\frac{1}{N-1}\sum_{i=1}^{N}(x_i-\bar{x})^2}, \ \text{or} \ \sqrt{\frac{1}{N}\sum_{i=1}^{N}(x_i-\bar{x})^2} \tag{2}$$

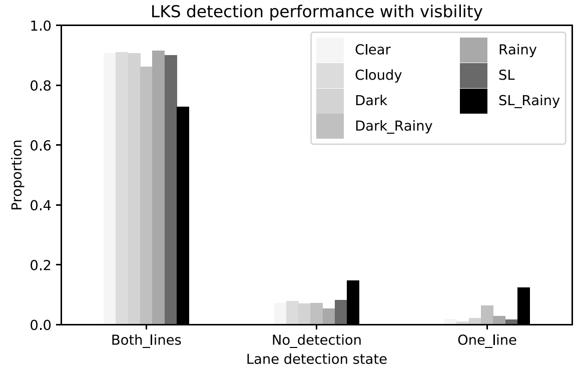
where,

$$\begin{aligned} x_i &= \text{The i-th value of lateral lane position} \\ \bar{x} &= \text{Mean lateral lane position of a sample} \\ N &= \text{Number of data points in the sample} \end{aligned}$$

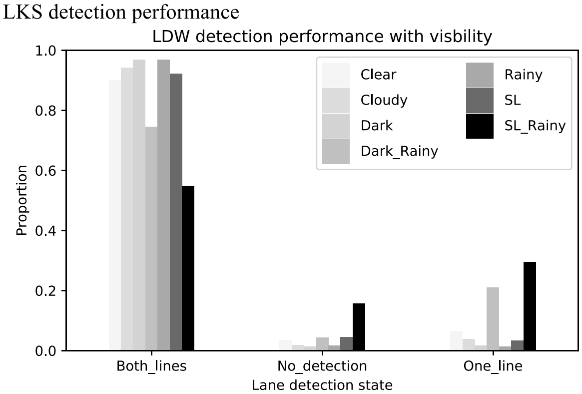
B. PERFORMANCE EVALUATION RESULTS

The indicators defined previously were used to evaluate the performance of the LKS and LDW system in different driving conditions. Firstly, the detection performance was evaluated in different visibility conditions. Figure 5 presents the results of the detection performance of LKS and LDW in the different visibility conditions. The calculated proportions were normalized with respect to the duration of driving in different visibility conditions. That is, driving in a certain visibility condition longer than in another has been normalized.

To a large extent, both systems detect both lane markings in different visibility conditions. However, there are some notable visibility conditions in which detection performance dropped significantly. For the LKS, the proportion of both lines detected has a large deterioration in the 'SL_Rainy' condition, and to a lower extent in the 'Dark_Rainy' condition. Besides that, the proportion of both lines detected seems to be stable across the other visibility conditions. For the LDW, the effects are more pronounced. In the 'SL_Rainy' condition, there was a major deterioration in the proportion of both lines detected and also in the 'Dark_Rainy' condition. The 'Dark' visibility condition had the highest proportion of both lines detected as compared to the other visibility conditions.



(a) LKS detection performance

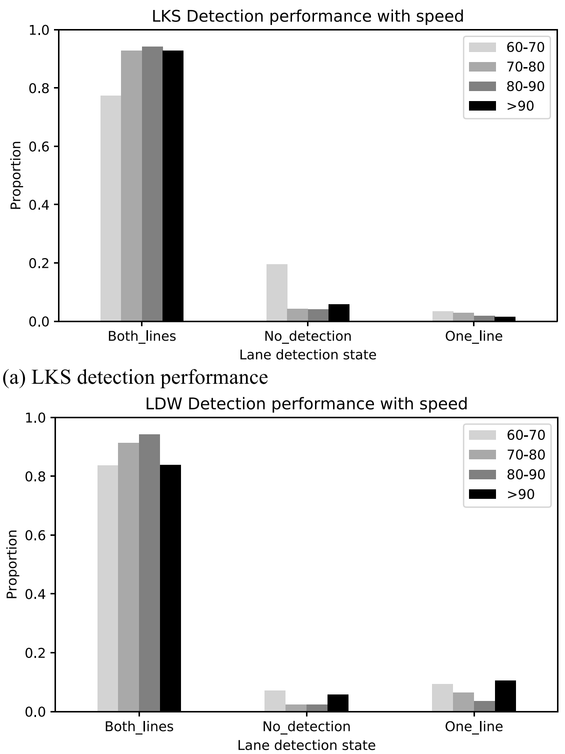


(a) LDW detection performance

FIGURE 5. Detection performance of LKS (upper) and LDW (lower) in different visibility conditions.

Figure 6 shows the detection performance at different driving speed categories. Here also the calculated proportions were normalized with respect to the duration of driving in different speed categories. Both lines were detected in most of the times. For both LKS and LDW, there seems to be a notable increase in the detection performance of both lines when the speed increases from the range of '60-70' to '70-80' category. For the LDW, the results indicate a reduction in the detection performance of both lines from the '80-90' to '>90' category.

As for lane keeping performance, the measured Lane Position was used to estimate the effect of different driving conditions. Figure 7 shows box plots of the measured Lane Position in different conditions of speed categories, curve type, and lane widths. The results indicate that travelling at higher speeds generally makes the LKS keep the vehicle more to the left of the lane center, except at speeds higher than 90 kmph category. The LKS also steers the vehicle more to the left side of the road while driving on left curves compared to when driving on straight sections or right curves. Finally, on narrow roads ($\leq$ 250 cm), the



(a) LKS detection performance

(b) LDW detection performance

FIGURE 6. Detection performance of LKS and LDW at different speed categories.

LKS steers the vehicle more towards the left side of the lane center compared to wider roads (> 250 cm).

IV. MODELLING LANE DETECTION AND LANE-KEEPING PERFORMANCE

Two regressions models were developed to estimate the effect of driving conditions on the performance of the lane assistance systems: a detection performance model, and a lane-keeping performance model. The detection performance for the LKS and LDW was estimated together as one system as the objective was to identify the effect that the driving environment has on detection performance of these camera-based lane assistance systems. The function of both systems in terms of detection was the same: to detect lane markings. Hence, it was not of interest to identify the effects on detection performance for the individual vehicle system models. The development of the regression models and the results are discussed below.

A. DETECTION PERFORMANCE MODEL

In the detection performance model, the predicted variable is the detection state, which could be both lines detected, one line detected, or no lines detected. Therefore, the predicted variable of the model is categorical. The possible predicting variables are visibility condition and speed category. As previously stated, the lane position was measured only for one test session where the lane width and type of curve were recorded. Hence, lane width and type of curve are additional predicting variables to be considered for the last test session. However, due to the few data points, these two variables were excluded.

The Pearson's Chi-Square tests were conducted to check the association of the predicting variables and the predicted variable. As both the predicted and predicting variables are categorical, Chi-square test was used. The Pearson's Chi-Square test result (X^2 (12, N = 99500) = 6460.58, p $<$ .001) confirmed that the visibility conditions and the detection states have a significant association. Also, the Pearson's Chi-Square test result (X^2 (6, N = 99500) = 3730.28, p $<$ .001) confirmed that the speed category and the detection states have a significant association. Therefore, the visibility condition and speed category were both included as predicting variables in the model.

As the predicted variable is categorical with three values and the predicting variables are also categorical, a Multinomial Logistic Regression model was chosen. Moreover, as different systems/vehicles would probably have inherently different levels of performance, therefore a mixed model with a random intercept was estimated. To build and estimate the model, the Generalized Mixed Linear Model (GMLM) in SPSS [36] was used. The model accounted for the impact of speed category and visibility conditions as Fixed effects, while a random intercept was included to capture the correlations between the observations of the same system. The reference predicted variable was set to 'No lines detected'. Table 3 presents the fixed coefficient estimates for the detection performance model.

For the Visibility conditions, all the estimates are significant at 95% and therefore also significantly different from the reference condition ("Dark" condition). The estimates are all negative, except for the "Rainy" condition. The interpretation is that as compared to the "Dark" condition, all other visibility conditions, except for the "Rainy" condition ("Rainy" only refers to daytime driving, as defined in Section II-C), have a lower probability of "Both lines" detected. This seemingly counterintuitive result is because almost all the "Rainy" conditions logged in the field test was "Light Rain". The visibility conditions can also be ranked (as compared to "Dark" condition), in a decreasing order of probability of "Both lines" detected as follows: "Rainy", "Cloudy", "Clear", "Streetlights", "Dark and Rainy", and "Streetlights and Rainy". The rainy night conditions are the "worst-performing", but the day rainy condition "best performing". This result, however, must be considered with the nature of rainy conditions during the test as light rain was largely prevalent during the day test drives. Almost all of heavy rain conditions were experienced during the nighttime driving.

For Speed categories, except for the 70-80 kmph category, the other categories have significant estimates at 95% and therefore also significantly different from the 80-90 kmph category. It also means that the 70-80 kmph category is not significantly different from the 80-90 kmph category. Driving at 60-70 kmph or >90 kmph, as compared to at 80-90 kmph,

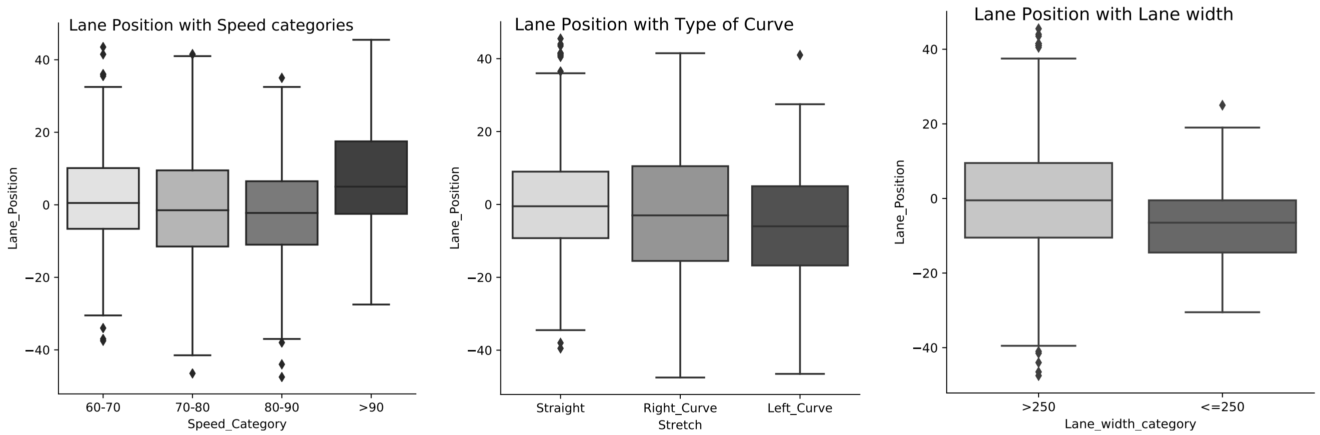


FIGURE 7. LKS Lane Position in different driving conditions.

TABLE 3. Fixed coefficient estimates for both lines detection—GLMM.

Model term	Coefficient estimate	Standard Error	95% CI	p-value
Intercept	3.818	0.423	[2.989, 4.648]	<0.001
Clear Visibility	-0.377	0.056	[-0.487, -0.267]	<0.001
Cloudy Visibility	-0.260	0.056	[-0.370, -0.151]	<0.001
Rainy Visibility	0.190	0.093	[0.008, 0.372]	0.041
Dark_Rainy Visibility	-0.544	0.063	[-0.668, -0.421]	<0.001
SL Visibility	-0.383	0.055	[-0.492, -0.275]	<0.001
SL_Rainy Visibility	-1.750	0.067	[-1.881, -1.619]	<0.001
Dark Visibility		*(Reference value)*		
60-70	-1.542	0.045	[-1.630, -1.454]	<0.001
70-80	-0.006	0.047	[-0.099, 0.087]	0.907
>90	-0.590	0.072	[-0.730, -0.449]	<0.001
80-90		*(Reference value)*		

has a lower probability of having "Both lines" detection. The chance of "Both lines" detection is more than twice at speeds higher than 90 kmph than at speeds of 60-70 kmph, with respect to 80-90. This may be partially explained because the LKS requires a minimum speed of about 65 kmph to start activation (however, it may be noted that while decelerating from higher speeds, the LKS remains active up to 60 kmph).

B. LANE-KEEPING PERFORMANCE MODEL

The commonly adopted norm is to use Mean Lateral Position as the indicator as it is useful to evaluate the lane-keeping performance on specific stretches of roads that are of a fixed length. In this study, however, the performance on a temporal level was of interest, as the objective was to evaluate the performance while driving in different environmental conditions. Therefore, it was decided to use the Lane Position as the predicted variable. The predicting variables considered were speed category, lane width, and type of curves. The model did not account for the visibility conditions as there were no changes in the visibility conditions during the selected test drive. As the predicting variables are all categorical, ANOVA test was used to examine which variables have a significant association with the Lane Position. Moreover, a multiple comparisons test was done to see if there were significant differences in the Lane Position between the categories of the variables. As the predicted variable was continuous, and the predicting variables were categorical, a Multiple Linear Regression model was chosen. The SPSS tool was used to build the regression model.

The ANOVA test showed that for speed categories, there were significant differences in the Lane Position between the 60-70 kmph category and the 80-90 kmph category (Lane position being 3.49 cm more left at 80-90 than at 60-70, significant at 95% confidence level), between 70-80 kmph with >90 kmph (Lane position being 8.57 cm more right at >90 than at 70-80, significant at 95% confidence level), and between 80-90 kmph with >90 kmph (Lane position being 9.80 cm more right at >90 than at 80-90, significant at 95% confidence level). Therefore, there was a significant effect of Speed on Lane Position.

For lane widths, it was observed that the only significant differences were between the <250 cm category and all other lane width categories (namely 250-270 cm, 270-290 cm, 290-310 cm, and >310 cm). The mean differences between the <250 cm category and the other categories were, respectively, −6.92 cm, −6.10 cm, −6.83 cm, and −7.95 cm, all being significant at 95% confidence level. The negative sign indicates that the lane position was more left on lane widths less than 250 cm than on all other measured lane widths. There was no significant difference of Lane Position between the other lane width categories. Therefore, there was a significant effect that a lane width has on Lane Position. Moreover, given the high insignificance of

TABLE 4. Lane-keeping performance model coefficient estimates.

Model term	Unstandardized B	Standard Error	95% CI	Significance
(Constant)	.799	0.545	[-0.270, 1.869]	0.143
Lanewidth_below_250	-6.125	1.367	[-8.807, -3.443]	<0.001
Lanewidth_above_250			*(Reference value)*	
Left Curve	-6.685	1.200	[-9.039, -4.330]	<0.001
Right Curve	-2.102	1.207	[-4.469, 0.266]	0.082
Straight section			*(Reference value)*	
60-70	2.092	1.224	[-0.309, 4.494]	0.088
80-90	-1.625	0.818	[-3.230, -0.020]	0.047
>90	8.095	2.551	[3.091, 13.099]	0.002
70-80			*(Reference value)*	

differences between the other categories, the Lane width category was redefined as a binary variable (either <= 250 cm, or > 250 cm).

For type of curve, there were significant differences in Lane Position between left curves and straight sections (Lane position being 6.39 cm more left on left curves than on straight sections, significant at 95% confidence level), as well as between left and right curves (Lane position being 4.27 cm more left on left curves than on right curves, significant at 95% confidence level). Therefore, there was a significant effect that the type of curve has on Lane Position.

Hence, to build the lane-keeping performance regression model, Speed, Lane width, and Type of curve are chosen to be included. Table 4 presents the fixed coefficient estimates for the lane-keeping performance model. The significant effects are only from Lane width_below_250, the Left_curve, the Speed_over_90, and the Speed_80-90.

Lane widths below 250 cm tend to make the vehicle steer in such a way that the Lane Position is about 6 cm more left than on roads having Lane widths over 250 cms. Also, driving on left curves makes the LKS to keep about 6.7 cms more left than on straight sections. Driving over 90 kmph tends the LKS to keep about 8 cm more right than 70-80 kmph (however, driving above 90 kmph was negligible during the field test session), and driving at 80-90 kmph tends the LKS to keep about 1.6 cm more left than 70-80 kmph.

V. PERFORMANCE CLASSIFICATION AND ODD LEVELS OF SERVICE

This section describes the classification of the measured performance indicators using defined thresholds. Then, the driving conditions are classified into different levels of service based on the performance evaluation.

A. PERFORMANCE EVALUATION THRESHOLDS

The performance of the lane assistance systems was measured using the indicators discussed earlier. For each of these indicators, thresholds were defined for classification: High, Medium, and Low Performance. This classification was done for each performance indicator separately. The performance thresholds for each indicator are shown in Table 5.

These thresholds were partly based on existing literature that used performance indicators for human driving. The 'average' performance measured using these indicators for human driving has been defined as 'High Performance'.

TABLE 5. Performance evaluation thresholds for the indicators.

Indicator	High Performance	Medium Performance	Low Performance
Percentage No Lines Detection	≤ 5%	>5%, ≤10%	>10%
Percentage Both Line Detection	>90%	≤90%, >85%	≤ 85%
MLP[1]	≤ ±2 cm	> ±2 cm, ≤ ± 4 cm	> ±4 cm
SDLP	≤ 15 cm	> 15 cm, ≤ 30 cm	>30 cm

[1]Mean and Median Lane Position

The medium and low-performance thresholds have been decided upon using reasonable logic and expectations. The performance of the lane assistance systems in different driving conditions was evaluated using the defined performance thresholds to derive the levels of service of the driving conditions (Operational Design Domain).

B. OPERATIONAL DESIGN DOMAIN LEVELS OF SERVICE

The driving conditions under which the corresponding level of performance was observed is used to formulate the levels of service of Operational Design Domain. First, the performance thresholds are used to evaluate the performance in the different driving conditions. Table 6 shows the performance thresholds applied to the detection performance in different visibility conditions and speed categories. Values in bold indicate high performance, those in italics indicate medium performance, and those underlined indicate low performance.

For the "Percentage Both Lines Detection" indicator, Clear, Cloudy, Rainy, and Dark are the "High Performance" driving conditions. Dark_Rainy and Streetlights are the "Medium Performance" driving conditions, and Streetlights_Rainy is the "Low Performance" driving condition. Previous discussion showed that all conditions are significantly different from the Dark condition in terms of Percentage Both Lines Detection. Concerning the speed categories, 60-70 kmph has "Low Performance", 70-80 kmph and 80- 90 kmph have "High Performance", and >90 kmph has "Medium Performance" for both the detection indicators. Previous discussion showed that only 70-80 kmph was not significantly different from the 80-90 kmph in terms of Both Lines Detection.

TABLE 6. Detection evaluation in visibility conditions and speed categories.

Driving condition	Both Lines Detection (%)	No Lines Detection (%)
Visibility category		
Clear	**90.3**	*5.6*
Cloudy	**92.4**	*5.4*
Rainy	**92.9**	**4.6**
Dark	**94.7**	**3.3**
Dark_Rainy	*82.0*	*6.3*
Streetlights	*90.0*	*6.6*
Streetlights_Rainy	61.6	15.4
Speed category (kmph)		
Clear	80.4	13.5
Cloudy	**92.1**	**3.3**
Rainy	**94.2**	**3.3**
Dark	*88.4*	*5.8*

For the "Percentage Both Lines Detection" indicator, Clear, Cloudy, Rainy, and Dark are the "High Performance" driving conditions. Dark_Rainy and Streetlights are the "Medium Performance" driving conditions, and Streetlights_Rainy is the "Low Performance" driving condition. Previous discussion showed that all conditions are significantly different from the Dark condition in terms of Percentage Both Lines Detection. Concerning the speed categories, 60-70 kmph has "Low Performance", 70-80 kmph and 80- 90 kmph have "High Performance", and >90 kmph has "Medium Performance" for both the detection indicators. Previous discussion showed that only 70-80 kmph was not significantly different from the 80-90 kmph in terms of Both Lines Detection.

Table 7 shows the performance thresholds applied to lane-keeping performance in different lane widths, types of curves, and speed categories. Again, values in bold indicate high performance, those in italics indicate medium performance, and those underlined indicate low performance.

Lane width less than or equal to 2.5 m cause the LKS to have "Low Performance", and lane width above 2.5 m cause the LKS to have "High Performance". While the Mean (and Median) Lane Position is significantly more Left on lane widths less than or equal to 2.5 m, the SDLP is lower than the SDLP on lane widths greater than 2.5 m. This suggests that the vehicle manufacturer intended this kind of lane-keeping performance. As for curves, the MLP on Left curves is significantly more left than on Straight sections and Right curves and is classified as "Low Performance". The MLP for right curves is not significantly different from Straight curves as previously discussed. Concerning the speed categories, only the >90 kmph sees "Low Performance". Earlier discussion already indicated the low duration of driving at >90, and hence may be due to other specific situational factors. All other speed categories see "High Performance" in terms of MLP.

TABLE 7. Lane-keeping performance evaluation (+ right bias; - left bias).

Driving condition	Mean Lateral Position (cm)	Median Lateral Position (cm)	SDLP (cm)
Lane width			
Less than or equal to 2.5 m	-6.69 (Left)	-6.50 (Left)	**10.97**
Greater than 2.5 m	**-0.21** (Left)	**-0.50** (Left)	**14.82**
Curve type			
Straight	**0.18** (Right)	**-0.50** (Left)	**13.72**
Left curve	-6.21 (Left)	-6.00 (Left)	*16.38*
Right curve	**-1.94** (Left)	*-3.00* (Left)	*18.11*
Speed category			
60-70	**1.50** (Right)	**0.50** (Right)	**14.92**
70-80	**-0.75** (Left)	**-1.50** (Left)	*15.09*
80-90	**-1.99** (Left)	*-2.25* (Left)	**13.13**
>90	7.82 (Right)	5.00 (Right)	*18.98*

TABLE 8. ODD levels of service.

Level of Service	Visibility condition	Speed category	Lane width	Type of curve
High Performance	Dark, Rainy, Cloudy, Clear	70-80, 80-90	>= 2.5 m	Straight section, Right curve
Medium Performance	Streetlights, Dark_Rainy	>90		
Low Performance	Streetlights_Rainy	60-70	< 2.5 m	Left Curve

These performance classification of different driving environments are combined to form the Levels of Service of the driving environment for those systems. Table 8 presents the Levels of Service of the ODD from the described classification.

VI. CONCLUSION AND DISCUSSION

This section first summarizes the results, then provides a reflection on the research method and the results of this study in comparison to the state of the art, and finally the research limitations.

A. SUMMARY

This research provides an insight into the driving environment factors that affect the performance of lane assistance systems such as the LKS and LDW. The performance evaluation was divided into detection performance and lane-keeping performance, where detection performance measured the ability of the system to detect lane markings, and lane-keeping performance measured the ability of the LKS to keep the vehicle inside the lane by automatic steering. The field test revealed the effect that visibility conditions and vehicle speed had on the lane detection performance. Driving at night in rain under streetlights resulted in the lowest lane detection performance. This is attributed to the very low visibility of lane markings in that driving condition, making it hard even for human eyes to see the markings. Driving at night in clear weather resulted in the highest detection performance.

This is explained by the high contrast difference between the lane marking and the pavement, which is more pronounced during the night. Speeds of 70-80 kmph resulted in the highest detection performance. These results showed that these systems detect “lines” in the driving environment, which does not necessarily need to be lane markings. In some instances during the field test, the LKS followed the pavement repair patchwork instead of the lane markings. In some other instances, it identified the high contrast difference between the pavement and the shoulder (mostly grass or soil/sand) as lane markings and was found to even drive very close to the shoulder.

The identified factors that affected the lane-keeping performance were lane width, type of curve, and speed. Lane widths less than or equal to 250 cm resulted in the LKS steering the vehicle more to the left of the lane center (and consequently closer to the median). Human drivers also display this behavior by steering away from the edge of the road when on narrow lanes. The low Standard Deviation of lane position on lane widths less than or equal to 250 cms seems as intended behavior by the vehicle manufacturer. The highest lane-keeping performance was observed on lane widths wider than 250 cms. Lane position measured on curves and straight sections showed that there is a significant difference in lane position between left curves and straight sections. The LKS steered the vehicle more towards the left side of the lane center when driving on left curves than on straight sections. This result cannot be explained without knowing the precise mechanism of lane-keeping adopted by the LKS. However, it may be interesting to note that human drivers also tend to keep to the inside of the curves (more left on left curves and more right on right curves). Although there was no significant difference between the right curves and straight sections, a significant difference may become evident with more data. The highest lane-keeping performance was observed on straight sections and the least on left curves. Finally, investigation of the effect of speed on lane-keeping performance showed that only speeds over 90 kmph resulted in low performance, however, the duration of driving at over 90 kmph was very small. All other speeds resulted in high performance.

B. REFLECTION ON METHOD

The field test proved to be invaluably useful. Field tests (including structured experiments) are an effective way to test the performance of a System as it results in getting the “ground truth”. It is possible to capture the actual effect of the driving environment on the performance of the systems. Driving with the two vehicles on the road gave rise to deep insights into their functioning, that may not be expected or predictable. For instance, the effect of lane width on lane-keeping performance could not have been predicted without knowing the algorithm of the LKS, which is generally not publicly available for commercial car manufacturers. The field test provided this important insight. While the field test was useful, there are limitations also with this method for performance evaluation. It may not be able to capture all the different possible scenarios that can be encountered by the systems. There is also heavy dependence on the extent and quality of the data collected. Besides, the results are valid for those specific scenarios and limited to the observations. Thus, there is little opportunity to gain a macroscopic understanding of the effects of the driving environment and to extend it to other Advanced Driver Assistance Systems (ADAS).

C. REFLECTION ON STATE OF THE ART

One of the most common expected changes in road infrastructure is that lane widths can be reduced as Lane Assistance Systems can keep consistently to a specific position in the lane without deviations [25], [29], [37]. The field test showed that the position of the vehicle on the lane is inconsistent. The position depends on different factors such as speed, type of curve, and width of the lane. It also depends on other potential factors that this research did not study (such as lane marking configuration, type, quality, pavement characteristics, shoulder type, median type). During the field test, the LKS failed to safely navigate some sharp curves. Moreover, different manufacturers have different algorithms that run the LKS, and they currently have no standard guidelines on what acceptable good performance is. Therefore, reduction in lane widths must be considered carefully after extensive study on the capabilities of the LKS, and certainly is not a measure that road authorities must implement immediately.

The pilot test conducted by [34] found that the Lane Keeping System cannot operate on lane widths less than or equal to 2.5 m, and it always can operate on lane widths greater than or equal to 2.75 m. It was argued that lane widths could not be reduced with the current automation level of vehicle development. However, the field test conducted in this research reveals that the LKS can operate in lane widths less than 2.5 m as well. However, the performance reduces from “High Performance” at lane widths greater than 2.5 m to “Low Performance” at lane widths of 2.5 m and narrower. Thus, field tests done using different automated systems result in different results.

Existing literature focuses on lane markings with respect to their quality and type [26], [27]. It is suggested that lane markings must be consistent and of good quality as inconsistent lane markings could confuse the Lane Assistance Systems. This is not very conclusive from this research, as the detection relies on finding a “line” on the road, as opposed to identification of a “lane marking”. Lane markings that are different in configuration are not expected to have a significant effect as long as they can offer a “line” as required. The suggestion of [27] to provide different types of lane markings to cause different driving behaviors based on the lane markings may take time. However, as detection methods advance to include identification of lane markings and corresponding changes to the algorithm of these systems, the configuration of lane markings and its effect on performance becomes important.

D. RESEARCH LIMITATIONS

Data collected of the visibility conditions were subjective as it was only noted if there was clear weather, rainy weather, or streetlights. Therefore, there was no data collection done for the objective visibility, such as lighting intensity of the streetlights. Thus, the results are suggestive of the effect that visibility conditions have on the performance but unable to precisely estimate the effect of the variability of the conditions themselves (such as different intensities of streetlights) on the performance. There was no data collected concerning the steering wheel, which could have been an important aspect. Thus, there was no data on the time of application of steering correction or the extent of steering correction. The differentiating between the human driver steering and the LKS steering was done by manual logging by the co-driver. So, there are bound to be some errors. However, the delays in logging change of steering control have been corrected. The detection performance is measured from the dashboard of the vehicles. This indicates what the vehicles "think" they see. There is not necessarily an implication on the quality of infrastructure that they are driving on. However, there were only a few instances when there was an actual mismatch between the detection state and the actual presence of lane markings, as determined by a manual check. The data on the position of the vehicle on the lane was available only for one test session due to time and resource constraints. Therefore, there could not be any insights into the effect of visibility conditions on the lane position. Also, the dataset was limited in size due to it being only one test session. Additional data could have resulted in more reliable results. Finally, as discussed previously, lack of lane marking type and quality could have provided better insights into the performance of the lane assistance systems.

VII. RECOMMENDATIONS

First, recommendations with respect to the research method are summarized, followed by the recommendations for future research, and finally recommendations for road authorities.

A. METHOD RECOMMENDATIONS

Firstly, understanding the actual mode of operation of the ADAS would be invaluable. However, the algorithms implemented in these systems are not publicly available. Therefore, it would be beneficial to collaborate with vehicle manufacturers resulting in insights that would make the results much more valuable. Secondly, more precise and reliable devices for measurements, such as CAN bus or LiDAR, must be used as opposed to video cameras for measurements (that this research used for measuring the distance to the lane marking). Steering wheel data is also highly recommended to be collected for evaluating ADAS such as LKS and LDW. This data would be useful in measuring indicators such as Steering Reversal Rate, Steering response time, and Number of steering reversals. These indicators are expected to add much more explanation to how the vehicle is driving. Finally, it is recommended to perform field tests with vehicles from different manufacturers, as done in this research, to account for market variability.

B. RECOMMENDATIONS FOR FUTURE RESEARCH

Further research could focus on including additional ADAS, such as Adaptive Cruise Control (ACC), therefore also going from Level 1 to higher levels of automation. It would be interesting to see how the vehicle drives when these ADAS are used individually and also in combination. It is also crucial to study the interaction between the human driver and the ADAS, and how that affects the performance of the vehicle in different road environments. Furthermore, the effect that surrounding traffic has on the performance of the ADAS-equipped vehicle and also vice-versa are very interesting research questions. In terms of data collected, it is recommended to collect lane marking quality and lane marking configuration data. This data would have added much value by providing much more insights into the factors that affect the performance of the ADAS. Future studies could find data on the roadside infrastructure useful as components along the road could affect the performance of the LKS and LDW systems. Furthermore, it is very interesting to develop mathematical models to gain a precise understanding of the effect of various factors on performance. At the same time, it should be noted that the validity and applicability of these models may be limited to the specific automated driving systems that are studied. Finally, it is also useful to perform similar studies not only on provincial roads but also on the national motorways and other road environments.

C. PRACTICAL RECOMMENDATIONS

The Automated Vehicle Safety Consortium [10] has laid out a best practice for defining the ODD that can be used by vehicle manufacturers. The defined driving environment features that are relevant to this research, such as lane width, posted speed, weather (cloudy, rainy, etc.) conditions, lighting conditions, and horizontal alignment, are specified. This research contributes to the next step, that is, measuring and classifying the performance of automated vehicles in different driving environments, and also providing a taxonomy to the ODDs with respect to levels of performance. Using a method similar to this research, road authorities can identify critical road sections that are expected to result in low performance of ADAS enabled vehicles, and set-up a prioritization system to make improvements to these road sections.

The European New Car Assessment Programme [11] uses a maximum permissible Distance-To-Lane-Edge (DTLE) value to evaluate lane assistance systems, meaning that the lane-keeping system must not permit the vehicle to cross the inner edge of the lane marking by the specified distance. The lane-keeping systems are evaluated using this criterion in a binary way (pass/fail). Such an evaluation is insufficient as it only focuses on the "final stage of failure". It is critical to understand the performance of these systems even when the vehicle is within the lane. Moreover, only straight road sections are included in the test protocol. As was seen in this research, curves are critical sections where lane assistance systems may fail. Hence, the testing protocols for such systems must include more driving environments, especially high-risk situations such as curves.

There must be a collaboration between the road authorities to share knowledge and experience to ensure consistency in terms of road design standards, and also in terms of outlook towards ADAS enabled vehicles. Furthermore, collaborations with vehicle authorities and also vehicle manufacturers are needed to agree on accepted performance standards for driving on the roads with mixed traffic. The taxonomy of the levels-of-service defined in this research is an effort towards this direction.

ACKNOWLEDGMENT

This research was hosted at the Provincie Noord-Holland (Province of North Holland) in The Netherlands. The Province of North Holland arranged and funded the field test and data processing.

REFERENCES

[1] B. Schoettle and M. Sivak, "A preliminary analysis of real-world crashes involving self-driving vehicles," Transp. Res. Inst., Univ. Michigan, Ann Arbor, MI, USA, Rep. UMTRI-2015-34, 2015.

[2] D. Han and K. Yi, "Evaluation of adaptive cruise control algorithms on a virtual test track," in *Proc. Amer. Control Conf.*, Minneapolis, MN, USA, 2006, pp. 5849–5854.

[3] Y. Li, Z. Li, H. Wang, W. Wang, and L. Xing, "Evaluating the safety impact of adaptive cruise control in traffic oscillations on freeways," *Accident Anal. Prevention*, vol. 104, pp. 137–145, Jul. 2017.

[4] V. Milanés and S. E. Shladover, "Modeling cooperative and autonomous adaptive cruise control dynamic responses using experimental data," *Transp. Res. C, Emerg. Technol.*, vol. 48, pp. 285–300, Nov. 2014.

[5] J. Ploeg, B. T. M. Scheepers, E. van Nunen, N. van de Wouw, and H. Nijmeijer, "Design and experimental evaluation of cooperative adaptive cruise control," in *Proc. 14th Int. IEEE Conf. Intell. Transp. Syst.*, Washington, DC, USA, pp. 260–265, 2011.

[6] J. Ploeg, A. F. A. Serrarens, and G. J. Heijenk, "Connect & drive: Design and evaluation of cooperative adaptive cruise control for congestion reduction," *J. Modern Transp.*, vol. 19, no. 3, pp. 207–213, 2011.

[7] W. J. Schakel, C. M. Gorter, J. C. F. de Winter, and B. van Arem, "Driving characteristics and adaptive cruise control? A naturalistic driving study," *IEEE Intell. Transp. Syst. Mag.*, vol. 9, no. 2, pp. 17–24, Apr. 2017.

[8] *Europe on the Move: Commission Completes Its Agenda for Safe, Clean and Connected Mobility*, European Commission-Press Release, European Commission, Brussels, Belgium, 2018.

[9] Society of Automotive Engineers, *(R) Taxonomy and Definitions for Terms Related to Driving Automation Systems for On-Road Motor Vehicles*, Standard J3016_201806, 2018.

[10] Automated Vehicle Safety Consortium, *AVSC Best Practice for Describing an Operational Design Domain: Conceptual Framework and Lexicon*, Standard AVSC00002202004, 2020.

[11] European New Car Assessment Programme. (2019). *Assessment Protocol–Safety Assist*. [Online]. Available: https://doi.org/10.1017/CBO9781107415324.004

[12] A. Das, A. Ghasemzadeh, and M. M. Ahmed, "Analyzing the effect of fog weather conditions on driver lane-keeping performance using the SHRP2 naturalistic driving study data," *J. Safety Res.*, vol. 68, pp. 71–80, Feb. 2019.

[13] Z. Chu, C. Wu, and N. Sepehri, "Automated steering controller design for vehicle lane keeping combining linear active disturbance rejection control and quantitative feedback theory," *Proc. Inst. Mech. Eng. I, J. Syst. Control Eng.*, vol. 232, no. 7, pp. 937–948, 2018.

[14] A. Montella, M. Aria, A. D'Ambrosio, F. Galante, F. Mauriello, and M. Pernetti, "Simulator evaluation of drivers' speed, deceleration and lateral position at rural intersections in relation to different perceptual cues," *Accident Anal. Prevention*, vol. 43, no. 6, pp. 2072–2084, 2011.

[15] J. Verster and T. Roth, "Standard operation procedures for conducting the on-the-road driving test, and measurement of the standard deviation of lateral position (SDLP)," *Int. J. Gen. Med.*, vol. 4, pp. 359–371, May 2011.

[16] P. Green *et al.*, "Typical values for driving performance with emphasis on the standard deviation of lane position: A summary of the literature," Univ. Michigan Transp. Res. Inst., Ann Arbor, MI, USA, Rep. UMTRI-2003-42, 2003.

[17] P. Li, N. Merat, Z. Zheng, G. Markkula, Y. Li, and Y. Wang, "Does cognitive distraction improve or degrade lane keeping performance? Analysis of time-to-line crossing safety margins," *Transp. Res. F, Traffic Psychol. Behaviour*, vol. 57, pp. 48–58, Aug. 2018.

[18] B. Mehler, B. Reimer, and J. F. Coughlin, "Sensitivity of physiological measures for detecting systematic variations in cognitive demand from a working memory task: An on-road study across three age groups," *Human Factors*, vol. 54, no. 3, pp. 396–412, 2012.

[19] B. Reimer, "Impact of cognitive task complexity on drivers' visual tunneling," *Transp. Res. Rec.*, vol. 2138, pp. 13–19, 2009.

[20] S. Glaser, S. Mammar, M. Netto, and B. Lusetti, "Experimental time to line crossing validation," in *Proc. 8th Int. IEEE Conf. Intell. Transp. Syst.*, Vienna, Austria, 2005, pp. 791–796.

[21] A. P. Tarko, "Use of crash surrogates and exceedance statistics to estimate road safety," *Accident Anal. Prevention*, vol. 45, pp. 230–240, Mar. 2012.

[22] E. Johansson *et al.*, *Review of Existing Techniques and Metrics for IVIS and ADAS Assessment (D2.2.1)*, document IST-1-507674-IP, AIDE, 2004. [Online]. Available: http://www.aide-eu.org/pdf/sp2_deliv_new/aide_d2_2_1.pdf

[23] J. Östlund *et al.*, *Driving Performance Assessment Methods and Metrics (D2.2.5)*, document IST-1-507674-IP, Adaptive Integrated Driver-Vehicle Interface Project (AIDE), 2005. [Online]. Available: http://www.aide-eu.org/pdf/sp2_deliv_new/aide_d2_2_5.pdf

[24] W. W. Wierwille, L. Tijerina, S. Kiger, T. Rockwell, E. Lauber, and A. C. J. Bittner, "Heavy vehicle driver workload assessment—Task 4: Review of workload and related research," Rep. DOT HS 808 467, 1996.

[25] H. Farah, S. M. J. G. 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*. Cham, Switzerland: Springer, 2018, pp. 187–197.

[26] "Future proofing infrastructure for connected and automated vehicles," Catapult Transport Systems, Milton Keynes, U.K., Rep. V1.1, 2017. [Online]. Available: https://s3-eu-west-1.amazonaws.com/media.ts.catapult/wp-content/uploads/2017/04/25115313/ATS40-Future-Proofing-Infrastructure-for-CAVs.pdf

[27] R. Huggins *et al.*, "Assessment of key road operator actions to support automated vehicles," Austroads Ltd., Sydney, NSW, Australia, Rep. AP-R543-17, 2017.

[28] P. Nitsche, I. Mocanu, and M. Reinthaler, "Requirements on tomorrow's road infrastructure for highly automated driving," in *Proc. Int. Conf. Connected Veh. Expo (ICCVE)*, 2014, pp. 939–940.

[29] P. Morsink, E. Klem, I. Wilmink, and M. de Kievit, "Zelfrijdende auto's: Ontwikkelagenda ZRA en wegontwerp," Haskoningdhv Nederland B.V., Amersfoort, the Netherlands, Rep. T&PBE6132R001D03, 2016.

[30] *Vehicle Fleet Penetrations and ODD Coverage of NRA-Relevant Automation Functions up to 2040*, document Deliverable 2.1, MANTRA, 2019.

[31] *Marking the Way Towards a Safer Future*, European Road Federation, Brussels, Belgium, 2013.

[32] C. Davies, "Pavement markings guiding autonomous vehicles—A real world study," in *Proc. Autom. Veh. Symp.*, 2016.

[33] *Roads That Cars Can Read: A Quality Standard for Road Markings and Traffic Signs on Major Rural Roads*, European New Car Assessment Programme (Euro NCAP) & European Road Assessment Programme (EuroRAP), Leuven, Belgium, 2011.

[34] A. García, and F. J. Camacho-Torregrosa, "Influence of lane width on semi-autonomous vehicle performance," *Transp. Res. Rec.*, vol. 2674, no. 9, pp. 279–286, 2020.

[35] H. Farah *et al.*, "An empirical analysis to assess the operational design domain of lane keeping system equipped vehicles combining objective and subjective risk measures," *IEEE Trans. Intell. Transp. Syst.*, early access, Feb. 3, 2020, doi: 10.1109/TITS.2020.2969928.

[36] *IBM SPSS Statistics for Macintosh, Version 25.0*, IBM Corporation, Armonk, NY, USA, 2017.

[37] T. Johnson, R. Chen, R. Sherony, and H. C. Gabler, "Driver lane keeping behavior in normal driving using 100-car naturalistic driving study data," in *Proc. IEEE Intell. Veh. Symp. (IV)* Gothenburg, Sweden, 2016, pp. 227–232.

NAGARJUN REDDY received the M.Sc. degree in civil engineering (transport and planning) from the Delft University of Technology, The Netherlands, where he is currently pursuing the Ph.D. degree with the Department of Transport and Planning. In his Ph.D. research, he is working on modeling human driver behavior in the presence of automated vehicles and the resulting behavioral adaptation.

HANEEN FARAH received the Ph.D. degree in transportation engineering from the Technion—Israel Institute of Technology in 2009. She is currently an Associate Professor with the Department of Transport and Planning and also the Co-Director of the Traffic and Transportation Safety Lab, Delft University of Technology (TU Delft). Before joining TU Delft, she was a Postdoctoral Researcher with the KTH-Royal Institute of Technology, Stockholm, Sweden. Her research interests include studying the implications of road infrastructure design on road user behavior and traffic safety. In her research, she combines her expertise in transportation engineering, with her curiosity in the fields of human factors and econometrics to study these connections. She is currently examining the implications of the advances in vehicle technology and automation on the road infrastructure design and road user behavior in the framework of several national and international projects.

YILIN HUANG received the Ph.D. degree in modeling and simulation (M&S) from the Delft University of Technology, The Netherlands, in 2013. From 2014 to 2016, she was a Postdoctoral Researcher working on prosumer behavior modeling in the smart grids. Since 2016, she has been an Assistant Professor of M&S with the Section of Systems Engineering and Simulation, Faculty of Technology, Policy and Management, Delft University of Technology. Her research mainly concerns M&S methodology with particular interests in data-driven and data analytics-related subjects, and interoperability and composability of dynamic models. The application domain of her research ranges from transportation and logistics, smart grid, and sustainable consumption to humanitarian innovations.

THIJS DEKKER received the M.Sc. degree in transport, infrastructure, and logistics from the Delft University of Technology in 2014. He is currently working as an Internal Consultant with the Province of Noord-Holland, The Netherlands. He is a specialist in road design, in particular road design to facilitate state-of-the-art vehicles.

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