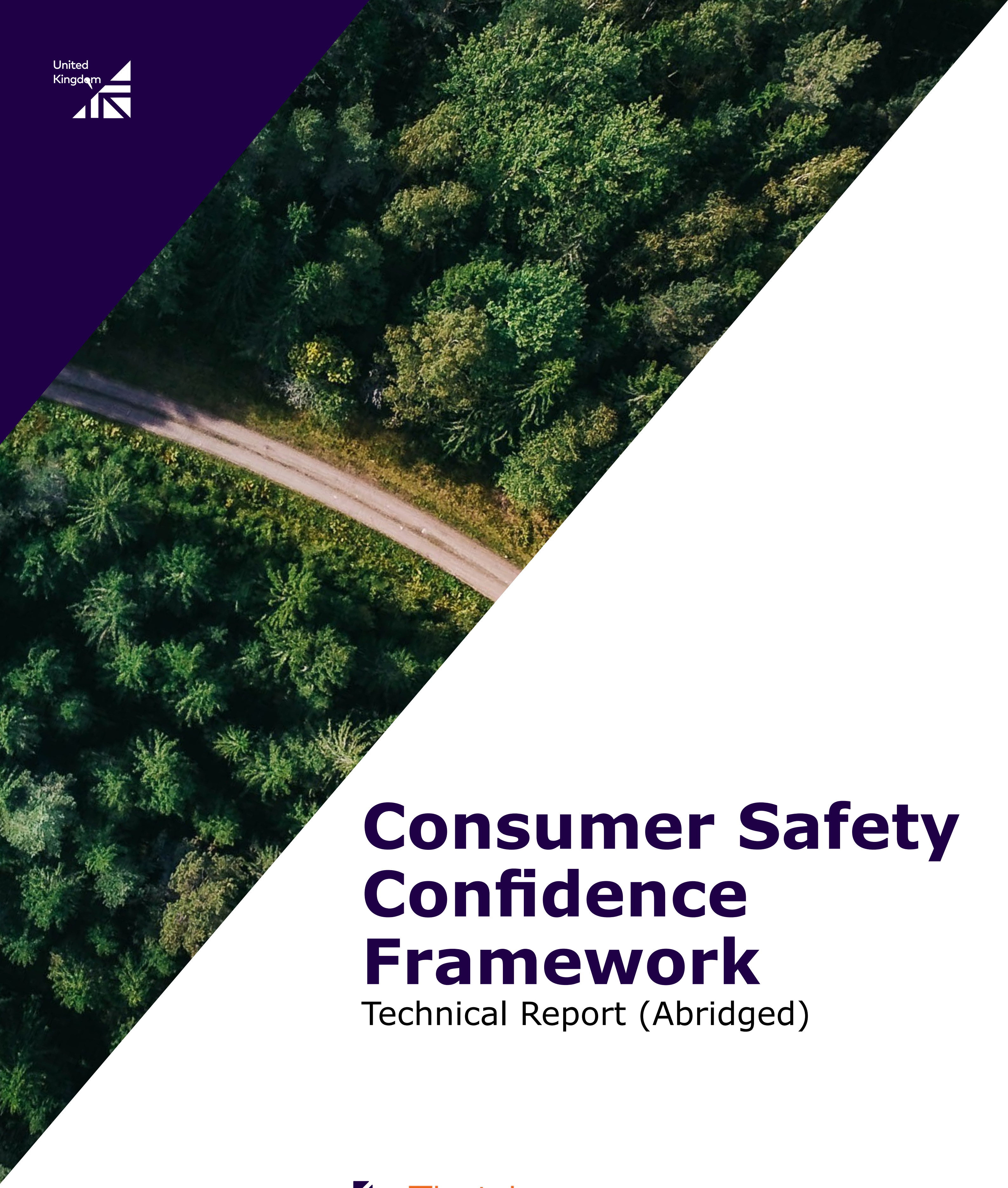
United
Kingdom



Consumer Safety Confidence Framework

Technical Report (Abridged)

ZENZIC
SELF-DRIVING REVOLUTION

Thatcham Research

WMG
THE UNIVERSITY OF WARWICK

A HORIBA COMPANY MIRA

UTAC

AESIN
AUTOMOTIVE ELECTRONICS INNOVATION

Contents

Disclaimer

This report has been produced by Thatcham Research under a contract with Zenzic-UK Ltd (Zenzic). Any views expressed in this report are not necessarily those of Zenzic. The information contained herein is the property of these organisations and does not necessarily reflect the views or policies of the customer for whom this report was prepared. Whilst every effort has been made to ensure that the matter presented in this report is relevant, accurate and up to date, Zenzic and/or any of the authors of this report cannot accept any liability for any error or omission, or reliance on part or all of the content in case of incidents that may arise during trialling and testing. In addition, Zenzic and/or any of the authors of this report cannot accept any liability for any error or omission, or reliance on part or all of the content in another context.

For further information on this report, please contact the Zenzic team at info@zenzic.io

Author
Tom Leggett – Thatcham Research
Max Calcroft – Thatcham Research

Glossary of Abbreviations, Acronyms & Industry Terms

ABD – Anthony Best Dynamics

ACC – Adaptive Cruise Control

AD – Assisted Driving

AEB – Automatic Emergency Braking

ALKS – Automated Lane Keeping System

CAM – Connected and Autonomous Mobility

CANBUS – Controller Area Network bus

CBAR – Combined Brake and Accelerator Robot

CCRb – Car to Car Rear braking

CCRm – Car to Car Rear moving

Challenging Vehicle – A moving vehicle which will interrupt the actions of the vehicle under test

DAU – Data Acquisition Unit

Ego – Another term for the vehicle under test from the Latin "I"

Euro NCAP – European New Car Assessment Programme

FCW – Forward Collision Warning

GPS – Global Positioning System

GST – Guided Soft Target

GVT – Global Vehicle Target

i-ACC – Intelligent Automatic Cruise Control

IPG - Ingenieurgemeinschaft Prof. Dr.-Ing. R. Gnadler GmbH. (Simulation platform)

LC – Lane Change

MRM – Minimum Risk Manoeuvre

ODD – Operational Design Domain

OEM – Original Equipment Manufacturer

SAE – Society of Automotive Engineers

SDL – Scenario Description Language

SR – Steering Robot

Static/Stationary Vehicle – A stationary vehicle obstructing the path of the vehicle under test

TP – Test Point

Vlat – Lateral Velocity; the average lane departure velocity achieved between starting and ending a lane change manoeuvre.

VT – Virtual Testing

PT – Physical Testing

THW – Time Head Way

TTC – Time to Collision

UiC – User in Charge

UNECE - United Nations Economic Commission for Europe

VUT – Vehicle Under Test

VM – Vehicle Manufacturer

WMG – Warwick Manufacturing Group

Executive Summary

Introduction

This project focuses on the development of proof of concept test protocols and procedures that lead to the development of a draft Consumer Safety Confidence Framework Rating. The draft framework will provide the building blocks for an expansive independent assessment program for automated vehicles.

Traditional physical testing procedures provide a high level of confidence and independent scrutiny of vehicle safety performance. As vehicles and their technology become more advanced, the behaviour and capabilities of automated technologies such as ALKS require a much broader testing methodology. Virtual testing allows for a wide range of scenarios to be efficiently conducted, reenforcing consumer confidence in the capabilities of ALKS. It is vital that consumers trust the virtual testing, therefore the results must be verified using traditional physical testing. This project looked at a comparison of independent virtual and physical testing, in terms of their result alignment and how that influenced the development of the proof-of-concept ALKS rating framework.

Scenario and Test Vehicle Identification

The chosen scenarios for testing were Cut-Out and Cut-In. Both scenarios are technically challenging assessments for vehicle object detection and identification, but additionally are visually easy to comprehend for both consumers and law makers alike. Most importantly, both the cut-out and cut-in scenarios are already established test scenarios both within regulation UNECE R157 as well as independent testing programmes such as Euro NCAP. Additionally, the burden for manufacturers to perform such tests is significantly reduced if they already perform similar tests for other regulatory or assessment purposes. Test plans for each scenario were created based upon UNECE R157 and the performance capabilities of the test vehicle.

<u>Cut-Out Scenario Definition:</u>

A lead vehicle is travelling in lane at 60 km/h. The test vehicle follows the leading vehicle at a set distance, controlled by the adaptive cruise control, matching the speed of the lead vehicle. At a defined point, the lead vehicle changes lane, to reveal a stationary vehicle in the path of the test vehicle. The test vehicle was then assessed on its ability to avoid or mitigate the collision with the stationary vehicle. The speed at which the lead vehicle changes lane varied for different test points.

<u>Cut-In Scenario Definition:</u>

The test vehicle is travelling in lane at 60 km/h. The challenging vehicle in an adjacent lane is traveling at a speed slower than the test vehicle. At a defined longitudinal distance between the front of the test vehicle and the rear of the challenging vehicle, the challenging vehicle will perform a lane change into the same lane as the test vehicle. The test vehicle was then assessed on its ability to avoid or mitigate the collision with the challenging vehicle. The speed at which the challenging vehicle changes lane varied for different test points.

The selected test vehicle was a 2021 Toyota C-HR GR Sport HEV CVT. No ALKS equipped vehicles were publicly available during this project. The tier-one supplier who supported the project from a technical standpoint designed the radar, camera and brake systems for the Toyota C-HR. This meant that they had access to the simulation models of the system, without having to directly interact with Toyota. This in turn allowed us to virtually simulate the performance of the vehicle from a completely independent standpoint. The separate models were assembled into a vehicle model where each sub-system was able to communicate with the other, and thereby provide a simulated response of an equivalent physical test vehicle. It should be noted that Toyota were not involved in this project.

Physical and Virtual Testing

The physical tests defined by the test plan were conducted at two CAM Testbed UK sites: HORIBA MIRA and UTAC. Additional testing was also performed by Thatcham Research to help define the test scenarios and establish the test vehicle performance capabilities. The virtual tests defined by the identical test plan were executed by the tier-one supplier, with support from Warwick Manufacturer Group (WMG), in the simulation platform IPG. The virtual model of the Toyota C-HR was developed and created in IPG’s CarMaker software.

The cut-out test scenario demonstrated a good alignment of results, despite the occurrence of object detection loss by the test vehicle during the simulated tests. Furthermore, the virtual testing was able to confirm that the +0.5 s test points were not the performance limit of the vehicle. Virtual results were seen to have a gentle initial response (comfort braking) to the stationary target, however object loss was observed by the slight increase in acceleration followed by sharp emergency braking. This initial period of comfort braking is likely a design choice of the adaptive cruise control system; expecting the driver to respond to the situation. This highlighted a limitation of using a proxy-ALKS vehicle for this scenario.

Mixed correlation between physical and virtual results was observed for cut-in, with the virtual results demonstrating a consistent test vehicle response able to avoid a collision with only low levels of braking. This may be due to the detection of the challenging vehicle in the simulation being under ideal circumstances. Whereas the physical testing was influenced by external factors such as the reflectivity of the challenging vehicle cutting in front of the test vehicle. However, generally both virtual and physical results showed gradual and sustained comfort braking as the vehicle cuts in.

As expected, virtual and physical test results were not perfectly aligned. The biggest discrepancies were found in the cut-in scenario where the physical testing produced no response from the test vehicle, whereas the virtual testing showed consistent test vehicle comfort braking and collision avoidance. This highlights the challenging nature of simulation comparison, as the environmental factors vary for each scenario that is tested. The combination of the challenging cut-in scenario, the way in which the test vehicle is designed to respond and the limitations of only using components of the vehicle model virtually, led to this discrepancy. If the simulation could use an entire vehicle model, that had been thoroughly verified and validated, the results may have been more representative. The cut-out scenario however provided good confidence in virtual testing, with its strong alignment in results and maximum acceleration values. Therefore, demonstrating the proof of concept of this project, that simulation can provide useful results for an independent rating assessment.

The conclusion was that virtual testing requires a highly verified and validated vehicle model in order to closely represent the results of physical testing. Despite the time constraint of this project, a vehicle model that was able to closely align with the physical test results was created. Further verification and validation would have likely provided even more representative results across the virtual and physical results. The most straight forward method of obtaining a highly verified and validated vehicle model, would be to work directly with the vehicle manufacturer. However, it is unlikely that the vehicle manufacturer would allow such access to their intellectual property for the purposes of virtual testing due to security concerns. Therefore, allowing vehicle manufacturers to execute the simulation testing themselves, using their own highly verified and validated models, negates this issue. This however emphasises the need to verify the results provided against physical testing. The crucial output from the virtual testing will be the performance indicators from each assessment.

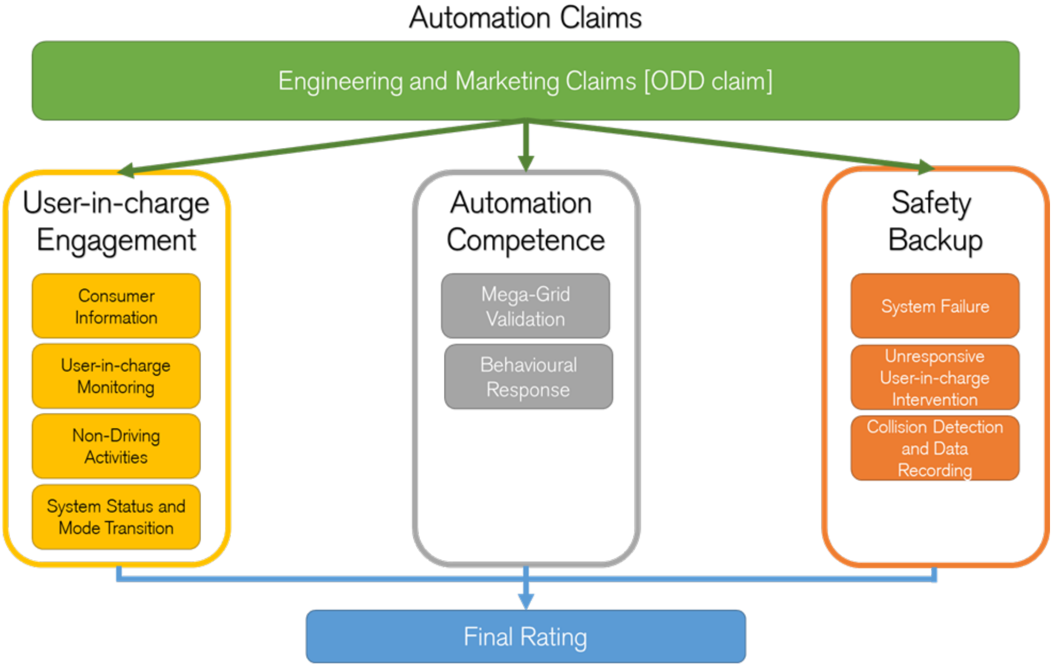
Framework Philosophy

The key philosophy of the ALKS Consumer Safety Confidence Framework is that the safety rating of the vehicle under test must be a function of its Operational Design Domain (ODD). Each ALKS equipped vehicle will have its own unique ODD, in which the defined limits of when it can operate are described. However, there is a need to verify the claimed ODD of the manufacturer against the measured performance of the vehicle under test. The framework begins with the Automation Claims, where the driving domain of the vehicle under test is defined against the assessment criteria. This ODD checklist (detailed further in this document) forms the basis of the assessment framework. This checklist will inform which tests are to be carried out both virtually and physically. Additional to the checklist, further information such as marketing media will be procured and assessed.

Vehicle manufacturers should be rewarded for providing not only a wide-ranging ODD, but also an accurately described one. This balance between claimed ODD and the measured assessment performance will feed into the three sections User-in-Charge Engagement, Automation Competence and Safety Backup. Which in turn will generate the over-all vehicle rating.

Framework Structure

Figure 1 Thatcham Research draft framework structure



Automation Competence has been divided into two parts: mega-grid validation and behavioural response. The mega-grid refers to large number of scenarios with a large number of parameters for each, creating a web of test points. The majority of these test points will be executed in a virtual testing environment. It is proposed that the responsibility of executing the virtual tests is given to vehicle manufacturer or tier one supplier. These virtual results will then be spot-checked using physical testing, to ensure confidence in the virtual testing results. The key scoring influence will be the actual measured performance from the physical testing. Behavioural response allows for more complex scenarios to be assessed, where the outcome of the test may not be as definitive as the mega-grid method. Ensuring that this rating system is modular and has the ability to expand beyond ALKS for other more advanced automated technology is crucial.

User-in-Charge Engagement recognises that understanding the limitations of ALKS is crucial to not only the safe operation of the system but also how readily consumers will adopt the new technology. If users overestimate or underestimate the capabilities of ALKS, it could result in a dangerous situation. Therefore, it is important to assess how vehicle manufacturers advertise and explain this technology to consumers.

Safety Backup acknowledges that the capabilities of ALKS requires numerous and highly advanced sensors such as radar, LiDAR and cameras. During real world driving one or more of these sensors may deteriorate over time or become damaged or blocked in adverse weather conditions. The capabilities of ALKS may be affected in these cases, and therefore it is important that the safe operation of the system is not diminished.

1. Background

1.1 Project Objectives

- Create a proof of concept draft of a Consumer Safety Confidence Framework for ALKS which:
    - Outlines the key areas interest which must be considered to allow the safe adoption of ALKS by consumers
    - Builds the foundations of a rating system based on the Operational Design Domain (ODD), safety performance and behavioural response of an ALKS equipped vehicle
    - Provide a mechanism by which virtual testing can be used in tandem with physical testing, in order to expand the independent testing rigor
    - Has a modular structure to allow expansion for future automated technology
- Create a technical report which:
    - Details the methodology and test execution which will feed into the creation of the framework
    - Provide confidence in the ability to use independent virtual testing to meaningfully assess the response of automated vehicle technology
- Demonstrate the use of CAM Testbed UK for conducting physical testing

1.2 Scope of Work

- Identify relevant test scenarios for ALKS
- Procure test vehicle
- Execute physical testing for the selected test scenarios
    - Create test plan using identified test scenarios
    - Perform the physical tests at two CAM Testbed UK sites
- Execute virtual testing for the selected test scenarios
    - Create vehicle model for use in virtual environment
    - Mirror test plan for use in virtual environment
    - Perform the virtual tests in the simulation platform
- Compare and analyse the virtual and physical results to provide guidance in the creation of the rating framework
- Provide a proof of concept draft structure for a modular rating framework

2. Project Partners

2.1 Thatcham Research – Project Lead

2.2 Warwick Manufacturing Group (WMG)

2.3 Automotive Electronic Systems Innovation Network (AESIN)

2.4 CAM Testbed UK – Physical Test Facilities

The UK is the global centre for the innovation and development of connected and self-driving vehicle technologies. CAM Testbed UK is the only place worldwide with the capability to safely take ideas from concept to development both virtually and physically, all within a 3-hour drive. The UK's comprehensive and integrated facilities are world-leading, with the cross-sharing of data and a collaborative way of working. HORIBA MIRA and UTAC are members of CAM Testbed UK.

3. Scenarios, Test Vehicle and VUT Performance Benchmarking

3.1 Vehicle Under Test (VUT) Selection

The vehicle used for testing was a 2021 Toyota C-HR GR Sport HEV CVT. The vehicle itself was not equipped with ALKS, however it was fitted with Toyota Safety Sense assisted driving technology to act as a proxy-ALKS vehicle for the purposes of this project. This assisted driving technology included radar guided Adaptive Cruise Control (ACC) which controls the speed of the vehicle as well as automatically adjusting the vehicle speed to maintain a safe distance from a leading vehicle. The radar system also controls the automatic emergency breaking (AEB), which will reduce the speed of the vehicle if the system considers that a collision is imminent.

The test vehicle was chosen for the following reasons:

The tier-one supplier who supported the project from a technical standpoint built and designed the radar, camera and brake systems for the Toyota C-HR. This meant that they had access to the simulation models of the system, without having to directly interact with Toyota.

This allowed us to virtually simulate the performance of the vehicle from a completely independent standpoint. It should be noted that Toyota were not involved in this project.

Although the supplier had access to the radar, camera and brake models, they did not have a complete "off the shelf" vehicle model which represents the entire vehicle system. Therefore, they had to build a comprehensive vehicle model that could be used in virtual testing, with the assistance of WMG.

3.2 Relevant Scenario Identification

Chosen Scenarios for Testing

The scenarios chosen were cut-out and cut-in. These scenarios provide a technical challenge to the VUT but are also very relevant to a potential situation that will be faced by an ALKS equipped vehicle.

A non-ego vehicle changing lane is a scenario in which an ALKS equipped vehicle would commonly experience within the operational design domain (ODD) of the system. The cut-out represents the most difficult circumstance of this, where the lead vehicle reveals a stationary vehicle in lane. The simplicity of this scenario is easy to comprehend visually for both consumers and law makers alike. This enables non-technical users of ALKS to understand the various situations that must be assessed to ensure safe adoption.

Additionally, both scenarios are a technically challenging assessment for object detection and identification. In cut-out, as the lead vehicle changes lane it slowly reveals the stationary vehicle, but it is very difficult for the VUT to differentiate the lead vehicle from the partially revealed stationary vehicle. This is where advanced identification and categorisation systems are required in order to provide a good response. For example, a basic system can only “look” at the vehicle ahead, whereas a more complex system can identify two or more vehicles in front. This allows a much earlier response to any potentially dangerous situations. The same concept can be applied for the cut-in scenario, where the VUT must be able to identify the side of a challenging vehicle moving into its path. This movement can be hard to identify with varying lateral velocities for radar only systems when there are several objects with varying lateral velocities as viewed in the reference frame of the VUT.

Finally, both the cut-out and cut-in tests are already established test scenarios both within independent testing programmes such as Euro NCAP and the regulation UNECE R157. Although the focus of this project is to create an assessment programme for consumers, vehicle manufacturers must also buy in to the concept of the framework. By starting with established test scenarios, manufacturers will be more inclined to participate and help with any future development of the assessment framework. Additionally, the burden for manufacturers to perform such test is significantly reduced if they already perform similar tests for other regulatory or assessment purposes.

Lateral Velocity Definition:

The term lateral velocity referenced in UNECE R157 and the Euro NCAP AD protocol differ and for clarity, any use of the term lateral velocity (Vlat) in the context of this document refers to: the average lane departure velocity achieved between starting and ending a lane change manoeuvre.

For example, a lateral movement from one lane into another of 3.6 m, performed by a vehicle in 3.6 seconds, would result in a Vlat of 1.0 m/s.

Cut-Out Scenario Definition:

The VUT follows the Lead Vehicle at an ACC distance setting equivalent to 2.0 seconds. The lead vehicle changes lane at a defined point, to achieve an average lateral velocity defined in table 3, to reveal a stationary vehicle in lane. The VUT will then be assessed on its ability to avoid or mitigate the collision with the stationary GST.

- Following distance (time-headway) will be set using the VUT ACC settings. It was found that the second setting (out of three possible settings), stabilised at approximately 1.8 to 1.9s. This allows the VUT to better represent a proxy-ALKS vehicle. The lane centering support will be activated on the VUT.
- Speed of both VUT and Lead Vehicle will be 60 km/h. The current version of UNECE R157 limits the maximum speed of an ALKS equipped vehicle to this speed.
- The ACC speed of the VUT will be set > 60 km/h to ensure that it maintains the set following distance to the Lead Vehicle.
- The lateral velocity of the Lead Vehicle is derived from the time taken to complete the lane change. The lane change start point distance was taken from the UNECE R157 data sheet in Figure 8 and then adjusted to ensure the lead vehicle moved 1.8m laterally to narrowly avoid a collision with the stationary vehicle.
- The lane change starts at a defined distance between the lead vehicle front bumper and the rear of the GST.

- A lateral velocity of < 0.4 m/s was deemed unrealistic and therefore excluded from the test points.

The main concern of the cut-out scenario was whether the steering robot would be able to achieve the desired lateral velocity (Vlat). Table 3 indicates the Vlat values greater than 2.0 m/s in red, as it was thought to be difficult to physically complete the lane change manoeuvre in less than 1.8 seconds. Physical on-track testing would be needed to verify whether these test points were feasible.

Cut-In Scenario Definition:

The VUT is travelling in lane at 60 km/h. The challenging vehicle in an adjacent lane travels at a speed slower than the VUT. At a defined longitudinal distance between the front of the VUT and the rear of the challenging vehicle, the challenging vehicle will perform at lane change into the same lane as the VUT. The VUT will then be assessed on its ability to avoid or mitigate the collision with the challenging vehicle.

- Following distance will be set using the VUT ACC settings. The setting will be the same as the cut-out to ensure a THW of 2.0 seconds to align with the UNECE R157 requirements.
- The GST will have different speeds depending on the scenario; 50 km/h, 40 km/h, 30 km/h and 20 km/h. These speeds are taken from the UNECE R157 requirements.
- The lateral velocity of the challenging vehicle is derived from the time taken to complete the lane change. The lane change start point distance was taken from the relevant UNECE R157 data sheet.
- The lane change starts at a defined distance between the VUT front bumper and the rear of the challenging vehicle.
- The lane change ends when the centre of the rear of the challenging vehicle first crosses the centre of the target lane.
- A lateral velocity of < 0.4 m/s was deemed unrealistic and therefore excluded from the test points.

The main concern of the cut-in scenario was whether the GVT would be able to achieve the desired lateral velocity (Vlat). This concern mirrors the cut-out but is worsened as the platform performing the lane change manoeuvre is the GST, which does not have the same turning agility of a regular vehicle. Physical on-track testing would be needed to verify whether these test points were feasible with the GST.

4. Physical Testing

4.1 HORIBA MIRA Physical Testing

Cut-Out Testing Conclusion

- HORIBA MIRA results achieved desired lane changes lane change start point and lateral velocities
- HORIBA MIRA vs Thatcham Research data showed very good alignment for +1.0s test point, including acceleration profiles and stopping distances
- HORIBA MIRA data suggested that the +0.5s test point was not the performance limit of the VUT

The tests were shown to accurately execute the correct lane change in accordance with the test definition for cut-out. The lead vehicle used by HORIBA MIRA was a Volvo XC90, which had notably better steering response than that of the Ford Fiesta used at Thatcham Research, despite being considerably heavier. For further testing after this project, consideration should be made to define the required lead vehicle to ensure greater consistency.

The +1.0s test points aligned very closely to the Thatcham Research data, indicating that the response of the VUT was reliable. The maximum deceleration values were consistent with that of Thatcham Research, however all test results exhibited breaking greater than -5.0 m/s², indicating that emergency braking was required to avoid a collision. HORIBA MIRA data suggested that the VUT was able to avoid a collision at the +0.5s test point. This contradicts the performance benchmark results from Thatcham Research and highlights the importance of conducting multiple tests to confirm performance results. For further testing at this test point multiple runs will be conducted to ensure consistent VUT response.

Cut-In Testing Conclusion

- HORIBA MIRA results achieved desired lane changes lane change start point and lateral velocities for the majority of test point
- Higher lateral velocities would be difficult to achieve using the currently available test equipment for remote guided targets
- HORIBA MIRA data demonstrated the VUT was unable to provide a consistent braking response, especially the +1.0s test points
- HORIBA MIRA results demonstrated a combination of comfort and emergency braking

The cut-in testing highlighted the physical limitations of the test equipment when attempting to achieve high lateral velocities. The GST is not designed for high manoeuvrability, and this was evident by the difficulty in achieving lateral velocities above 1.5 m/s. However, below this it was able to achieve the desired test points accurately and reliably. Regardless, the lane change of the challenging vehicle did occur at the correct longitudinal distance.

The VUT response was varied across the test points. For both +1.0 s scenarios the VUT struggled to reliably brake for the challenging vehicle and often required driver intervention to avoid a collision. This highlights how difficult this scenario is for a system to detect, as often they are designed to identify the front and rear of a vehicle. In this manoeuvre the side of the challenging vehicle is the most prominent object, which this system found difficult to identify and respond to.

For both +2.0 s scenarios the VUT response was much more reliable and demonstrated a mixture of comfort and emergency braking. For the majority of these test points the VUT was able to avoid a collision with comfort braking alone. This is due to the relative speed between the VUT and challenging vehicle being comparatively low (10 km/h and 20 km/h respectively), and therefore only minimal braking is required. The challenge for the VUT is being able to identify the challenging vehicle before a collision occurs.

The performance limitations of the VUT were demonstrated in this scenario. As mentioned previously this vehicle is not an ALKS equipped vehicle, and therefore cannot be expected to perform at the same level. Initial testing by Thatcham Research showed that +1.0s scenarios resulted in some braking response, but the HORIBA MIRA testing demonstrated that only reliable braking was achieved at +2.0s.

4.2 Physical Testing at UTAC

Cut-Out Testing Conclusion

- UTAC results achieved desired lane change start point and lateral velocities
- Comparing UTAC and Thatcham Research data showed very good alignment for +1.0s test point, including acceleration profiles and stopping distances
- The UTAC data confirmed that the +0.5s test point was not the performance limit of the VUT

The tests were shown to accurately execute the correct lane change in accordance with the test definition for cut-out. The lead vehicle used by UTAC was a Mercedes-Benz A-Class coupe, which had notably better steering response than that of the Ford Fiesta used at Thatcham Research. For further testing after this project, consideration should be made to define the required lead vehicle to ensure greater consistency. The +1.0s test points aligned very closely to the Thatcham Research data, indicating that the response of the VUT was reliable. The maximum deceleration values were consistently lower than that of Thatcham Research, however all test results exhibited braking greater than -5.0 m/s^2, indicating that emergency braking was required to avoid a collision. The UTAC data confirmed the findings of the HORIBA MIRA results; the VUT avoided a collision at the +0.5s test point. This contradicts the performance benchmark results from Thatcham Research and highlights the importance of conducting multiple tests to confirm performance results.

5. Virtual Testing

5.1 Virtual Model Test Set-up

The virtual model of the Toyota C-HR was developed and created in IPG's CarMaker software. This process involved selecting the vehicle dynamics including its relevant sub-models such as braking, steering, powertrain and suspension. All relevant parameters were then identified and parameterised for all sections of the model from vehicle to environment to controller. Simulation data was then tuned against reference data and the simulation was then executed, producing the relevant results and visualisation of the desired test runs.

Within the simulation environment the ARS510 radar sensor controls the braking response, with the MFC431 camera used for additional classification and identification of objects. It should be noted that the primary function of the camera is to provide functions such as lane keep assist rather than emergency braking response. The primary braking response is controlled by the radar sensor. After initial parametrisation from CCRm results (section 3.5) the brake model was further refined to reduce oscillation in results. The root cause was found to be an error in the feedback between the braking model and the requested braking. The brake model provided a larger feedback response than was truly being achieved, which resulted in the response algorithm reducing braking prematurely before correcting and reapplying the brakes. This is what caused the oscillation in the data seen in Figure 44.

Any braking greater than -5.0 m/s^2 was considered emergency braking within the testing and anything less, defined as comfort braking.

The values required for a programmed lane change for both cut-out and cut-in were calculated in CarMaker. Parameters such as vehicle speed, lateral distance required for lane change and the time in which the lane change needed to occur were input as variables. CarMaker then calculated the required lateral velocity to complete the lane change within the specified lane change duration. The distance at which this lane change began was input for each test point based on the initial datasheet for cut-out and cut-in and was triggered precisely once the vehicle reaches this point.

Due to the nature of the simulation, multiple runs of the same parameter setup would result in exactly the same data output. Therefore, conducting multiple runs of the same scenario within the simulation would only provide duplicated data results. For each test point only one simulation execution was required.

5.2 Virtual Testing Comparison Conclusion

- Cut-out tests showed good alignment between the virtual and physical results
- Stopping distances were smaller for virtual testing
- Maximum decelerations were higher for virtual testing
- Cut-in tests showed varied alignment between the virtual and physical results
- Virtually tested +1.0 s scenarios (TP1-5 and 11-16) generally resulted in collision avoidance, this was not observed in the physical testing results

- Virtually tested +2.0 s scenarios (TP6-10 and 17-22) resulted in collision avoidance, this was observed in the physical testing results

The cut-out test scenario demonstrated a good alignment of results, despite the object detection loss of the VUT in the simulation. Furthermore, the virtual testing was able to confirm that the +0.5 s test points were not the performance limit of the vehicle. Physical test results were generally more erratic in nature compared to virtual. With the Thatcham Research data, a spike is seen in acceleration around 6m before the static vehicle. This may be due to brief object loss via the system where it momentarily loses track of the static vehicle ahead, but it was not possible to confirm this as access to the VUT CANBUS was not available during this project. Virtual results were seen to have a greater initial response to the stationary target, however object loss as previously mentioned can be seen at an earlier stage by the slight increase in acceleration followed by emergency braking to around -11.0 m/s². This initial period of comfort braking is likely a design choice of the ACC system. As it is not ALKS, the ACC system is expecting the driver to respond to the situation, this initial period of comfort braking is designed to encourage a reaction whilst also giving the driver more time to act. This highlights a limitation of using a proxy-ALKS vehicle for this scenario.

Mixed correlation between physical and virtual results was observed for cut-in, with the virtual results demonstrating a consistent VUT response able to avoid a collision with comfort braking alone. This may be due to the detection of the challenging vehicle in the simulation being under ideal circumstances. Whereas the physical testing was influenced by external factors such as the reflectivity of the GVT. The current GVT revision F, which was used for testing has been certified to represent the radar signature of a small vehicle front and rear. The side of the GVT has not yet been certified to wholly represent the side of a vehicle which may influence the ability of the VUT to detect the challenging vehicle as it performs a lane change. However, generally both virtual and physical results showed gradual and sustained comfort braking as the vehicle cuts in. Virtual braking profile is initially harsher but still below -5 m/s² of emergency braking.

As expected, virtual and physical test results were not perfectly aligned. The biggest discrepancies were found in the cut-in scenario where the physical testing produced no VUT response, whereas the virtual testing provided consistent VUT comfort braking and collision avoidance. This highlights the challenging nature of simulation comparison, as it varies for each scenario that is tested. The combination of the challenging cut-in scenario, the way in which the VUT is designed to respond and the limitations of only using components of the vehicle model virtually, led to this discrepancy. If the simulation could use an entire vehicle model, that had been thoroughly verified and validated, the results may have been more representative. The cut-out scenario however provided good confidence in virtual testing, with its strong alignment in results and maximum acceleration values.

Additionally, the crucial output from the virtual testing will be the performance indicators from each assessment. These are typically collision avoidance and maximum deceleration and are considered the dynamic response of the vehicle. As long as confidence can be provided that these results are replicated across virtual and physical, the integration of virtual testing in a rating framework will be successful.

It should also be noted that the weather conditions and track environment were not taken into consideration for this testing. All of the physical tests were performed at UNECE R130 compliant lane marked areas, and the virtual environment was matched to this requirement. However, other physical features such as road furniture, surface friction or slight inclines were not taken into consideration. Likewise, the weather parameters for each physical test were recorded but not implemented into the simulation. This was because the nature of the VUT response is mainly determined by the radar and not the camera. Weather for every run of the physical testing was not severe enough as to affect the performance of the radar, and therefore was not included in the simulation testing.

6. Framework Development

6.1 Summary

The purpose of the physical and virtual comparison was to establish whether independent virtual testing could provide reliable results, that could be used in a rating system. Although the virtual test results had some limitations, generally there was good alignment between the physical and virtual testing, and thereby providing confidence that independent virtual testing can provide valuable results in the determination of vehicle performance. Further work is required to expand this confidence across more test scenarios and a variety of test vehicles, but the proof of concept has been established.

Using Euro NCAP Assisted Driving protocol created by Thatcham Research as the initial foundations, the ALKS Consumer Safety Confidence Framework was created. The work detailed within this report fed into the creation of the "Automation Competence".

In the context of this framework, the ODD represents the operating environment within which ALKS can perform the dynamic driving task. There may be functional requirements for such automated systems to safely operate which are not included in the ODD stated in this framework.

8.2 Framework Philosophy

The key philosophy of the ALKS Consumer Safety Confidence Framework is that the safety rating of the VUT must be a function of its Operational Design Domain (ODD). Each ALKS equipped vehicle will have its own unique ODD, in which the defined limits of when it can operate are described. However, there is a need to verify the claimed ODD of the manufacturer and measured performance of the VUT.

For example, a vehicle manufacturer that claims an ODD that can operate across many different conditions e.g. heavy rain, but when the performance of the VUT is independently assessed, it is found that it cannot operate safely in heavy rain.

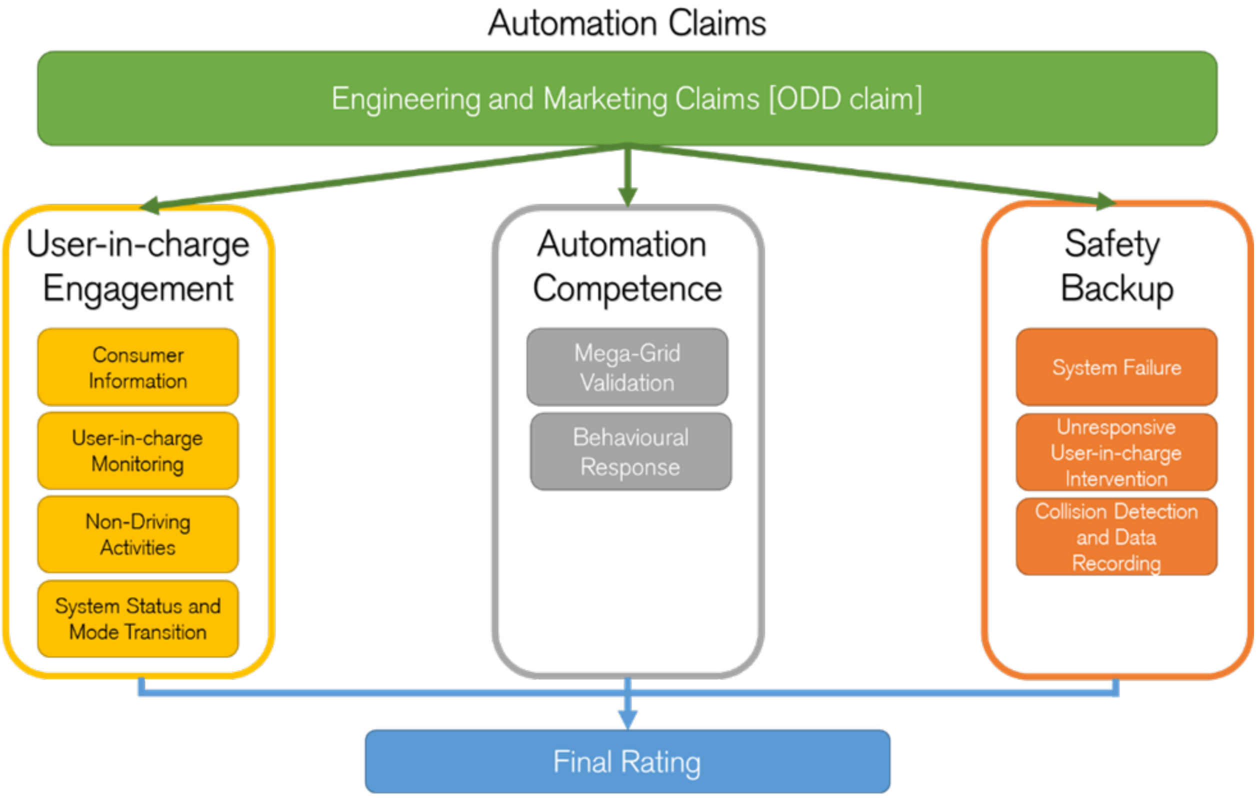
If the rating system is solely based on claiming a wide ODD, in this example the VUT could receive a high rating. However, verifying the claims of the capabilities of the system would result in a low rating. Therefore, the importance of verifying the manufacturer claims against real world performance is vital to provide consumers with confidence in an ALKS equipped vehicle.

A rating system that is based solely on real word performance from a claimed ODD could, conversely, result in manufacturers claiming a narrow ODD. Using the same example, claiming that the VUT cannot operate in rain would result in no testing to verify this claim. Therefore, the additional importance of encouraging a wide but also accurate ODD is the key philosophy of this framework.

The framework begins with the Automation Claims, where the driving domain of the vehicle under test is defined against the assessment criteria. This ODD checklist forms the basis of the assessment framework. This checklist will inform which tests are to be carried out both virtually and physically. Additional to the checklist, further information such as marketing media will be procured and assessed.

It is recognised that each ALKS vehicle will have different capabilities and vehicle manufacturers should be rewarded for providing not only a wide-ranging ODD, but also an accurately described one. This balance between claimed ODD and the measured assessment performance will feed into the three sections User-in-Charge Engagement, Automation Competence and Safety Backup. Which in turn will generate the over-all vehicle rating.

Figure 2 Thatcham Research Draft Framework Structure



6.3 Automation Claims

The ODD definition or automation claims of the ALKS equipped vehicle can be seen as a check list of what conditions the system can operate within. This check list of desired ALKS capability will be provided to the VUT manufacturers, who will indicate the conditions that VUT can operate within. This indicated or claimed ODD will then limit the maximum score achievable by the system. For example, a system that can only achieve 70% of the ODD will be limited to a maximum final score of 70%. This claimed ODD will then be assessed against the actual performance of the VUT, using a combination of virtual and physical testing. The ODD taxonomy is based on PAS 1883:2020.

Engineering Claims

The ODD is classified into three attributes:

- Environment – weather and atmospheric conditions
- Scenery – non-movable elements of the operating environment
- Dynamics – movable elements of the operating environment

Figure 3 ODD Structure and sub-attributes

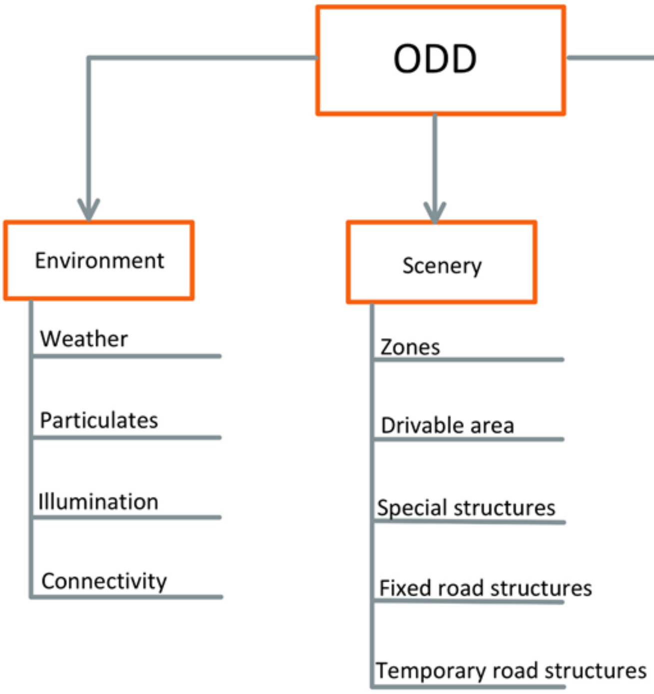


Table 1 ODD Structure in Tabular format

Attribute	Sub-attribute 1	Sub-attribute 2	Capability	Score
Environment	Illumination	Day	Yes	65%
		Night	No	
	…	…	…	
Scenery	Drivable Area Geometry	Straight roads	Yes	75%
		Curves (<1/500 m radius of curve)	Yes	
	Lane dimensions	≥3.6 m width	Yes	

		<3.6 m width (temporary road marking)	No	
	…	…	…	
Dynamics	Traffic	Presence of emergency vehicles	Yes	75%

Environmental conditions play an important role in influencing the safe operation of automated vehicle technology. They also generally present as one of the biggest challenges to successful implementation, as the variety of such conditions varies greatly across geographical locations. The environmental conditions have the potential to impact all ALKS functions from perception and planning to actuation control, as they might impact visibility, sensor fidelity, vehicle manoeuvrability due to changing road surface conditions, and communication systems.

Scenery is less of a focus for ALKS due to the constraints of road types where this technology is permitted to operate. ALKS can only be activated under certain conditions on roads where: vulnerable road users are prohibited, and which are equipped with a physical separation that divides the traffic moving in opposite directions and prevent traffic from cutting across the path of the vehicle. This means that ALKS should only operate on dual carriage ways or motorways. Therefore, any ODD definitions for junctions and roundabouts are not applicable to ALKS.

Dynamic elements refer to the vehicle under test and other actors such as traffic. ALKS is primarily a traffic chauffeur system which will therefore be influenced by the density, volume and flow rate of traffic. Additionally, the system's ability to identify the type and size of other vehicles will be important for to ensure a safe VUT response.

Marketing Claims:

As well as the engineering ODD claims, the marketing claims of the VUT must also be assessed. Many consumers will not look at technical documents to understand the capabilities of an ALKS equipped vehicle and will instead rely upon marketing content and information. Therefore, it is vitally important that this marketing material is also assessed, to ensure that the engineering and marketing claims align, but also that the marketing claims are not misleading. For example, a non-technical user may not understand why the ALKS system does not function above 60 km/h (37 mph), and thereby assume a fault is present.

6.4 Automation Competence

The automation competence will compare the ODD claims to the measurable performance of the ALKS equipped vehicle. This has been divided into two parts: mega-grid validation and behavioural response. Mega-grid validation was the core idea behind this project and is the reason why independent virtual testing had to be undertaken.

The mega-grid refers to large number of scenarios with a large number of parameters for each, creating a web of test points. The ALKS equipped vehicle will be responsible for the entire driving task (within the ODD) and therefore many more scenarios must be assessed, in order to provide confidence that the technology can respond safely. However, this mega-grid would increase the test burden significantly to the point where it may not be possible to physically verify each test point. Therefore, these test points will be executed in a virtual testing environment. These virtual results will then be spot-checked with physical testing, to ensure that the virtual testing results can be trusted. The key performance metrics of the mega-grid will align with current requirements e.g. collision avoidance, speed reduction and maximum deceleration values.

The behavioural response allows for more complex scenarios to be assessed, where the outcome of the test may not be as definitive as the mega-grid method. Ensuring that this rating system is modular and has the ability to expand beyond ALKS for other more advanced automated technology is crucial. The behaviours exhibited by the VUT will be constrained by rules such as the Highway Code, where responses are conditional. For example, Rule 135 of the Highway Code contains the statement: "…In congested road conditions do not change lanes unnecessarily…". There is not a definition for a congested road provided, therefore the automated system will need to infer and decide whether or not a lane change manoeuvre is appropriate. The final complete framework will need to define what a good and bad behavioural response will be, and will be based on human driver characteristics, framework philosophy and the rules of the road.

6.5 User-in-Charge Engagement

As mentioned in the beginning of Section 7, the marketing claims of the manufacturer are of equal importance to the engineering claims. Consumers will get most of their information about a vehicle from marketing, not technical documents. It is therefore important that consumers are given the correct information and that it is easily assessable.

The driver of an ALKS vehicle is no longer responsible for the driving task when ALKS is active. They become a user-in-charge (UiC) and are permitted to perform activities that do not relate to the safe operation of the vehicle, for example consuming content on the in-car display. The activities that are and are not permitted will be defined by regulation but may vary between countries. It is vital that consumers are aware of what activities they can and cannot participate in for the same implementation of automated technologies.

The UiC may participate in non-driving activities, but with ALKS they are required to resume control of the vehicle within a specified time. Therefore, the ALKS equipped vehicle must monitor the UiC to ensure they are available to resume control, once a command is issued. The capability of UiC monitoring systems may vary between manufacturers and some may be able to prevent prohibited activities e.g., sleeping.

System Status and Mode transition represent the need to better understand how the UiC will be kept in the loop when ALKS is active and once a transition demand has been presented. It must be clear to the UiC when they are able to participate in non-driving activities but also when they must respond to resume control of the vehicle.

6.6 Safety Backup

The capabilities of ALKS requires numerous and highly advanced sensors such as radar, LiDAR and cameras. During real world driving one or more of these sensors may deteriorate by age or become damaged or blocked in adverse weather conditions. The capabilities of ALKS may be affected in these cases, and therefore it is important that the safe operation of the system is not diminished. Understanding how the system responds to certain errors will be covered by regulation, but further expansion and requiring specific information to be provided the driver will result in greater reassurance in the event of a system error or failure.

The implementation of the Minimum Risk Manoeuvre (MRM) will differ between manufacturers, and with updates to UNECE R157, additional capabilities such as lane changing may be possible. Therefore, understanding the intention of the VUT whilst performing an MRM is critical. Additionally, features such as E-call will likely offer greater assurance to consumers in the event of an emergency.

Requirements for collision detection and data recording beyond the regulation will provide additional confidence for first responders, collision investigators, and vehicle insurers. Encouraging manufacturers to offer beyond what is required allows the supporting industries of automation to have confidence in the safety of the systems.

6.7 Scoring

The key philosophy of this framework is to encourage a wide and accurate ODD, demonstrating the technical capabilities of an ALKS equipped vehicle. However, a key factor in the consumer uptake and confidence will be their understanding of this technology. Therefore, the importance of user-in-charge education and ultimately engagement is crucial to the successful deployment of ALKS.

User-in-Charge Engagement and Automation Competence share equal weighting of 40% each. Safety Backup, although important for systems failure, will likely be heavily determined by factors such as regulation both locally and internationally. And therefore represents 20% of the score weighting.

Table 2 Framework Point Allocation

User-in-Charge Engagement	**Proportion of Points**
Consumer Information	30%
User-in-charge Monitoring	30%
Non-Driving Activities	20%
System Status and Mode Transition	20%
Total	40%

Automation Competence	**Proportion of Points**

Mega-Grid Validation	75%
Behavioural Response	25%
Total	40%

Safety Backup	Proportion of Points
System Failure	40%
Unresponsive User-in-charge Intervention	40%
Collision Detection and Data Recording	20%
Total	20%

6.8 Scenario Description

In order for the vehicle manufacturer to execute the virtual tests according to the assessment specification, each scenario must be provided in an opensource language format. This format must also be usable across any simulation platforms and human readable.

Scenario Description Language (SDL) provides a two-level abstraction to detail the environment, weather and dynamics of each scenario. Level 1 utilises a structure natural language format which is easy for non-technical end users to comprehend. Level 2 syntax sits at the logical scenario level and contains more detailed information and utilises a formal machine readable format. Upon the generation and formatting of scenarios, they are then stored within the Safety Pool™ scenario database. This database is freely assessable by vehicle manufacturers and tier one suppliers.

7. Suggested Further Work

7.1 Development of a Full Independent Consumer Rating Scheme

It is recommended that the framework developed in this project be expanded upon to produce a full protocol for an independent consumer rating scheme. This project laid the foundations for a deeper investigation into the various aspects needed to safely implement automated vehicles on UK roads.

This next project should be focused on ALKS and the need to ensure its safe and assured deployment in the UK, through the development of thorough and well-defined test protocols and procedures. Further targeting of adoption into independent bodies such as Euro NCAP will provide greater confidence in safety for consumers. This will also continue to set the trend for the UK to be a leader in automated and autonomous vehicle testing.

Additional focus will be the alignment of the rating scheme with UK regulation, Highway Code amendments and any work undertaken by organisations such as the Vehicle Certification Authority. Alignment of all these factors is key for a successful and valuable independent rating framework.

Key Considerations for Future Work

To ensure that the framework is robust and representative, the following considerations must be included in future work for the development of the rating scheme:

- The capabilities of a true ALKS vehicle and future capabilities of automated technologies
- Real world accident data to be used to inform development of additional scenarios
- Data from other testbeds/studies informing testing/regulation
- Definition on the desired response of automated vehicles in certain scenarios i.e. what good and bad behaviour looks like
- Development of the desired ODD for ALKS vehicles, what should be the minimum requirements in the context of this framework for a “good” ODD.
- System improvement over time due to machine learning, artificial intelligence and over-the-air updates.
- Definition of the test parameters to allow consistency of testing across test facilities, both virtually and physically. For example, definition of a lane change start, definition of the lead vehicle for certain scenarios. Consideration of executing multiple runs of each test point to provide statistical analysis.
- Continued alignment with requirements for automated systems such as regulation and independent bodies e.g. Euro NCAP. The landscape of automated vehicle capabilities is constantly changing and ensuring alignment across these parties will allow the safe adoption of this technology.

Table of Figures

Table of Tables

ZENZIC

SELF-DRIVING REVOLUTION

For the full technical report, please contact info@zenzic.io

zenzic.io