VERIFICATION OF CONNECTED AND AUTONOMOUS VEHICLES

VeriCAV – Evaluation of Automated Driving Systems in Simulation

Final Summary Report

March 2021

Contents

1. **Welcome** 3
   - Innovation need 3
   - Scenario testing in Simulation – key for Automated Vehicle Testing 3
   - Landscape for Simulation Verification and Validation 4
2. **Introducing VeriCAV** 5
3. **VeriCAV Framework Overview** 6
4. **Test Generator** 8
5. **Smart Actors** 10
   - Data Sources 11
   - Modelling Approaches 12
   - Evaluation 12
   - VeriCAV Actor Interface 13
6. **Test Oracle** 14
   - Occupant Comfort Oracle Metric & Virtual Reality Assessment 15
7. **Search Space Optimisation** 17
8. **Modular Architecture** 18
   - Flexible to choice of Simulation Tool 18
   - Using and Collaborating with Industry Standards 19
9. **VeriCAV Engagement** 20
10. **Moving beyond VeriCAV** 22
   - Acknowledgements: 22

VeriCAV VERIFICATION OF CONNECTED AND AUTONOMOUS VEHICLES · A HORIBA COMPANY MIRA · CATAPULT Connected Places · aimsun. · UNIVERSITY OF LEEDS

1. Welcome

In January 2019, a consortium from academia and industry commenced the VeriCAV project. This innovative, collaborative project had the goal of creating a framework to allow efficient testing of Automated Driving Systems (ADS) in simulation. Two and a quarter years later, this summary report describes how this goal has been met, outlines the tools developed, and highlights the key findings.

Innovation need

Automated vehicles can transform the way we use transport, unlocking time currently spent driving, reducing accidents from human error, and improving mobility for those who can't easily drive themselves. To achieve this vision, automated vehicles need to be **tested robustly** to ensure their safety before deployment. The complexity and variability of driving in a public road environment makes overcoming this hurdle a significant challenge, however one increasingly important technique involves testing a range of **scenarios in simulation.**

Testing in simulation allows developers to have:
- Full control of test conditions,
- Ease of testing within a safe environment,
- The ability to run and repeat large numbers of tests in parallel.

The challenge in simulation is ensuring realism whilst avoiding the need to formulate and evaluate each test scenario manually.

This is the challenge VeriCAV addresses.

Scenario testing in Simulation – key for Automated Vehicle Testing

It is currently an open challenge to prove that an automated vehicle will always take the safest course of action in every situation it may encounter. It is becoming increasingly accepted that a **combination of test methods** across **different test environments** is required. Methods include: simulation, vehicle-in-the-loop, closed-road testing and public road trials. All these methods can be enhanced and linked together using scenario-based testing.

Scenario-based testing: when a test programme is based on a series of test scenarios. A scenario can be described by road information, stationary objects, movable objects and their movements and environment conditions. Testing in this way, rather than to a list of specific functional requirements, allows challenging and/or standardised sets of scenarios to be the focus of automated vehicle testing without having to go through the extended periods of routine, unchallenging driving that exist in physical on-road testing.

Landscape for Simulation Verification and Validation

The current state of the art in virtual verification and validation is characterised by both its breadth (there are many manufacturers of software at differing Technology Readiness Levels) and depth (with efforts ranging from commercial software, all the way through to material that forms the public-facing aspect of major R&D initiatives, such as PEGASUS[1] in Germany). However, the field is still relatively immature, with many offerings containing only partial integration between elements of the overall verification and validation solution.

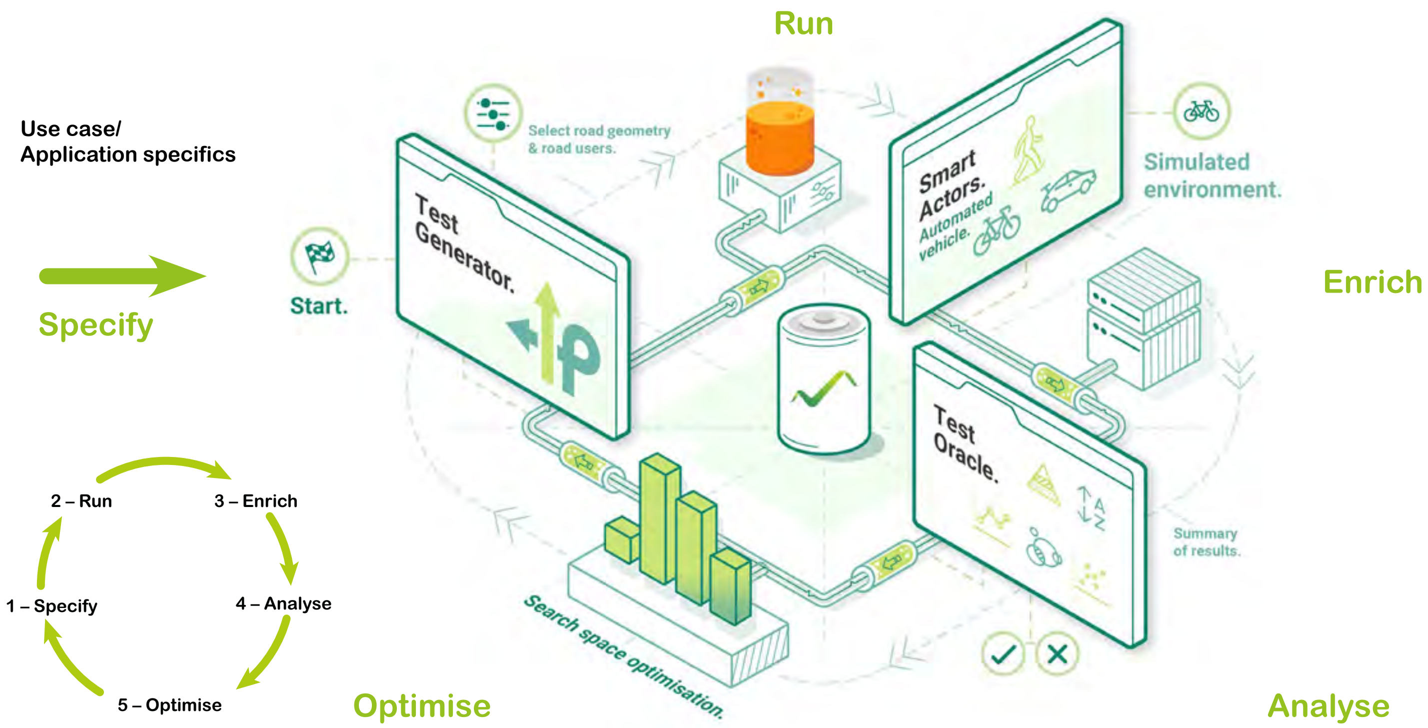
Progress within the UK has accelerated recently following focussed funding from Centre for Connected and Autonomous Vehicles (CCAV)[2] under the CAV Simulation tranche of projects in 2018 (of which VeriCAV is one).

Centre for Connected & Autonomous Vehicles

1 https://www.pegasusprojekt.de/en/home
2 https://www.gov.uk/government/organisations/centre-for-connected-and-autonomous-vehicles

2. Introducing VeriCAV

VeriCAV (Verification of Connected and Autonomous Vehicles) has been a multi-million pound, 27 month collaborative research project involving a partnership of four organisations: HORIBA MIRA (as industry lead), Connected Places Catapult, the University of Leeds, and Aimsun. The goal was to create a framework to allow efficient testing of Automated Driving Systems (ADS) in simulation.



VeriCAV partner organisations:

HORIBA MIRA	HORIBA MIRA, the industry lead partner, led on the test generation, test evaluation, search space optimisation, and integration with the ADS.
CATAPULT Connected Places	Connected Places Catapult led on project management, marketing, systems requirements, simulation platform integration, and verification and validation of the test framework.
UNIVERSITY OF LEEDS	University of Leeds delivered vehicle and pedestrian actors with intelligent behaviour to populate the simulation environment. These are known as smart actors.
aimsun	Aimsun advised on issues relating to the wider context of traffic simulation.

3. VeriCAV Framework Overview

Figure 1 - Assessing the performance of an automated vehicle in pulling out into a gap in a roundabout scenario

VeriCAV focused on the use-case of a central testing framework to be used by automated vehicle developers and approval authorities. Within the context of this use-case, the project addressed some of the key challenges associated with virtual testing of highly automated vehicles[3]:

Reducing the amount of human effort required in the creation, running and analysis of effective test scenarios in simulation
Improving the realism of simulation actor's behaviour.
Establishing reusable interfaces between the simulation and ADS components

3 For more complete descriptions of project challenges see https://vericav-project.co.uk/wp-content/uploads/2020/07/VeriCAV_Introduction_Paper_2020.pdf

The **VeriCAV** project has developed an integrated platform to allow ADS to be tested in simulation. As automated vehicle technology becomes increasingly sophisticated (and the UK more prepared for driverless cars on public roads), the industry needs to explore ways to ensure autonomous vehicles operate safely.

However, safety evaluation is laborious and complex, and real-world testing can be impractical and incomplete – simulation gets around these problems and opens the door to evaluating multi-layered and uncommon situations. The VeriCAV framework will also mean significantly improving test efficiency when evaluating countless driving scenarios, and at the same time it can replicate the behaviour and actions of actors in a realistic and consistent manner.

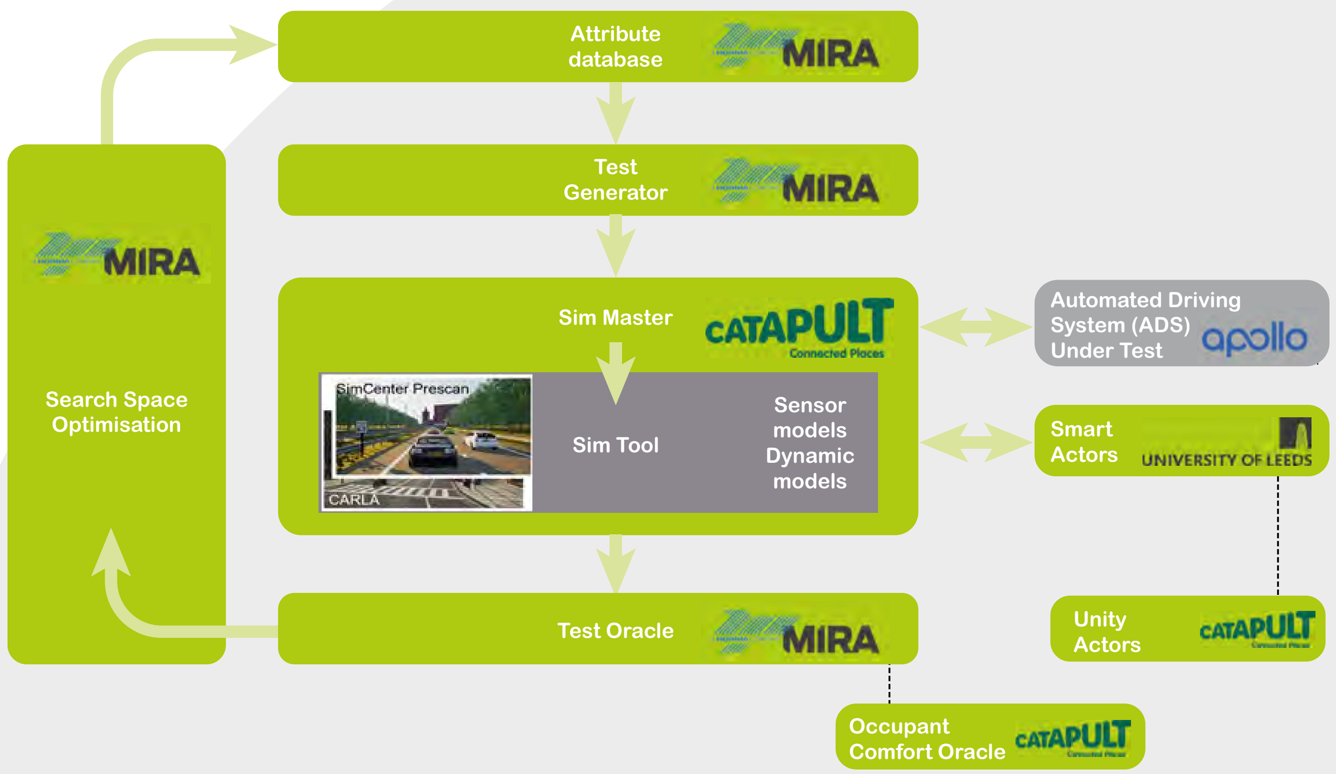
The VeriCAV system is made up of several sub-systems developed by the consortium partners

- **Test Generator** to produce scenarios to test,
- **Simulation Master** to interface a simulation tool to the rest of the system,
- **Test Oracle** to analyse test results,
- **Search Space Optimisation** to intelligently select the next set of test parameters

Additionally, the VeriCAV system architecture includes components that can be exchanged modularly:

- **Simulation Tool** in which the virtual scenario unfolds
- **Smart Actor Controller** to provide realistic actor behaviour
- **System under Test** i.e. the ADS

As the ADS navigates the scenario, it is rated by the oracles on its safety, progress though the scenario and the etiquette it demonstrates to other road users. This evaluation data can then be used to improve the driving performance of the ADS.



4. Test Generator

As Advanced Driver Assistance Systems and ADS features increase in complexity and scope, the relevant scenario space rapidly becomes too large for exhaustive test coverage to be a realistic objective; instead, a robust sampling mechanism is required to ensure an adequate number and distribution of scenarios are assessed. With a wide variety of different implementations and Operational Design Domain, each system-under-test will have unique failure modes and we therefore expect a fixed scenario library approach to be incomplete and inefficient.

VeriCAV has developed a generative approach to scenario definition. “Logical” scenario parameters, which encode the different variables to be tested and their ranges of potential values, are stored in a database such that they can be combined in many different permutations. Users configure the **Test Generator** to initialise the parameter space to their exact test requirements.

The VeriCAV system then samples specific “concrete” test cases from within these ranges in an ASAM OpenSCENARIO v1.0 format[4] for compatibility with the widest range of simulation tools. The samples are determined by the **Search Space Optimisation** function described later in this report.

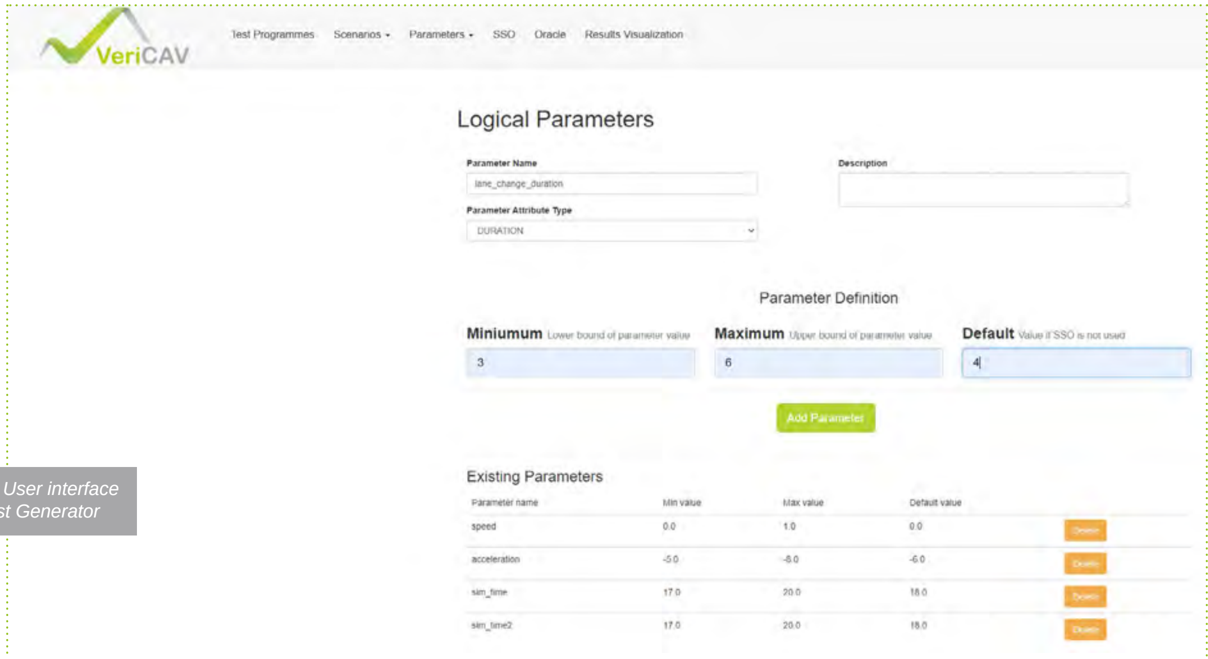


Figure 2 - User interface for the Test Generator

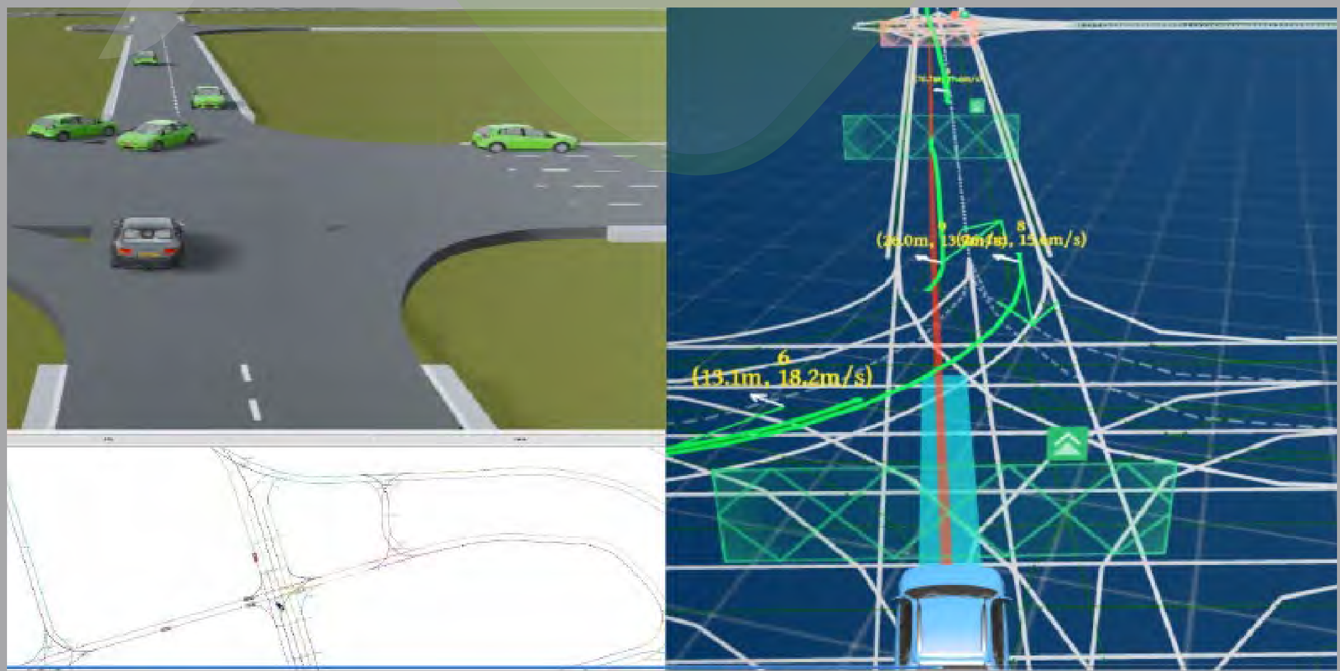
4 ASAM OpenSCENARIO being a standard for describing the dynamic content in scenario-based testing

A key aspect of any test scenario is the interactions between the system-under-test and other road users; therefore, manoeuvres can be defined with a variety of triggering conditions that initiate events dynamically within the scenario. These conditions are also parameterised using “logical” ranges for sampling purposes.

Road networks are stored in an ASAM OpenDRIVE format and are currently loaded from a user-defined library; however, the automated generation or manipulation of map files is a logical extension to this work based on the same sampling principles.

5. Smart Actors

VeriCAV also introduces Smart Actors. These actors address the major limitation in the ability to model multi-agent interactions at the smaller scale, where the detailed dynamics of phenomena like human situational awareness and inter-agent communication cannot be neglected. The outcomes of closed loop simulations will be strongly affected by the interaction between the ADS and the individual Smart Actors.



Scenarios running in Prescan (top left), Smart actors (bottom left), Apollo ADS internal world model for decision making (right)

Implementing Smart Actors required the selection of the appropriate modelling approaches, developing parameterised models and using an appropriate data set and calibration technique to calculate the model parameters.

Four different modelling approaches were chosen. Each have recently been used to explain vehicle interactions at the smaller scale but not originally designed for traffic simulation. Within VeriCAV, the quality of the traffic generated by the different approaches could be compared for the first time.

The models included a **Cognitive model** based on previous work performed at the University of Leeds on the European Commission funded interACT project, a **Machine Learning model** based on extensions of the state of the art found in literature, a **Choice Model**, based on advanced models developed in the Choice Modelling Centre at the University of Leeds, and finally **Game Theory** models based on enhancements of the state of the art available in the literature.

The scenarios were: **Pedestrian crossing, Pedestrian crossing the road at an unsignalised location; T-junction, Vehicle turning right across traffic, Roundabout** and **Motorway merge.**

Data Sources

The Pedestrian Crossing and T-junction data was collected in the interACT project[5]. The pedestrian road crossing data was collected both with real-world data collection as well as in controlled experiments at the University of Leeds HIKER simulator. The T-junction data was collected in the real world only[6].

For the Roundabout scenario, the data was taken from the rounD dataset[7]. Motorway merge data was taken from the I-80 portion of the Next Generation Simulation (NGSIM) project collected by the US Department of Transportation[8].

Additional data of pedestrians crossing while distracted was collected in the **HIKER simulator** as part of VeriCAV. 60 participants: 30 males and 30 females, ages 18-68, participated by crossing the road in Virtual Reality while facing multiple approaching vehicles with varying time gaps. In addition to a baseline scenario, the participants were asked to cross the road while performing a visual task on a phone, performing a verbal task while wearing an audio headset, and crossing under time pressure. The results of the experiment capture important changes in the gaps accepted by pedestrians in the various conditions as well as differences in crossing speed and safety margin.

5 https://www.interact-roadautomation.eu/
6 (Merat, N., Lee, Y. M., Markkula, G., Uttley, J., Camara, F., Fox, C., Dietrich, A., Weber, F. and Schieben, A. (2019). How Do We Study Pedestrian Interaction with Automated Vehicles? Preliminary Findings from the European interACT Project. Road Vehicle Automation 6, 21-33. doi: 10.1007/978-3-030-22933-7_3).
7 https://www.round-dataset.com/
8 https://www.fhwa.dot.gov/publications/research/operations/06137/.

Modelling Approaches

The **Cognitive models** are based on evidence accumulation models, also known as drift diffusion models,

"which models two-choice decision tasks as a biased random walk towards two opposing decision thresholds … and extends it in the first instance by allowing time-varying inputs [and] further extended … by suggesting that several of these decision units can be interconnected" [9]

The **Machine Learning model** is based on Long-Short Term Memory (LSTM), an instance of Recurrent Neural Networks. The LSTMs incorporate manoeuvre aware social pooling and take as input the prior trajectories and speeds of the ego vehicle and surrounding vehicles and predict the speed and trajectory of the ego vehicle including lane changes.

The **Choice Model** is based on a critical gap-acceptance model where drivers' decisions are based on the specific traffic conditions related to each available gap. For the roundabout case, significant factors included: ego vehicle speed, circulating vehicle speed, vehicles in the conflict zone, distance to the conflict zone, vehicle type, time-to-collision and vehicle entry lane.

The **game theoretic** work compared the performance of existing game-theoretic models. The model parameters used were based on those provided in the literature and testing was performed using the I-80 NGSIM road layout so that results could be compared to the real world.

Figure 4 - Highly Immersive Kinematic Experimental Research (HIKER) pedestrian lab at University of Leeds

9 (Dietrich, A. et al (2019) Preliminary description of psychological models on human-human interaction in traffic .

Evaluation

The evaluations for all the models were performed looking at vehicle through-put and conflict-crash analysis performed via the Surrogate Safety Assessment Model (SSAM)[10]. The number of produced conflicts and crashes is examined together with the severity of time-to-collision (TTC), post-encroachment time (PET) and deceleration rate (DR) measures. Specific results and models are presented in a separate report and in additional papers that have been submitted for publication.

VeriCAV Actor Interface

The actor interface used by the Smart Actors was delivered by another actor controller module that provided scripted actor behaviour. This module made use of an external OpenSCENARIO engine, ***Environment Simulator Minimalistic (esmini)***[11]***,*** to update actors. By implementing an external OpenSCENARIO engine the VeriCAV framework is still able to support simulation tools with less support for OpenSCENARIO. Actors controlled via this module included passenger cars, pedestrians, and cyclists - with additional actor types available via the interface but not implemented.

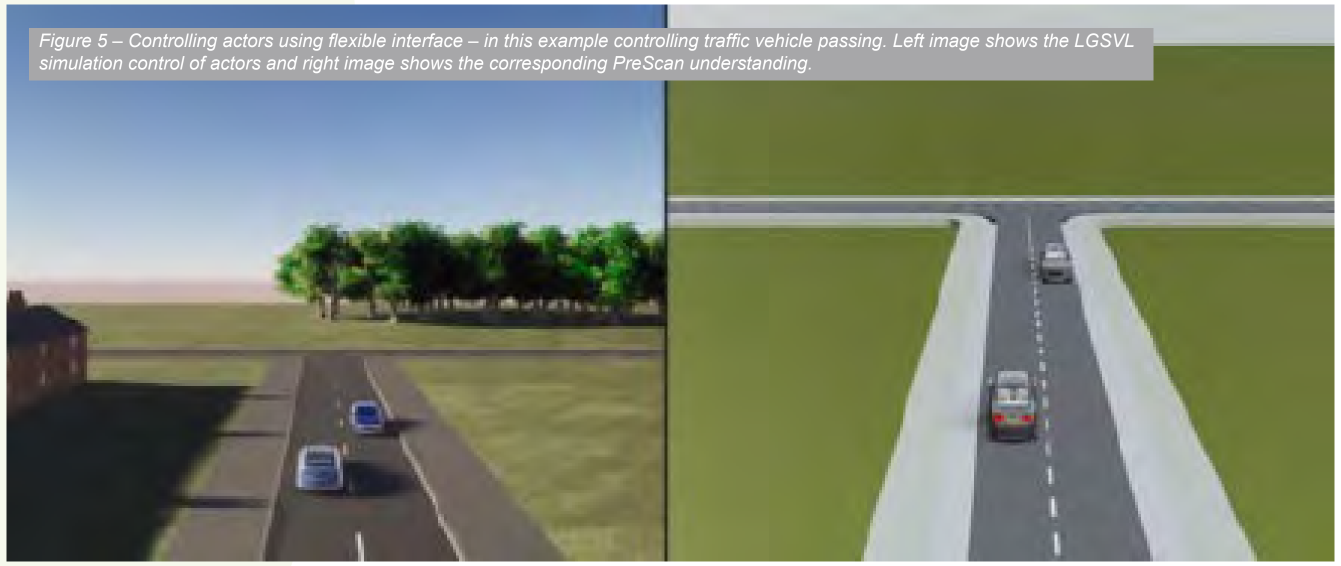


Figure 5 – Controlling actors using flexible interface – in this example controlling traffic vehicle passing. Left image shows the LGSVL simulation control of actors and right image shows the corresponding PreScan understanding.

To further demonstrate the flexibility of the actor interface, an additional source of actor control was implemented using the open-source simulation tool ***LGSVL***[12] with actor control built-in. This tool was integrated with the VeriCAV framework using the existing actor interface and allows different actor behaviour models to be used. With actor behaviour being a critical component of scenario-based testing, it has been demonstrated that the VeriCAV framework is well-suited to accommodate future developments in this area.

10 Pu, L., Joshi, R., & Energy, S. (2008). Surrogate Safety Assessment Model (SSAM)--software user manual.
11 https://github.com/esmini/esmini/
12 https://www.lgsvlsimulator.com/

6. Test Oracle

Developments in simulation and test automation allow large volumes of scenarios to be executed but the resulting test data must be analysed accurately and efficiently to assess system performance and inform overall test coverage.

Within VeriCAV we differentiate between metrics and oracles as follows. **Metrics** are functions and statistics computed directly from the simulation “ground truth” records. Example metrics include lane position, vehicle acceleration relative to neighbouring road users and time to collision (TTC). These are typically evaluated per time step of the simulation.

Test oracles are calculated from a set of metrics for a holistic assessment of different aspects of vehicle performance such as safety, progress or comfort. To make this assessment meaningful, test oracles should score relative to the scenario conditions, to determine whether the observed behaviours and performance were appropriate in that **context.** Within VeriCAV, each oracle is calculated globally as a single score per scenario to inform the **Search Space Optimisation** function.

A three-step design pattern for oracle computation is described within the project:

- **Scenario decomposition** – The main attributes and events that occurred are identified via the test records to ensure that the subsequent analysis is correctly attributed.
- **Sensitivity stage** – A first filter of the metrics identifies all occurrences where performance deviates from an ideal response.
- **Specificity stage** – A second filter further evaluates the data from the previous stages to rationalise, contextualise or otherwise explain away “justifiable” anomalies in the system performance, so as to focus attention on the remaining instances of highest concern.

4 ASAM OpenSCENARIO being a standard for describing the dynamic content in scenario-based testing

Validation of any test oracle remains an open and challenging topic, however two mitigations are proposed. Firstly, VeriCAV oracles assess performance in relative terms to prioritise rather than replace manual analysis, the learning from which can be used iteratively to further refine the metrics and oracles. Secondly, a set of project scenarios (including highway and roundabout test cases) were physically executed on the HORIBA MIRA proving ground to provide a second source of training and validation data for the oracles.

Occupant Comfort Oracle Metric with Virtual Reality Assessment

One VeriCAV oracle metric predicted **Vehicle Occupant comfort** using only the simulated scenarios as input. This metric focused on **five primary factors:**

1. Acceleration (longitudinal and lateral)
2. Jerk (longitudinal and lateral)
3. Headway to vehicles ahead (in the same lane)
4. Proximity to vehicles in the adjacent lane (both oncoming and overtaking)
5. Gaps between vehicles when joining a new carriageway (T-Junction and roundabouts)

Each factor was evaluated to determine the thresholds for being comfortable, feeling normal, and uncomfortable. The algorithm scored the ADS based on exposure time to the threshold categories, applying a negative penalty that increased with discomfort.

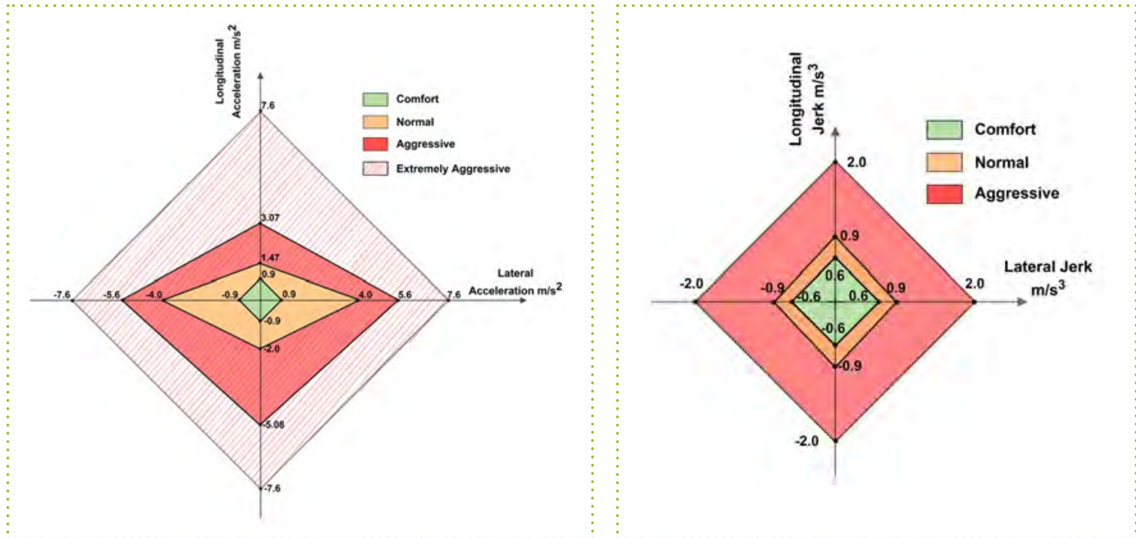


Figure 6 Examples of the thresholds showing the boundaries defined for Acceleration and Jerk (longitudinal and lateral)

Remote user trials were delivered remotely using Oculus Quest 2 VR headsets. Multiple driving scenarios were created for these user trials:

1. Motorway (joining and exiting)
2. Roundabout (with and without vehicles)
3. Major road (passing oncoming vehicle and slowing down for slower vehicle ahead)
4. T-Junction (with and without vehicles)

These scenarios involved creation of road geometry (from OpenDRIVE), adding 3D models (road markings, buildings, trees, signs), as well as vehicles movements (using OpenSCENARIO). The ADS was then tasked with navigating through the scenario. Each scenario was translated into a 360-degree video with a driver seated perspective. Indicator lights, plus audio for road and drivetrain noise, were added to improve the feeling of immersion.

Trial participants were exposed to the scenarios and pressed a button when they felt uncomfortable. After each scenario, participants answered a questionnaire to identify which of the comfort factors had contributed to their overall perceived level of comfort.

The user trials successfully showed that when the scenario behaviour exceeds the algorithm's 'Normal' thresholds, i.e. moving into 'Aggressive', participants were pressing the trigger button to indicate being uncomfortable.

Virtual reality offers researchers a consistent baseline experience and the ability to record human physiology and responses in a cost-effective and efficient way. This innovative user trial method offers significant potential for use in future automated vehicle user trials.

Figure 8 - Driver viewpoint 360 degree scenario used in the Virtual reality assessment

7. Search Space Optimisation

VeriCAV seeks to make efficient use of simulation effort by directing testing to the regions most relevant to the specific system-under-test. Two objectives are drawn from this: firstly, tests must be distributed, with a least some minimum density, to mitigate the risk posed by sparse coverage across the whole parameter search space and to ensure no "corners" are left unexplored; secondly, further test evidence must be accumulated to support a statistical conclusion to the test activity.

These objectives represent significant technical challenges given the high dimensionality of the search space for even simple driving scenarios, the potential for non-linear, transitional, responses to small changes in initial conditions and the complex interdependencies and correlations which may exist between parameters.

The **Search Space Optimisation** capability developed within VeriCAV has four distinct phases of operation, each of which feeds back to the test generator for simulation execution:

- **Initial sampling** – A preliminary set of tests is generated to achieve a well distributed minimum level of coverage
- **Confidence mapping** – The data from the initial test scenarios are used to characterise a surrogate model which predicts the system performance across all points in the search space. An associated confidence distribution directs further testing to regions of uncertainty in the model. The outcomes of these subsequent tests are used to optimise the surrogate, hence increasing confidence in its predictions.
- **Exploitation sampling** – Once the surrogate model has been refined, attention focuses on boundary regions around performance thresholds to ensure that system failures or degradation are well understood, and regions of deterioration are sampled with increased resolution.
- **Conclusion** – The data and models generated throughout the previous steps provide statistical evidence of overall system performance and residual risk.

8. Modular Architecture

A key objective for the VeriCAV project was to develop a robust, modular, and flexible testing framework. There was therefore a specific focus on the interfaces between software modules developed by the consortium partners to ensure they met the VeriCAV requirements but also considered future use. Interfaces were identified in the functional architecture of the framework and working groups were setup to manage each interface.

Flexible to the Simulation Tool used

The VeriCAV framework was designed to be agnostic to the simulation tool used at its core. To achieve this, a simulation wrapper concept was developed which provided an interface from a given simulation tool to the rest of the VeriCAV framework. Simulation wrappers were implemented for ***Siemens SimCenter Prescan*** and the ***CARLA*** simulation tool.

Figure 9 - CARLA (left) and PreScan (right) running the same scenarios generated by the VeriCAV framework.

As with the simulation tool, it is also beneficial to be able to change elements of the framework as required and so all the sub-systems within the VeriCAV framework are designed to be modular. This also allowed development teams, distributed across the consortium partners and working remotely, to work without requiring the full VeriCAV framework to test with.

Using and Collaborating with Industry Standards

The challenges of virtual verification of Automated Vehicles are significant and will require collaboration to solve. In this regard, the partners in the VeriCAV consortium have engaged with **ASAM**[13], who manage the development of the OpenSCENARIO and Open Simulation Interface (OSI) standards; the **MUSICC**[14] project, an open database of scenarios for scenario-based testing of AVs, and other CAV simulation projects funded by the Centre for Connected and Autonomous Vehicles (CCAV).

13 https://www.asam.net/
14 https://cp.catapult.org.uk/project/multi-user-scenario-catalogue-for-connected-and-autonomous-vehicles/

9. VeriCAV Engagement

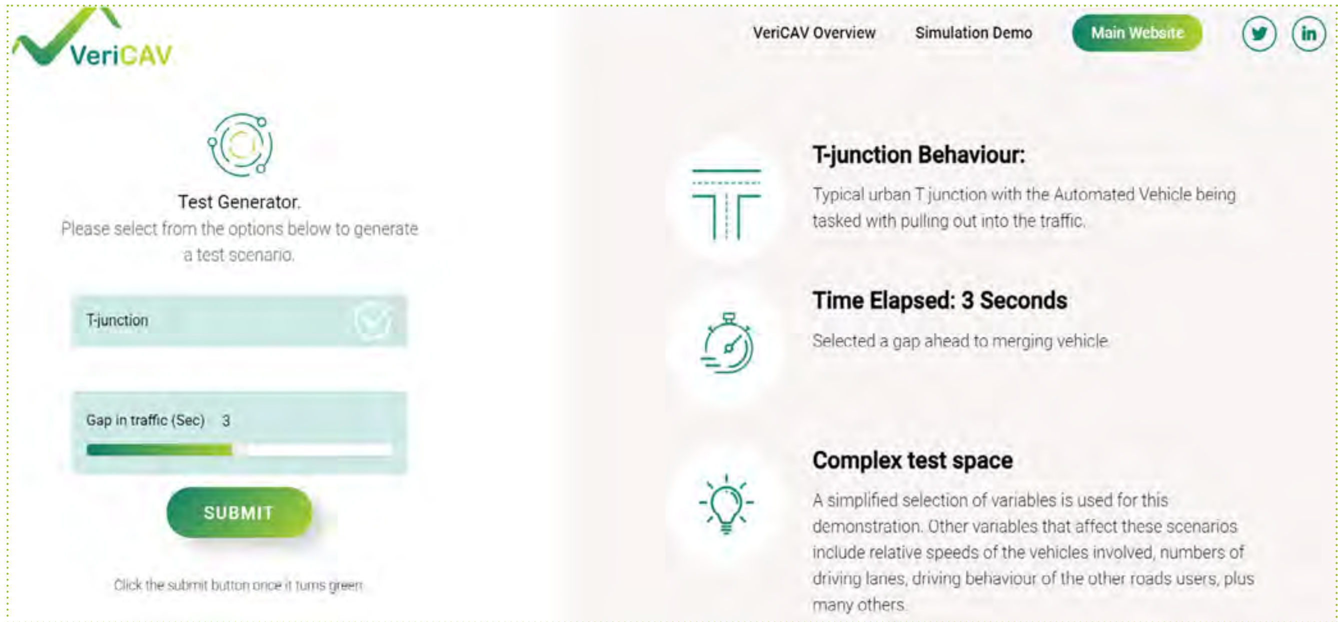
Throughout the VeriCAV project the consortium have presented at multiple industry and academic events - both physically and virtually. Including:

Centre for Connected and Autonomous vehicles (CCAV) Summer CAV cohort event, 2019
IMechE - Automated and Autonomous Vehicles: Overcoming engineering Challenges for Future Mobility, Horiba Mira, 2019
Cenex Low Carbon Vehicle (LCV) conference (both 2019 & 2020)
Zenzic CAM Roadmap launch, 2019
ITU Digital, Budapest, 2019
Auto AI Berlin, 2019
CAV@CSC – Culham Science Centre, 2019
European Transport Conference, 2019
Australian CAV Summit, 2019
RoboSoft: Software Engineering for Robotics, Royal Academy of Engineering, 2019
Advanced Engineering, Birmingham, 2019
Zenzic hosted VIP ITS/CAM Delegation from Japan visiting Connected Places Catapult, 2019
L3Pilot Summer School for Automation Driving Implementation and Testing, 2020
SafeComp Conference on Computer Safety, Reliability and Security, 2020[15]
6th Symposium Driving Simulation (SDS) , 2020
Centre for Connected and Autonomous vehicles (CCAV), CAV Sim cohort, 2020
Open CAV Cohort Jan 2021
VeriCAV Final Project Showcase, March 2021[16].

15 Test Framework and Key Challenges for Virtual Verification of automated vehicles: the VeriCAV project https://hal.archives-ouvertes.fr/hal-02931723
16 https://cp.catapult.org.uk/event/vericav-final-showcase/

To promote understanding of the project, an ***interactive section of the VeriCAV website*** was established as an accessible introduction to the key functions of VeriCAV.

Figure 10 - Interactive website - allowing visitors to generate and visualise small-scale example scenarios



In March 2021, VeriCAV concluded with a successful **Virtual showcase** to publicly share the research and innovation delivered by the consortium.

The consortium have published a number of reports accessible from the **Downloads section of the VeriCAV website:** ***https://vericav-project.co.uk/*** - with more papers planned to be published throughout 2021.

10. Moving Beyond VeriCAV

VeriCAV has achieved its goals of developing a framework for testing of Automated Vehicles in simulation and has laid the foundations for further innovations in this complex and important topic.

The partners are keen to discuss the methodology and results with interested parties to identity potential collaboration opportunities and build upon the capabilities developed as part of VeriCAV.

The consortium partners are continuing to actively work in this area. Any enquiries regarding VeriCAV or future activities should be addressed to: **vericav-project@cp.catapult.org.uk**

Find out more by visiting our website and visiting our Twitter & Linkedin pages:

www.vericav-project.co.uk

@VeriCAVProject

Linkedin.com/company/vericav-project

Acknowledgements:

VeriCAV was only achieved with the hard work and innovation from all those involved from within the consortium partners: HORIBA MIRA, Connected Places Catapult, University of Leeds, and Aimsun.

Thanks for the support from Innovate UK and Centre for Connected and Autonomous Vehicles.

VeriCAV has proved test concepts using Siemens SimCenter PreScan. Thanks to the Siemens PreScan team and Advanced SimTech for technical support during this project.

Thanks also to the CARLA open-source simulator community and the esmini team.

VeriCAV

VERIFICATION OF CONNECTED AND AUTONOMOUS VEHICLES

A HORIBA COMPANY MIRA

CATAPULT Connected Places

aimsun.

UNIVERSITY OF LEEDS

Centre for Connected & Autonomous Vehicles

UKRI Innovate UK