PEER REVIEW PAPER

This paper has been peer-reviewed and published in a special edition of Traffic Injury Prevention 24(S1), by Taylor & Francis Group. The complete paper will be available on the Traffic Injury Prevention website soon. To access ESV Peer-reviewed papers click the link below
https://www.tandfonline.com/toc/gcpi20/24/sup1?nav=tocList

A PRAGMATIC APPROACH TO SAFE OPERATION FOR DRIVERLESS SHUTTLES DURING DEVELOPMENT

Matthias Strauß
Christopher Pinke
Continental Autonomous Mobility Germany GmbH
Germany

Danilo da Costa Ribeiro
Wolfgang Schramm
Continental Autonomous Technologies GmbH
Germany

Paper Number 23-0072

ABSTRACT

Driverless shuttles are a modern vehicle platform designed to operate autonomously, constituting a very promising building block of future mobility solutions. At Continental, there is a long experience with these types of vehicles. In this paper, some of this experience regarding operating such vehicles during development is shared. In particular, the focus is on how a safe operation can always be ensured. To this end, a release process for such an operation, how a pragmatic safety assessment can be done and some of the peculiarities of driverless shuttles are presented.

INTRODUCTION

With recent technological advances, fully autonomous vehicles become more and more realistic. Driverless shuttles, aiming at SAE L4 autonomy, are expected to be one of the major pillars of future mobility and an ideal entry point for fully autonomous mobility solutions.
The absence of a human driver in the chain-of-effects of L4 systems changes the demands on the actual vehicle architecture and layout. Driverless shuttles, see Figure 1, are consequently designed for driverless operation without traditional instrumentation like brake pedal or steering wheel. Comparable in capacity to a small bus, the shuttles are normally capable to drive at lower speeds of up to 20 kph and are operated in private or sub-urban areas, providing first and last mile services. Notwithstanding this restricted operational design domain (ODD), they complement existing mobility solutions like taxis, busses, or trains, especially when operated in an on-demand service. Hence, shuttle systems are an ideal candidate to introduce driverless systems to a wider usage. For more information, please refer to [1], [2], or [3].

Figure 1. A driverless shuttle, developed during the CUbE activities at Continental.

Currently, driverless shuttles are operated predominantly in pilot applications. Direct feedback is retrieved from users of the service at this early stage. For recent examples see [4], [5], or [6]. In parallel, the technology matures as more and more field experience is collected. Continental has a long and successful history connected to driverless shuttles. Within the Continental Urban Mobility Experience (CUbE) activities such systems are developed and operated in Germany, USA, China, and Japan. This includes operations on public roads in mixed traffic and even pedestrian zones. Here, a deep understanding of the specifics of driverless shuttle systems has been developed, especially regarding their safety aspects and how such systems can be released for an intended use. See e.g. [7], [8], [9] as well as [10] for more information.
While established manufacturers like EasyMile [11] and Navya [12] already have a certain market penetration, recently more and more companies have focused on driverless shuttles, e.g. Zoox [13], Cruise [14] or ZF [15]. However, compared to traditional automotive business, existing shuttle systems are typically not automotive grade yet and are built on smaller scales. In addition, most deployments still require a safety operator to be on board as the technology itself is not yet capable of operating without human supervision, although this is obviously the major focus of development of shuttle manufacturers.
Due to the non-traditional design of the shuttles, also the development of such vehicles varies in some aspect from traditional vehicles. Any driverless system must fulfill the highest safety standards as a human is not available anymore as a fallback level. Established safety analysis methods like HARA [16] and FMEA [17] are often not sufficient for analyzing such complex systems.
Addressing this gap, we would like to share our approach and our insights in the field of driverless shuttles to benefit the development of safe and reliable driverless shuttle solutions. In this paper, we present a pragmatic approach regarding the safety argumentation of prototypes for driverless shuttles. The intent is to provide a general viewpoint on this topic which can serve as a guideline for the safe development and operation of such systems.
To this end, we present a generic shuttle architecture and investigate which effects the failures of the various system components can have, up to a level that is relevant for a development prototype. We expand this generic view by providing insights into a concrete prototype implementation in one of our CUbE vehicles, covering brake, battery, and steering components as well as the full chain-of-effect of autonomous driving. We elucidate measures to fulfill safety requirements based of the safety analysis.
This analysis can serve as a predecessor for more sophisticated and established methods mentioned above while enabling the operation of the prototype in a safe manner.

A RELEASE PROCESS FOR DRIVERLESS VEHICLES DURING DEVELOPMENT

Central part of operating a prototype vehicle is a defined release process. This process can vary depending on the specific development of interest. For the development of driverless shuttles, or more general highly automated vehicles, the process must cover the assessment and acceptance of all hazards posed by the system of interest, especially when trialing and operating in a public environment.
In recent years, driverless vehicle trials have been based solely on exception permits. Only recently, the worldwide first L4 law has been passed in Germany, see [18], [19], or [20], providing a solid legal basis for the development and homologation of L4 systems. However, for the application of said law in actual development activities, details still must be clarified, see [21].
The major steps of such a release process are depicted in Figure 2, following a corresponding Continental internal prototype process. It involves the clear definition of the system of interest and its intended usage, the thorough analysis of said usage and finally the approval of the intended usage. All of these steps are based on various assessment results. This process was applied during the CUbE activities and will be the frame of the content presented in the subsequent chapters.

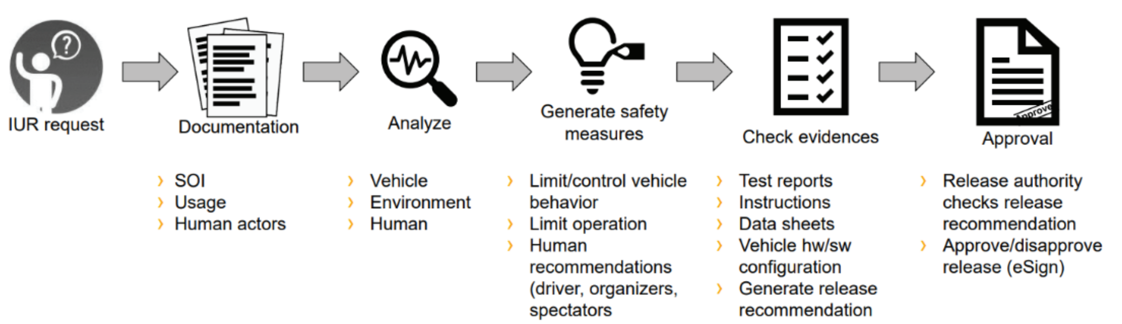


Figure 2. Exemplary Process for an intended usage release (IUR), showing the major steps towards a safe operation of the prototype vehicle.

ANALYSIS OF THE INTENDED OPERATION

An important building block to release a development system is the description of the system of interest and the intended operation. For the latter, the following topics needs to be analyzed:

- System environment:
  - What is the operational environment or domain in which the system is being used?
  - What are the actors involved and what are their respective roles?
- System context:
  - With which people or systems does the system interact or exchange information?
- System behavior
  - What is the intended system behavior…
    - …within the environment and with respect to the other traffic participants?
    - …with respect to user interaction?

In the following, the analysis of the operational environment is limited to the intended route to keep the problem space manageable. Using modern satellite or aerial pictures, digital maps, on-site inspections, and similar data sources, valuable information about the road topology, expectable traffic participants and constraints to the operation are gathered. This information allows to split the route into smaller sections, like connecting roads and intersections or other areas of interest This creates a sort of scenario catalogue of the intended operation. An example is given in Figure 3.

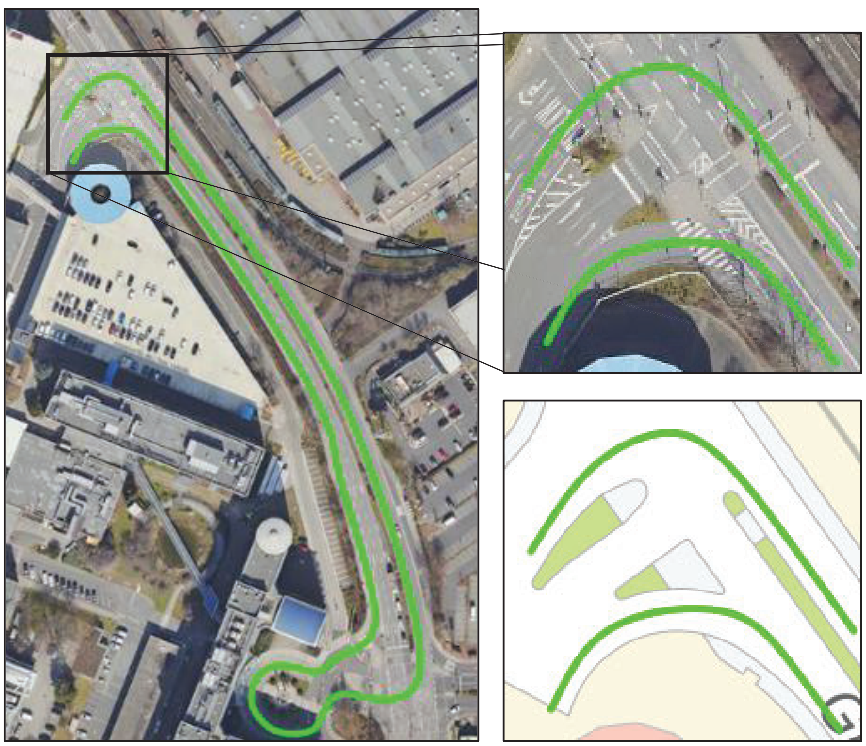


Figure 3. Aerial picture of a route (green), including route detail and extracted relevant scenario. Material taken from https://geoportal.frankfurt.de/.

Scenes and situations outside the route are not considered at this stage and are not part of the immediate challenge at hand.

Complexity within the operational environment and the shuttle's system functionality can be increased over time as the development of the shuttle system continues. This includes certain limitations the system has at a certain point in time. To challenge these limitations, it is then up to the development team to align on the appropriate measures. This is done together with the security and safety experts and risk owners. Measures can be manifold, e.g., performance limitations, additional supportive helping systems, external systems, or even organizational measures.

The work results from the Pegasus Project [22] are respected and considered for analysis of the operational environment and systematic identification of scenes and scenarios. One can describe the operational environment and the ODD in close alignment to the ASAM OpenX Standards [23], especially OpenDrive and

OpenScenario as well as ISO 34502 [24]. This can be done via the proposed six-layer model as described in the Pegasus Method [22]:

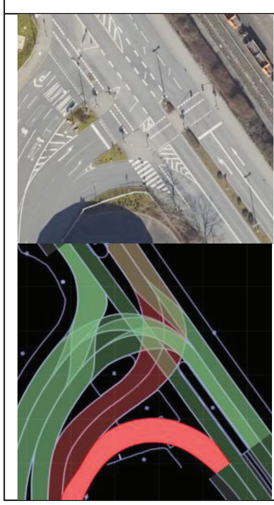
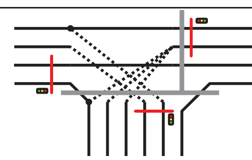
- Layer 1: Street Layer
- Layer 2: Traffic Infrastructure
- Layer 3: Temporal Modifications
- Layer 4: Moveable Objects
- Layer 5: Environment Condition
- Layer 6: Data and communication

and by abstracting the level of detail for scenario descriptions into functional, abstract, logical, and concrete scenarios.

Structuring the environment analysis into these layers helps to focus on certain topics in the team discussions with the safety engineers and the function experts. To simplify the problem space further, the route where the shuttle will be operated is sliced into smaller segments. This allows systematical identification of the road topology (Layer 1), whether the system needs to support perception of traffic infrastructure, like Traffic Signs or Traffic Lights, (Layer 2), or if there are temporal modifications like construction sites at the time of the intended operation (Layer 3). The segments are analyzed and provided to the team in form of a story board table with some easy understandable annotations.

An exemplary simplified description of one road-segment of the route storyboard from Figure 3 is shown in Table 1. With this information, the function designers, security, and safety engineers can identify hazards and hazardous situations according to the intended use. From these results, potential safety measures are derived to ensure a safe operation. These might include instructions for the safety operator, limiting the system performance, providing external support for the safety operator, controlling the environment, or others. This is described in more detail below.

Table 1.
Route-Storyboard element example

Scene and Layout	Maneuvers, Activities and Abilities	Topology and Objects	Traffic Participants
	Follow Lane: - Straight - Curve Stop at crosswalk Stop at stop line Follow preceding vehicle Perceive traffic light (and status) Detect TPOs (other vehicles and VRUs) TPO prediction	 Traffic sign Traffic light Crosswalk	Cars Buses Trucks Motorcycles Cyclists Pedestrians

SAFETY MANAGEMENT SYSTEM

The complexity and novelty constituted by highly automated vehicles requires more than just the functional safety analysis well established in the automotive industry. With the goal to ensure a high level of safety, not only at the development phase, but also during the vehicle's entire lifecycle a Safety Management System (SMS) was established, following [25]. Figure 4 shows an overview of the SMS in operation.

As can be seen in Figure 4, the SMS was designed to cover all the phases of the Safety Assessment activities from the development phase (starting after the system definition until the entry-into-service – EIS) until the product's end of life. According to the SMS Framework proposed by International Civil Aviation Organization (ICAO) and used widely at the aviation industry [25], the SMS has four main pillars/components: Safety Policy, Safety Risk Management, Safety Assurance and Safety Promotion. The scope of this paper is the development activity, not the organization. Therefore, special attention will be given to Safety Risk Management and Safety Assurance only and these two components are described below in more detail.

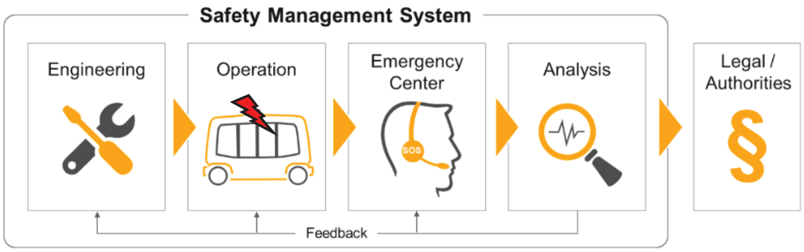


Figure 4. Overview of Safety Management System

Safety Risk Management

This SMS component's activities take place before the entry-into-service and are responsible to ensure that the operation is safe enough to start. When the operational concept is ready, the safety analysis has the required inputs and can be carried out. It includes a top-down and a bottom-up approach.

The top-down approach is a hazard-centric analysis based on the System-Theoretic Process Analysis (STPA) [26]. This, after assessing the operation, defines the accidents (or losses, as denoted in the context of STPA) and hazards present and search for causal factors. Those could not only be triggered by technical reasons, but also from human/systems interface and specification/performance insufficiencies. Figure 5 provides a more detailed overview of the Safety causality chain used in the safety assessment.

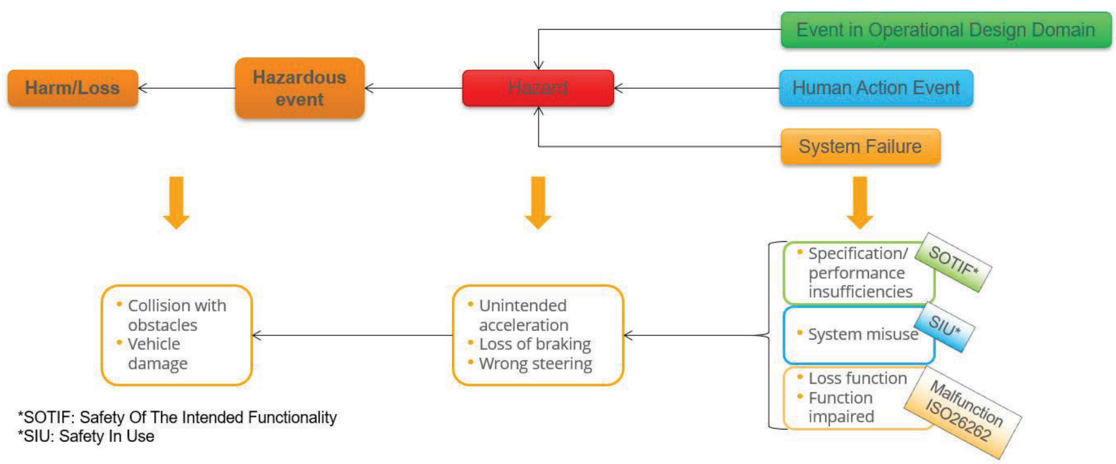


Figure 5. Safety causality chain used during safety assessment.

For the scope of this paper and the generic architecture described below, a simplified bottom-up approach called "Component Failure and Effect Analysis" (CFEA) is presented. It is based on the Failure Modes and Effects Analysis (FMEA) method [17] and, as the name says, has the goal to identify the effects of some critical component failures. In this pragmatic way, however, various insights about driverless systems can already be collected. Especially, it can be applied nicely to prototype systems such as the one presented below.

After defining the causal factors for every hazard, a set of safety requirements is defined, implemented, and verified.

Safety Assurance

Considering the operation of highly automated vehicles contains a high number of unknown (and possibly unsafe) scenarios, see ISO 21448 [27], a special focus was given to the safety assurance part of the SMS. Support is provided to the people involved in the operation facing an unexpected situation, but also to learn from mistakes.

This was achieved through the establishment of an Emergency Response Plan. It created a communication channel from the people involved in the operation to an emergency team that could support and give instructions anytime. It was not limited to operators and vehicle occupants, but also applying to vulnerable road users. Additionally, any safety-critical event would be registered in a database, enabling the engineering team to mitigate any issue that caused this event. The safety assessment can then be updated, and the previously unknown hazard is now mitigated.

GENERIC DRIVERLESS SHUTTLE ARCHITECTURE

Having outlined the steps towards the operation of driverless shuttle prototypes, the basic vehicle architecture of driverless shuttles will be presented in a generic way in the following. The focus will be on the elements related to driving only, excluding higher-level functionality like environment perception or human machine interaction (HMI). The architecture consists of several key elements which are independent of the specific type and model. This is depicted in Figure 6.

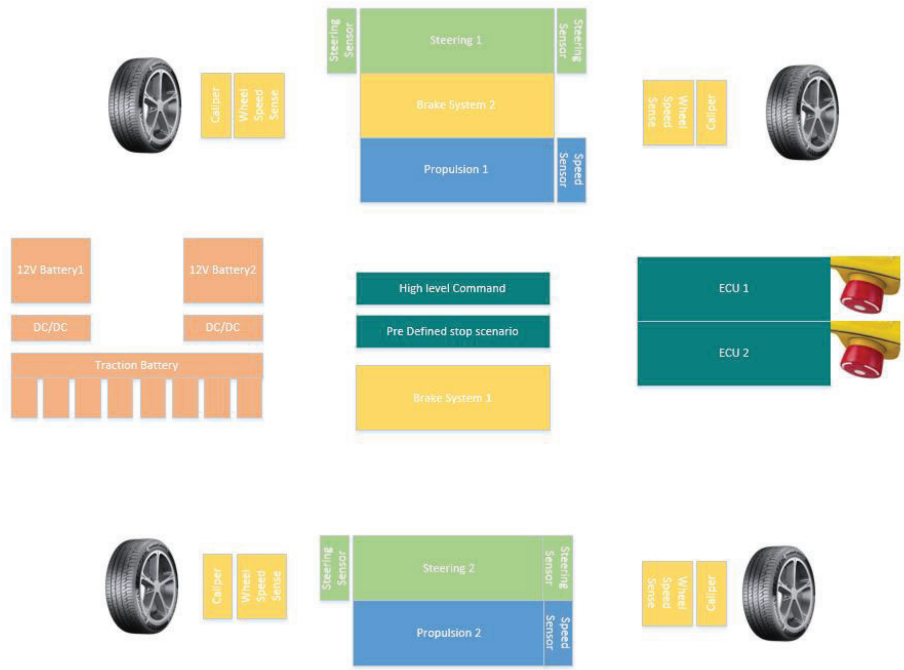


Figure 6. Generic Architecture of driverless shuttles

Obviously, most of the elements are the same as in traditional automobiles. However, without a human in the loop, there are different requirements regarding fallback levels, e.g., the redundant layout of the ECUs. Also, the supervision of the proper functionality of each element must be automatized as well. In addition, certain components must be redundant, e.g., the brake system. A somewhat special case are the emergency buttons installed inside the vehicle. They allow the passengers to trigger a Minimum Risk Maneuver in case of an emergency that the vehicle did not recognize (a so-called Emergency Stop or E-Stop).

In general, each element in the architecture can fail because of four distinct reasons: Loss of power supply, loss of communication, internal failure (mechanical or electrical), as well as a wrong action or command send to the element. Any failure of one element can induce other failures or complications in the system. Hence, the driverless system must be realized such that in each failure case for any key element, it is ensured that the vehicle reaches a safe state. For currently available commercial systems and in particular for prototype systems, it is very likely that the system cannot recover from that failure. In that case, an according mitigation must be in place. Until now a common solution is to install supervision by a human safety operator.

But even in case of human supervision, not all failures can be intercepted by the safety operator, which should be illustrated briefly in the following. In case the safety operator is standing, all failures resulting in high jerks and high longitudinal or lateral accelerations can lead to situations in which the operator may not be able to reach the emergency button to stop the vehicle. In addition, in a well-functioning system, it could be expected that the safety operator is somewhat distracted, leading to additionally increased reaction time.

Hence, for certain failures like described above, the vehicle must reach a safe state automatically and corresponding actions must be initiated. These depend on the criticality of the failure. For driverless shuttles, the most common reaction to a failure is the E-Stop. Depending on the vehicle speed at the failure event, one can roughly distinguish three possible E-stop scenarios:

- No steering torque and decelerate to standstill with variable braking deceleration (speeds < 15kph)
- Stop in ego lane. Following a pre-defined steering angle and deceleration profile (<70 kph in inner cities)
- Stop on emergency lane (higher speeds and on highways or rural roads).

Within the scope of the FCEA, we are looking at the worst-case scenario that can result of each component failure. In addition, also on this high-level, the severity and probability of the induced hazard can be evaluated. This should be illustrated in the explicit analysis of selected components for a specific prototype in the next chapter.

APPLICATION TO A PROTOTYPE SYSTEM

In the following, some details will be provided about one specific shuttle prototype developed at Continental, the so-called CUbE-3 vehicle (see Figure 7). The intended usage of the vehicle was to operate it on routes like the one shown in Figure 3. Accordingly, the aforementioned steps towards a release of the vehicle will be briefly summarized, with a focus on the safety aspects, including the FCEA.

Figure 7. The CUbE-3 prototype vehicle.

After analyzing the route and creating the according scenario catalogue, the top-down safety analysis comprised of the analysis of the Safety-in-Use, i.e., analysis of all expected human-machine interactions, and the ODD. For the Safety-in-Use analysis, every human actor with potential contact to the operation was listed. These were, among others, safety operator and passengers. Using the STPA method, these actors were included in the vehicle's control structure and their interfaces were assessed. In this way, every possible unsafe control action was identified, and safety requirements were generated to mitigate accidents. For example, the safety operator needs a certain time to detect a hazard, make a decision, and execute it. This reaction time was used to calculate the minimal distances to objects, such that the operator can bring the vehicle to a safe state at any time. The vehicle was required to operate only within these limits. Another example considers a situation in which the vehicle actuation receives ambiguous commands or commands beyond a defined safety limit. To ensure that the safety operator has the full authority over the vehicle, a hierarchy was created in which the safety operator's commands have highest priority and overrule other commands. Also, it was ensured that the safety operator is the only one that has access to the vehicle controls next to the high-level software during operation. Corresponding documents were created with instructions for anyone interfacing the operation. These must be handed out to and respected by the involved people.

After that, the ODD was analyzed and any hazard that can occur during the vehicle operation was identified and mitigated. This included, for example, areas subject to parked vehicles and people crossing the road. The parked vehicles may obstruct the vehicle`s sensors as well as the safety operator`s view and may lead to pedestrians being detected very late while they are crossing the road. To mitigate this, the vehicle speed was reduced in these areas and the safety operator was instructed to pay special attention when driving through these areas. Similar measures were applied to areas in which a higher concentration of pedestrians close to the road or crossing it is expected.

After the safety assessments mentioned above, the vehicle itself was in focus. The vehicle architecture is depicted in Figure 8. The basis for the vehicle was an EasyMile EZ10 Gen2 vehicle. The EZ10 is an autonomous minibus, which has a capacity for ten people and can reach a speed of up to 20kph. The vehicle was heavily modified, keeping only lights, air conditioning, doors, air spring system and safety lasers from the original equipment. All other components, including battery, propulsion, steering, brake system, and high-level sensors were replaced. After the modifications of the vehicle (see Figure 8), it was able to drive up to 30kph. With respect to the generic driverless shuttle vehicle architecture shown above (see Figure 6), the component "ECU1" was realized using a rapid prototyping ECU manufactured by Continental. This will be referred to

simply as “ECU” in the following. For the intended usage of the vehicle, it was acceptable not to be able to continue driving in case of a failure of ECU1 and the component “ECU2” was not realized. In case of failure of ECU1, the brake system would directly trigger an emergency stop. Running mainly the high-level software, the component “High Level Commands” was realized by a PC. The component “pre-defined stop scenario” was a stop in ego lane as described above.
In the following, a brief analysis of selected components is provided according to the FCEA mentioned above, including the mitigations for identified hazards.

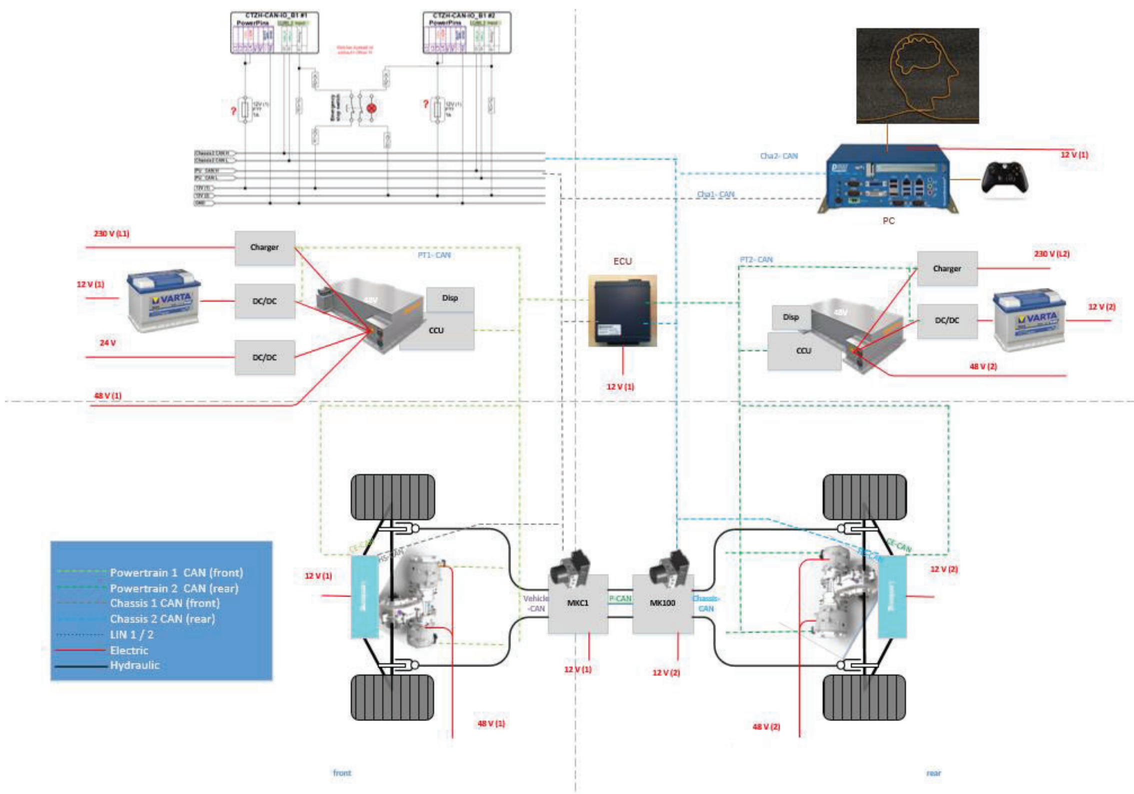


Figure 8. E/E architecture of the CUbE-3 prototype vehicle.

Main Brake System
The main brake system is the MKC1 of Continental, having three main roles: Deceleration of the vehicle, providing anti-lock braking (ABS) or traction control interfaces (TCS) as well as supervision of vehicle commands send to the brake system. Regarding the latter, in case there are any failures in timing, checksum, or the values itself, the brake system can degrade the shuttle system and bring the vehicle to a stop automatically. Using a hydraulic brake system brings the vehicle to a safe stop and standstill, which cannot be guaranteed just by braking with electric motors. For electrical engines, in comparison, there is a chance that the vehicle even gets faster or starts driving backwards if there is a failure in the fallback control loop.

Secondary Brake System
The secondary brake system is the Continental MK100, and it is the fallback system for the main brake system. Assuming that the fallback level is not active very often, not the whole functional range is needed. For example, TCS could be disabled in a degraded mode. But basic functions like ABS or Electronic Brakeforce Distribution (EBD) are still needed, as otherwise the vehicle could get instable during emergency maneuvers with high decelerations. Also, regular self-tests are required for the fallback level as otherwise there is the risk of a sleeping failure.
Another requirement for the fallback brake system is that it should be able to deliver higher decelerations than in normal driving mode. This results from the detection time of the failure. If a deceleration is requested and the first brake system fails, the shuttle will still be moving, and consequently further decrease the distance to the object it is braking for. With that also the required deceleration to stop in front of the object increases. For example, if the maximum deceleration in the driving mode is 3.5m/s² the fallback level should at least reach 5m/s². In addition, the maximum jerk and pressure build up time play a big role in such a scenario.

Steering

In the CUbE-3 vehicle, two steering systems were integrated, one on the front and one on the rear axle. Both are designed the same way, such that they can reach the same steering angle. In that way, the vehicle is already equipped with a redundant steering system. Nevertheless, there are failure modes which are different to simple "loss of steering": Self-steering and misalignment. The latter typically has a mechanical reason, which can be regarded as quite improbable, as steering systems are designed in a very safe way. And in addition, the misalignment is assumed to be only very small in that case.
On the other hand, self-steering is a very dangerous failure case. The available driving path is in most cases very narrow and most of the time there is oncoming traffic to be expected, e.g., in inner cities. Hence, already a deviation of 1,5m to the target track can lead to severe incidents. With an assumed lateral acceleration of 6,5 m/s^2, 1,5m are reached within 0,7s. This is clearly below the reaction time of a safety operator. Anyway, even an immediately started emergency brake maneuver could not prevent a crash in this situation.
Therefore, two safety features were implemented, one regarding the self-steering and one regarding the maximum steering gradient. The self-steering detection feature checks whether the steering sensors deliver valid data and whether the actual steering angle is close to the requested one. On the other hand, the steering angle gradient is limited, with the limitation also being supervised. Both supervisions are capable of directly triggering a full stop of the vehicle. With that, the reaction time of the driver is taken out of the loop.

48 Volt Power Supply

The CUbE-3 vehicle has been equipped with two 48V power supplies. In that way that the vehicle is capable to apply engine torque even if one battery fails. As the intended vehicle speeds are below 30 kph and it was operated only on normal inner-city roads, a safe stop in case one battery fails was an acceptable safety requirement. As it is not a critical failure if one 48V power supply fails, a moderate deceleration was chosen to stop the vehicle.

Propulsion

The propulsion system of the CUbE-3 vehicle consists of two engine packages. Every engine package consists of two Belt Driven Starter Generators (BSGs) with 7kW each, as well as a differential. The maximum acceleration of the vehicle, however, was limited to 3,5 m/s² to protect standing passengers. Every propulsion system was connected to its own 48V Battery system, which in turn allowed the continuation of the mission if one 48V system fails, as described above.

DC-to-DC converters

The vehicle was equipped with automotive-grade DCDC converters. These are controllable via CAN and can transform energy from 12V to 48V and from 48V to 12V. They have an own supervision and shut off the DCDC functionality if a failure was detected on one of the board nets, while still communicated this error on CAN. In that way, a redundant information on the state of the 48V board nets (one from the battery and one from the DCDC converters) was available. In addition, there was also redundant information on the state of the 12V board nets, one from the DCDC converters and one from the Brake System. Whenever there was a failure in the board nets or a mismatch of information, a soft stop was triggered.

12 Volt power supply

Most of the actuation and computation devices are feed by 12V. To achieve redundancy, two 12V batteries were installed in the vehicle. This is rather a brute-force method, but sufficient for the scope of this prototype. A long-term solution would be to use only one battery and in addition install a supervision, which would decouple the battery from the board net in case it drains too much current.

High Level command and supervision (PC + ECU)

Within the software stack running the vehicle there was a distinction between high- and low-level software. The high-level software contains object detection, environmental modelling, prediction, vehicle maneuver and motion planning. The lower-level software consists of control algorithms and vehicle interface handling.
The lower-level software was again split into two parts, one executed on the PC for the vehicle control algorithms and one running on the ECU to safeguard the actuator commands and to ensure correct interface handling.
In case of an error in the vehicle commands from the PC, these were overwritten, and a brake and steer straight command was sent to the actuators. This was very meaningful, because in case of high acceleration, high deceleration or high steering requests, the movement of the cabin becomes so severe that the safety operator would not be able to push the emergency button.
Also, the actuators had safety mitigations. If the commands sent from the ECU were corrupted, the steering was set to inactive (and with that steering straight) and the brake system was building up brake pressure.

In any of these cases, a recording was triggered, covering at least the last sixty seconds. The recording was then analyzed, and the incident was discussed by the engineering team with the safety experts in a safety management meeting, as defined in the SMS mentioned above.

Emergency buttons
The two emergency buttons are read-in in a redundant way by two ECUs. To be safe against shortcuts, the emergency stop button voltage was set to be 8V. As soon as it would drop below 5V or increase above 11V an emergency stop was triggered.

Door System
The door system, originating from busses, was designed to stay closed when a failure occurs. While a bus has more than one door which can serve as emergency exit, this is not valid in the case of the shuttle. Therefore, the doors were equipped with an emergency button, which disabled the whole door.
On the other hand, it is possible that this emergency button is pressed during normal operation. Therefore, a safety mechanism was implemented which triggers a soft stop of the vehicle in that case.

With these analysis results at hand, the CUbE-3 vehicle was successfully released and operated. Similar steps were performed for other prototype systems at Continental, covering different use cases as well as different vehicle platforms.

CONCLUSIONS AND PERSPECTIVES

In this paper, it was outlined how prototypes of driverless shuttles can be used during development, either towards a complete shuttle system or also for specific components or functions thereof. We argued that a certain release process should be installed in order to safely prepare and operate such prototypes. We presented a generic view on driverless shuttles in terms of their basic architecture. A safety assessment was then performed using one specific prototype vehicle developed and operated at Continental as a hands-on example. Here, we combined both top-down and bottom-up analysis steps to provide a pragmatic way to release such a vehicle. For the bottom-up analysis, we present a pragmatic approach called Component Failure and Effect Analysis (CFEA).

The considerations presented here should not substitute existing and established methods but should complement those within the scope of advanced engineering or development activities in general, in which sometimes a pragmatic approach delivers more value in a faster way. The results presented here may well be used as baseline for further developments, such that driverless shuttles become more and more mature in the near future.

REFERENCES

[1] ERTRAC Working Group: “Connectivity and Automated Driving", *Connected, Cooperative and Automated Mobility Roadmap,* European Road Transport Research Advisory Council (ERTRAC), 2022.

[2] Department for Transport and Department for Business, Energy & Industrial Strategy: Centre for Connected and Autonomous Vehicles, *Connected & Automated Mobility 2025: Realising the benefits of self-driving vehicles in the UK,* Crown, 2022.

[3] U.S. Department of Transportation, *Automated Vehicles 3.0: Preparing for the Future of Transportation 3.0,* 2018.

[4] Rhein-Main-Verkehrsverbund Servicegesellschaft mbH, „Probefahrt in die Zukunft,“ 2022. [Online]. Available: https://www.probefahrt-zukunft.de/.

[5] Verkehrsbetriebe Hamburg-Holstein, „RealLabHH: Hamburg testet die Mobilität von Morgen,“ 2021. [Online]. Available: https://vhhbus.de/reallabhh-reallabor-hamburg/.

[6] W. Davis, „Toyota e-Palette autonomous vehicles to be rolled out 'within the next few years',“ 2021. [Online]. Available: https://www.drive.com.au/news/toyota-e-palette-autonomous-vehicles-to-be-rolled-out-within-the-next-few-years/.

[7] A. Hartmannsgruber, J. Seitz, M. Schreier, M. Strauss, N. Balbierer und A. Hohm, „CUbE: A Research Platform for Shared Mobility and Autonomous Driving in Urban Environments,“ in *IEEE Intelligent Vehicles Symposium*, Paris, 2019.

[8] C. Pinke, A. Hohm und M. Grießer, „Industrialized technology building blocks for on-demand autonomous shuttles in urban and sub-urban use cases enabling mass-market scales, affordability, automotive reliability and safety,“ in *27th ITS World Congress*, Hamburg, 2021.

[9] F. Lotz, R. Grewe und C. Pinke, „A “Common Core” Architecture as an Enabler for Cross-Platform Autonomous Driving,“ in *Apringer ATZ - Automatisiertes Fahren 2022*, Wiesbaden, 2022.

[10] Continental, „Autonomous in Continental's and EasyMile's Robo-Taxi at the International Motor Show (IAA),“ 2019. [Online]. Available: https://www.continental.com/en/press/press-releases/robo-taxi/.

[11] I. Lunden, "EasyMile raises $66M for its autonomous people-and-goods shuttles," 2021. [Online]. Available: https://techcrunch.com/2021/04/28/easymile-raises-66m-for-its-autonomous-people-and-goods-shuttles/.

[12] Business Wire, „Navya Sells 8 Self-Driving Shuttles in the US to Autonomous Mobility Provider Beep,“ 2022. [Online]. Available: https://www.businesswire.com/news/home/20220911005052/en/.

[13] Freethink Team, „Inside Zoox: The robot vehicle totally changing transportation,“ [Online]. Available: https://www.freethink.com/series/hard-reset/autonomous-vehicle.

[14] A. J. Hawkins, „Cruise starts pre-production of its autonomous shuttle thanks to $5 billion from GM,“ 2021. [Online]. Available: https://www.theverge.com/2021/6/15/22534945/gm-credit-cruise-autonomous-vehicle-origin-production.

[15] Sustainable Bus Editorial Team, „ZF autonomous shuttle presented in UK,“ [Online]. Available: https://www.sustainable-bus.com/news/zf-autonomous-shuttle-presented-uk/.

[16] ISO 26262:2011 Road vehicles - Functional safety, 2018.

[17] AIAG; VDA, FMEA Handbook, 2019.

[18] German Federal Government, *Straßenverkehrsgesetz (StVG) (German Road Traffic Act),* 2021.

[19] German Ministry for Transportation and Digital Infrastruktur, „Verordnung zur Genehmigung und zum Betrieb von Kraftfahrzeugen mit autonomer Fahrfunktion in festgelegten Betriebsbereichen (Autonome-Fahrzeuge-Genehmigungs-und-Betriebs-Verordnung - AFGBV),“ 2022. [Online]. Available: https://ec.europa.eu/growth/tools-databases/tris/index.cfm/de/search/?trisaction=search.detail&year=2021&num=344&dLang=DE.

[20] German Federal Government, „FAQ Autonomous Driving,“ [Online]. Available: https://www.bundesregierung.de/breg-en/news/faq-autonomous-driving-1916398.

[21] German Federal Motor Transport Authority, „Erprobungsgenehmigung gemäß § 1i StVG in Verbindung mit § 16 AFGBV,“ 2022. [Online]. Available: https://www.kba.de/DE/Themen/Typgenehmigung/Autonomes_automatisiertes_Fahren/Erprobungsgenehmigung/erprobungsgenehmigung_node.html.

[22] Pegasus Project, „Pegasus,“ [Online]. Available: https://www.pegasusprojekt.de/en/.

[23] Association of Standardization of Automation and Measuring Systems, „ASAM Homepage,“ [Online]. Available: https://www.asam.net/.

[24] ISO 34502:2022 Road vehicles - Test scenarios for automated driving systems - Scenario based safety evaluation framework, 2022.

[25] Internatioal Civil Aviation Organization (ICAO), Safety Management Manual (SMM) - Doc 9859 AN/474, 2013.

[26] N. G. Leveson und J. P. Thomas, „STPA Handbook," March 2018. [Online]. Available: https://psas.scripts.mit.edu/home/get_file.php?name=STPA_handbook.pdf.

[27] ISO 21448:2022 Road vehicles - Safety of the intended functionality, 2022.

PEER REVIEW PAPER

This paper has been peer-reviewed and published in a special edition of Traffic Injury Prevention 24(S1), by Taylor & Francis Group. The complete paper will be available on the Traffic Injury Prevention website soon. To access ESV Peer-reviewed papers click the link below
https://www.tandfonline.com/toc/gcpi20/24/sup1?nav=tocList

RESEARCH OF LSTM MODEL FOR VEHICLE CONTROL SYSTEM OF AUTOMATED DRIVING SYSTEMS (ADS)

Kaisei, Honda Tasuku, Yamazaki Tomohiro, Baba Shoko, Oikawa
Shibaura Institute of Technology, Human Machine System Laboratory

Toshiya, Hirose
Shibaura Institute of Technology, Department of Engineering Science and Mechanics

Paper Number 23-0108

ABSTRACT

Research Question/Objective: This study aims to construct a long short-term memory (LSTM) model of the vehicle control system for automated driving systems (ADSs) that does not cause annoyance or distrust. Furthermore, this study investigates the effect of LSTM hyperparameters on model accuracy. A survey showed that certain drivers did not use levels 1 and 2 of the ADS function because they were annoyed with the driving behavior of the ADS-controlled vehicle. Although the driving behavior of the ADS-controlled vehicle causes distrust in passengers, it cannot effectively enable safe driving. This study focuses on a novel vehicle control method that reduces annoyance and distrust in passengers and contributes to the safe operation of ADSs. These control methods involve the application of a long short-term memory (LSTM) model that learns long-term time-series data. This system enables the construction of ADS control algorithms from LSTM models based on personalized driver operations during ordinary driving.

Methods and Data Sources: LSTM models were constructed for highway driving in the following three driving scenarios. Scenario-1: following a preceding vehicle, Scenario-2: passing a preceding vehicle at low speed with lane change, Scenario-3: a sudden lane change by a vehicle in the passing lane in front of the vehicle. The effect of LSTM hyperparameters on the accuracy of the LSTM model was investigated for each driving scene. The data of these models were sourced from an experiment using a driving simulator conducted to determine driver behavior.

Results: The results verified the accuracy of the model that simulated the driving operation of the driver. The model accuracy was improved by setting LSTM hyperparameters. In Scenario-1, the number of units, learning rate, and the number of epochs affected the coefficient of determination. The coefficient of determination tends to be particularly high for a large number of units. In Scenario-2, unlike Scenario-1, a large number of units was not required to obtain a high coefficient of determination. The coefficient of determination did not change with the epoch. In Scenario-3, similar to Scenario-1, the number of units, learning rate, and epoch affected the coefficient of determination, whereas the coefficient of determination decreased at epochs above 800.

Discussion and Limitations: In each scenario, the hyperparameters affecting the accuracy were different. A limitation of this study is that it focuses on the driver model. The LSTM model applying ADS was evaluated.

Conclusions: For the ADS control algorithm (SAE Levels 3, 4, and 5), we constructed LSTM models that reflect the characteristics of personalized drivers. The results showed that the LSTM hyperparameters affecting the

coefficient of determination tended to differ among different scenarios. In the future, evaluation of the effectiveness of the LSTM model when applied to ADS is necessary. Novel control systems for ADS with LSTM models contribute to the development of ADS system design.

INTRODUCTION

In recent years, vehicles that use automated driving systems (ADSs) and other automated driving technologies have become increasingly popular. According to a survey of the global market for ADSs, the number of vehicles equipped with advanced driver assistance systems (ADAS)/ADS is expected to reach 79,153,000 units by 2030 [1]. Although ADS contributes to reducing a driver's driving burden by replacing the driver and improving safety, there are potential problems that need to be solved in the future. According to a survey of driver feelings regarding ADAS functions widely used currently, more than 54% of those who own ADAS-equipped vehicles believe that ADAS functions conversely increase the possibility of accidents, and 70% of drivers have turned off ADAS functions [2]. The cause is attributed to the low personal adaptability of the system to each driver that significantly affects the function unacceptability [3][4]. This distrust of the ADAS can also be applied to ADS. For example, when following a preceding vehicle, the driver may be annoyed by the system's frequent acceleration and deceleration despite a sufficient distance, or the driver may distrust the system when the system does not make such a decision even though the preceding vehicle may be traveling at a low speed, and the driver intends to overtake it. To solve these problems, considering personal adaptability in automatic driving control systems is necessary [5][6]. The purpose of this study is to develop a long short-term memory (LSTM) model of the vehicle control system for ADS that does not cause annoyance or distrust and to investigate the effect of LSTM hyperparameters on model accuracy. Therefore, we propose an algorithm for a new personalized vehicle control system that contributes to safe driving for ADS, and we use a driver model based on LSTM to construct this system. Two methods are available for constructing driver models. The first method constructs driver models corresponding to various driving scenarios by modeling the driver's driving behaviors, such as cognition, decision-making, and operation. The second method involves constructing and integrating driver operation models that fit each driving scenario. The second method integrates several driver operation models that are constructed for each driving scenario. Driving scenarios include free driving, following, lane change, and merging. Operational models are easier to construct than cognitive and assessment models, and numerous previous studies have been conducted [7][8][9]. In addition, by limiting the target driving scenario, the driver's cognitive information can be identified. In this study, driver models were constructed using the latter method. In this study, several personalized driver models were developed to demonstrate their feasibility of constructing personalized models. The driving scenarios targeted in this study include car-following, overtaking, and cut-in behaviors. In modeling each scenario, the effect of the LSTM hyperparameters on accuracy was investigated.

METHODOLOGY

Driving Experiment Using a Driving Simulator

Driving experiments were conducted using a driving simulator (DS) to obtain data regarding a driver's ordinary driving behavior. The experiment was conducted using a DS, as shown in Figure 1. A six-axis sway device equipped with electric actuators was used to simulate the sensation of driving a DS. The six-axis sway device can simulate the pitch motion caused by the driver's gas pedal and brake operations and the roll motion caused by the steering operation. The upper part of the six-axis sway device includes a turntable that can rotate ±180° and simulate yaw motion. The screen is cylindrical with a radius of 2.5 m, and six projectors are used to simulate the traffic scene in all directions. The experiment was conducted on five male participants in their 20s (average age: 21 years). The experimental conditions include three driving scenarios: car-following, overtaking, and cut-in. The participants were informed to drive under each condition as they typically do, to obtain their driving characteristics. All these scenarios were simulated in an environment similar to a standard Japanese highway. This study was approved by the Ethics Committee of the Shibaura Institute of Technology.

Figure 1. Driving Simulator.

<u>Car-following Scenario</u>

An overview of the car-following scenario is shown in Figure 2. In Figure 2, the red, blue, and black vehicles represent the ego, preceding, and other vehicles, respectively. Participants boarded and drove the ego vehicle; the preceding vehicle traveled in front of the ego vehicle. The ego vehicle travelled at 100 km/h and followed the last car in a line of cars traveling at 40–60 km/h caused by a traffic accident until the congestion was cleared. Each scenario spanned approximately 2 min and 30 s. The drivers were asked to drive in the scenario 10 times to obtain the training data. During the experiment, we obtained the ego vehicle's velocity, acceleration, velocity relative to the preceding vehicle, and the headway distance from the preceding vehicle. Driving behaviors of the sequence of drivers from the beginning to the end of the journey were modeled.

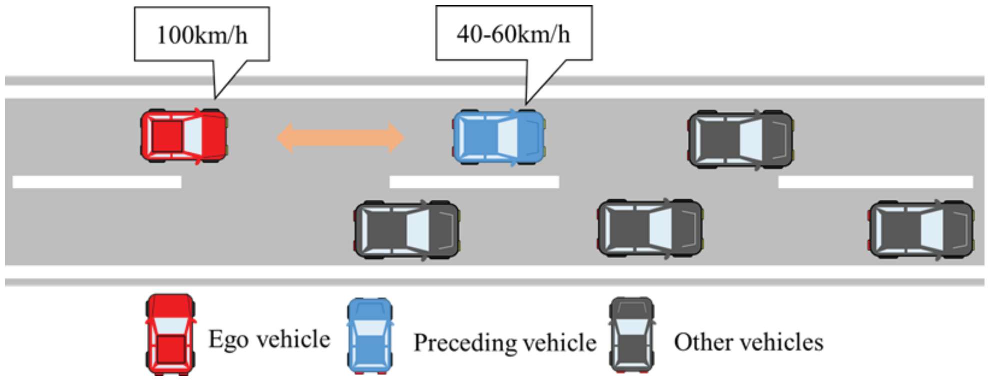


Figure 2. Overview of the car-following scenario.

Overtaking Scenario

An overview of the overtaking scenario is shown in Figure 3. In Figure 3, the red, blue, and green vehicles represent the ego, the preceding vehicle in the travel lane, and the preceding vehicle in the overtaking lane, respectively. The participants boarded and drove the ego vehicle, and the preceding vehicle was placed in front of the ego vehicle. The green vehicle in Figure 3 travels in the passing lane at a speed of 100 km/h. The ego vehicle traveled at 100 km/h, approached the preceding vehicle traveling at approximately 80 km/h, and decelerated eventually. The driver subsequently changed lanes to the overtaking lane, passed the vehicle, and changed lanes back to the original lane, while focusing on the vehicle in the overtaking lane. Each scenario spanned approximately 1 min and 10 s. The driver was asked to drive the scenario 10 times to obtain training data. In the experiment, we obtained the ego vehicle's steering angle, acceleration, relative velocity, and headway distance relative to the preceding vehicle and the relative velocity and distance relative to the vehicle in the overtaking lane. The model targeted the period spanning from the time the driver turned on the blinker and performed the overtaking maneuver until he returned to his original lane.

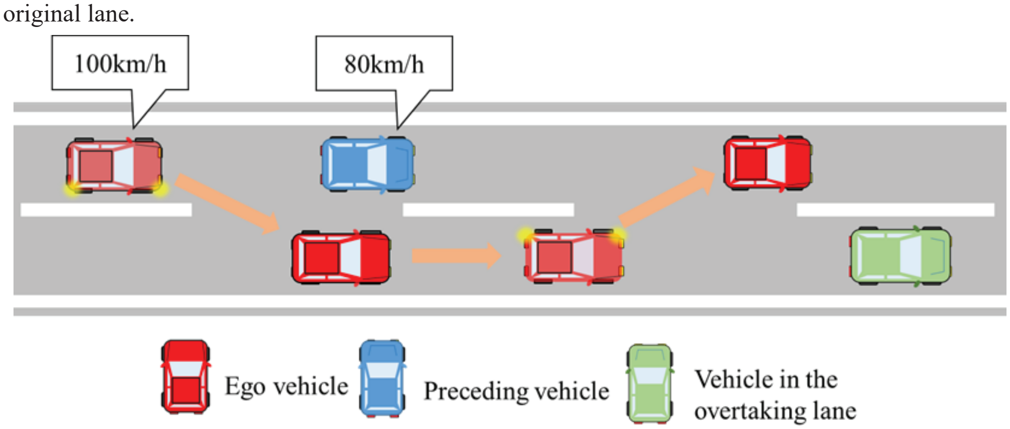


Figure 3. Overview of the overtaking scenario.

Cut-in Scenario

An overview of the cut-in scenario is shown in Figure 4. In Figure 4, the red, blue, and black vehicles represent the ego, cut-in, and other vehicles, respectively. In this scenario, when the ego vehicle was traveling at 70–80 km/h, a cut-in vehicle passed it at 100 km/h from the passing lane on the right side and cut in front of it. The distance between the ego vehicle and the preceding vehicle in the travel lane at the time of the cut-ins was set at intervals of 1 m from 7 m to 13 m. The driver was asked to run the scenario 10 times to obtain the training data. During the experiment, the driver acquired the ego vehicle's velocity, acceleration, relative velocity, and headway distance relative to the cut-in vehicle. The model targeted 5 s after the cut-in vehicle cut in.

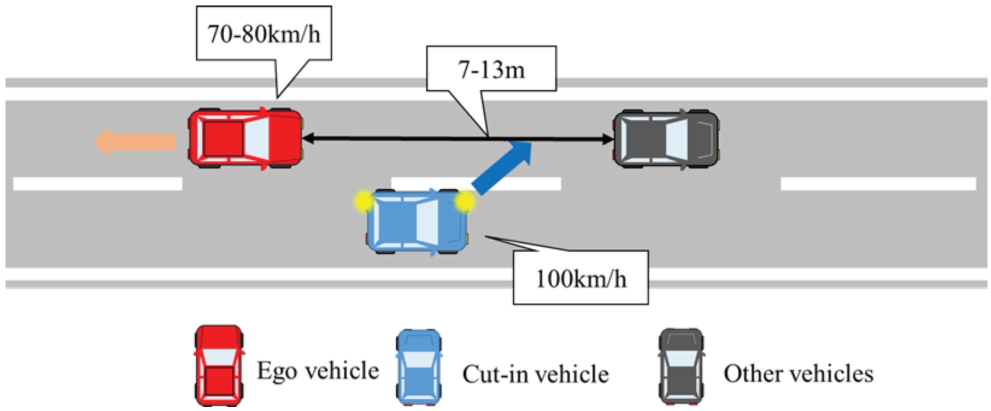


Figure 4. Overview of the cut-in scenario.

Novel Personalized Vehicle Control System

The proposed LSTM-based ADS control algorithm is illustrated in Figure 5. The data source stores information such as the relative position and relative velocity of the vehicle with respect to the road user (vehicles, pedestrians, bicycles, etc.) and the shape of the road on which the vehicle travels in the experiment. The data source can utilize a connected vehicle system, the V2X (Vehicle-to-Everything) system [10], in which information can be mutually obtained through communication between the vehicle, road user, and infrastructure. The lower layer of the data source is a scenario classification model that classifies driving scenarios. Based on the information obtained from the data source, the system classifies driving scenarios corresponding to the current vehicle environment and applies a driver model that is appropriate for that scenario. By personalizing the scenario classification model, it is possible to adapt it to each driver (for example, whether and when to change lanes). After classification, the driver model was provided with the necessary input data. As mentioned previously, these driver models were constructed for each driving scenario. In this study, driver models were constructed for car-following, overtaking, and cut-in scenarios, as an example. A personalized driver model is a position- or acceleration-based model that considers the driving characteristics of each driver. Because the personalized driver model outputs the vehicle position and acceleration/deceleration, it cannot control the vehicle directly. The control mechanism (controller) calculates the acceleration stroke, brake stroke, and steering amount based on the predicted data (acceleration/deceleration and position) output by the driver model and controls the vehicle. The calculated operation amounts are applied to the vehicle model to obtain the necessary information for the next control step that is subsequently fed back into the data source. Consequently, performing personalized ADS control corresponding to various time-varying driving scenarios is feasible. Based on the premise of this vehicle control system, we discuss the construction of a personalized driver model using LSTM.

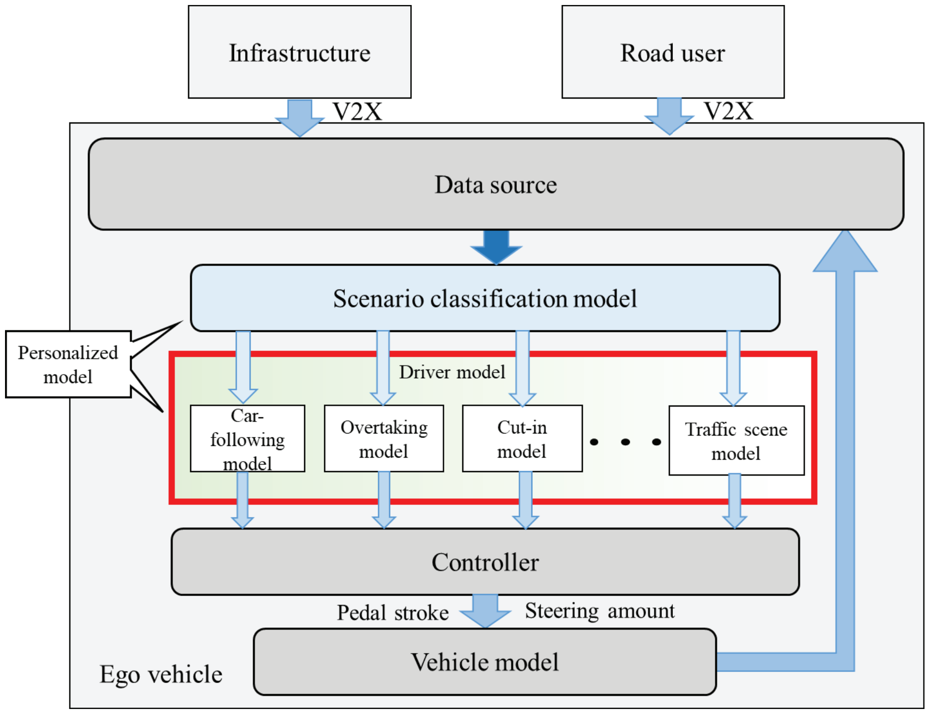


Figure 5. Proposed ADS control algorithm.

Method of Constructing Personalized Driver Model

We constructed personalized driver models using LSTM. Figure 6 shows the construction flow of the personalized driver model. LSTM is a type of recurrent neural network that exhibits a superior processing capability for time-series data. Figure 7 shows the personalized driver models constructed in this study. The inputs for the car-following and cut-in models include the velocity of the ego vehicle, relative velocity, and headway distance to the preceding or cut-in vehicle. The inputs for the overtaking model include the steering angle of the ego vehicle, relative velocity, and headway distance to the preceding vehicle in the travel lane, and the relative velocity and headway distance to the vehicle in the overtaking lane. The output of the car-following and cut-in models is longitudinal acceleration. However, the outputs of the overtaking model were longitudinal and lateral accelerations considering the driving maneuvers. In the LSTM, the number of hidden units, initial learning rate, and number of epochs were set as parameters for training. Table 1 shows the parameter setting. The numbers of hidden units were 25, 50, 100, 150, and 200. The initial learning rates were 0.001, 0.005, 0.01, 0.015, and 0.02. The number of epochs was 400, 600, 800, 1000, 1200, and 1400. In other words, LSTM was trained 150 times for each scenario and for all the participants. To ensure reasonable impact of the parameters on model accuracy, we fixed the initial values that were randomly output. The learning rate was set to decay by 0.999 per epoch. Data obtained from the driving experiments using a driving simulator were divided into 80% training data and 20% test data to construct the model. To confirm the feasibility of the model, a highly accurate driver model was constructed based on 150 training sets, and acceleration simulations were conducted.

Table 1. LSTM learning parameters

Parameters	Value
Hidden units	25, 50, 100, 150, 200
Initial learning rate	0.001, 0.005, 0.01, 0.015, 0.02
Epochs	400, 600, 800, 1000, 1200, 1400

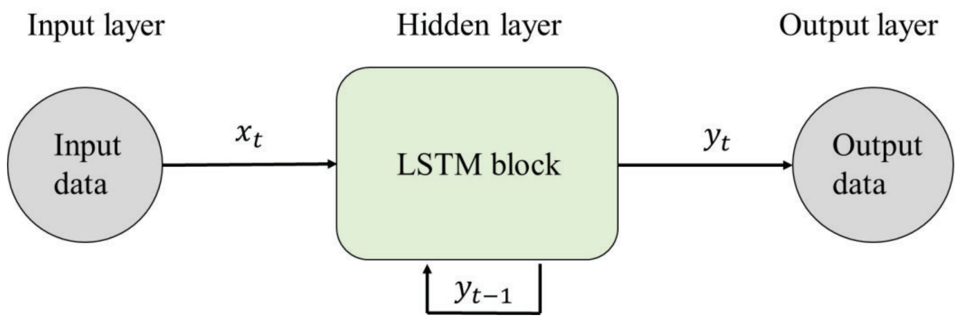


Figure 6. Personalized driver model construction flow.

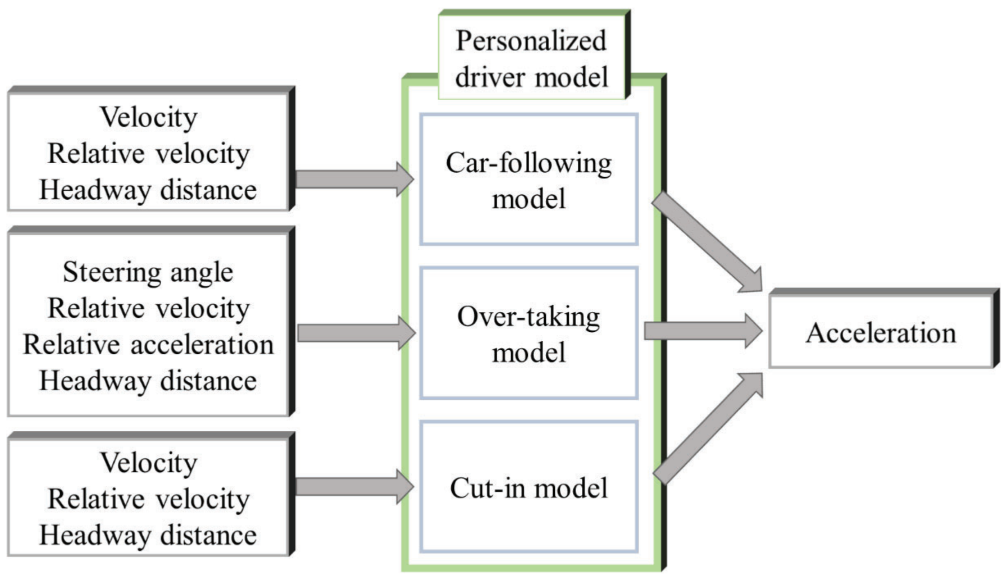


Figure 7. Personalized driver mode .

Analysis Method

To investigate the impact of LSTM hyperparameters on modeling accuracy, LSTM trained 150 scenarios and calculated the coefficient of determination from the test results. The formula for calculating the coefficient of determination is given by Equation (1). The closer the coefficient of determination is to 1, the better the model fits the test data and the higher the model accuracy.

$$R^2 = 1 - \frac{\sum_{i=1}^{n}(y_i - y_i')^2}{\sum_{i=1}^{n}(y_i - y_i'')^2} \qquad \text{Equation (1)}$$

(y_i: measured acceleration; y_i': predicted acceleration; y_i'': average value of the measured acceleration)

Acceleration simulations were conducted using highly accurate driver models that were constructed based on a study using the above coefficients of determination. Simulation results were evaluated using the root mean squared error (RMSE). The formula for calculating the RMSE is shown in Equation (2). The smaller the RMSE, the smaller the error, and the higher the prediction accuracy.

$$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(y_i - y_i')^2} \qquad \text{Equation (2)}$$

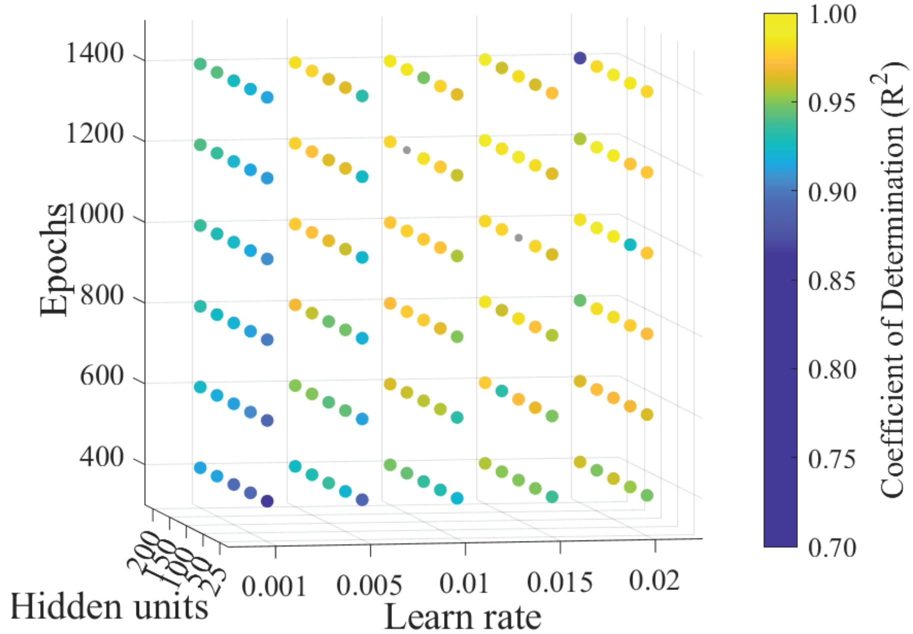
(y_i: measured acceleration; y_i': predicted acceleration)

RESULT

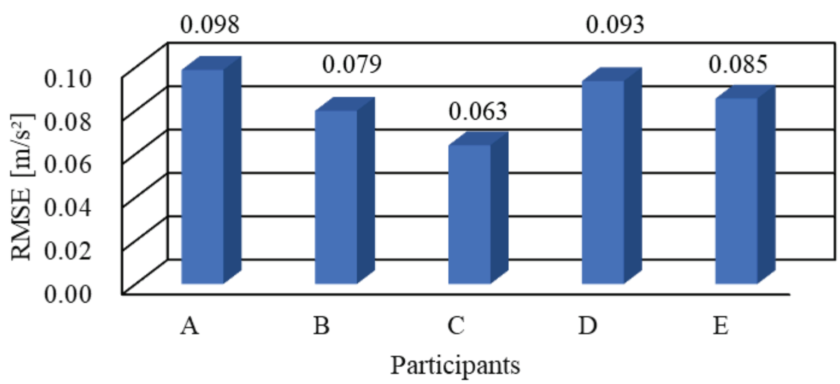
The results of the coefficients of determination for the driver model constructed with 150 hyperparameters for each scenario are described below.

Car-following Model

Figure 8 (a) shows a three-dimensional plot of the relationship between the average coefficient of determination and the hyperparameters for all participants in the car-following model. The color of the plotted points changes with the value of the coefficient of determination, with a yellow scheme for points closer to 1 and a blue scheme for points further away from 1. The larger the number of units, learning rate, and epoch, the larger the coefficient of determination and the more accurate the model. In particular, the coefficient of determination tended to be higher when the number of units was larger. However, when the number of learning epochs was high for a large number of units, as in the case of 200 units, 1400 epochs, and a learning rate of 0.02, the expressiveness of the model became exceedingly high, resulting in overlearning. Figure 8 (b) shows the model with the highest accuracy (lowest RMSE) from 150 training results for each participant in the experiment. The acceleration simulation results showed low errors with RMSE values of less than 0.1 [m/s²], and the car-following behavior could be modeled with high accuracy.



(a) Average coefficient of determination for all participants in the car-following model.



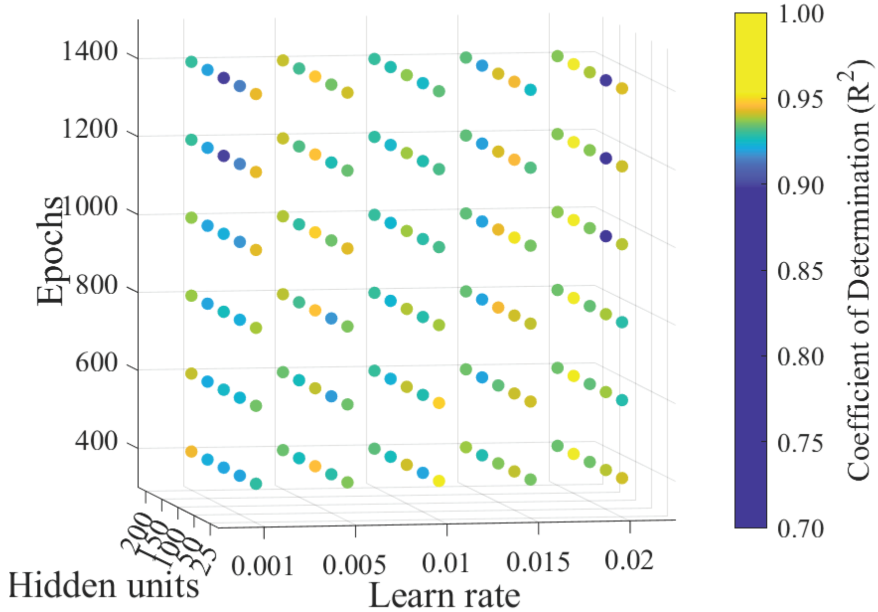
(b) RMSE values with the highest accuracy of acceleration simulation for each participant.

Figure 8. Car-following model construction results.

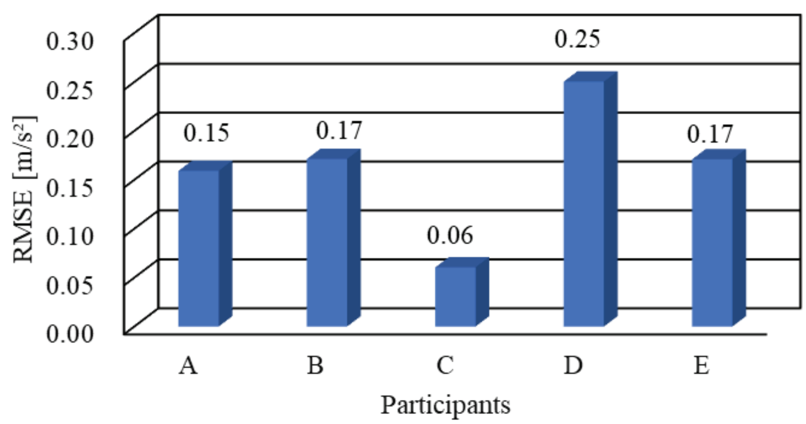
Overtaking Model

Figure 9 (a) and 10 (a) show three-dimensional plots of the relationship between the average coefficient of determination and the hyperparameters for all participants in the longitudinal and lateral directions of the overtaking model, respectively. As shown in Figure 9 (a), in the longitudinal direction, the coefficient of determination changed significantly with changes in the number of units and learning rate; nonetheless, no specific trend was observed. However, the coefficient of determination did not change significantly with the number of epochs. In particular, a high coefficient of determination independent of epoch was obtained when the number of units was 150 and the learning rate was 0.02. As shown in Figure 10 (a), in the lateral direction, as in the longitudinal direction, the coefficient of determination was unaffected by changes in the number of epochs. Figure

9 (b) and 10 (b) show the model with the highest accuracy (lowest RMSE) among the results simulated from 150 learning results. In both the longitudinal and lateral directions, Participant C exhibited a smaller RMSE for the simulation results, whereas Participant D exhibited a larger RMSE for the simulation results.

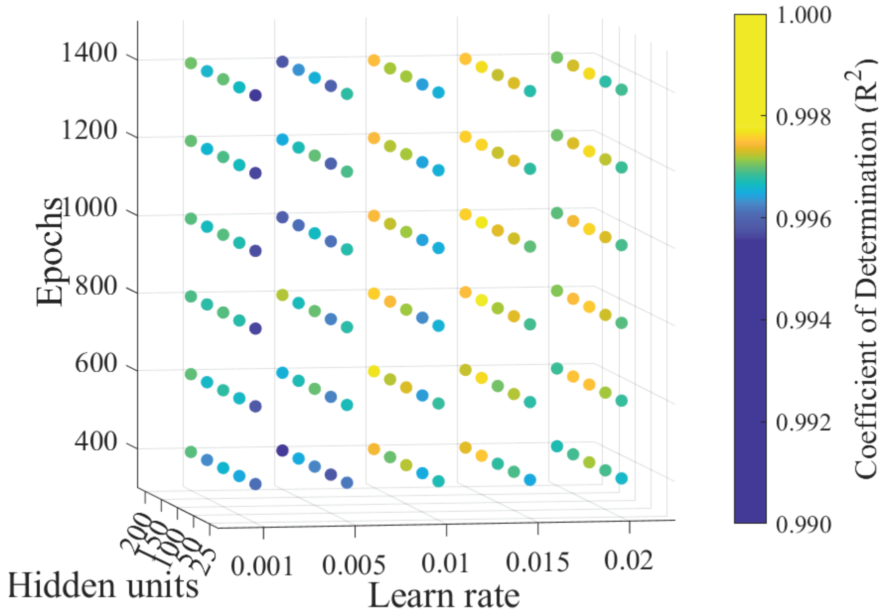


(a) Average coefficient of determination for all participants in the overtaking model.

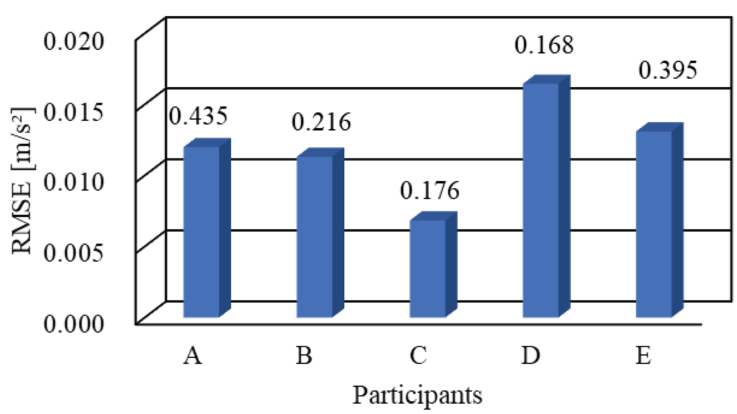


(b) RMSE per experiment participant.

Figure 9. RMSE values with the highest accuracy of acceleration simulation for each participant.



(a) Average coefficient of determination for all participants in the overtaking model (lateral).

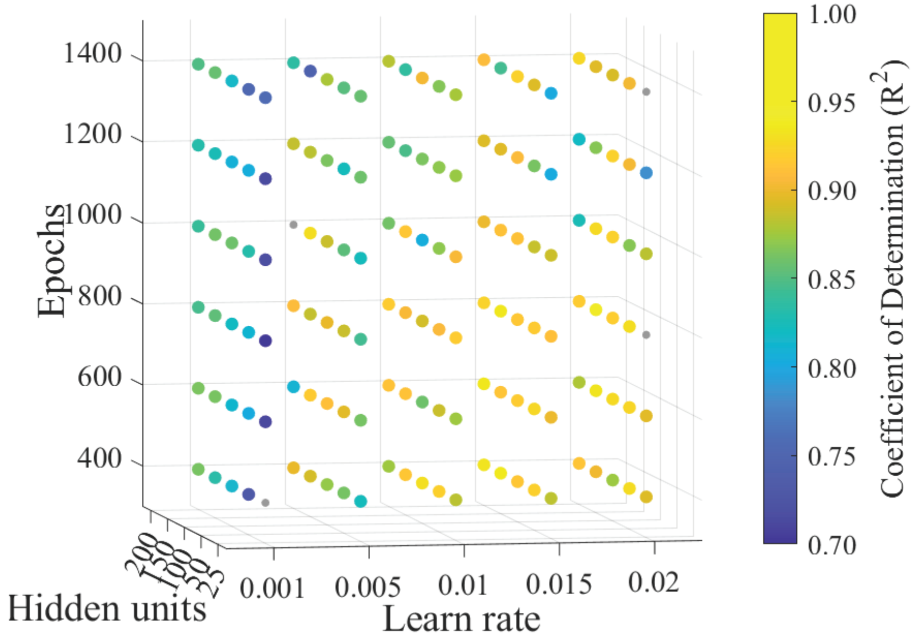


(b) RMSE values with the highest accuracy of acceleration simulation for each participant.

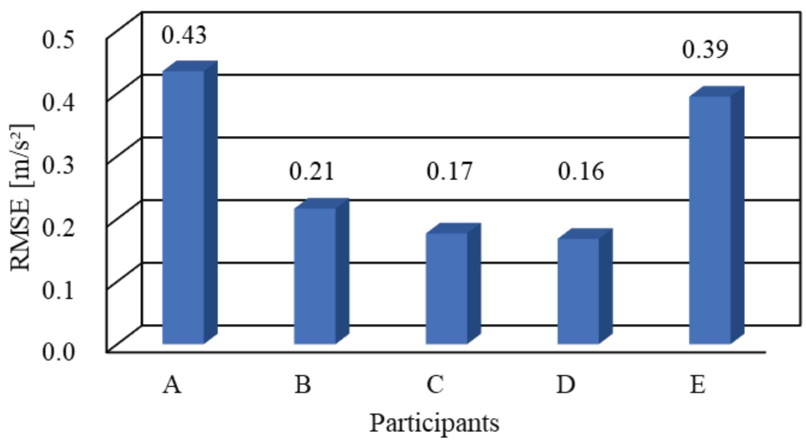
Figure 10. Results of constructing an overtaking model (lateral).

Cut-in Model

Figure 11 (a) shows a three-dimensional plot of the relationship between the average coefficient of determination and the hyperparameters for all participants in the cut-in model. The coefficient of determination tended to increase with the number of units, learning rate, and epoch. In particular, stable, high coefficients of determination were obtained for epochs below 800. Figure 11(b) shows the model with the highest accuracy (lowest RMSE) from 150 training results for each participant. As per the acceleration simulation results, participants B, C, and D were characterized by small RMSEs of 0.21, 0.17, and 0.16 [m/s²], respectively, while participants A and E were characterized by large error RMSEs of 0.43 and 0.39 [m/s²].



(a) Average coefficient of determination for all participants in the cut-in model.



(b) RMSE per experiment participant.

Figure 11. RMSE values with the highest accuracy of acceleration simulation for each participant.

DISCUSSION AND LIMITATION

The hyperparameters affecting the coefficients of determination tended to differ among the 150 modeled scenarios in this study. Below, we discuss the LSTM hyperparameters that affect the coefficient of determination for each model and the RMSE for each participant in the experiment.

Car-following Model

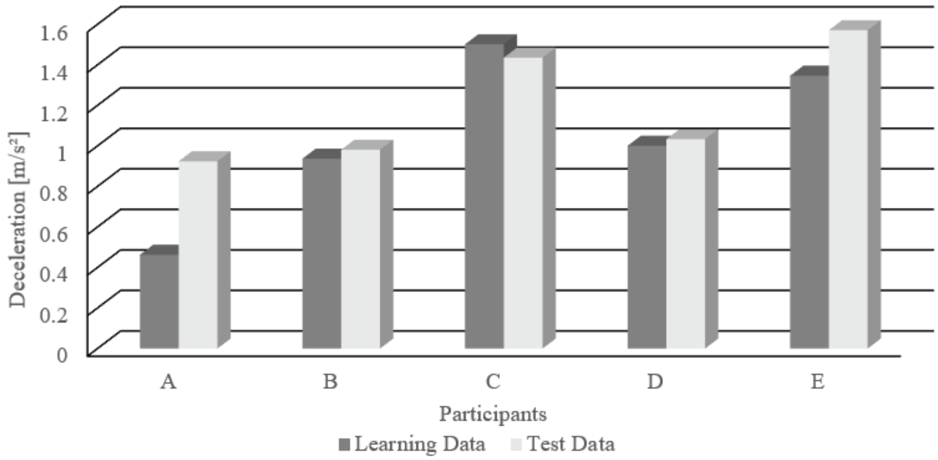
For the car-following model, the number of units, learning rate, and epoch affected the coefficient of determination. The car-following scenario in this experiment features a more complex acceleration/deceleration behavior and various patterns than the overtaking and cut-in scenarios, and a large number of units are required to represent the driver's driving characteristics. The RMSE values were less than 0.1 [m/s²] for all participants in the experiment, indicating that the experiment adequately reproduced the driving behavior of individual drivers.

Overtaking Model

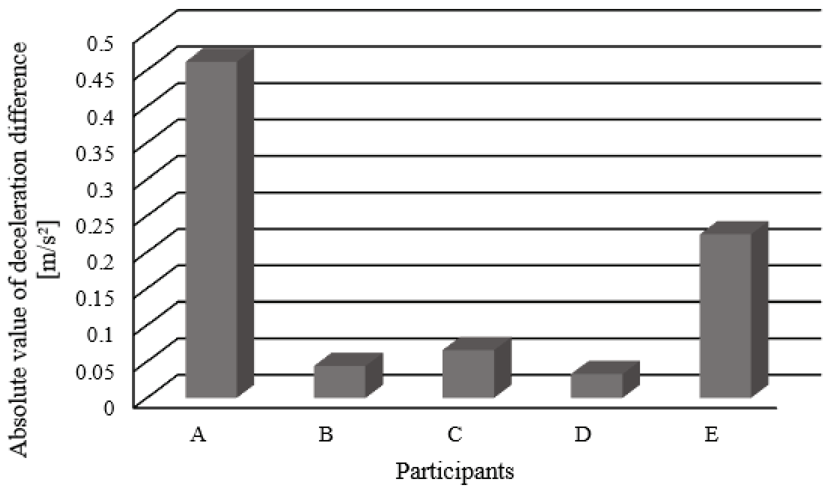
In the longitudinal direction, the overtaking model afforded high coefficients of determination, independent of the number of epochs in the case of 150 units and a learning rate of 0.02. In the lateral direction, high coefficients of determination were obtained independent of the number of epochs in the case of 100 or more units and a learning rate of 0.01 or more. The coefficient of determination does not change with the number of epochs. In machine learning, the accuracy of the model increases as the training progresses; however, after a certain level of training, the accuracy of the model decreases owing to over-training. As shown in Figure 11 (a), in one case, the accuracy of the cut-in model decreased at epochs above 800, owing to overlearning. As shown in Figures 9 (a) and 10 (a), the coefficient of determination did not change with the epoch, suggesting that the parameter set for the number of learning epochs was insufficient for identifying the trend. The acceleration simulation results for each participant in the experiment showed that participant D exhibited a slightly larger RMSE than the other participants that may be attributed to the fact that Participant D performed the overtaking maneuver at a different time from the other participants in the driving simulator experiment. The other participants in the experiment started to change lanes after the vehicle in the overtaking lane passed their vehicle, whereas participant D changed lanes before the vehicle in the overtaking lane passed his/her vehicle. In this study, the intention to begin lane change was not considered, and only the overtaking maneuver was targeted; therefore, the input data were not optimized. In the future, the scenario classification model shown in Figure 5 can be used to personalize lane-change decisions, and the accuracy is expected to be improved by examining the optimal driver model input.

Cut-in Model

In the cut-in model, the number of units, learning rate, and epoch affected the coefficient of determination; however, unlike the car-following model, the coefficient of determination decreased when the epoch was exceedingly large. A decrease in the coefficient of determination was observed for epochs greater than 800. This may be because the driver's driving behavior in the cut-in scenario requires only braking in this study, and thus a highly accurate model could have been constructed even at low epochs. The results of the acceleration simulations for each experimental participant showed that participants B, C, and D were characterized by smaller RMSEs, while participants A and E were characterized by larger RMSEs. A feature of data-driven models is that their prediction accuracy depends on the training and test data. The reason for the large RMSE for participants A and E in the experiment may be that they did not have sufficient training data. Figure 12 (a) shows the average acceleration of the training and test data for each participant during the experiment. Figure 12(b) shows the difference between the average accelerations of the training and test data for each participant during the experiment. The difference in average acceleration between the training data and test data was large for participants A and E, indicating that the training data used in this study were not sufficient to cover the acceleration range of the test data.



(a) Average acceleration of training and test data for each participant.



(b) Difference between the training and test data.

Figure 12. Comparison of training and test data for each participant.

A limitation of this study is that it focuses on the driver model, as shown in Figure 5. We constructed a driver model for following, overtaking, and cut-in. Nonetheless, there are other scenarios that need to be validated as well; examples include merging and driving on curved roads. In addition, in the future, we intend to evaluate the effectiveness of the constructed LSTM driver model when applied to the proposed ADS control algorithm.

CONCLUSIONS

For automated driving technologies such as advanced driver assistance systems (ADAS) and automated driving systems (ADS) to be accepted by drivers, the elimination of annoyance and distrust is important. Therefore, this study proposes an algorithm for a novel personalized vehicle control system for ADS. The proposed system includes a personalized driver model constructed using LSTM. A driver model is constructed for each driving scene. In this study, a driver model was constructed for car-following, overtaking, and cut-in behaviors, and the effect of LSTM hyperparameters on the model accuracy was investigated. The results are as follows.

1. For the car-following model, the number of units, learning rate, and epoch affected the coefficient of determination. The coefficient of determination tends to be particularly high for a large number of units.
2. Unlike the CF model, the overtaking model does not require a large number of units for realization of a high coefficient of determination. The coefficient of determination did not change with the epoch.
3. The cut-in model, similar to the car-following model, showed that the number of units, learning rate, and epoch affected the coefficient of determination; however, the coefficient of determination decreased at epochs above 800.

Acceleration simulations were performed using the most accurate 150 models constructed. Highly accurate simulation results were obtained by optimizing the input and training data acquisition methods and the LSTM hyperparameters. Future evaluation of the effectiveness of the driver model when applied to the ADS is necessary. We expect that the new control system for ADS using the LSTM model proposed in this study will contribute to the development of an ADS system design.

REFERENCES

[1] Yano Research Institute Ltd. 2022. “Possibilities and Future Prospects for Automated Driving Systems.”

[2] CCC Intelligent Solutions Inc. 2021. “CRASH COURSE Midyear Report.”

[3] Jianqiang, W. Lei, Z. Dezhao, Z. Keqiang, L. 2013. “An Adaptive Longitudinal Driving Assistance System Based on Driver Characteristics.” IEEE Transactions on Intelligent Transportation Systems, 14-1,6236181, 1-12

[4] Cai, J., Jiang, H., Wang, J. 2022. “Implementation of the Human-Like Lane Changing Driver Model Based on Bi-LSTM." Discrete Dynamics in Nature, and Society, 9934292

[5] Awan, H.H., Pirdavani, A., Houben, A., Adnan, M., Brijs, T., Tom Brijs. 2019. “Impact of perceptual countermeasures on driving behavior at curves using driving simulator” Traffic Injury Prevention vol.20 93-99

[6] Marsden, G., McDonald, M., Brackstone, M. 2021. “Towards an understanding of adaptive cruise control.” Transportation reserch part C, emerging technologies vol.9, issue 1, 33-51

[7] Hamada, A., Oikawa, S., Hirose, T. 2022. “Construction of personalized driver models based on LSTM using driving simulator,” SAE Technical Paper, doi:10.4271/2022-01-0812.

[8] Woo, H., Sugimoto, M., Wu, J., Tamura, Y., Yamashita, A., Asama, H. 2018. “Trajectory prediction of surrounding vehicles using LSTM network”. IEEE Intelligent Vehicles Symposium (IV).

[9] Chen, Y., Hu, C., Wang, J. 2019. “Human-Centered Trajectory Tracking Control for Autonomous Vehicles with Driver Cut-In Behavior Prediction,” IEEE Transactions on Vehicular Technology, 68-9, 8758867,8461-8471

[10] Ning, L., Nan, C., Ning, Z., Xuemin S., Jon, W. M., 2014. “Connected Vehicles: Solutions and Challenges” IEEE Internet of Things Journal, vol. 1, issue 4, 289 – 299

CAPABILITY-BASED ROUTES FOR DEVELOPMENT, TESTING AND OPERATION OF SAFE AUTOMATED VEHICLES

Moritz Lippert
Institute of Automotive Engineering, Technical University of Darmstadt
Germany

Hermann Winner
Former professor for Automotive Engineering, Technical University of Darmstadt
Germany

Paper Number 23-0121

ABSTRACT

Ensuring safety of automated vehicles (AVs) within their operational design domain (ODD) is essential for a release. In some parts of the ODD, ensuring safe operation is more challenging, requiring more sophisticated driving capabilities. For example, the same intersection requires different capabilities depending on the selected turn, i.e., if driving right, left, or straight ahead. To guarantee safe operation, only route sections for which capabilities for safe driving are available and validated should be selected. So far, the direct relationship between routes within ODDs and the driving capabilities of AVs has not been explicitly addressed. This paper presents for the first time an approach to identify routes with driving requirements that do not exceed driving capabilities of AVs. To this end, this approach builds on the Behavior-Semantic Scenery Description (BSSD), which links behavioral demands for AVs directly to the scenery as a central element of the ODD. Based on the BSSD, route-based behavioral requirements are derived. Geometric characteristics of the scenery are used to specify driving requirements and driving capabilities that can be matched as a function of route and developed matching criteria. This matching is integrated into a conventional route planner, which as a result determines routes that are drivable based on the driving capabilities of an AV. The application to a real road network shows that the identification of capability-based routes is generally possible. Different intersections demand different requirements and lead to different routes. Nevertheless, several challenges are discussed that need to be overcome for a real-world application for development, testing, and operation of AVs.

INTRODUCTION

According to the current state of research and technology, the development and release process of automated vehicles (AVs) is a wide field of unsolved problems. One core challenge is the safety validation of these vehicles, which can no longer be performed using a conventional statistical approach [1] due to the high complexity of the overall system. Therefore, alternatives are needed to partially replace the statistical safety validation and thus reduce the overall effort in the development of AVs. One approach, which is additionally reinforced by various norms and standards such as ISO/TR 4804 [2] or ISO 21448 [3], is the safety-by-design approach. With the help of this approach, safety-relevant aspects are to be explicitly addressed in the development from the beginning, so that the final proof is facilitated. The goal of a safety or functional validation is to prove that the AV behaves according to the functional specification and does not exhibit any safety-critical deviations from this behavior. In this context, the functional scope is defined in the operational design domain (ODD), which describes the operational conditions for which an AV is specified and designed to function [4]. As the only final published standard for the specification of ODDs so far, PAS 1883 [5] defines the three main components *scenery*, *environmental conditions* and *dynamic elements* to describe an ODD. As a core element, the scenery describes all non-movable elements of the ODD. According to Ulbrich et al. [6], the scenery includes, for example, the road and lane network, but also stationary objects such as traffic lights or curbs. Thus, the scenery describes the space for the vehicle motion and behavior.

The proof of safe function and operation is thus closely linked to the defined scenery. In order to fulfill its function, an AV requires driving capabilities that can cope with the driving requirements arising from the scenery within the ODD. Many problems regarding development and safety validation of AVs are considered in related work, but the essential minimal task of AVs is often neglected - accomplishing an actual route within the ODD. The route forms the central element of AVs' mission accomplishment, accurately mapping the interdependencies between driving capabilities and location-based, different driving requirements of the ODD. Different routes potentially demand different driving capabilities that must be proven. Therefore, it seems reasonable to design the development, testing, and operation of AVs based on routes within the ODD.

For this purpose, it is necessary to establish a relationship between the scenery to be driven on and the resulting driving requirements as well as driving capabilities. The dependency between driving requirements and specific scenery areas within the ODD is explicitly addressed for the first time in a previous work [7] of the author team of this paper. We show how behavioral requirements for AVs can be derived based on routes within the scenery. For deriving the behavioral requirements, the Behavior-Semantic Scenery Description (BSSD) is used, which was co-developed by the same author team [8, 9]. The BSSD represents the behavioral limits imposed on AVs based on the scenery and applicable traffic rules enabling a route-based analysis of the ODD. Based on this previous work, the present paper develops a method for identifying capability-based routes. The driving requirements of these routes do not exceed the driving capabilities of selected AVs - no matter if already released or still under development. These routes and their associated driving requirements as well as required driving capabilities can be used for both the development process and operation of safe AVs.

This paper is structured as follows. First, the fundamental basics for a route-based scenery analysis based on the previous work are explained. Subsequently, an overall concept for the identification of capability-based routes is developed and implemented. Using a real road network as a basis, the developed concept is applied and evaluated. Finally, the remaining need for research is derived. A detailed description of the related work is intentionally not provided in this paper, as it can be found in the recently (co-)published papers of this author team [7-9].

FUNDAMENTALS

The dependency between scenery and resulting driving requirements for AVs is already described by Lippert and Winner [7]. The basis is the externally observable behavior of an AV while driving through a scenery. Based on this behavior, it can basically be determined how safe or unsafe an AV behaves during operation. This behavior is significantly influenced by the scenery, which imposes behavioral limits on an AV when combined with the applicable traffic rules. For example, a stop sign means that an AV must stop at the associated stop line before driving any further. It does not matter whether or not priority is given to another traffic participant during the stop. With the help of the BSSD, these behavioral limits are explicitly linked to the scenery and represented in a map. In this way, route-based behavioral requirements can be derived that result from the behavioral limits of the BSSD. The behavioral demands serve the present work as a basis for identifying the capability-based routes. Therefore, the BSSD and subsequently the derivation of route-based behavioral requirements is explained in the following.

Behavior-Sematic Scenery Description [8, 9]

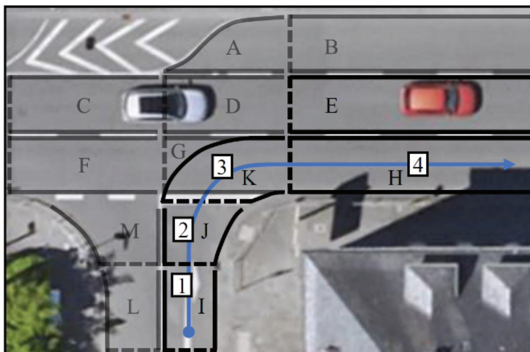
The BSSD represents the legal behavioral limits based on scenery and traffic rules. These so-called *behavioral demands* are represented using directional *behavior spaces*, which usually describe a lane segment and within which the behavioral demands do not change. A behavior space is always described with four behavioral attributes that reflect the behavioral demands: *Speed* (S), *Boundary* (B), *Reservation* (R), and *Overtake* (O). Each of these attributes has additional properties that concretize the behavioral demands. Speed describes any limitations on travel speed. Boundary describes limitations when crossing the boundaries of the behavior spaces. Behavioral demands regarding priority and residence within the behavior spaces is described by the reservation attribute. Permission or prohibition of overtaking is represented by the attribute overtake.

For illustration, a behavior space is concretely considered using a real scenery. Figure 1 shows an aerial view of a real scenery in Darmstadt, Germany, in the upper left. The remaining part of the figure is ignored for the explanation of the behavior space. The behavior spaces are labeled with capital letters. Each behavior space has exactly one longitudinal (entry) boundary (black dashed line) and two lateral (exit) boundaries (black solid line). For each behavior space, there is an inverted behavior space that represents the behavioral demands for the opposite direction of travel (only one direction shown here). The principle of the behavior space is explained using behavior space K, which represents a right turn at a T-intersection with a stop sign. A speed limit of 30 km/h applies in this area. This limit is stored in the speed attribute. Before entering the behavior space, the vehicle must stop due to the stop sign. Therefore, the attribute boundary contains the behavioral demand *stop* for the longitudinal boundary. Laterally, the behavior space should not be left when turning, so the behavioral demands for the lateral boundaries prohibit crossing. For example, between behavior spaces H and E is a dashed lane marking, so the behavioral demands for these behavioral boundaries allow passing at this point. Behavior space K may only be traversed if priority is given to traffic participants approaching from F and C. Therefore, the attribute reservation shows *externally-reserved* for the road user types *motor vehicle* and *bicyclist* potentially coming from C and F (these *reservation links* are not shown here). Overtaking is allowed in the behavior space, so the attribute overtake is set to the value *yes*. In the same way, all other behavior spaces are defined. The concrete, remaining properties of the four behavioral attributes are not further relevant in the context of this work, which is why they are not discussed in more detail.

Route-Based Behavioral Requirements [7]

Lippert and Winner show how behavioral requirements are derived based on BSSD. For this purpose, they first define the terms *global* and *local behavioral requirements*. Global behavioral requirements apply everywhere regardless of the scenery and are, for example, requirements regarding collision avoidance (e.g. [10]). Local behavioral requirements, on the other hand, have a local scope and are scenery-specific. Thus, they do not apply everywhere, but only at specific locations within a scenery. The derivation of the behavioral requirements refers only to the local behavioral requirements, which is why they are simply called behavioral requirements in the following. According to Lippert and Winner, the complete behavioral requirements result only from a concatenation of the behavior spaces. Thus, lane-accurate routes are considered that concatenate the behavior spaces of the BSSD and thus generate a sequence of the different behavioral demands. The concatenation of the different behavioral demands can lead to new behavioral demands depending on the *transition* (*longitudinal* or *lateral*). Based on the behavioral spaces concatenated within a route, behavioral requirements are derived for each behavioral space and the resulting behavioral demands.

Figure 1 shows the resulting behavioral requirements for an example concatenation of behavior spaces. The lane-accurate route leads from behavior space I to behavior space H. A total of four behavior spaces are traversed, of which behavior space K represents the right turn considered earlier. The concatenated behavior spaces E_i are numbered sequentially. For each behavior space E_i, the associated transition $T_{i-1,i}$, the behavioral demands D_i of the BSSD, and the behavioral demands $D_{i-1,i}$ resulting from concatenation are shown. For each behavior space, a speed limit of 30 km/h applies. For the transitions between the behavior spaces, only when entering E_3 there is a condition that the AV must stop before proceeding. As described in the previous example, the AV must additionally give priority to motor vehicles and cyclists at E_3 and must not obstruct them. From the concatenation it additionally follows, among other things (not shown here), that the traffic participants entitled to priority must be indicated that priority is actually granted to them. In the other behavior spaces, the AV does not have to grant priority, which is why the reservation attribute shows *own-reserved* (own). Overtaking is allowed in any behavior space (yes). In the lower part of the figure, the resulting behavioral requirements are listed and marked with a cross according to their validity in each behavior space. This brief example is used to understand the route-based behavioral requirements as a basis for capability-based routes, and therefore it is not elaborated further.



E_i	$i = 1$	$i = 2$	$i = 3$	$i = 4$
$T_{i-1,i}$		longitud.	longitud.	longitud.
D_i	S: 30 km/h R: own O: yes	S: 30 km/h B: allowed R: own O: yes	S: 30 km/h B: stop R: ext. O: yes	S: 30 km/h B: allowed R: own O: yes
$D_{i-1,i}$		-	R: indication of giving priority	-

Requirements of E_i	$i = 1$	$i = 2$	$i = 3$	$i = 4$
SR1: The AV shall not exceed the maximum permissible speed limit.	x	x	x	x
BR1: The AV shall stop at the longitudinal boundary before proceeding.			x	
RR1: The AV shall not obstruct traffic participants with reservation entitlement for the space.			x	
RR1.1: The AV shall indicate in advance by adjusting the driving speed reasonably that it will give priority to traffic participants who have priority.			x	

Figure 1. Example for behavior spaces and corresponding route-based behavioral requirements based on [7]. SR: Speed Requirement; BR: Boundary Requirements; RR: Reservation Requirement.

CONCEPT OF MATCHING REQUIREMENTS AND CAPABILITIES

The main goal of this work is to identify routes that can be accomplished by AVs based on their driving capabilities. Consequently, the driving requirements of the route must not exceed the driving capabilities of the vehicles. In order to identify an exceeding of driving capabilities it is necessary that they can be matched with the driving requirements of the route. This matching determines whether the route can be mastered by an AV or not. In order to match, the requirements and capabilities must be compatible with each other. This means that for each driving requirement there must also be a corresponding driving capability. It is important that the driving capabilities of the AVs can be proven. According to the state of the art, the proof of driving functions is typically achieved with the help of various tests [11]. Therefore, it is assumed that test certificates exist for driving capabilities that have been tested and thus proven. If a driving capability has been successfully tested and proven, a test certificate exists for this driving capability. Whether the driving capabilities really meet the driving requirements is determined by the matching process. This process requires matching criteria that determine a match based on appropriate metrics. The following argumentation results from these considerations:

Let DR_i be the set of necessary driving requirements in order to drive in a concatenated behavior space E_i. Furthermore, let $\bigcup_{m=1}^{n_{\mathrm{DC}}} DC_m$ be the superset of $n_{\mathrm{DC}} \in \mathbb{N}$ sets of driving capabilities of an AV that is proven with a corresponding superset of test certificate sets $\bigcup_{m=1}^{n_{\mathrm{DC}}} TC_m$ and compatible with DR_i. Then an AV shall only be allowed to drive in E_i if at least one set of proven driving capabilities DC_m matches the set of driving requirements DR_i. A match of DR_i and DC_m is determined with a matching process based on matching criteria.

Therefore, for matching to be possible, the following conditions must be met:

- For each set of driving requirements DR_i of a concatenated behavior space E_i, there is at least one compatible set of driving capabilities DC_m.
- The set of driving capabilities DC_m must be provable using tests in order to provide an associated set of test certificates TC_m for AVs.
- There are matching criteria for identifying matches between driving requirements of DR_i and driving capabilities of DC_m.

Consequently, driving requirements and driving capabilities must be defined in a way that allows for matching. This matching is performed based on matching criteria. Thus, the identification of drivable behavior spaces is enabled. Specifications for driving requirements and driving capabilities as well as matching criteria are developed below.

Specification of Driving Requirements

So far, only behavioral requirements have been derived based on the BSSD. These behavioral requirements can be used for matching with driving capabilities, but there is a problem with the level of abstraction. If only the behavioral information from the BSSD is used, potentially many driving requirements are identical, even though significantly different requirements are imposed on the dynamic driving task (DDT) and thus on the driving capabilities. A simple example of this is the behavioral requirements of the behavioral attribute reservation (cf. Figure 1). The requirements only demand that priority can be granted to certain types of road users from certain areas and that the granting of priority is indicated accordingly. These behavioral requirements potentially apply to many different intersections, since no reference to the geometry of the scenery has been made so far. Geometric information such as position, dimensions, or orientation of relevant areas would solve this problem. For example, right-before-left priority intersections with different relative orientations of their intersection arms would be distinguishable. Based only on the behavioral requirements, they would be the same.

Thus, the goal is to identify necessary geometry-based specifications for the driving requirements that form a basis for matching with driving capabilities. For this purpose, the relevant scenery properties are divided into specification categories. The specification category serves as a container for the concrete specifications of behavior spaces. For the demonstration of the methodical procedure of this work, the specification and subsequent work steps are performed and described on the basis of the reservation behavioral requirements. The reservation behavioral requirements are chosen because they have a high complexity compared to other behavioral requirements. Therefore, they are particularly suitable for demonstrating the procedure. The reservation behavioral requirements RR1 and RR1.1 (cf. Figure 1) are considered in the following from the DDT perspective referring to a considered behavior space with these requirements.

Probably the most obvious specification category is the type of traffic participant. This information already exists explicitly within the behavior space in the BSSD, but must still be included in the specification. Otherwise, it would not be explicitly specified that a corresponding driving capability must meet this specification. Different road user types require different capabilities of an AV, as they must not simply be recognized as a dynamic object, but must necessarily be classified according to the reservation as well. Reservation requirements necessitate this classification, as it determines which traffic participants are entitled to a reservation.

The speed limit of the traffic participants entitled to reservation is also relevant for the specification of the requirements. The capability to perceive these traffic participants must be tested and demonstrated based on different speeds of movement. It may well make a difference whether traffic participants are potentially approaching at 30 km/h or 50 km/h. In this case, a different behavior is demanded of the complete automation chain, which must be explicitly demonstrated. For this purpose, it is additionally necessary to include the speed limit of the AV before entering the behavior space under consideration. Different relative speeds between AV and the other traffic participants require explicit proof and therefore explicit specification for the same reason.

The perception of traffic participants entitled to reservation is only successful if they are also sensed in the relevant areas of the scenery. For this purpose, the direction of origin of the traffic participants entitled to reservation is explicitly stored in the BSSD. Based on the BSSD map, the absolute positions of the linked areas are thus available. However, only the direction of origin is indicated by the BSSD and not the entire area of origin potentially to be monitored. It is necessary for the specification to define the position of these relevant origin areas. Thereby, the information about the road course from the respective direction of origin should not be lost, so that potential paths along which reservation-authorized traffic participants move can be represented.

According to the above mentioned requirements an AV shall not travel through the considered behavior space if any other traffic participant with reservation-entitlement is present. This applies both to traffic participants who are already in the behavior space under consideration and to traffic participants who want to enter the space. In addition, if an AV is in the space itself, it must leave the space as soon as possible. Differences in the fulfillment of these requirements arise from the geometry of the considered behavior space. Different lengths of the behavior space mean different distances that the AV must travel through. But also different curved shapes of the behavior space possibly influence the driving behavior of the AV. Therefore, it is necessary that these geometric properties are part of the requirements specification.

In addition to the geometry of the behavior space under consideration, the geometry of the preceding behavior space(s) is also relevant. Depending on which curvature is present in the previously concatenated behavior spaces, for example, the AV will approach the considered behavior space with a different orientation. Depending on the orientation, the relevant perceptual areas for the relevant traffic participants differ. However, since based on the requirement specification it is not yet specified how exactly the vehicle aligns in the behavior spaces, the orientation of the AV cannot be part of the requirement specification. Rather, the orientation of the behavior spaces must be considered. The design and proof of the specific driving capabilities can thus be unrestricted, so that the actual orientation of the AV in the behavior space is defined in the development process.

Finally, if the behavior space under consideration is highly curved, as is the case with behavior spaces for turning, occlusion may occur. Depending on the position of potentially present planted areas, walls or buildings, the area to be monitored might not be completely visible or only visible at a late stage. If such occlusion is present, the driving behavior must be adjusted so that the reservation requirements are not violated. For example, depending on the type of occlusion, it may be necessary for an AV to slowly move into the considered behavior space so that the relevant areas can be observed. These cases have to be specified explicitly, since an extra proof has to be provided accordingly.

Overall, the following specification categories result for the reservation requirements RR1 and RR1.1:

- Type of relevant traffic participant
- Speed limit of relevant traffic participant
- Geometry and position of relevant area of origin
- Geometry and position of considered behavior space
- Geometry and position of relevant area of preceding behavior space(s)
- Speed limit of relevant preceding behavior space(s)
- Geometry and position of relevant area of occlusion

Specification of Driving Capabilities

What do the capabilities look like to enable matching? Simple driving capabilities that are tailored exactly to the driving requirements may be used. In this way, a capability always meets exactly one requirement. The following generic example illustrates the relationship between driving requirement and driving capability:

- Driving requirement: *The AV shall/ shall not perform a certain action under certain conditions.*
- Driving capability: *The AV is capable of performing/ not performing a certain action under certain conditions.*

The advantage of this very direct matching is that the capabilities fit the requirements in every case. There is no need for reasoning that assigns different capabilities to single requirements. This leads to the fact that the driving capabilities can be addressed by different vehicle-specific solutions. However, the proof of the capabilities must then be vehicle specific. In this way, driving capabilities are defined universally and uniformly without excluding specific technical solutions or developments. For this reason, this direct assignment appears not only intuitive but also practicable with regard to a uniform specification of driving capabilities.

The alternative to this approach is to further decompose the driving capabilities. It is possible to break down the driving requirements to subsets of the DDT. The result is a set of capabilities that contribute to the main capability being met at the behavioral level. Various approaches are suitable for decomposing the capabilities. Basically, at the beginning of the decomposition, the decision has to be made how fine granular to decompose. In order to remain as abstract and generic as possible, the Sense-Plan-Act paradigm [12] is suitable, since it only provides for a decomposition of the DDT on three layers. For a more detailed decomposition, which allows a deeper analysis of the capabilities, the six layers of functional decomposition according to Amersbach and Winner [13, 14] or skill graphs introduced by Reschka et al. [15] are suitable.

However, these approaches have two fundamental drawbacks in terms of matching requirements to capabilities. First, the complexity of the functional relationships between the individual sub-capabilities increases with each step of the decomposition. Second, there is a certain arbitrariness in the choice of decomposed sub-capabilities. Both phenomena lead to the fact that selected capabilities can no longer be assigned to requirements in a simple way. Another problem is that the choice of partial capabilities often already assumes technical solutions, such as concrete vehicle setups. This would lead to the need to specify different capabilities for different AVs. However, the present approach aims to be as solution-neutral as possible and thus independent of a concrete vehicle specification. This increases the universal applicability of this approach. A solution to these problems is not known to the authors of this work. Nevertheless, if these problems are mitigated or even eliminated by appropriate approaches, capability decomposition would be another option for the matching process.

Due to the difficulties pointed out for the decomposition of driving capabilities, the method of driving capabilities analogous to driving requirements presented before is chosen in the present work. This is done by reformulating the requirements into capabilities as shown. In order to ensure that the comparison between driving requirements and driving capabilities contributes to a statement about the drivability of the behavior spaces, matching criteria must be defined that are as clear as possible. For this purpose, the aforementioned considered reservation requirements are chosen, which have a high potential for route-specific differences. Thus, the reservation requirements RR1 and RR1.1 are selected. For these requirements, the analogous reservation capabilities (RCs) are formulated and the derived specification categories are assigned:

- RC1: The AV is capable of avoiding obstructions of traffic participants with reservation entitlement for the behavior space.
- RC1.1: The AV is capable of indicating in advance by adjusting the driving speed reasonably that it will give priority to traffic participants who have priority.

Matching Criteria

For the matching between the driving requirements RR1 and RR1.1 and the driving capabilities RC1 and RC1.1, each specification category is considered individually. To ensure that the capabilities meet the requirements, the matching criteria of all specification categories must be met. To illustrate the reasoning, a X-intersection in a 30 km/h speed zone is considered in Figure 2. In Germany, the right-of-way rule "right-before-left" applies at such intersections. This means that traffic participants coming from the right from the perspective of a vehicle entering the intersection have priority. Additionally, left-turners must generally give priority to oncoming traffic. The *geometry and position of considered behavior space* as well as the *geometry and position of relevant area of occlusion* will be neglected in the following. The remaining specification categories are sufficient for an evaluation of the overall approach in a first implementation, since sufficient variations are to be expected. Additionally, areas reserved by pedestrians are not considered. The focus of the matching criteria is placed on areas of origin for motor vehicles, bicyclists and rail vehicles.

The specification categories of the reservation requirements have dependencies that must be considered in the nomenclature of the matching criteria. The following variables are introduced and partially shown in Figure 2a:

- For each externally-reserved behavior space E_i, there exist $n_{\text{orig},i} \in \mathbb{N}$ areas of origin $\left(A_{\text{orig},i,k}\right)_{k=1,2,\ldots,n_{\text{orig}}}$ (orange areas).
- Each area of origin $A_{\text{orig},i,k}$ is associated with a set of traffic participant types $P_{i,k}$.
- For the set of traffic participant types $P_{i,k}$ assigned to an area of origin $A_{\text{orig},i,k}$, there is a maximum speed limit $v_{\text{lim,orig},i,k}$ (speed limits of road user types within an area rarely differ).
- The relevant speed limit $v_{\text{lim,pre},i}$ of the AV for approaching the considered behavior space E_i is assigned to the relevant area of preceding behavior space(s) $A_{\text{pre},i}$ (blue area).

<u>**Type of traffic participant**</u> The matching criterion for the type of traffic participant is based on a nominal scale. This criterion is satisfied only if a successfully proven set of traffic participant types P_{proof} related to the associated proven area of origin matches the required set of traffic participant types $P_{i,k}$. The following matching criterion results:

$$P_{\mathrm{proof}} = P_{i,k} \qquad \text{(Equation 1)}$$

Possible traffic participant types are *motor vehicle*, *pedestrian*, *bicyclist*, and *rail vehicle*. Therefore, the following closed set is defined for the road user types: $P \in \{\text{motor vehicle, pedestrian, bicyclist, rail vehicle}\}$.

Speed limit of relevant traffic participant For the speed limit of relevant traffic participant types, the maximum speed limit within the set of traffic participant types $P_{i,k}$ is chosen. The following assumption is made: If an AV has a successful proof of granting priority to a traffic participant coming from $A_{\mathrm{orig},i,k}$ with a speed limit $v_{\mathrm{lim,orig,proof}}$, then it is able to grant priority even with equal or lower speed limits $v_{\mathrm{lim,orig},i,k}$ of traffic participants. The following matching criterion results:

$$v_{\mathrm{lim,orig,proof}} \geq v_{\mathrm{lim,orig},i,k} \qquad \text{(Equation 2)}$$

Speed limit of relevant preceding behavior space(s) For the speed limit within the relevant area of preceding behavior space(s) $A_{\mathrm{pre},i}$, the maximum speed limit $v_{\mathrm{lim,pre},i}$ among these behavior space(s) is chosen conservatively. The following assumption is made: If an AV has a successful proof of the required capabilities with a speed limit $v_{\mathrm{lim,pre,proof}}$ for approaching, then the proof is valid even for equal or lower speed limits $v_{\mathrm{lim,pre},i}$. The following matching criterion results:

$$v_{\mathrm{lim,pre,proof}} \geq v_{\mathrm{lim,pre},i} \qquad \text{(Equation 3)}$$

Geometry and position of relevant area of preceding behavior space(s) and relevant area of origin The matching criteria for the geometries and positions of the different areas are considered together. Basically, the relevant areas of origin must always be considered relative to the relevant area of preceding behavior space(s). This is because an AV approaches the considered behavior space within the area of preceding behavior space(s) and meanwhile already has to execute the DDT to grant priority. Accordingly, the relative positions of the relevant areas of origins to the AV's approach are crucial. One way of matching successfully proven and required combinations of the relevant areas is to superimpose the areas based on an equal reference system. Thus, an overlap of the matched areas can be identified. The same principle of this matching can be applied using a geometric parameterization of the relevant areas. The advantage here is a simpler and more efficient identification of the geometries of the areas as well as the matching itself. Therefore, with regard to the application of the matching criteria, a geometric parameterization of the relevant areas is performed. For this purpose, the following assumption is made: *The relevant road areas can be approximated by rectangles*.

Intersections are scenery components that predominantly contribute to externally-reserved behavior spaces. In urban areas, intersections and junctions are designed so that the associated road arms are straight with sufficient distance to the intersection. This must be taken into account during the design and construction of roads by ensuring that all intersection accesses (in the sense of sufficient distance) are identifiable in good time [16, p. 109]. The extensive implementation of this layout principle can be easily confirmed by looking at suitable aerial images, such as those from Google Earth [17]. Additionally, it is noticeable that the lane or road widths do not change significantly in the areas around the intersections. Based on these findings, the assumption made is retained. However, it must be assumed that there are exceptions that are not correctly represented due to this assumption (cf. Discussion of Results).

With this assumption, the following simplifications result:

- The alignment of the relevant areas is determined by longitudinal and lateral offsets and a constant angle relative to each other.
- The geometry of the relevant areas need only be specified by a constant width. The length is no longer needed, since if the alignment - and thus the course - of the area is known, the length along this course can be chosen according to the associated speed limits for a proof. The proof thus confirms the driving capability regardless of the length of the area.
  Note 1: In the case of a non-rectilinear course of the areas, the length is relevant because the alignment changes along the course.
  Note 2: This simplification does not apply to the identification of occlusion, since the complete geometry of the relevant area is required for this process.

Therefore, the relevant areas can be parameterized as shown in Figure 2b. Each area is characterized by a width:

- Width $w_{\mathrm{orig},i,k}$ of relevant area of origin $A_{\mathrm{orig},i,k}$
- Width $w_{\mathrm{pre},i}$ of relevant area of preceding behavior space(s) $A_{\mathrm{pre},i}$

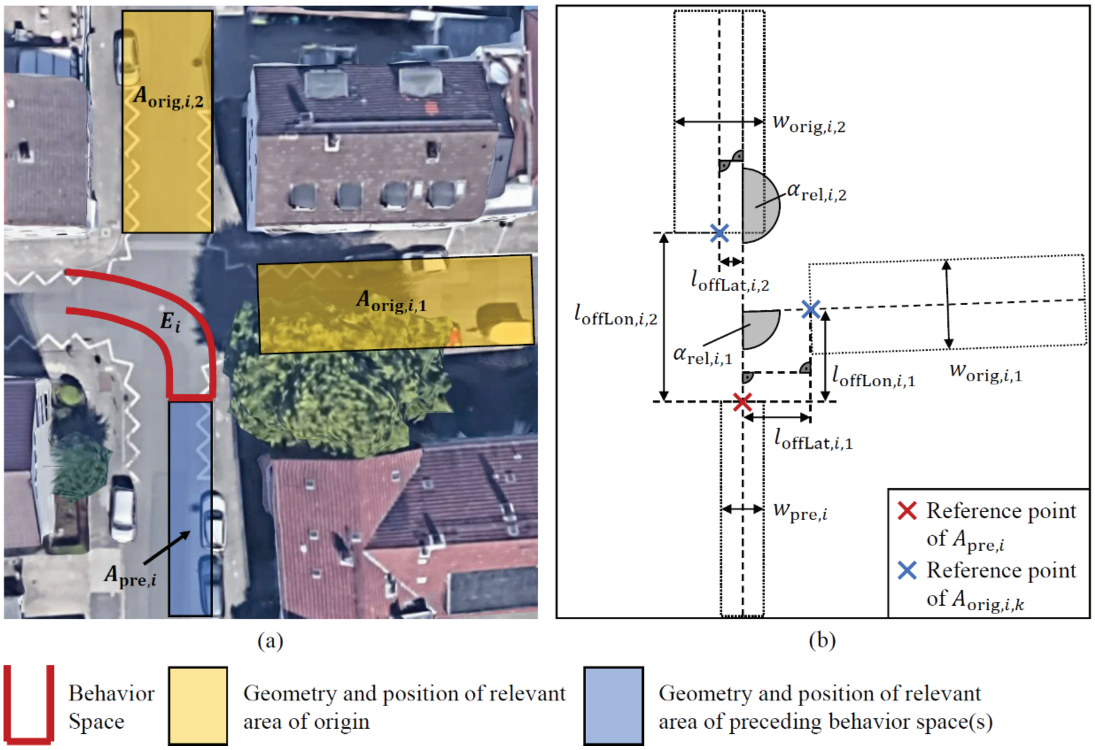


Figure 2. Specification for externally-reserved behavior spaces (Aerial image ©2022 Google/AeroWest).

The alignment of the areas is defined using reference points and relative angles. The reference points result from the geometric analysis of the scenery based on the BSSD map. They mark the end of the areas in the direction of the intersection. They form the crossing point between the virtual center line of an area and the transition line to the intersection area. All areas are positioned relative to the reference point of the relevant area of preceding behavior space(s) (red cross). In this way, all relevant areas are represented relative to the approach area of the AV. For this purpose, the reference points of these areas (blue cross) are defined using the longitudinal and lateral offsets $l_{\text{offLon},i,k}$ and $l_{\text{offLat},i,k}$ relative to reference point (red cross) and virtual centerline of $A_{\text{pre},i}$. The orientation of the areas is defined by a relative angle $\alpha_{\text{rel},i,k}$ describing the angle between the virtual centerlines of $A_{\text{pre},i}$ and $A_{\text{orig},i,k}$. With the help of this parameterization, a matching is made possible, for which the matching criteria are defined below.

Basically, all parameters are matched. Only if all required parameters are within the tested parameter spaces (index: proof), the capabilities are considered sufficient. The following assumptions are made:

- *Widths of the areas*: If the AV has a successful proof for a certain width, then this proof is also valid for smaller widths. This assumption results from the consideration that a smaller area to be observed or driven on within the proven range is equally covered as a subset.
- *Offsets and relative angles between the areas*: If the AV has a successful proof for a range of values of these parameters, then this proof is valid for all values within this range. Also in this case, the different parameters within the range of values are a subset of the proven values.

The following matching criteria result:

- Width of $A_{\text{pre},i}$:

$$w_{\text{pre,proof}} \geq w_{\text{pre},i} \qquad \text{(Equation 4)}$$

- Width of $A_{\text{orig},i,k}$:

$$w_{\text{orig,proof}} \geq w_{\text{orig},i,k} \qquad \text{(Equation 5)}$$

- Longitudinal and lateral offset between $A_{\text{pre},i}$ and $A_{\text{orig},i,k}$:

$$l_{\text{offLon,proof,min}} \leq l_{\text{offLon},i,k} \leq l_{\text{offLon,proof,max}} \qquad \text{(Equation 6)}$$

$$l_{\text{offLat,proof,min}} \leq l_{\text{offLat},i,k} \leq l_{\text{offLat,proof,max}} \qquad \text{(Equation 7)}$$

- Relative angle between $A_{\text{pre},i}$ and $A_{\text{orig},i,k}$:

$$\alpha_{\text{rel,proof,min}} \leq \alpha_{\text{rel},i,k} \leq \alpha_{\text{rel,proof,max}} \qquad \text{(Equation 8)}$$

CAPABILITY-BASED ROUTING

In this section, the elaborated matching concept is used to identify capability-based routes. Capability-based routes are routes with driving requirements that do not exceed the driving capabilities of the AVs. Thus, the driving requirements of the scenery potentially lead to routes that deviate from conventionally planned routes. Therefore, a novel route search must be developed that takes into account this new criterion of *capability-exceeding* of AVs. Conventional route planners do not accomplish this so far, but are still suitable as a basis for this approach.

Concept of Capability-Based Route Search

A route planner searches for routes based on selected optimization criteria [18]. Conventional optimization criteria are the shortest or even the fastest route. For this purpose, the road network is typically divided into edges and nodes. Nodes are equivalent to intersections or junctions and represent the connection points of edges. Edges represent the individual road segments that are connected via the nodes. For the planning of routes within traffic networks, however, this division between nodes and edges is not always appropriate. In order to consider concrete turns or lane changes and thus lane-accurate paths within the network, this division is reversed. In so-called *edge based routing*, finer granular road sections (e.g. lane sections) become nodes and pairs of adjacent road sections become edges [19, p. 127]. In this way, lane-accurate route planning is enabled. Depending on the selected criterion, the edges of the road network are weighted with different costs based on a cost function. From a starting node to a destination node, there are different paths depending on the network size, alternating nodes and edges. As a result, the combination of edges with the lowest total cost represents the optimal route with respect to the selected criterion.

The BSSD road network does not necessarily consist of a graph that is suitable for routing. For the capability-based route search, a preprocessing is necessary to transform the BSSD road network into a graph. In this process, the lateral and longitudinal connections of the individual behavior spaces are explicitly represented as edges and the behavior spaces themselves as nodes. In this conversion process, the explicit BSSD information is lost. However, this loss of information is intentional, since the road network should be reduced to the minimum necessary information for efficient route search. With the help of a cost function the explicit BSSD information is transferred into the edge weighting. The result is a routing graph that enables explicit routing based on the weighted edges without exceeding the driving capabilities of AVs. For the identification of capability-based routes, a conventional route search algorithm is needed in addition to the weighted routing graph. This searches for the route with the lowest total costs within the weighted routing graph based on a starting point and a destination point. Since the route search algorithm is state of the art, the focus in this section is on the generation of the weighted routing graph. This is crucial for route identification since it defines the routing cost. The route search algorithm simply sums up the costs based on the routing graph and selects the edge combination with the minimum costs. The following modules are developed to create the routing graph:

- *Capability-Based Cost* module: Calculates the edge weights of the routing graph based on the matching results (further described in the following).
- *Matching* module: Matches driving requirements with existing driving capabilities based on the defined matching criteria (not further described in the following).
- *Requirement Generation* module: Generates the driving requirements of the concatenated behavior spaces based on the defined specification (not further described in the following).

The principle of finding an optimal route from origin to destination also applies to capability-based route search. A new optimization criterion is needed for the intended function of identifying routes that are feasible for AVs. This does not mean that the conventional criteria must be discarded. Even if a new criterion is used based on the new search function, the determined route should still be optimal with respect to conventional vehicle navigation. Accordingly, the route found should be the shortest or fastest possible despite further optimization criteria, for example. The basis of the new route planning should therefore be based on conventional route planning, so that these criteria and cost functions can be adopted. For the new criterion, however, a different or adapted cost function is required to weight the edges. In principle, there are two extreme forms of edge weighting. A weight can become minimal in the optimal case, i.e., theoretically assume a value of zero. The other extreme is an infinitely high weight assigned to edges that are maximally far from an optimum based on the evaluation of the cost function. Depending on the cost function, all other values are conceivable within these extremes. Negative costs are not allowed. From the previous findings, it is clear that the new cost function must be based on the previously presented matching of driving requirements and driving capabilities. Since the matching is performed for concatenated behavior spaces, it is suitable for *edge based routing*. The matching can either fail - driving capabilities are exceeded - or succeed - driving capabilities are not exceeded. Therefore, a cost function must be defined for the transitions between concatenated behavior spaces, following the matching concept.

The cost function to be defined must produce only two values. In the case of exceeding, the cost must be maximum so that the behavior space under consideration is excluded from the planning. In the case of a successful match, the cost must be minimal. In order to identify a successful match, the matching criteria of the specification categories of driving requirements and driving capabilities are applied. Thus, the matching cost function $c_{\text{match},i}$ for a concatenated behavior space E_i is defined as a function based on the set of driving requirements DR_i and all available, certified sets of driving capabilities $\bigcup_{m=1}^{n_{\text{DC}}} DC_m$ as follows:

$$c_{\text{match},i} = f\left(DR_i, \bigcup_{m=1}^{n_{\text{DC}}} DC_m\right) \qquad \text{(Equation 9)}$$

The matching cost function has the following range of values according to the considerations:

$$c_{\text{match},i} = \begin{cases} 0 & \text{for not exceeding the driving capabilities} \\ \infty & \text{for exceeding the driving capabilities} \end{cases} \qquad \text{(Equation 10)}$$

To avoid that the matching cost function $c_{\text{match},i}$ interferes with the conventional cost function $c_{\text{conv},i}$ in an undesired way, the cost functions have to be separated unambiguously. This means for the total cost function $c_{\text{tot},i}$:

- As long as driving capabilities are not exceeded, the total cost function $c_{\text{tot},i}$ is determined by the conventional cost function $c_{\text{conv},i}$.
- Once exceeding is identified, the total cost function $c_{\text{tot},i}$ is determined by the matching cost function $c_{\text{match},i}$.

Using the binary range of values of the matching cost function $c_{\text{match},i}$, these requirements can be realized via a simple addition of the cost functions. This results in the following total cost function $c_{\text{tot},i}$ for E_i:

$$c_{\text{tot},i} = c_{\text{conv},i} + c_{\text{match},i} \qquad \text{(Equation 11)}$$

Since the routing graph is generated and weighted in a preprocessing as described above, it is practical to remove edges with an infinite weight directly in this process. Thus, the route search algorithm does not have to visit these edges at all resulting in a more efficient calculation. This is taken into account in the implementation. For very large road networks with frequently updated data, an on-the-fly calculation of the edge weights within the iterations of the routing algorithm would also be suitable. This way, the entire road network would not always have to be preprocessed. Since the road network considered for the implementation in this work is rather small, this approach is not pursued further.

Implementation

The described concept of capability-based route planning is implemented based on the high-definition map framework Lanelet2 [20] for which a BSSD map instantiation was developed by Lippert et al. [9]. Besides providing a map format, another advantage of this framework is the availability of a comprehensive C++ software library for handling Lanelet2 map data [21]. The following software modules are used from this library as a basis for capability-based route planning:

- *lanelet2_core*: Basic module for handling Lanelet2 maps as well as all related primitives like points, linestrings, or lanelets (= lane sections). Extensive functions for geometric calculations are provided.
- *lanelet2_io*: Module for reading and writing Lanelet2 maps.
- *lanelet2_traffic_rules*: Module for interpreting selected traffic rules in Lanelet2 maps such as passability or speed limits of lanelets based on country and road user type. This module is used in this work as input for the generation of a routing graph, since traffic passable lanelets must be known for this purpose.
- *lanelet2_routing*: Module for route search within Lanelet2 maps. This module generates a routing graph based on conventional optimization criteria such as shortest or fastest path. Within the routing graph the optimal route is subsequently searched. It is possible to modify and extend the cost function as desired. To ensure that only passable lanelets are used from the point of view of traffic rules, the *lanelet2_traffic_rules* module is included in addition to the cost function to generate the routing graph. Thus, a basic function is already given, which generates routes that are correct from the traffic point of view based on the conventional optimization criteria.

For the handling of the BSSD data within the Lanelet2 framework the *BSSD data handler* is used, which was developed especially for this purpose. When reading the maps, this handler parses the BSSD data and makes it available to the software environment according to the generic BSSD structure [9]. In this way, arbitrary queries regarding all BSSD information are made possible. The *capability-based cost*, *matching* and *requirement generation* modules are implemented as part of the *lanelet2_routing* module.

Application in a Real-World Road Network

For the demonstration of capability-based route search based on a BSSD map, a road network specifically selected for this purpose is used. The road network covers large parts of Darmstadt's city center and includes both large, multi-lane roads and small residential streets. In doing so, the network connects specific intersections that were selected based on their different characteristics in terms of behavioral demands and geometry. These intersections therefore form the key component of the network for the following application.

Specific sets of reservation capabilities are first defined as the basis for the routing application. These capabilities are exemplary and serve to demonstrate the approach developed in this work. Since many intersections do not require priority to be granted when turning right, the capabilities are defined based on parameter values for straight and left turns. Therefore, when traveling on the road network under consideration, some intersections do not result in reservation requirements. However, as soon as an intersection is to be crossed straight ahead, it may be necessary to give priority from the right in $30\ \mathrm{km/h}$ speed zones. Therefore, the first capability set $DC_{\mathrm{reserv,1}}$ is designed to grant priority from the right ($\alpha_{\mathrm{rel,proof,min|max}} = 80°|100°$). This set is valid for reservation-entitled road user types motor vehicle and bicyclist and speed limits of $30\ \mathrm{km/h}$. As shown in the considered data, right is not necessarily equal to right. Thus, $DC_{\mathrm{reserv,1}}$ is extended by a second capability set $DC_{\mathrm{reserv,2}}$, which includes an additional angle range for traffic from the right ($\alpha_{\mathrm{rel,proof,min|max}} = 50°|70°$). In order to use the road network even more efficiently, another capability set $DC_{\mathrm{reserv,3}}$ is defined, which specifically includes granting priority from the front ($\alpha_{\mathrm{rel,proof,min|max}} = 170°|190°$). Compared to the previously defined capability sets, the longitudinal and lateral offsets differ. This is necessary because the offsets depend on the angular orientation of the areas. Last, the capability set $DC_{\mathrm{reserv,3}}$ is extended to a new set $DC_{\mathrm{reserv,4}}$ so that rail vehicles are covered in addition to motor vehicles and bicyclists. Additionally, the covered speed limits are increased to $50\ \mathrm{km/h}$. The resulting driving capability sets including the defined specification parameters are summarized in Table 1.

Table 1. Specified reservation capabilities (MV: motor vehicle | B: bicyclist | RV: rail vehicle).

Parameter	$\boldsymbol{DC_{\mathrm{reserv,1}}}$	$\boldsymbol{DC_{\mathrm{reserv,2}}}$	$\boldsymbol{DC_{\mathrm{reserv,3}}}$	$\boldsymbol{DC_{\mathrm{reserv,4}}}$	**Unit**
$v_{\mathrm{lim,pre,proof}}$	30	30	30	50	km/h
$w_{\mathrm{pre,proof}}$	4	4	4	4	m
P_{proof}	$\{\mathrm{MV, B}\}$	$\{\mathrm{MV, B}\}$	$\{\mathrm{MV, B}\}$	$\{\mathrm{MV, B, RV}\}$	-
$v_{\mathrm{lim,orig,proof}}$	30	30	30	50	km/h
$l_{\mathrm{offLon,proof,min\vert max}}$	$3 \mid 10$	$3 \mid 10$	$10 \mid 30$	$10 \mid 30$	m
$l_{\mathrm{offLat,proof,min\vert max}}$	$3 \mid 10$	$3 \mid 10$	$-10 \mid 0$	$-10 \mid 0$	m
$\alpha_{\mathrm{rel,proof,min\vert max}}$	$80 \mid 100$	$50 \mid 70$	$170 \mid 190$	$170 \mid 190$	°

Based on the defined driving capabilities, the capability-based route search is applied to the present road network. In order to demonstrate the functionality of the implemented route search, the start and destination points are defined with respect to diverse route options. For this purpose, the route planning should have different route options that require different driving capabilities. The following route requests are performed (driving capability set m corresponds to $DC_{\mathrm{reserv},m}$):

- Route 1: Shortest route without driving capability sets
- Route 2: Shortest route with driving capability set 1
- Route 3: Shortest route with driving capability set 1 and 2
- Route 4: Shortest route with driving capability set 1, 2 and 3
- Route 5: Shortest route with driving capability sets 1, 2, 3, and 4

Figure 3 shows the road network with selected start and destination points and the calculated routes. For a clear representation of the different routes, the road network is shown twice. The presented road network is colored yellow and the detailed modeled intersections within it are highlighted in magenta. Start and destination points are shown with corresponding markers. As a reference for the calculated routes, the shortest route based on the conventional optimization criterion *shortest path* of the Lanelet2 algorithm is shown in green. The focus will be on the intersections modeled in detail, so *all other (non-magenta) intersections* in the road network will be handled *without* a capability-based matching. Route sections without the modeled intersections therefore always represent the shortest path based on the conventional optimization criterion.

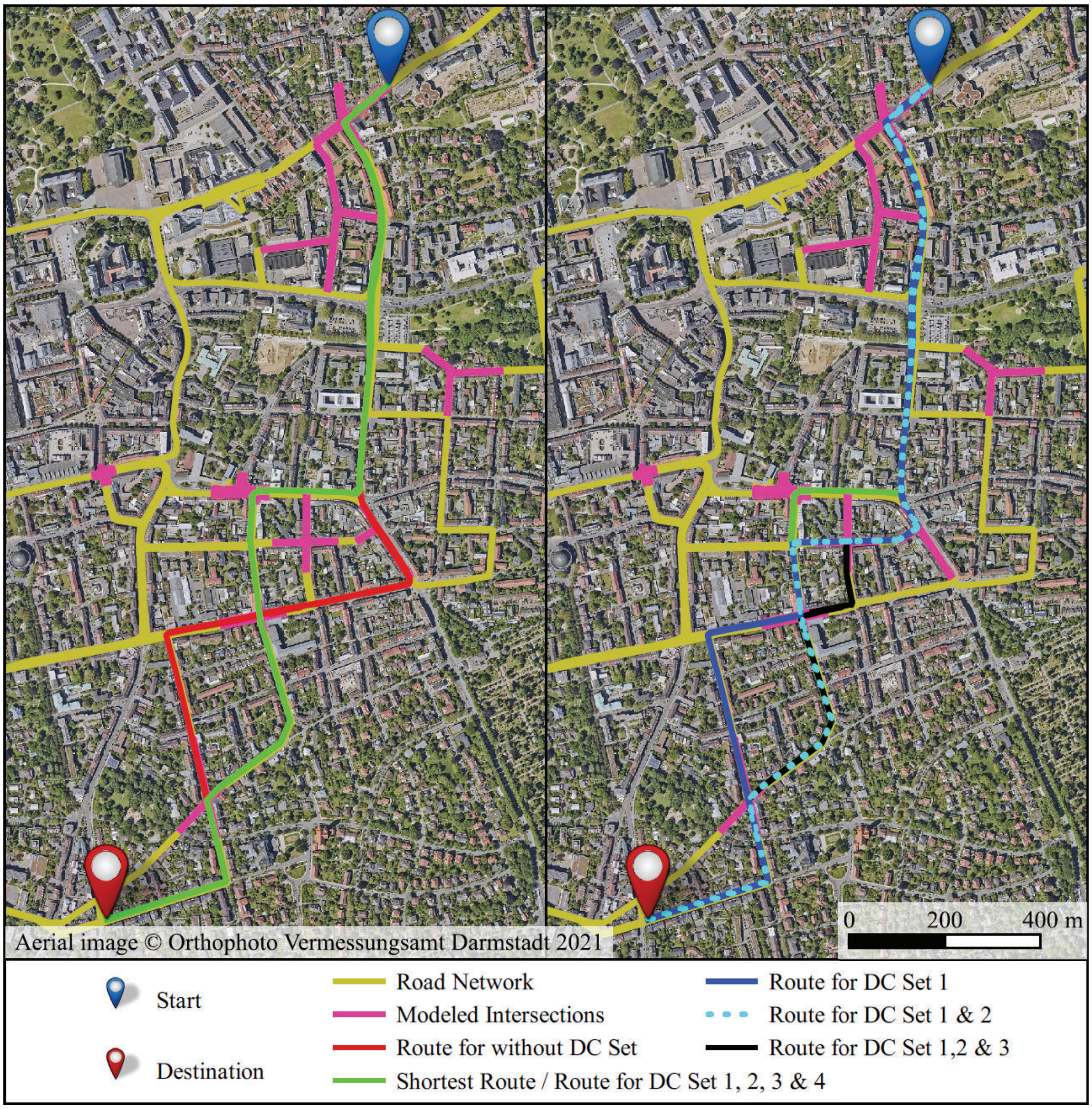


Figure 3. Routing results for the defined driving capability sets.

Route 1: This route is colored red and is partially below the green route as it follows the same course in these sections. The red route only contains behavior spaces that have no driving requirements in terms of reservation. Thus, all externally-reserved behavior spaces are excluded in the generated routing graph. The route therefore only passes through intersections where the AV has priority. At almost all intersections, routing is therefore straight ahead or to the right. Of particular note is the first intersection in the route where a left turn is made. This is possible without the appropriate capability sets because protected left turns are provided at this intersection (green turn arrow at the traffic signal). The last intersection on the route is in a residential area where the right-before-left rule applies. Nevertheless, it is allowed to cross this intersection because the intersection arm on the right in the direction of the route is a one-way street. Since the one-way street is for both motor vehicles and bicyclists and leads away from the intersection, no priority needs to be given.

Route 2: The blue route is based on driving capability set 1, which gives priority to motor vehicle and bicyclists in 30 km/h speed zones from the right. Therefore, the priority road within this route is exited into a residential area, as this path is shorter than that described in the red route. When turning right from the priority road, priority does not need to be granted and priority in the residential area is covered by the capability set. Finally, the residential area is exited by a right turn at an intersection with traffic signals. From this point on, the blue route again coincides with the red route.

Route 3: If the driving capabilities of the blue route are extended by capability set 2, an even shorter route succeeds. The dotted route colored in cyan avoids a detour at the end of the route by crossing the penultimate intersection with traffic lights straight ahead. This crossing requires no special capabilities, but a left turn is made at the intersection after it. However, the associated behavior space requires that priority can be granted from the right from a previously uncovered angular area. The new capability set 2 gained adds this missing angular range to the overall driving capabilities.

Route 4: If the previous driving capabilities are extended by set 3, the even shorter black route is found. The new capability set now allows driving in externally-reserved behavior spaces that require granting priority from the front. This enables left turns at many intersections. Accordingly, compared to Route 2 or 3, it is now possible for the AV to make a left turn in a residential area as well as at a major intersection. Despite the intersection being controlled by a traffic signal, the AV must yield priority from the front because left turns are not protected. Thus, although this intersection is geometrically similar to the first intersection in the route, different driving capabilities are required despite the traffic signal.

Route 5: With the driving capabilities available, Route 4 is not far from the shortest route shown in green. In order for the shortest route to be mastered by the AV, further driving capabilities are lacking. Part of the green route is on a road with rail traffic. The left turn required with this Start-Destination combination must also cross rail traffic. This means that in addition to motor vehicles and bicyclists, the AV must also give priority to rail vehicles from the front. Furthermore, both the associated behavior space and the linked areas of traffic participants with priority are in a 50 km/h speed zone. Driving capability set 4 covers these new requirements, so that this part of the network can also be driven on. If the AV possesses all defined capability sets, it is thus able to travel the shortest route.

Discussion of Results

With the application of the implemented route planning, it is shown that the identified routes are dependent on driving requirements and driving capabilities. Based on the reservation requirement specifications developed in this work, the route planning performs a comparison with the driving capabilities defined for the road network at hand. It is shown that the route planning only calculates routes with driving requirements that do not exceed the available driving capabilities. The performed variation of the available driving capabilities supports this result. Nevertheless, the demonstration initially only shows that capability-based routes can be identified based on the available specifications (for RR1 and RR1.1), but not whether these routes do not exceed the actual available driving capabilities of a real AV. Besides the omitted specifications for the remaining behavioral requirements, the following sources of errors and general uncertainties potentially prevent or complicate a real-world application of the approach.

Map-Based Errors A digital map is necessary for the derived approach. Regardless of map format, it is possible that the map data is incorrect. The data may contain inaccuracies with regard to *geometry*. These include simple offsets as well as shape errors or incorrect angles of individual map elements. Another source of errors is incorrectly *labeled data*. This primarily affects the validity of the behavioral demands of the BSSD, resulting in incorrect driving requirements, for example. In addition, the *actuality* of the map data is a potential source of error. A road network is often subject to many changes, such as road works. Changed scenery elements or temporary changes that are not stored in the map lead to errors of the approach presented here. In the road network modeled in this work, there are no noticeable geometry or label errors. This was verified using highly accurate and georeferenced aerial imagery. In addition to the aerial imagery, the accuracy of the labeled information was partially verified with the help of on-site inspections and publicly available map data. However, especially with regard to the actuality of the data, errors cannot be completely excluded.

Model-Based Errors The aforementioned sources of error related to the maps used are independent of the specification approach chosen in this work. However, errors also arise based on the models used. For the purpose of developing an initial implementation, the simple assumption was made that the areas to be specified (A_{orig}, A_{pre}) are rectangles. However, based on this assumption, errors may occur if the mentioned areas deviate significantly from a rectangular shape in reality. The routing results presented are based only on intersections where there is little to no deviation of the rectangle approximation from reality. For a higher accuracy of the approximation in general, it is possible to assume other shapes as a basis for the specification. However, it should be noted that the specification of driving requirements is always a trade-off. On the one hand, the real world should be approximated as closely as possible so that the specification is as close to reality as possible. On the other hand, the level of abstraction of the driving requirements should be high enough so that common test certificates can be created for similar behavior spaces. Thus, the closer the specification is to the real world, the more difficult it becomes to harmonize similar behavior spaces. Furthermore, a detailed modeled world requires many more specification parameters than shown in this work. This is another drawback to modeling the real world in too much detail.

General Uncertainties In addition to the aforementioned sources of error that directly lead to a faulty specification, there are uncertainties in the overall approach that need to be critically questioned. Uncertainties refer to circumstances that may cause capability-based route finding to be infeasible in the manner presented in this work, regardless of implementation. A well-known problem in the field of automated driving is proving *completeness*. The approach described here is also affected by this, since it is not possible to identify with certainty whether potentially important information is lost during a development step. Thus, although the presented approach is systematically structured, it cannot be spoken of as being complete without further ado. The choice of *specification* represents another uncertainty. It has not been conclusively clarified whether the identified specification parameters are suitable, on the one hand, for mapping the real requirements of the routes and, on the other hand, for providing proof with test certificates. Finally, the *matching criteria* are subject to a number of assumptions. For example, for all specified widths of the different areas, it is assumed that the proof of a certain width automatically implies the proof of lower widths. It has not been demonstrated that the assumptions apply to a real-world proof. For this reason, not only the suitability of the specification, but also the suitability of the matching criteria must be proven for a realization.

OUTLOOK

The present work shows that the identification of capability-based routes is generally possible. This also implies that they can potentially be used for development, testing, and operation of safe automated vehicles. However, based on the discussion of the results, the following challenges need to be overcome before this approach may be used in real-world application.

Map creation and update: A correct map is essential for the realization of capability-based route search. Particularly important is the fidelity of the geometry, the labeled context knowledge (e.g., types of geometry elements), and the actuality of the data.

Proof of specification: The suitability of the identified reservation requirement and capability specifications must be proven (the same applies to yet missing specifications). This will require performing and evaluating tests in simulations and the real world. The specification is falsified if an AV successfully tested for a particular driving capability set is unable to drive a route segment with equally specified driving requirements. Based on the results, it may be necessary to adjust the specifications.

Proof of matching criteria: The suitability of the identified marching criteria must be proven (the same applies to yet missing criteria). This requires demonstrating that an AV with certified capability sets is actually capable of driving the route segments identified as non-exceeding. If route sections identified as passable exceed the driving capabilities of the AV, the matching criteria are falsified and must be revised.

Creation of test certificates: A general challenge becomes the creation and proof of test certificates for capability sets. The tests must prove that an AV actually possesses the capabilities that are certified using the test certificates. This challenge is closely coupled with the proof of specification and matching criteria.

ACKNOWLEDGEMENT

This research is accomplished within the project "UNICAR*agil*" (FKZ 16EMO0286). We acknowledge the financial support for the project by the Federal Ministry of Education and Research of Germany (BMBF).

REFERENCES

[1] Wachenfeld, Walther; Winner, Hermann: The Release of Autonomous Vehicles, in: Maurer, Markus; Lenz, Barbara; Winner, Hermann; Gerdes, J. Christian (PUB): Autonomous Driving: Technical, Legal and Social Aspects, Springer, 2016

[2] International Organization for Standardization: ISO/TR 4804:2020: Road vehicles - Safety and cybersecurity for automated driving systems - Design, verification and validation, 2020

[3] International Organization for Standardization: ISO 21448:2022: Road vehicles - Safety of the intended functionality, 2022

[4] Society of Automotive Engineers: SAE J3016: Taxonomy and Definitions for Terms Related to Driving Automation Systems for On-Road Motor Vehicles, 2021

[5] British Standards Institution: PAS 1883:2020: Operational design domain (ODD) taxonomy for an automated driving system (ADS) - Specification, 2020

[6] Ulbrich, Simon; Menzel, Till; Reschka, Andreas; Schuldt, Fabian; Maurer, Markus: Defining and Substantiating the Terms Scene, Situation, and Scenario for Automated Driving, in: 2015 IEEE 18th International Conference on Intelligent Transportation Systems (ITSC 2015), pp. 982–988, 2015
[7] Lippert, Moritz; Winner, Hermann: How to Derive Behavioral Requirements for Automated Driving from a Behavior-Semantic Scenery Description, in: 14. Uni-DAS e.V. Workshop Fahrerassistenz und automatisiertes Fahren: 09.–11.05.2022, pp. 217–229, 2022
[8] Glatzki, Felix; Lippert, Moritz; Winner, Hermann: Behavioral Attributes for a Behavior-Semantic Scenery Description (BSSD) for the Development of Automated Driving Functions, in: 2021 IEEE International Intelligent Transportation Systems Conference (ITSC), pp. 667–672, 2021
[9] Lippert, Moritz; Glatzki, Felix; Winner, Hermann: Behavior-Semantic Scenery Description (BSSD) of Road Networks for Automated Driving, URL: https://arxiv.org/abs/2202.05211, 2022
[10] Shalev-Shwartz, Shai; Shammah, Shaked; Shashua, Amnon: On a Formal Model of Safe and Scalable Self-driving Cars, URL: http://arxiv.org/pdf/1708.06374v6, 2017
[11] Junietz, Philipp; Wachenfeld, Walther; Klonecki, Kamil; Winner, Hermann: Evaluation of Different Approaches to Address Safety Validation of Automated Driving, in: 2018 21st International Conference on Intelligent Transportation Systems (ITSC), pp. 491–496, 2018
[12] Nilsson, Nils J: Principles of artificial intelligence, Springer Science & Business Media, 1982
[13] Amersbach, C.; Winner, H.: Functional Decomposition, An Approach to Reduce the Approval Effort for Highly Automated Driving. In: 8. Tagung Fahrerassistenz, 22.-23. November, München, 2017
[14] Amersbach, C.; Winner, H.: Functional Decomposition - A Contribution to Overcome the Parameter Space Explosion during Validation of Highly Automated Driving In: Traffic Injury Prevention sup1, Issues 20, pp. 52–57, 2019
[15] Reschka, Andreas; Bagschik, Gerrit; Ulbrich, Simon; Nolte, Marcus; Maurer, Markus: Ability and skill graphs for system modeling, online monitoring, and decision support for vehicle guidance systems, in: 2015 IEEE Intelligent Vehicles Symposium (IV), pp. 933–939, 2015
[16] Forschungsgesellschaft für Straßen- und Verkehrswesen: Richtlinien für die Anlage von Stadtstraßen (RASt), 2006
[17] Google LLC: Google Earth, URL: https://earth.google.com/, 2022, visited on 07/14/2022
[18] Bast, Hannah; Delling, Daniel; Goldberg, Andrew; Müller-Hannemann, Matthias; Pajor, Thomas; Sanders, Peter; Wagner, Dorothea; Werneck, Renato F.: Route Planning in Transportation Networks, in: Kliemann, Lasse; Sanders, Peter (PUB): Algorithm Engineering: Selected Results and Surveys, Springer International Publishing, 2016
[19] Delling, Daniel; Sanders, Peter; Schultes, Dominik; Wagner, Dorothea: Engineering Route Planning Algorithms, in: Lerner, Jürgen; Wagner, Dorothea; Zweig, Katharina A. (PUB): Algorithmics of Large and Complex Networks: Design, Analysis, and Simulation, Springer Berlin Heidelberg, 2009
[20] Poggenhans, Fabian; Pauls, Jan-Hendrik; Janosovits, Johannes; Orf, Stefan; Naumann, Maximilian; Kuhnt, Florian; Mayr, Matthias: Lanelet2: A high-definition map framework for the future of automated driving, in: 2018 IEEE Intelligent Transportation Systems Conference, pp. 1672–1679, 2018
[21] FZI Forschungszentrum Informatik: Lanelet2 GitHub repository, URL: https://github.com/fziforschungszentrum- informatik/Lanelet2, 2018, visited on 11/01/2022

TOWARDS MODELING DRIVER PERFORMANCE WITHIN CRASH-RELEVANT SCENARIOS AS VIRTUAL REFERENCE FOR THE SAFETY OF AUTOMATED VEHICLES

Hendrik Weber
Lutz Eckstein
Institute for Automotive Engineering, RWTH Aachen University
Germany

Felix Beringhoff
RWTH Aachen University
Germany

Johanna Josten
fka GmbH
Germany

Paper Number 23-0126

ABSTRACT

Technological advancements have shown the viability of Automated Driving (AD) and have created high expectations on its benefits – especially in terms of safety. An important step for the introduction of AD on public roads is providing an acceptable proof of AD's positive risk balance compared to today's traffic consisting of human-driven vehicles. Simulation of scenarios has become an essential tool for such analyses, since field operative tests have been shown infeasible as only means for such proof. Nonetheless, data is needed from which to derive human driver behavior as a reference within simulated scenarios. This paper presents an approach for modeling human driver behavior within defined scenarios to serve as a reference for AD.

As a fundamental step to establish a suitable reference, we outlined the architecture of a parameterizable model of driver performance within crash-relevant scenarios, in which the driver model switches from a continuous control to a reactive behavior. The structure is based on well-established concepts like abstraction levels for the driving tasks, cognitive processes, and steps within information processing.

A decision tree-like structure serves as guidance for the modularization of the driver reaction within different scenarios, which allows creating modules of decision-making processes as well as implementation of possible reactions within a scenario.

To show the feasibility of the architecture and modules, and to demonstrate the applicability of the model, we conducted a driving simulator study of a scenario with a vehicle crossing from the right. Within the scenario, we varied the configuration of the potential crash (ego striking and borderline case) as well as apply two values of the available time to react. The study follows a within-subject design with 24 participants. The observed reaction choice, time and intensity were measured and then used to parameterize the driver model.

Braking was the most frequently observed driver reaction, while potential crash configuration apparently influences the reaction choice. The observed driver behavior was in line with assumptions based on the state of art, which were used for the initial architecture and decision making of the developed driver model. Re-simulating the scenario with the parameterized model led to a similar frequency of crashes as in the simulator study.

The experiment provided evidence that the driver model is built on reasonable assumptions for structuring the decision-making process and modeling dependencies between situational variables and reaction parameters. Due to sample characteristics such as age, the gathered parameters cannot serve as a general reference. However, it is not expected that a more diverse sample will disprove the assumptions for the model architecture.

The theoretical considerations for modeling the decision-making process and its dependency on situational variables make apparent which complexity lies within modeling driver reactions.

The proposed model for driver performance within crash-relevant scenarios aims to serve as a reference to prove the positive risk balance of AD. It provides a clear path for the establishment of a general reference model. Yet, the paper shows that the establishment of a baseline for all relevant scenarios comes with a tremendous effort and complexity.

OBJECTIVE

Over the last years, Automated Driving (AD) has been one of the most prominent research topics within the automotive industry, which was driven by advancements in sensing technology and improved processing power. Many players – some of which are traditional automotive OEMs integrating AD technology in their products and some startups or technology companies from other fields – have been showcasing the latest stage of demonstrator vehicles to be operating well on public roads. Few companies (e.g., Mercedes [1]) have released SAE Level 3 technology (see [2]), which allows automated operation within operating conditions at whose limits the driver is required to retake control of the vehicle. Other companies have instantiated SAE Level 4 operation, where no driver is required in defined areas (e.g., [3]). Still, no roll out of higher automation has been achieved at a larger scale.

One challenge, which automated driving is facing, is verifying the safety of AD technology, such that it is actually societally acceptable to have automated vehicles on public roads. Part of this challenge lies within the question, what level of safety should be achieved to be societally acceptable. Some authorities, for instance in Germany [4], have formulated the need for AD to achieve a positive risk balance compared to human drivers. Companies like Waymo have formulated a reference for safety in the form of a "Non-Impaired, Eyes ON the conflict" driver [5].

It has been shown, that proving the positive risk balance by means of on-road testing is economically infeasible, as even with simplifying assumptions a number of billions of km needs to be driven [3], [4]. An alternative approach for proving the positive risk balance lies within a scenario-based approach, which has for instance been established by the German PEGASUS Project [7]. A scenario-based approach allows focusing on Verification and Validation (V&V) activities that are particularly relevant for safety and thus promised an overall reduction in the testing effort.

Multiple definitions of the term *scenario* in the scope of automated driving have been proposed. A general overview of definitions related to scenarios for the validation and verification of automated driving systems (ADS) has been established in ISO 34501 [8]. It defines a **scenario** as the *sequence of the scenes integrated with the ADS(s)/subject vehicle(s), and its/their interactions in the process of performing (a) certain Dynamic Driving Task(s).* The exact separation of abstraction layers of scenarios is less relevant for this paper. Within the following, we will use the term **driving scenario**, which was established in the scope of prospective effectiveness assessment in the L3Pilot project [9]: *A driving scenario is a short period of driving defined by its main driving task (e.g., car following, lane change) or triggered by an event (e.g., an obstacle in the lane).* This definition can to a large extent be seen synonymous to the more abstract definition of **scenario category** in ISO 34501: *set of scenarios that share one or more characteristics.* Scenario-based testing activities can consist of a combination of tests in open fields, controlled field tests and simulation-based testing.

The challenge that comes with a scenario-based approach for verifying safety of automated driving is the need for representing the human driver as reference within the scenarios tested. Crash data provide a data source of driver reactions within defined safety-relevant scenarios, yet it needs to be noted that each accident contains only one driver reaction, which may represent statistical outliers of driver behavior. Moreover, the collection of accident data represents a tremendous effort to cover all the scenarios relevant for the safety validation of automated driving. Studies in simulators or controlled fields allow collecting the behavior of drivers within multiple defined scenarios, but the efforts associated with such studies also appear not to be economically feasible for verifying the safety of automated driving in completeness. An alternative lies within collecting a limited set of data on driver behavior and to establish a model of a driver that can be applied within simulation. Such models are referred to as **driver performance models** and can be built on scientific findings on driver behavior within safety-relevant situations and parameterized with data coming from various data sources.

In the following, an approach for modeling scenario-dependent driver performance data based on findings of reaction patterns within crash-relevant situations will be presented. The model is aimed at providing a modular framework for possible driver reactions within a driving scenario. The modularization should create ease of use and motivate reuse of software modules modeling driver behavior. Moreover, the model should allow multiple reactions within a concrete driving scenario based on probability distributions. The choice of reaction, the reaction time and the reaction intensity should be modeled in a way to make them dependent on situational parameters of the driving scenario.

RELATED WORK

The creation of a driver performance model depends on a detailed understanding of how a driver perceives and processes information and executes the driving task. Though driver behavior is still subject to research and will continue to be for some time, there are certain findings which can be used as the basis for driver modeling. In the following, fundamental findings on driver behavior will be presented and then put in the context of driver modeling. After that, we will focus on driver behavior modelling within crash-relevant situations, highlight existing models of driver performance and present relevant applications.

Structuring the driving task

Donges provides a fundamental structure of the driving tasks, which is not only used as reference of driver behavior but also as reference for the architecture of advanced driver assistant systems (ADAS) or automated driving systems [10]. Donges divides the driving task into three levels: navigation as the highest level, which comprises the selection of the correct route in either known or unknown surroundings, guidance as second level which includes tasks such as choosing the correct lane or executing a turning maneuver, and stabilization as the lowest level, which encompasses lane keeping or the control of the distance to the lead vehicle.

Another important classification of driver behavior is the classification of goal-oriented actions by Rasmussen [11]. He defines three categories of behavior: Knowledge-based behavior holds actions on which a person must actively apply and transfer knowledge from his/her experience to a situation which he/she has not experienced sufficiently often before. Rule-based behavior is applied in situations that occur often and in which a person can directly select an action based on rules learned from experience. Skill-based behavior is used for activities where no conscious control process needs to be applied as the situations or activity are well known to a person.

Table 1 provides a mapping between the levels of the driving tasks by Donges and Rasmussen based on [11].

Table 1.
Relations between levels of driving task [10] and classification of goal-oriented behavior [12] as given in [11]

	Knowledge-based	Rule-based	Skill-based
Navigation	x		
Guidance	x	x	x
Stabilization			x

Driver Modeling

A first use case for driver modelling was the application within traffic flow simulation, where the focus is less on simulating safety-relevant scenarios as an interaction of a small number of vehicles, but rather simulating entire traffic networks to investigate network capacities and congestion phenomena. The primary focus for such models is on modelling the following behavior of drivers. Most prominent models in this field are the Intelligent Driver Model (IDM) [13] or the driver model by Wiedemann [14], which finds its application in different simulation tools. While the IDM models driver behavior by means of a differential equation whose parameters can be set to mimic different driver attitudes, the model by Wiedemann aims at representing driver physiological processes, which are based on principles of looming and unconsciously keeping a desired headway. Parameters within the model point directly to interpretable parameters, like desired speed, desired time gap or gap at standstill.

More sophisticated approaches aim at modeling the driver as a complex control flow of a vehicle. Many approaches consider the three levels of the driving task by Donges [10], which splits driving into the navigation, the guidance and the stabilization layer. Klimke presents a driver model – also referred to as agent model – which follows these basic assumptions in its internal structure [15]. The model is intended to realize different tasks of maneuvers like the adaptation to a new speed limit. For this, a comprehensive architecture is presented, which covers all levels of the driving tasks and primary driver inputs (steering, throttle, braking), secondary inputs (e.g., lighting) as well as communication, and vehicle and system setup. The model was implemented as C++ class and published on GitHub.

A model aimed at covering the entire driving task is presented by [16], which aims at recreating physiological aspects of driver information processing. The model uses a stochastic process for modeling the driver's gaze behavior and some internal information processing. These may cause the model to perceive relevant information too late or not at all such that it may create crash-relevant situations or even crashes. This way, the model can be applied for effectiveness evaluation of automated driving systems, by either creating a baseline of simulated traffic or creating surrounding traffic for a single or multiple automated vehicles.

Driver behavior in crash-relevant situations

Everyday driving can be modeled as controller behavior (e.g., vehicle following and lane keeping), as for everyday-driving primarily skill-based or rule-based processes are of relevance. In contrast to everyday driving, crash-relevant situations are situations which the driver does not experience regularly. The driver needs to quickly select an action and implement it to avoid the collision with another object or to run off the road. The driver models which focus on the driver reaction to avoid a collision in crash-relevant situations are often referred to as driver performance models. Erbsmehl states, that the primary parameters for modeling the driver response in the crash-relevant situation are the reaction time and the reaction intensity [25].

When reaction times are analyzed or modeled, it is important to consider how they are defined. A high-level classification of the relevant elements of the overall reaction time is provided by Green [17]. Green defines the mental processing time – which in itself consists of *sensation*, *perception,* and *response selection and programming* –, the *movement time*, and the *device response time* [17]. If for instance a driver's reaction to a stimulus is observed by means of the deceleration of a vehicle, all mentioned elements of the reaction time will be executed until a deceleration of the vehicle can be observed. When modeling this deceleration, the device response time is typically not part of the driver model, but of the vehicle model.

Multiple studies have investigated the relation between reaction time and time-to-collision (TTC). Many of those show a positive correlation between TTC and reaction time of a driver, e.g., [18], [19] and [20], while [21] state an increase in reaction time at very small TTC. Apart from reaction time, the reaction choice of the driver model is particularly of relevance for crash avoidance or mitigation, which is especially relevant at conflicts at intersections, where evasive steering is a relevant option for the driver. Multiple studies have investigated, whether drivers execute a same direction swerve (SDS) or an opposite direction swerve (ODS) in crossing conflicts, with an SDS as pictured in Figure 1. The SDS is in [18] referred to as a *swerve into danger*: the driver executes an avoidance maneuver away from the other vehicle but at the same time into the direction in which the other vehicle is moving. From a consideration of the kinematics of the situation, in some of these situations a collision would have been avoidable with higher odds by braking and steering in the other direction. Weber et al. analyze whether driver execute a typical standard reaction within a scenario with another vehicle crossing the driver's path by means of driving simulator studies and accident analysis and found that driver's typically perform an SDS, which is assumed to be caused by a reflex action [19]. In a further study, they found that the probability of a standard reaction increases with lower time-to-arrival (TTA) values, which can be seen equivalent to TTC. Moreover, with lower TTA, less drivers showed a reaction consisting only of braking [18].

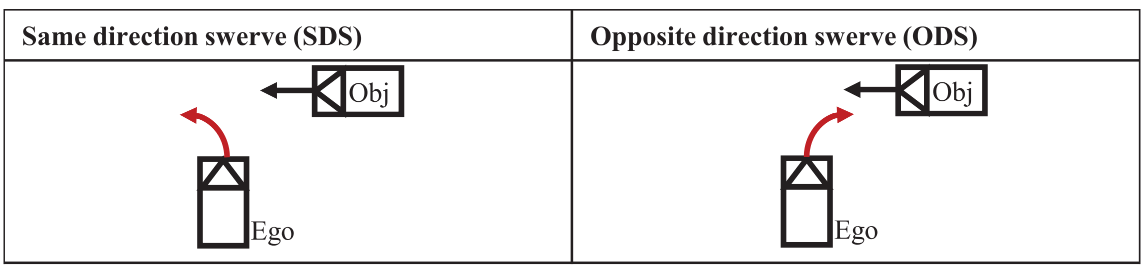


Figure 1. Definition of same direction swerve (SDS) and opposite direction swerve (ODS) based on [20]

Hu et al. also study driver behavior in crossing conflicts finding that an SDS-behavior is present in the scenario. How likely an SDS reaction is, depends, however, on the parameters of the scenarios [21] [22]. They show that the priority level (PL) is an important influence on the driver reaction. The PL expresses in a continuous number, how close one vehicle is to leaving the conflict zone before the other one enters it, as shown in Figure 2. The chance of an SDS, which Hu et al. consider as an irrational decision, increases with negative PL, while increased urgency of the reaction also increase the likelihood of an SDS.

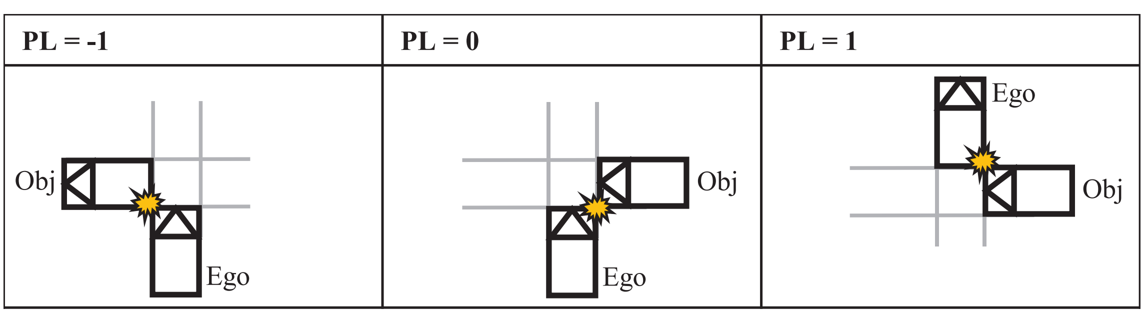


Figure 2. Characteristic values of the priority level (PL) as defined in [21]

Li et al. analyze different crash-relevant scenarios at intersections, where SDS behavior was also observed in a left turn across path (LTAP) scenario. Drivers executing a SDS had a significantly lower reaction time compared to drivers executing ODS [23].

Modelling evasive driving maneuvers

Based on physiological considerations, Lee showed that the control variable for braking is the derivative of the time-to-collision ($\dot{\tau}$), which drivers keep constant within a braking maneuver [26]. Yilmaz and Warren confirm the findings by Lee and further study the control process of braking, in which the driver shows a non-continuous behavior [27]. The brake pedal pressure is kept constant for a certain time until adjusted, based on $\dot{\tau}$. The $\dot{\tau}$-theory by Lee [26] in combination with the finding by Yilmaz and Warren [27] can be used to model driver behavior within braking situations, as applied by Roesener [28], where the braking application is modeled through a combination of an initial open-loop reaction continued by closed-loop control behavior.

Jurecki's and Stanczyk's [24] model for a straight crossing path scenario is one of the few models for modeling drivers' steering reactions in crash-relevant intersection scenarios. They model the steering wheel angle as a function of the lateral distance to the obstacle and the reaction time [18]. Their model is validated by comparing the modeled steering wheel trajectories with driving data from subject tests on a test track. They state that it is possible with their steering model to represent different driver behavior patterns using different parameterizations. However, a subdivision of the drivers into groups depending on their driving style is necessary [24]. A functional dependency of the parameters to modulate the reaction course does not seem to exist, which is why the use of predefined group-individual parameters is suggested [25].

A general statement regarding the steering behavior of drivers in crash-relevant maneuvers is given in [29]. According to the analyses, drivers tend to show an open-loop behavior, consisting of several, interrupted bell-shaped steering angle corrections [29]. The assumption of open-loop steering responses in lane changes on highways is presented in [28] and confirmed by a comparison with a closed-loop approach as a more accurate modeling approach.

Driver performance models and their application

Although driver behavior is still to a large extent subject to research, driver performance models have been used in different types of prospective safety assessments. The P.E.A.R.S. consortium [25] contributed to ISO technical report [26] which defines a method for prospective safety assessment. Within a prospective assessment, the baseline scenarios are needed, in which the system under assessment is expected to increase the safety. Baseline approach A as defined in [26] uses real world baseline scenarios, which may be reconstructed crashes or crash and near-crash situations from naturalistic driving data, also referred to as cases. In approach A these scenarios are used directly as baseline, such that the driver reaction is taken from the real case. Approach B modifies real world cases, which also makes it possible to derive more than one scenario from one real world case. Approach C generates synthetic scenarios, which are not directly linked to real world cases but uses data from real world cases or other driving data, such as driver behavior for generating synthetic cases.

While the modified scenarios are close enough to the real world for approach B, such that driver behavior in those cases is suitable as reference, approach C requires a driver model to produce baseline scenarios to serve as baseline within the prospective assessment. In [28], Fahrenkrog et al. presents a safety impact assessment study of an ADS for motorways which utilizes the model presented in [16]. The model uses a stochastic process for modeling the driver's gaze behavior, which in addition to some internal information processing, may cause the model to perceive relevant information too late or not at all such that it may create crash-relevant situations or even crashes. This way, the model can be applied for effectiveness evaluation of automated driving systems, by either creating a baseline of simulated traffic and creating surrounding traffic for a single or multiple automated vehicles. Within traffic scenarios spanning a stretch of motorway containing many vehicles, the model is used to represent the mechanisms leading to crash relevant situations as well as the driver reactions in crash-relevant situations. In cases where the ADS avoids a collision, it is still possible that a preceding vehicle collides with a heavily decelerating ADS.

Erbsmehl and Schebdat apply a simulation approach of driving scenarios with a driver performance model which uses distributions of reaction times and distributions of reaction intensity [29]. The model is applied in an assessment of a warning system.

Roesener et al. present a study where a traffic simulation with a driver model that explicitly models causation mechanisms for crash-relevant situations is combined with simulations of driving scenarios, where a driver performance model within crash-relevant driving scenarios is used as baseline [30]. The model implements results from a driving simulator study on the reaction time and reaction intensity to represent reference driver behavior within the simulated driving scenarios. In [9], the approach applied in [28] was applied for motorway scenarios and combined with a re-simulation of relevant real-world cases. For an urban ADS, the approach from [30] was refined and applied without the simulation of traffic scenarios.

In [31], Roesener builds upon the method established in [30] but uses a more sophisticated driver performance model to simulate the impact in rear-end and cut-in scenarios. The detailed model is presented in [32] and makes use of the findings from [33], but models the initial driver reaction as an open loop reaction. The model for evasive steering was built on the finding from [34] and uses low-level open-loop steering impulses to create the closed loop behavior.

Bärgman et al. perform a comparison of different driver performance models as references for the safety benefit of an integral safety systems [35]. One of the investigated models used a simple brake control sub-model with a constant jerk and the others use a linear relation between to the inverse of the optical time-to-collision, while for all the maximum deceleration is sampled from the same distribution. Apart from the first model, which uses no glance behavior model, all models use different models for the driver's glance behavior. The simulations show that the choice of model can have a significant effect on the safety benefits simulated.

Waymo present an alternative approach for modelling reaction times by implementing a belief update process in which the simulated driver initiates a reaction based on perceived violations to his prior belief [5]. This approach resolves two issues which driver performance models typically have in common. First, the update process takes into account an evolving traffic situation and thus allow to model the situation-dependence of the reaction time. Second, the modeling approach allows to define a stimulus in a scenario, even if no distinct event can be specified, which is only possible in controlled studies and often cannot be applied when using naturalistic data. The model is applied to rear-end crashes and near-crashes from the SHRP2 where the prior belief is modeled based on the looming principle (see [36]).

The UN ECE Regulation for the ALKS applies the modeled driver performance of a skilled human driver for deriving thresholds which situations are preventable or unpreventable [37]. The driving scenarios simulated are a cut-in, a cut-out with a slower vehicle in front of the vehicle cutting out, and the lead vehicle decelerating. The driver reaction is identified by means of a risk perception time, a delay in decision and a jerk time until the maximum deceleration. The regulation identifies clear definitions of the stimulus within the scenario. Simulation results show which configurations of the driving scenarios can still be considered avoidable.

METHOD

Modelling driver behavior is essential when carrying out a prospective effectiveness assessment of automated driving technology. When baseline cases are not taken from real world cases, where real driver behavior resulted in the crash or crash-relevant situation, driver behavior needs to be modeled within a synthetic scenario. While certain approaches like [16] use holistic models where safety-relevant situations are also caused by the driver model (e.g., by modeling a driver's inattentiveness), other approaches use predefined safety-relevant scenarios in which the opponent induces the crash-relevant situation and the vehicle under test must react to avoid a collision (e.g., as in [30]).

In the latter approach, driver behavior needs to be modeled specifically within the driving scenarios under investigation. It thus needs to be ensured, that the model used as reference within the baseline scenarios can produce a sufficiently realistic reaction to the safety-relevant situation. For this, the actual action of the driver needs to be implemented, which may consist of an open-loop action such as an initial brake application and a closed-loop control action, e.g., adjusting brake pressure during the course of the reaction. Apart from modelling the control action, their dependence on parameters within the driving scenario needs to be modeled (e.g., drivers may apply a stronger initial reaction in a situation with greater urgency). Moreover, multiple different driver reactions may be possible within a scenario. The reaction choice of the virtual driver may be modelled stochastically by choosing different controller implementations, while the probabilities of the different reaction choices may also depend on situational parameters. It should be noted that both of the approaches for using synthetic driving scenarios have their strengths and weaknesses. In general, modelling driving behavior remains a great challenge as the inner workings of a human being during driving are complex and affected by non-driving-related tasks.

The vastness of driving scenarios which need to be analyzed and possible options for driver reactions within a driving scenario result in a great overall complexity of aspects to be modeled. It is thus beneficial to instantiate one central driver model consisting of multiple interconnected modules which can be parameterized individually. Parameterizations should be easily defined and it should be possible to store and exchange them. This enables and motivates the reuse of models across multiple driving scenarios. In that way, it needs to be set for an assessment which modules should be used and how these should be parameterized.

In this work, we present a parameterizable driver model, whose architecture enables modularization and separate parameterization of model components, which in consequence eases reuse of model components creating a greater consistency across individual assessments and explainability of the model in the reporting of the setup, results and limitations of a prospective effectiveness evaluation. The remainder of this paper follows the following process steps:

0. Design architecture for a parameterizable driver performance model *
1. Identifications of the application of the model by
    a. Selecting a relevant driving scenario
    b. Selecting influencing scenario parameters
    c. Selecting suitable models for reaction
2. Plan and execute driving simulator study with a defined sample of drivers
3. Evaluate study
    a. Check if main design decisions for driver performance model are supported by study *
    b. Check if scenario parameters influence driver reaction within scenario
4. Use study as input for parameterization of model
    a. Decide for modeling of dependencies between scenario parameters and driver reaction
    b. Derive distributions for stochastic driver reaction
5. Re-simulate scenario of study with parameterized model
6. Validate results using study data (ideally with test data set)
7. *Apply model within prospective safety assessment*
   (The overall assessment needs to be validated as well)

These process steps serve as structure for this paper and explain how the model were to be applied in an effectiveness assessment. Steps marked with asterisk (*) are only relevant for this paper's scope and need not be executed in an application of the model: Step 0 presents the model architecture design, which is not a task to be repeated per application. Step 3.a validates this structure as part of the overall presentation of the model, supporting the overall design decision enabling reuse of the model. Step 7 encompasses the actual application of the model, which is out of scope of this paper. This application should follow the guidelines established by the P.E.A.R.S. consortium (see [26] and [27]), which also gives guidelines for a proper validation of the assessment. Apart from the steps listed above, this paper presents a discussion of the model, its setup und intended use.

DESIGN OF MODEL

The overall model architecture consists of two main design principles: A three-dimensional modularization of the driving task and a decision tree realizing the driver's reaction choice. The three dimensions along which the driving task is structured are:

1. The different layers of the driving tasks as defined by Donges [10]
2. The classification of target-oriented actions by Rasmussen [12]
3. The information processing chain divided into *perception*, *cognition* and *action*

These three dimensions of the driver reaction within crash relevant scenarios create a Rubik's Cube-like structure for the model. However, not all combinations of the categories create meaningful sub-processes. Such that not all of the elements of the Rubik's Cube are occupied by a software module which encapsulates physiological or statistical models recreating the driver behavior within a given driving scenario. The overall structure is presented in Figure 3. The navigation task typically only consists of knowledge-based processes, in areas where a driver needs to actively navigate, except for the sensory perception of environmental information, which is considered as a skill based-perception process. Within safety relevant scenarios, navigation is often not relevant. Only in special cases, additional workload by the navigation task may affect the driver's reaction but has not yet been subject to extensive research. Thus, navigation is only considered for completeness.

Skill-based processes typically do not require complex cognitive processes, such that they are modeled as sequences of perception and action. Skill-based reactions may be present as both, guidance, or stabilization tasks. Rule based actions are particularly relevant for the guidance layer as they encompass evasive maneuvers. For these, perception, cognition, and action are relevant processes, such that they are considered as modules within the driver model.

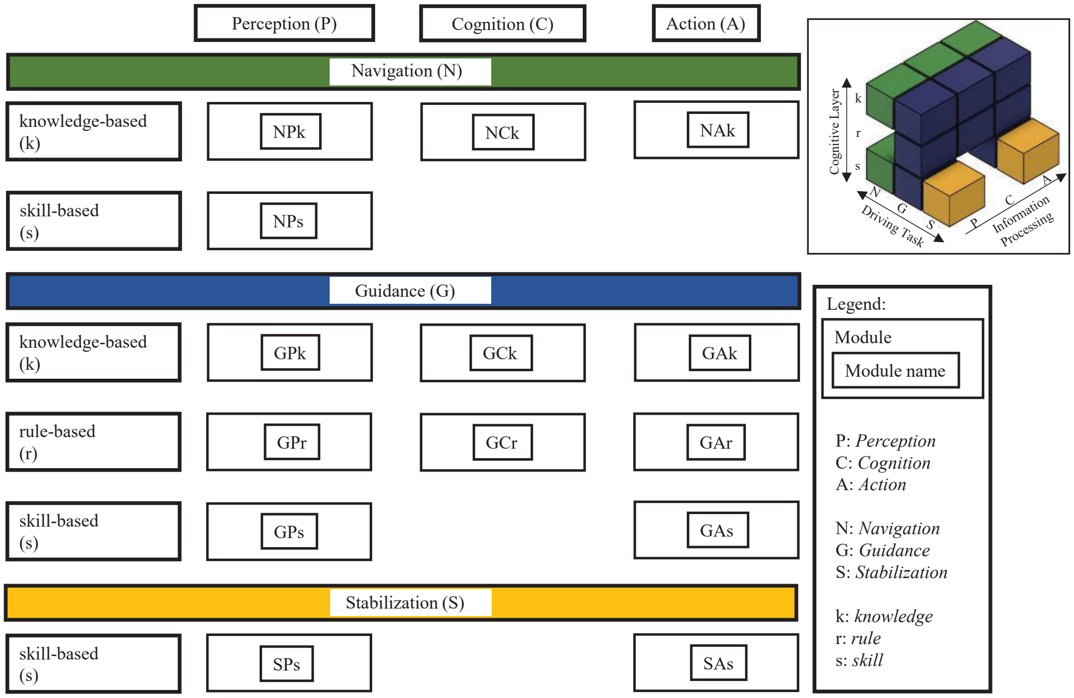


Figure 3. Generic architecture of the driver performance model

The model is implemented in Python and provides empty modules within its default configuration. The model uses OSI [38] as input data stream from the simulation environment. The model's output consists of the pedal positions and the steering wheel angle. At the time of writing, no harmonized driver model output interface existed, such that the output interface was implemented to suit the proprietary driver C-Struct of the simulation tool *Virtual Test Drive* by *Hexagon*, which was used as the tool for development and validation of the model. For an application these empty modules can be replaced by user-defined modules which can be set up as shown in Figure 4.

```
def GCk(Settings, Parameters, driver, ego):
    """Container for Guidance-Cognition-knowledge function calls"""
    import sys
    import os
    generic_duplicate_path \
        = os.path.join(Parameters.Parameter_generic_path, \
            "Guidance","G_Cognition")
            sys.path.append(generic_duplicate_path)
    ##################################
    # Import your functions below:
    #
    from cognition_stimulus_understanding \
        import cognition_stimulus_understanding
    [...]
    # Processing of perceived environment and checking for call for action
    driver = cognition_stimulus_understanding(Parameters,driver,ego)
    # Get reaction type:
    driver = cognition_decision_making_module(Parameters, driver)
    # Get reaction times:
    driver = cognition_decision_making_reaction_times(Parameters, driver)
    # Get reaction intensities:
    driver = cognition_reaction_intensity(Parameters, driver)
    [...]
    ##################################
    return driver
```

Figure 4. Python skeleton implementation module of driver reaction

The structure presented above does not yet provide a structure for the possible driver reactions. Thus, additionally to the structure presented, different options for driver reactions need to be defined. For this, we structure possible driver reaction types – in the following called RTYPE – by means of a decision tree. The root node of the decision tree – the high-level RTYPE – considers whether a longitudinal (1xx) or lateral reaction (2xx) is executed or a combination of both (3xx). For the intervention a distinction is made, whether lateral or longitudinal acceleration is increased (11x resp. 21x) or decreased (12x resp. 22x), which is considered as mid-level RTYPE

Combined reactions are more complicated since according to [39], no truly parallel reactions exist. For this, we consider the different combinations of increase or decrease of lateral or longitudinal acceleration (31x – 34x) and add whether lateral or longitudinal intervention was first initiated (e.g., 31.x-Long for an increase in longitudinal acceleration, followed by an increase in lateral acceleration). A further distinction can be made on the level of the low-level RTYPE, which considers which and how driver control units (i.e., accelerator pedal, brake pedal, steering wheel) were used to influence the vehicle dynamics. The definitions of the high- and mid-level RTYPES are given in Table 2 and a decision tree leading to these is presented in Figure 5.

Table 2.
Definition of reaction types (RTYPE)

RTYPE	**Definition**
1xx	**Reaction influences longitudinal vehicle dynamics (longitudinal reaction)**
11x	Increase of longitudinal acceleration
12x	Decrease of longitudinal acceleration
2xx	**Reaction influences lateral vehicle dynamics (lateral reaction)**
21x	Increase of lateral acceleration
22x	Decrease of lateral acceleration
3xx	**Reaction influences longitudinal and lateral vehicle dynamics (combined reaction)**
31x-Long	Increase of longitudinal acceleration + Increase of lateral acceleration
31x-Lat	Increase of lateral acceleration + Increase of longitudinal acceleration
32x-Long	Increase of longitudinal acceleration + Decrease of lateral acceleration
32x-Lat	Decrease of lateral acceleration + Increase of longitudinal acceleration
33x-Long	Decrease of longitudinal acceleration + Increase of lateral acceleration
33x-Lat	Increase of lateral acceleration + Decrease of longitudinal acceleration
34x-Long	Decrease of longitudinal acceleration + Decrease of lateral acceleration
34x-Lat	Decrease of lateral acceleration + Decrease of longitudinal acceleration
4xx	**No reaction**
40x	Unchanged vehicle dynamics

The selection of the driver reaction in a single simulation run is based on a probabilistic decision tree, in which probabilities can be modeled in dependence of situational parameter (see example in Figure 7). The distinction into typical and not typical reactions refers to the finding by Weber [40].

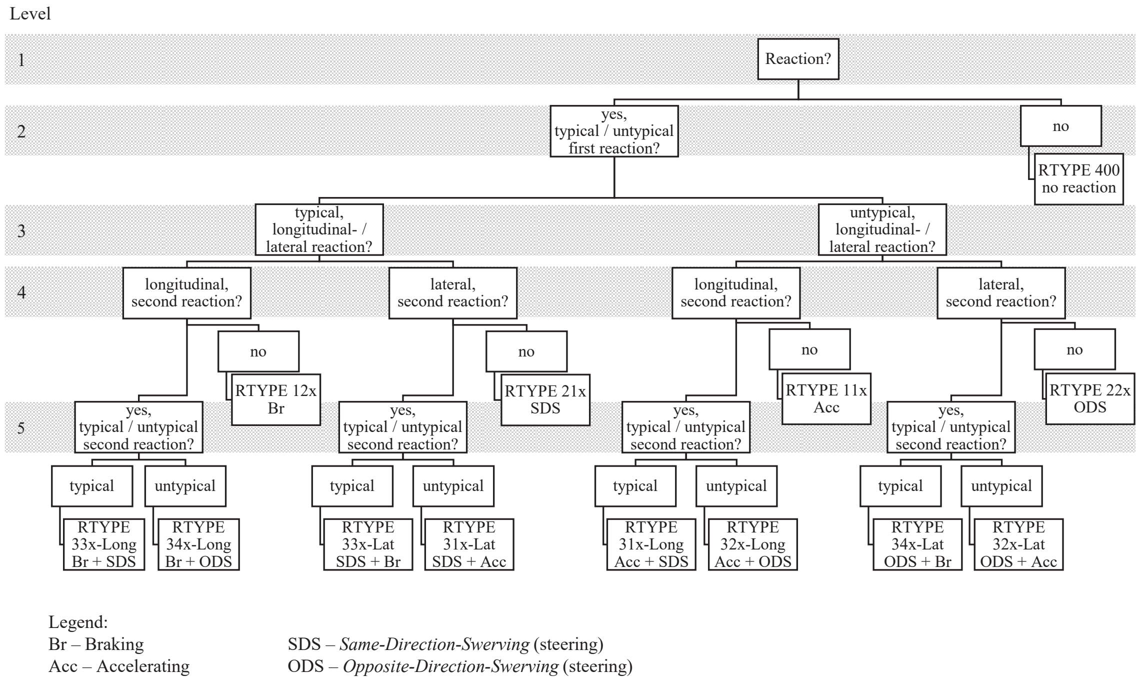


Figure 5. Decision tree-like model of the stochastic decision-making process of the driver performance model

APPLICATION OF MODEL

Our aim is to derive a specific driver performance model for crash-relevant urban scenarios. Following the previously disclosed method, we first select a logical urban scenario and identify its characteristic description parameters, before implementing reaction algorithms.

Selection of scenario

Road accident statistics from Germany of the year of 2019 show that most of the crashes reported to the police happen in urban areas. Especially junctions including driveways are crash prone with more than 50 % of these crashes happening there [41]. Research from the Intersection 2020 project [42] identifies *straight crossing paths (SCP)*, *left turn across path – opposite direction*, and *left turn across path – lateral direction* as most relevant intersection car-to-car scenarios to focus on the enhancement of road safety at intersections. We chose to focus on straight crossing paths accidents with an opposing vehicle challenging from the right hand side of the ego vehicle (see Figure 6).

Relevant scenario parameters

From literature and related studies on driver behavior in crash-relevant scenarios, the time-to-collision and the projected crash constellation hypothetically influence the driver's reaction. Therefore we chose the initial time-to-conflict-point (TTCP) and the initial priority level (PL) as the two variables for our model. The start of the crash-relevant scenario is defined by the object vehicle becoming visible after being obstructed by a vehicle (Figure 11). Both values are measured at the time when the object vehicle becomes visible from the view point of the ego vehicle. The TTCP is calculated by the distance to conflict point (DTCP) and the current velocity of the ego vehicle. The PL is calculated by the following equations:

For $TTCP_{Obj} - TTCP_{Ego} < 0$: $$PL = \frac{TTCP_{Obj} - TTCP_{Ego}}{TTCP_{Obj,exit} - TTCP_{Obj}}$$ Equation (1)

For $TTCP_{Obj} - TTCP_{Ego} > 0$: $$PL = \frac{TTCP_{Obj} - TTCP_{Ego}}{TTCP_{Ego,exit} - TTCP_{Ego}}$$ Equation (2)

For $TTCP_{Obj} - TTCP_{Ego} = 0$: $$PL = 0$$ Equation (3)

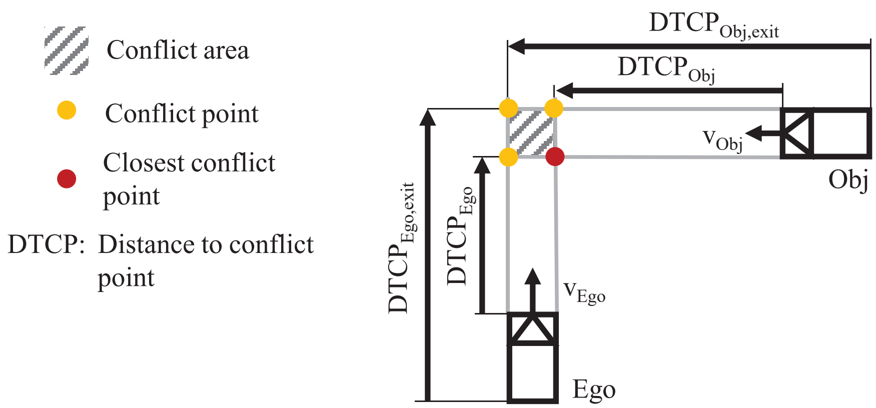


Figure 6. Description of the conflict

Implementation of model reactions

The process of human perception is modelled in a very simplified and ideal way in our driver performance model. Calculations required on the software side in relation to the perception of the environment are also assigned to the modules of perception. The perception process (incl. recognition) of the driver performance model consists of the detection of objects, the verification of the visibility of the objects, the computation of the scenario description variables TTCP and PL and the verification whether the visible object generates a request for action. To determine a reference time for the driver reaction, it is necessary on the software side to detect the time of occurrence of a perceptible crash-relevant situation. For this purpose, in the case of a suddenly appearing object, the earliest possible time is determined from which a point of the object would be visible to a driver without occlusion. The perception functions are distributed over the modules SPs, GPs, GPk and GCk (see Figure 3).

The reaction choice is modelled by the decision tree presented in Figure 5. The likelihood of a reaction choice depends on the scenario parameters TTCP and PL (see Table 7). The decision tree is located in module GCk.

For the reaction times, a dependency on the urgency of the scenario and the reaction type is considered. The reaction times are modeled by a truncated normal distribution around the mean of the measured reaction times of a driver population. For reaction times, 0 s or, for second reactions, the reaction time of the first reaction is set as the lower limit and the normal distribution is truncated accordingly. Reaction times are control unit specific and stochastically determined when simulating the driver performance model.

Driver responses to the steering wheel and pedals are modeled by reaction models from the literature, which are then fitted and parameterized to the data from our driving simulator study. For reaction intensity, we formed five discrete groups of reaction intensities for accelerator pedal actuation, braking, and steering: very low, low, medium, high, very high. For each reaction intensity, we parameterized the respective models with corresponding data from the study that are in the respective reaction intensity group. A reaction intensity is then stochastically determined specifically for the actuator when simulating the driver performance model and depends on the reaction time. The determination of a reaction intensity and reaction time is located in module GCk.

Regarding the modelling of the brake pedal stroke in crash-relevant situations, the results from [32] show that the assumption of an open-loop action of the driver provides good results. Based on this, we also model the longitudinal responses (step on the accelerator pedal or braking) by an open-loop response. The step on a (brake) pedal is modeled by the first-order transfer response of a system. During the simulation, the pedal position y at time n+1 is calculated discretely in time, considering the size of the time step dt, the parameters T and K, and the command variable u according to Equation (4).

$$y_{n+1} = \left(1 - \frac{dt}{T}\right) \cdot y_n + K \cdot \left(\frac{dt}{T}\right) \cdot u_n \qquad \text{Equation (4)}$$

For modelling the steering reaction of the drivers, we use an adapted formula based on the steering model of Jurecki [24]. Due to unsatisfactory results when fitting the model to the measurement data of the driving simulator study explained later, we extended the calculation formula by one term. This term is equivalent to an open-loop steering response, which is only time-dependent and not dependent on the lateral distance to the object. This extension goes along with the results from [43], according to which steering reactions are open-loop reactions. The adapted calculation formula can be found in Equation (5), where the parameter K_6 and the command variable u are introduced, which can be used to directly parameterize the maximum steering wheel angle (the reaction strength) during the steering reaction. The first two terms of Equation (5) resemble the discrete-time transfer behavior of a first-order system. The third term establishes a dependency of the steering wheel angle δ_{sw} on the lateral distance to the object and – by limiting $y_{lat,rel}$ – influences only the steering angle retraction.

$$\delta_{sw,n+1} = \left(\left(1 - \frac{dt}{W_4}\right) \cdot \delta_{sw,n} \cdot \frac{\pi}{180} + K_6 \cdot \left(\frac{dt}{W_4}\right) \cdot u_n + W_5 \cdot \left(\frac{dt}{W4}\right) \cdot y_{lat,rel,n}\right) \cdot \frac{180}{\pi} \qquad \text{Equation (5)}$$

$$\text{with } y_{lat,rel} = (y_{pos,obj} - y_{pos,ego}) + y_{offset} \qquad \text{Equation (6)}$$

$$y_{lat,rel} = \max(0, y_{lat,rel}) \qquad \text{Equation (7)}$$

$$y_{lat,rel} = min(0, y_{lat,rel}) \qquad \text{Equation (8)}$$

The steering and pedal reactions are implemented in the module SAs.

The parameterization of the driver reaction can be achieved using probabilistic trees in a JSON file as presented in Figure 7.

```json
"node_on_reaction": {
    "required": ["branches", "weights"],
    "properties": {
        "branches": ["typical_reaction", "untypical_reaction"],
        "weights": {
            "independent_var": {
                "name": "ttcp",
                "val": [1.43,2.10]
            },
            "weights_branch_typical_reaction": [22,24],
            "weights_branch_untypical_reaction": [2,0]
        }
    },
    "node_on_typical_reaction": {
        "required": ["branches", "weights"],
        "properties": {
            "branches": ["long", "lat"],
            "weights": {
                "independent_var": {
                    "name": "ttcp",
                    "val": [1.43,2.10]
                },
                "weights_branch_long": [22,22],
                "weights_branch_lat": [0,2]
            }
        }
    },
    ...
}
```

Figure 7. JSON-Example for parameterization of driver reaction

SIMULATOR STUDY TO PROVE BASIC ASSUMPTIONS AND FOR PARAMETERIZATION

In the summer of 2020, we conducted a driver simulator study to proof the general assumptions on which the decision tree-like structure of the driver performance model is built up on and for the parametrization of the model. In general, the driver simulator study is utilized to answer the following research questions:

RQ1: How do drivers react in a crash-relevant intersection scenario where another vehicle crosses the first-person path from the right?

From literature research, we learned about reactions drivers show in crash-relevant scenarios. Therefore, we aim to answer the research question, whether the drivers show hypothetical typical first reactions. In our selected straight crossing path scenario at an intersection, these hypothetical typical first reactions are braking or a same direction swerving as steering reaction. In brief, the RQ2 and related hypothesis are:

RQ2: Do human drivers show typical first reactions?

Furthermore, we want to observe whether the scenario description parameters TTCP and PL influence the driver's reaction choice. For that, we formulate RQ3 and related hypothesis 2.1 and 2.2 as follows:

RQ3: Does the frequency distribution of the reaction types change when the TTCP or the PL are changed?

Hypothesis 3.1) There is a difference in frequency distributions of response types when the PL is changed.

Hypothesis 3.2) There is a difference in frequency distributions of response types when the TTCP is changed.

The driving simulator study was conducted in a static driving simulator with a 360° screen surrounding the ego vehicle at fka GmbH / the Institute of Automotive Engineering (ika). We invited 25 volunteers to join our study. One participant aborted the study after the test drive due to simulator sickness. Therefore, 24 valid data sets were obtained. Due to the Covid-19 pandemic, the participant pool was limited to employees of ika. Participants did thus not receive compensation for taking part in the study. The study took about 45 minutes to complete.

Of all 24 participants, three were female and 21 male, with a mean age of 27 years (SD=4.63; range: 20 to 37 years). On average, the participants obtained their driver's license 9.5 years ago and drove 8,290 km in the last year (SD=6975.46 km). Minimum mileage within the last year was 70 km/year and maximum mileage was 25,000 km/year.

Each participant filled in a data protection and participant information sheet and conducted a familiarization drive of about 6 minutes prior to the first measured drive to get used to the simulator. For the test drives, the participants were told to follow a lead vehicle with 50 km/h on a straight urban road, which crosses several intersections with green phased traffic lights. Each drive persisted for about 3 to 5 minutes until a crash-relevant scenario occurs. The participants were not briefed about the crash-relevant scenario they experienced in the drives. In each of the 5 separate urban drives, participants experienced one of the described scenarios (repeated-measures design, Figure 8). During the third drive, the participants experienced a scenario in which a challenging vehicle crossed from the left. This third drive was used to make the crash-relevant scenario more unpredictable to the participants. We randomized the order of the four relevant scenarios between the participants.

The four scenarios in which a vehicle crosses from the right differ in the values of the two independent variables TTCP and PL. The first TTCP value is chosen so that it is physically possible to avoid a collision by pure braking. For this purpose, a reaction time of 1.34 s is assumed for the first scenario, based on the time from the "Cologne model" [44] and a braking acceleration of -9 m/s² is assumed [44]. This results in a TTCP of 2.11 s. For the second TTCP condition, the reaction time is halved. In the neutral PL condition, both vehicles would arrive at the same time at the conflict area and collide with the front edges of the vehicles assuming that both vehicles travel with an unchanged trajectory. In the negative PL-condition, the ego vehicle would strike the most rearward side area of the object vehicle with 100 % overlap, which leads to a PL value of -0.71 considering the vehicle dimensions of both vehicles. Under ideal conditions, assuming that the ego vehicle is driven by the participant with exactly 50 km/h in the middle of the road, the object vehicle would travel 35.2 km/h. This is in the middle of the range of usual values in SCP crashes according to [42]. The exact starting position and velocity of the object vehicle is calculated online and adapted to the participant's vehicle velocity and position so that the independent variable values are realized. However, due to the method used to trigger the object vehicle (via the Simulation Control Protocol of Virtual Test Drive), the independent variable values were not exactly met during the study which led to a mean deviation of -0.011 s of the targeted TTCP condition and +0.05 of the targeted PL condition.

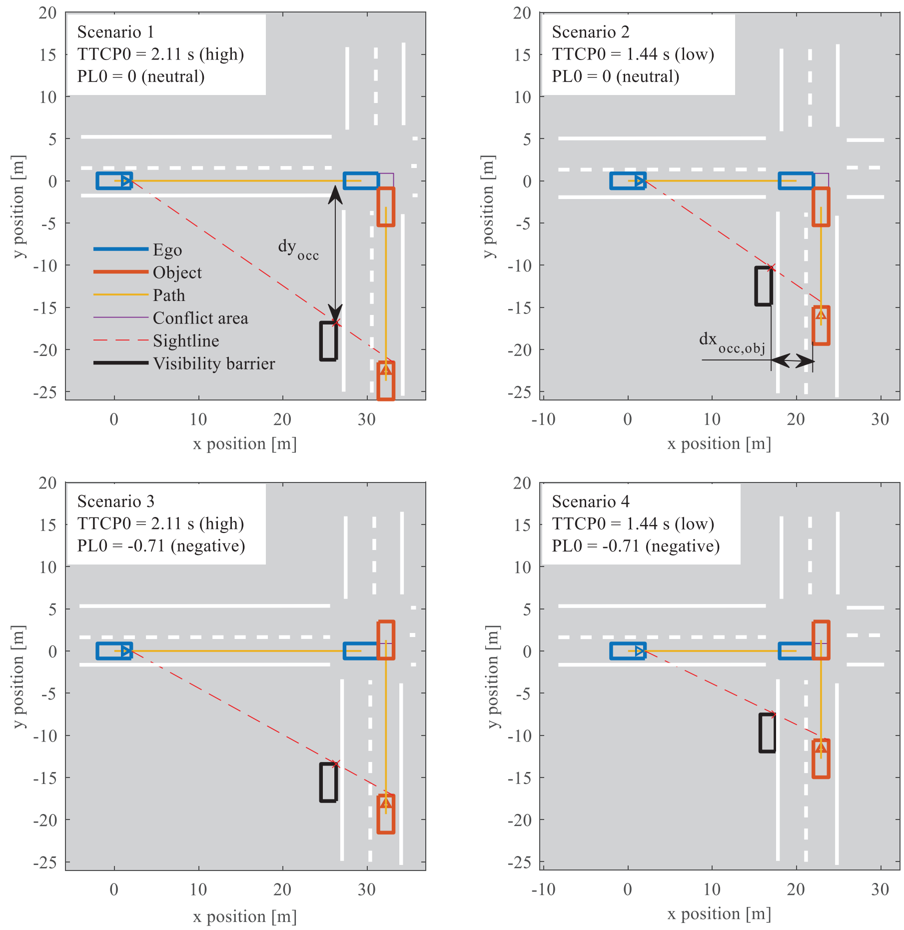


Figure 8. Scenarios used in the driver simulator study

Table 3.
Scenario parameters

	Independent variables		**Further scenario parameters**		
Scenario	TTCP0 [s]	PL0 [-]	theoretical v_{obj} [km/h]	dy_{occ} [m]	$dx_{occ,obj}$ [m]
1	2.11	0	35.2	-16.81	5
2	1.44	0	35.2	-10.29	5
3	2.11	-0.71	35.2	-13.39	5
4	1.44	-0.71	35.2	-7.53	4.5

The inputs on the steering wheel and pedals of each participant were traced and analyzed regarding reaction type, reaction time and intensity. Additionally, the participants answered a questionnaire after each drive to obtain subjective data, e.g., about the perceived criticality of the situation.

The following should be noted regarding the RTYPE classification of the study data:

- Actuator reactions that obviously occur after the crash-relevant situations have been thwarted or occur after the collision are not considered.
- In two cases, the accelerator pedal is depressed before the accelerator pedal is released and the brake pedal is depressed. This reaction pattern is not provided for in the RTYPE classification. It is classified here as RTYPE 11x, since in both situations the PL is increased and thus the initial accelerator pedal

depression has the greater effect than the brake pedal depression. This classification leads to a poor reproducibility of this same driver response.

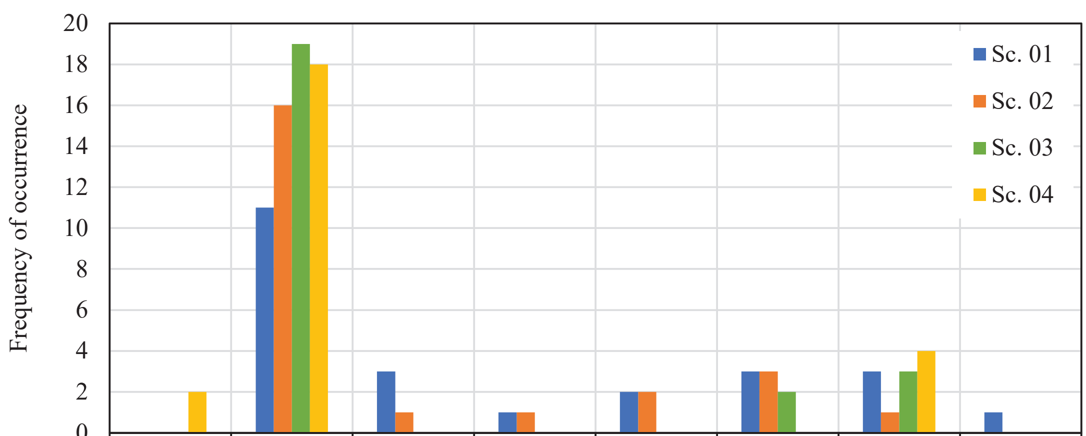
- In five cases, it appears that the driver is applying the brake pedal with the left foot as the accelerator pedal and brake pedal overlap for a short time. This very brief overlap is treated as if the accelerator pedal is released before the brake pedal is released.
- In one case, the driver first steers slightly in one direction and shortly afterwards steers more strongly in the other direction. Here, the steering response of the higher strength is used as the basis for the classification.
- In one case, the onset of braking and the onset of steering occur "simultaneously." In this case, the longitudinal reaction is assumed to be the initial reaction since the accelerator pedal was previously released.
- In two cases, the accelerator pedal is depressed during braking. Due to the very similar accelerator pedal and brake pedal travel, the accelerator pedal was probably inadvertently depressed with the same foot. The accelerator pedal is ignored for the classification.

RESULTS

The driving simulator study executed was used to confirm some of the assumptions the architecture of the driver performance model is based on. As following steps, the results were used to parameterize the model for the scenarios also investigated within the study. The parameterized model was then validated by comparing the frequency of crashes and impact speed with the outcomes of the study executed.

Driving Simulator Study

Figure 9 shows the frequency of occurrence of reaction types divided by scenario type. Among the scenarios, the frequency of occurrence of reaction types and the variability of reactions by the participants varies. The participants most frequently showed a single braking reaction (12x) along all scenarios. Steering to the right (ODS) only occurred after braking and never occurred as single reaction. Steering to the left (SDS) occurred after and before braking, after accelerating as well as a single reaction. Accelerating as single reaction occurred two times in Scenario 4. In Scenarios 1 and 2 (neutral PL), a higher variability in reaction types can be observed than in Scenarios 3 and 4 (negative PL). In Scenarios 1 and 2 (neutral PL), the participants steered sixteen times to the left (SDS) and four times to the right (ODS). In Scenarios 3 and 4 (negative PL), the participants steered two times to the left (SDS) and seven times to the right (ODS). Scenario 4 is the only scenario in which the initial reaction is always a longitudinal reaction.



	11x	12x	21x	31x-Long	33x-Long	33x-Lat	34x-Long	40x
Sc. 01	0	11	3	1	2	3	3	1
Sc. 02	0	16	1	1	2	3	1	0
Sc. 03	0	19	0	0	0	2	3	0
Sc. 04	2	18	0	0	0	0	4	0

Reaction types

Figure 9. Frequency of occurrence of reaction types in each scenario of the study, colored arrows indicating the effect of the reactions on the vehicle dynamics (green: increased acceleration, red: decreased acceleration (typically braking), blue: unchanged longitudinal acceleration)

Regarding our RQ2 about typical first reactions (braking or SDS steering (to the left)) of drivers, we analyzed the observed reaction types. From all of the 95 drives in which a reaction was shown, the participants showed a hypothetical first reaction in 90 of the rides. A Pearson Chi²-Test regarding the incidence of hypothetical typical (braking: [12x, 33x-Long, 34x-Long], SDS steering [21x, 33x-Lat]) vs. untypical first reactions in each scenario supports the hypothesis that the drivers show a hypothetical first reaction in all four scenarios ($X^2(1)$=19,174, p<.001 (Sc.1: n=23); $X^2(1)$=20,167, p<.001 (Sc.2: n=24); $X^2(1)$=16,667, p<.001 (Sc.4: n=24); (Sc.3: n=24) follows consequently, given that only typical first reactions were observed).

With the help of the frequency distributions of the reaction types per scenario, research question 3 is examined. The differences between the two PL conditions lie in the variety of reaction types and in the frequency of occurrence of individual reaction types. For example, pure braking reactions are observed 10 times (27 to 37) less frequently in total in the neutral PL condition than in the negative PL condition. In addition, seven different reaction types occur in total in the neutral PL condition, while four different reaction types can be observed in the negative PL condition. Thus, the results are consistent with hypothesis 3.1. Between the TTCP conditions at the same PL, there is a clear difference only for the frequency of pure braking reactions between scenario 1 and 2 (11 to 16). Based on the data, only a small influence of the TTCP on the driver's choice of action can be assumed. By the fact that almost every driver in the post-survey stated to have perceived the different temporal criticality between the scenarios, this small difference may nevertheless actually be due to the changed criticality. It should be noted that in the studies of [18] and [20], a discontinuity in the frequency course of braking and steering responses over a changed TTCP was observed. In Weber's study, braking reactions increased again at highly crash-relevant events, after the frequency of these had previously decreased with decreasing TTCP [18]. It is possible that the TTCP values in the present study were chosen in an unfortunate way to investigate the hypothesis, so that this discontinuity is reflected in the present study. For hypothesis 3.2, further research is needed at this point.

Parametrization of Driver Performance Model

The study's results are used to parametrize the previously described generic driver performance model. The frequency distribution of shown reaction types is used to parametrize the decision tree (see Figure 5) while the reaction times and reaction intensity are obtained from the traces of the pedal and steering wheel inputs.

In the study, driver behavior was recorded under two different PL and two different TTCP values and their combinations.

Since the literature review and our study show that drivers tend to change their behavior under different scenario conditions, this dependence of the probability of a reaction pattern is also reflected in the parameterization of the decision tree. We considered the influence of the scenario conditions, TTCP and PL, in different ways:

The influence of the priority level is discretized by creating two decision trees (see Figure 12) for different priority level ranges: one tree for the range of PL [-1, -0.4] and one tree for the range of [-0.4, 0.4]. During simulation, the appropriate decision tree is selected based on the priority level perceived by the perceptual module. It would also be possible to overlap ranges. A decision tree is then randomly selected in the overlapping area.

Up to five decisions are made within the decision tree. The probability of a decision is modeled using data from the driving simulator study. The probability is modeled per priority level at two support points of the TTCP. If the perception module of our driving performance model perceived a TTCP between the support points, the probability values at the point of perceived TTCP would be obtained by interpolation.

The assumption that driver behavior can be discretized into groups of similar PLs was not explored in our study. Rather, this assumption was made due to the lack of other experimental data. Also, the question of whether interpolation between TTCP support points is a valid procedure was not investigated in our driving simulator study neither, but would need to be investigated in follow-up studies. Our choice to incorporate the two scenario parameters differently in the parameterization of the decision tree served to demonstrate the possibilities of parameterization. However, we cannot claim that the assumptions made are correct.

In our model, the reaction time depends on the selected reaction type and on the TTCP perceived by the perception module of the driver performance model. This results in a small amount of data available in the study as a basis for the parameterization of reaction time, depending on the reaction type and TTCP condition. In order not to reduce the amount of data further, we refrain from distinguishing the PL condition, but without being able to claim that the PL has no influence on the reaction time.

From the study data, the average reaction time and the standard deviation are determined in order to model a normal distribution of the reaction time. The normal distribution is modelled as a truncated normal distribution so that implausible reaction times cannot arise. Implausible reaction times would exist if the reaction time takes on a negative value or if the reaction time of the second reaction is less than that of the first reaction. If a TTCP value between the two experimental conditions were perceived by the driver performance model, both the mean and the standard deviation would be interpolated. A reaction time is then determined probabilistically.

The conversion time between accelerator pedal release and brake pedal depression is parameterized as 0.2 s for all applicable reaction types. This means that the accelerator pedal is released 0.2 s before the parameterized brake reaction.

The reaction itself is modeled using the algorithms presented previously (Equation (4) – Equation (8)). To be able to model different reaction intensities, up to five reaction intensities per control element intervention are distinguished.

That is, a distinction is made between whether, for example, the brake pedal is depressed very lightly, lightly, moderately, strongly, or very strongly. A parameter set is determined for each of these reaction intensity groups.

The parameterization of the reaction intensity is carried out in six steps:

1. aggregation of the study data with respect to identical controller responses (see Table 10)
2. definition of a characteristic value for the reaction intensity, e.g., maximum pedal position
3. definition of the limits of the reaction intensity groups (see Table 9)
4. set up a linear regression model of the reaction intensity over reaction time
5. determine the probability of occurrence of a reaction intensity group using the linear regression model
6. fitting the control algorithms with the measured data of a reaction intensity group (see Figure 13 and Figure 14).

In the driver performance model, a reaction intensity of the respective control unit is then determined probabilistically as a function of the reaction time. Depending on the reaction intensity group, the respective parameter set for the reaction algorithms.

Results from scenario re-simulation with the parameterized model

As a means of validation for how the model works in applications in effectiveness evaluation, the scenarios that have been simulated in the simulator study have been simulated also with the parameterized model to compare, whether the driver reaction produced by the model produces similar outcomes as observed in the study. For this, we compared the frequency of collisions in the simulated scenario from 100 repetitions with the stochastic model with that from the simulator study. Figure 10 shows that the frequency of collisions is comparable between the study and the simulated driver response.

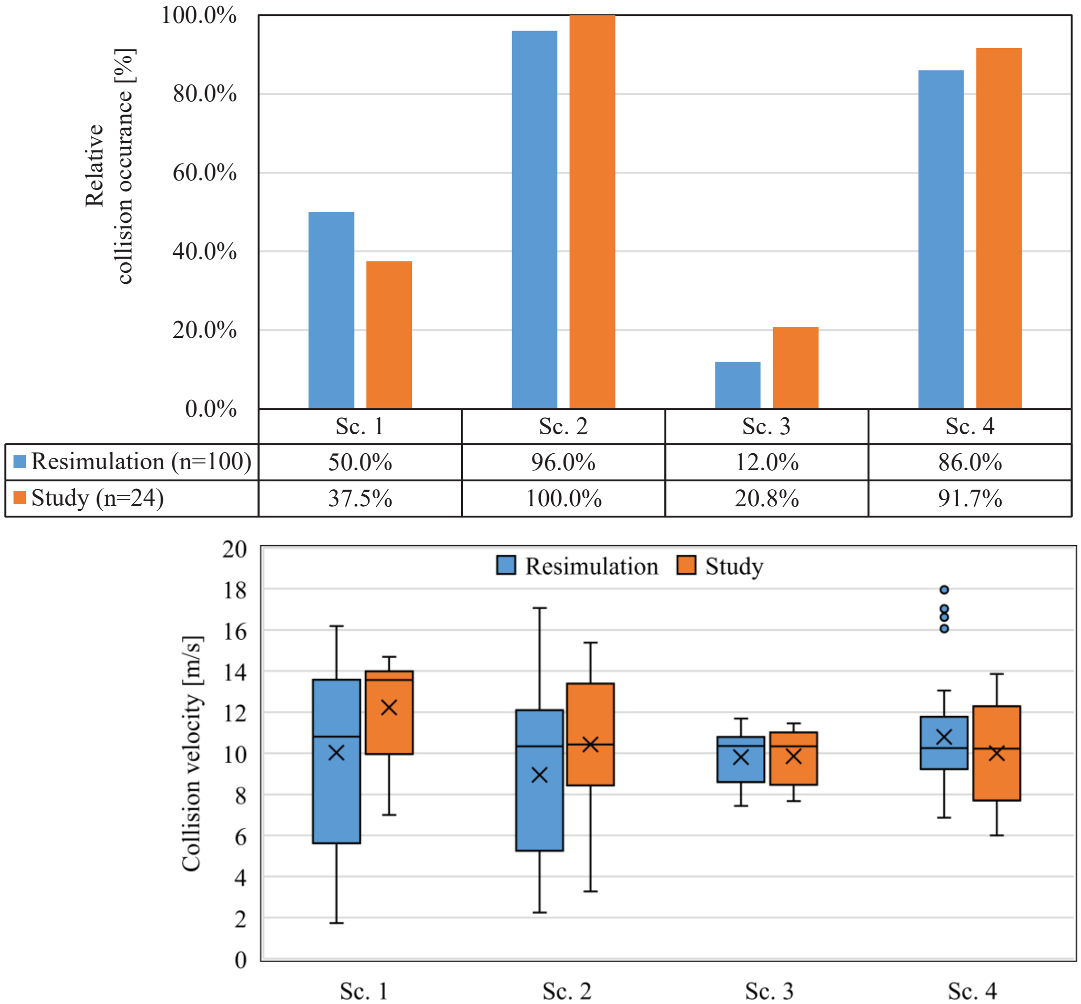


Figure 10. Comparison of the collision occurrence and velocity between the re-simulation and study

For scenarios, in which collisions occurred, we compared the speed at impact of the ego vehicle. For scenarios 3 and 4 (negative priority level) resulting collision speed show comparable distributions. For scenarios 1 and 2 (neutral priority level) distributions of impacts speeds show some deviations, while it needs to be noted, that these scenarios show an overall larger spread in the outcomes of the driver reaction.

DISCUSSION

The model presented in this paper provides a framework for integrating driver behavior within a toolchain for effectiveness assessment of automated driving systems. It supplies a container for processes defining the action decision, the reaction time, and the magnitude of the reaction.

As in principle, the model will follow the complexity of the collection of scenario categories which was used in the effectiveness assessment. Scenario-based evaluations as such run the risk of the curse of dimensionality. While evaluation of systems operating on motorways may allow the evaluation of a rather simple framework of driving scenarios, comprehensive assessments of urban operating domains can result in a wide variety to be tested. It depends on the assessment approach and overall sample size of scenarios whether each scenario to be simulated needs to be simulated multiple times to cover different possible driver reactions.

A general advantage within scenario-based testing is the possibility to focus on certain scenarios, i.e., the most relevant ones, either in terms of overall risk, frequency of occurrence or possible consequences. The driver model framework presented is intended to focus on the hypothetical typical reaction, thus it may be justified to parameterize and simulate not all physically possible driver reactions.

An issue lies within the generalizability of driver reactions: In the study presented, 2x2 scenario configurations were analyzed between participants. The results showed that there was a difference in driver reaction between the simulated scenarios. In principle, the parameters obtained for these scenarios could be extrapolated linearly to cover more possible parameter configurations, but especially in terms of driver reaction, linear extrapolation does not appear plausible. To parameterize a model, that can cover all parameter combinations of the scenario to be studied, a large-scale study would need to be executed and statistical methods to be applied to parameterize the model for the entire value range of the scenario. To apply the model within other scenarios, an analysis is needed to see which scenario parameters have most influence on the driver reaction, either based on literature or on dedicated studies of the scenario.

Even though there is the danger of the curse of dimensionality, applying a modular model as presented, would enable sharing the findings of individual studies to gradually cover the most relevant scenarios, such that over time, driver modeling of such scenarios consists of less and less black spots and holistic assessment of e.g., level 4 automated driving systems will be feasible.

Similar to driver performance models in general, the proposed model faces the challenge, that for the situations simulated a stimulus for the reaction needs to be defined (for further elaboration see [5]). This is possible in experimental conditions, like in the study performed, where the object vehicle was occluded while the ego vehicle entered the intersection. The object vehicle becoming visible could easily be used as the stimulus for determining the reaction time. In other situations, this may not be possible. For instance in a rear-end scenario with the leading vehicle already in deceleration and then applying a sudden stronger deceleration, the brake lights could not be used as stimulus for the driver response. The model proposed in [5] could provide solution to this challenge and could be integrated as rule-based perception model on the guidance layer (GPr). While [5] discusses, how the evolving situation can affect the driver's reaction choice – which was also one of the core assumptions of our model, though achieved through simpler metrics referencing a distinct stimulus – it does not yet provide a solution how the reaction choice may be implemented. As shown in literature (e.g., [22]) as well as is the study presented, parameters of the situation have an effect on the reaction choice of the driver. The choice of reaction can of course have great impacts on the outcome of the crash-relevant situation. As long as no model exists, which can implement both, a reaction time independent of the stimulus as well an influence of situational parameters on the action decision, it needs to be carefully considered, which aspect may be more relevant to the targeted evaluation.

CONCLUSIONS AND FUTURE WORK

This paper presents an architecture and an application of a generic driver performance model. The model used a fundamental structuring of the driver reaction by means of the levels of the driving task by Donges [10], the structuring of goal-oriented activities by Rasmussen [12], and the general information processing chain of a driver. Using a structure as presented and standardized interfaces as OSI [38] will allow an easy exchange of model components and an easier documentation of the baseline simulations within a predictive effectiveness assessment, which would ultimately result in a greater acceptance and comparability of results.

As next steps, the model needs to be parameterized for more driving scenarios, using greater samples to cover the entire driver population in terms of age, gender or driving experience. The parameterization of different scenarios can be driven by its intended applications, e.g., by focusing first on scenarios relevant for motorway ADS. The modularization of the model and the separation of modeling and parameterization would allow an easy reuse of components and parameter sets established. Moreover, more complex models for perception and cognition could be integrated depending on the use case. A tradeoff should be considered between a comprehensive modeling of the actual perception and cognition processes and a more straightforward though simplified approach using distributions of reaction times.

In principle, the application of the model faces the same pitfalls as scenario-based testing for ADS in general, but will also profit from its advancements. Research in scenario-based testing tries to identify the most relevant scenarios for an ADS such that an efficient yet comprehensive evaluation process can be established, for instance using a scenario framework as presented in [45]. This reduction of the scenario-space will also reduce the effort for parameterizing the driver performance model. Given that the variety of driver reactions requires the modeling of stochastic processes, each concrete scenario needs to be simulated multiple times, increasing simulation efforts. At the same time – once a concrete scenario has been simulated using the model – the results can be reused multiple times, in contrast to the V&V process of an ADF where each system iteration needs a repetition of simulations. Furthermore, findings on the typical driver reactions within crash-relevant scenarios (see [40]) may provide an acceptable reduction in driver reactions to be considered.

Within future work, we will establish a connection between the driver performance model presented and models that cover normal driving phase. A suitable model for this is the model by Klimke, as it follows a similar structuring of the driving task [15]. This may allow a comprehensive model that includes traceable crash causation mechanism which can also be used to induce crash-relevant situations. This way, the model could also be used within a traffic simulation-based assessment, where the causation of a crash does not need to be scripted within the scenario but is a result from the interaction of driver models. This approach would be similar to the approach in [16] but at the same time allows a more detailed modelling of driver reaction within the crash relevant scenario.

REFERENCES

[1] Kraftfahrt-Bundesamt, *KBA erteilt erste Genehmigung zum automatisierten Fahren (KBA grants first persmission for automated driving),* 2021.

[2] SAE International, "J3016 - Taxonomy and Definitions for Terms Related to Driving Automation Systems for On-Road Motor Vehicles," SAE International, 2021.

[3] Waymo, "Waymo One - Ride-hailing service," Waymo LLC, Mountain View, CA, USA, 2022.

[4] Ethikkomission Automatisiertes und Vernetztes Fahren, "Bericht Juni 2017," Federal Ministry for Digital and Transport, Berlin, Germany, 2017.

[5] J. Engström, S.-Y. Liu, A. Dinparastdjadid and C. Simoiu, "Modeling road user response timing in naturalistic settings: a surprise-based framework," 18 August 2022.

[6] N. Kalra and S. M. Paddock, "Driving to safety: How many miles of driving would it take to demonstrate autonomous vehicle reliability?," *Transportation Research Part A: Policy and Practice,* vol. 94, p. 182–193, December 2016.

[7] H. Winner, K. Lemmer, T. Form and J. Mazzega, "PEGASUS - First Steps for the Safe Introduction of Automated Driving," *Road Vehicle Automation 5,* p. 185–195, 26 June 2018.

[8] ISO, ISO/DIS 34501 - Road vehicles - Terms and definitions of test scenarios for, Vernier, Geneva: ISO, 2022.

[9] A. Bjorvatn, Y. Page, F. Fahrenkrog, H. Weber, E. Aittoniemi, E. Lehtonen, H. Per, A. Silla, J. Bärgman, M. Borrack, S. Innamaa, T. Itkonen, F. Malin, K. Pedersen, M. Schuldes, H. Sintonen, T. Streubel, H. Thierry, W. Hagleitner, J. Hiller and G. Torrao, "Deliverable D7.4 - Impact Evaluation Results," L3Pilot Project, 2021.

[10] E. Donges, "Aspekte der aktiven Sicherheit bei der Fuehrung von Personenkraftwagen," *Automobil-Industrie,* vol. 27, p. 183–190, 1982.

[11] E. Donges, "Fahrerverhaltensmodelle," in *Handbuch Fahrerassistenzsysteme*, H. Winner, S. Hakuli, F. Lotz and C. Singer, Eds., Wiesbaden, Springer Vieweg, 2015, p. 17–26.

[12] J. Rasmussen, "Skills, rules, and knowledge; signals, signs, and symbols, and other distinctions in human performance models," *IEEE Transactions on Systems, Man, and Cybernetics,* Vols. SMC-13, p. 257–266, 1983.

[13] M. Treiber, A. Hennecke and D. Helbing, "Congested traffic states in empirical observations and microscopic simulations," *Physical Review,* pp. 1805 - 1824, Aug 2000.

[14] R. Wiedemann, Simulation des Straßenverkehrsflusses, Schriftenreihe des Instituts für Verkehrswesen, 1074.

[15] J. Klimke, D. Becker and L. Eckstein, "System Design of an Agent Model for theClosed-Loop Simulation of Relevant Scenariosin the Development of ADS," in *29th Aachen Colloquium Sustainable Mobility*, 2020.

[16] A. Fries, F. Fahrenkrog, K. Donauer, M. Mai and F. Raisch, "Driver Behavior Model for the Safety Assessment of Automated Driving," in *2022 IEEE Intelligent Vehicles Symposium (IV)*, 2022.

[17] M. Green, ""How Long Does It Take to Stop?"Methodological Analysis of Driver Perception-Brake Times," *Transportation Human Factors,* vol. 2, p. 195–216, 2000.

[18] S. Weber and B. Farber, "Driver Reactions in Critical Crossing Situations," in *2015 IEEE 18th International Conference on Intelligent Transportation Systems (ITSC)*, Piscataway, 2015.

[19] S. Weber, "Standardreaktionen in kritischen Fahrsituationen: Studien mit dem Vehicle in the Loop zur Untersuchung des Reaktionsverhaltens bei Vorfahrtsmissachtung von rechts," München, 2015.

[20] J. M. Hankey, D. V. McGehee, T. A. Dingus, E. N. Mazzae and W. R. Garrott, "Initial Driver Avoidance Behavior and Reaction Time to an Unalerted Intersection Incursion," *Proceedings of the Human Factors and Ergonomics Society Annual Meeting,* p. 896–899, 1996.

[21] M. Hu and Y. Li, "Drivers' avoidance patterns in near-collision intersection conflicts," in *2017 IEEE 20th International Conference on Intelligent Transportation Systems (ITSC)*, Piscataway, 2017.

[22] M. Hu and Y. Li, "Same-Direction-Pattern of Drivers' Avoidance Behaviour in Critical Intersection Situations–Field Accident Data Analysis and Driving Simulation Study," in *2017 IRCOBI Conference*, 2017.

[23] X. Li, A. Rakotonirainy and X. Yan, "How do drivers avoid collisions? A driving simulator-based study," *Journal of Safety Research,* vol. 70, p. 89–96, September 2019.

[24] R. Jurecki and T. L. Stanczyk, "Driver model for the analysis of pre-accident situations," *Vehicle System Dynamics,* vol. 47, p. 589–612, 2009.

[25] R. Jurecki and T. Stańczyk, "Driver model for the analysis of pre-accident situations - Similarities and differences of result on a track and in a simulator.," *Archiwum Motoryzacji, Archives of Automotive Engineering,* pp. 149-172, 2011.

[26] Y. Page, F. Fahrenkrog, A. Fiorentino, J. Gwehenberger, T. Helmer, M. Lindman, S. Puch, F. Martin, U. Sander and P. Wimmer, "A comprehensive and harmonized method for assessing the effectiveness of advanced driver assistance systems by virtual simulation: the PEARS initiative," in *The 24th International Technical Conference on the Enhanced Safety of Vehicles (ESV)*, 2015.

[27] 7. ISO/TC 22/SC 36/WG, "ISO/PDTR 21934: Road vehicles — Traffic safety analysis — Prospective safety systemassessment by virtual simulation," 2017.

[28] F. Fahrenkrog, L. Wang, C. Rösener, J. Sauerbier and S. Breunig, *Deliverable D7.3 //: Impact analysis for supervised automated driving applications,* AdaptIVe Project, 2017.

[29] C. Erbsmehl and H. Schebdat, "Statistical driver model for accident simulation: Using a statistical driver model for benefit estimation of advanced safety systems with warning interfaces," in *6th International Conference on ESAR 2014*, 2015.

[30] C. Roesener, J. Sauerbier, A. Zlocki, L. Eckstein, F. Hennecke, D. Kemper and M. Oeser, *Potenzieller Gesellschaftlicher Nutzen durch Zunehmende Fahrzeugautomatisierung,* Bergisch Gladbach: Bundesanstalt für Straßenwesen, 2019.

[31] C. Roesener, "A traffic-based method for safety impact assessment of road vehicle automation - PhD Thesis," RWTH Aachen University, Aachen, 2020.

[32] C. Roesener, M. Harth, H. Weber, J. Josten and L. Eckstein, "Modelling Human Driver Performance for Safety Assessment of Road Vehicle Automation," in *21st IEEE International Conference on Intelligent Transportation Systems (ITSC 2018)*, 2018.

[33] E. H. Yilmaz and W. H. Warren, "Visual Control of Braking: A Test of the $\dot{\tau}$ Hypothesis," *Journal of Experimental Psychology: Human Perception and Performance,* vol. 21, p. 996–1014, 1995.

[34] O. Benderius, "Modelling driver steering and neuromuscular behaviour - Doctoral Thesis," Chalmers University of Technology, Göteborg, 2014.

[35] J. Bärgman, C.-N. Boda and M. Dozza, "Counterfactual simulations applied to SHRP2 crashes: The effect of driver behavior models on safety benefit estimations of intelligent safety systems," *Accident Analysis & Prevention,* vol. 102, p. 165–180, May 2017.

[36] D. N. Lee, "A theory of visual control of braking based on information about time to collision," *Perception,* vol. 5, p. 437–459, 1976.

[37] UN ECE, UN Regulation No. 157 - Uniform provisions concerning the approval of vehicles with regard to, Geneva, Switzerland, 2021.

[38] T. Hanke, N. Hirsenkorn, C. van-Driesten, P. Garcia-Ramos, Schiementz, M. S. and E. Biebl, "A generic interface for the environment perception of automated driving functions in virtual scenarios," 2017.

[39] H. Pashler, "Dual-task interference in simple tasks Data and theory," *Psychological Bulletin,* vol. 116, p. 220–244, 1994.

[40] S. Weber, K. Blum, A. Ernstberger and B. Färber, "Standard Reactions – Driver Reactions in Critical Driving Situations," in *6th International Conference on Applied Human Factors and Ergonomics (AHFE 2015) and the Affiliated Conferences, AHFE 2015*, 2015.

[41] Statistisches Bundesamt Destatis, Verkehrsunfälle - Fachserie 8 Reihe 7 - 2019, Wiesbaden: Statistisches Bundesamt Destatis, 2020.

[42] M. Wisch, A. Hellmann, M. Lerner, T. Hierlinger, V. Labenski, M. Wagner, H. Feifel, O. Robescu, P. Renoux and X. Groult, "Car-to-car accidents at intersections in Europe and identification of Use Cases for the test and assessment of respective active vehicle safety systems," in *26th International Technical Conference on the Enhanced Safety of Vehicles (ESV)*, Washington, D.C., 2019.

[43] O. Benderius and G. Markkula, "Evidence for a fundamental property of steering," in *Proceedings of the Human Factors and Ergonomics Society*, Vols. 2014-January, Los Angeles, SAGE Publications, 2014, p. 884–888.

[44] M. Burckhardt, Reaktionszeiten bei Notbremsvorgängen, Köln: Verlag TÜV Rheinland GmbH, 1985.

[45] H. Weber, J. Bock, J. Klimke, C. Roesener, J. Hiller, R. Krajewski, A. Zlocki and L. Eckstein, "A framework for definition of logical scenarios for safety assurance of automated driving," *Traffic injury prevention,* vol. 20, pp. S65-S70, June 2019.

APPENDIX

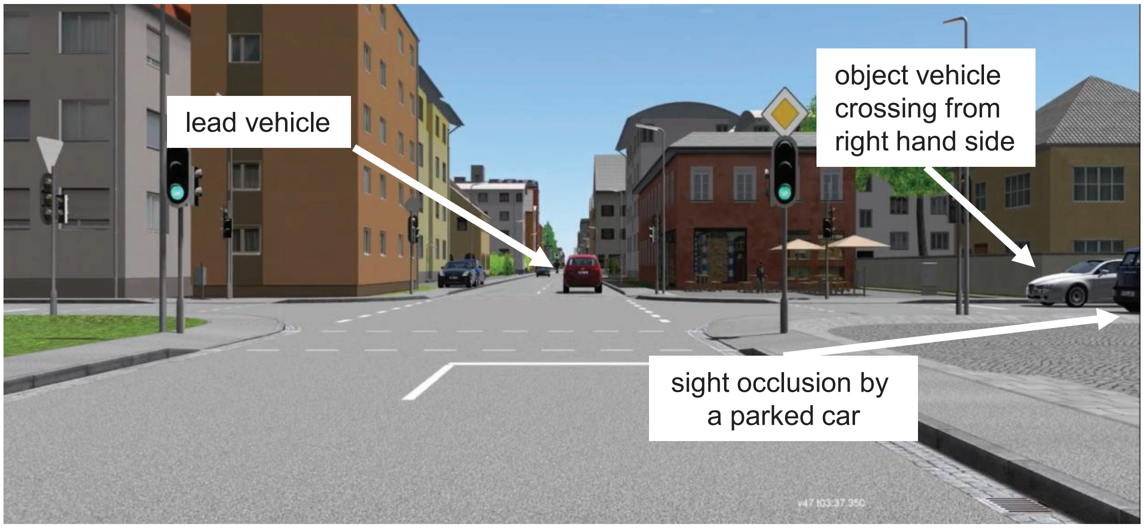


Figure 11. Annotated screen picture of scenario no. 3 of the study

Table 4.
Low-level RTYPE classification (1 of 3)

No.	Effect	Mid-level code	Control unit usage sequence	Control unit action	Low-level code
	RTYPE 1xx: longitudinal reaction				**1xx**
1	increase longitudinal acceleration	**11x**	A	Ai	11-1
2		**11x**	B	Bd	11-2
3		**11x**	A-B	Ai Bd	11-3
4		**11x**	B-A	Bd Ai	11-4
5	decrease longitduinal acceleration	**12x**	B	Bi	12-1
6		**12x**	A-B	Ad Bi	12-2
7		**12x**	B-A	Bi Ad	12-3
8	undefined longitudinal acceleration	1xx	A-B	Ai Bi	1x-1
9		1xx	B-A	Bi Ai	1x-2
10		1xx	A-B	Ad Bd	1x-3
11		1xx	B-A	Bd Ad	1x-4
	RTYPE 2xx: lateral reaction				**2xx**
12	increase lateral acceleration	**21x**	S	Sl	21-1
13		**21x**	A-S	Ad Sl	21-2
14		**21x**	S-A	Sl Ad	21-3
15	decrease lateral acceleration	**22x**	S	Sr	22-1
16		**22x**	A-S	Ad Sr	22-2
17		**22x**	S-A	Sr Ad	22-3
	RTYPE 3xx: combined reaction (1 of 3)				**3xx**
18	increase longitudinal acceleration + increase lateral acceleration	**31x-Long**	A-S	Ai Sl	31-1-Long
19		**31x-Long**	B-S	Bd Sl	31-2-Long
20		**31x-Long**	A-B-S	Ai Bd Sl	31-3-Long
21		**31x-Long**	B-A-S	Bd Ai Sl	31-4-Long
22		**31x-Long**	A-S-B	Ai Sl Bd	31-5-Long
23		**31x-Long**	B-S-A	Bd Sl Ai	31-6-Long
24		**31x-Lat**	S-A	Sl Ai	31-1-Lat
25		**31x-Lat**	S-B	Sl Bd	31-2-Lat
26		**31x-Lat**	S-A-B	Sl Ai Bd	31-3-Lat
27		**31x-Lat**	S-B-A	Sl Bd Ai	31-4-Lat

Legend
A: Accelerator
B: Brake pedal
S: Steering
i: increase
d: decrease
l: left, r: right
Long: longitudinal
Lat: lateral

Table 5.
Low-level RTYPE classification (2 of 3)

No.	Effect	Mid-level code	Control unit usage sequence	Control unit action	Low-level code
	RTYPE 3xx: combined reaction (2 of 3)				**3xx**
28	increase longitudinal acceleration + decrease lateral acceleration	**32x-Long**	A-S	Ai Sr	32-1-Long
29		**32x-Long**	B-S	Bd Sr	32-2-Long
30		**32x-Long**	A-B-S	Ai Bd Sr	32-3-Long
31		**32x-Long**	B-A-S	Bd Ai Sr	32-4-Long
32		**32x-Long**	A-S-B	Ai Sr Bd	32-5-Long
33		**32x-Long**	B-S-A	Bd Sr Ai	32-6-Long
34		**32x-Lat**	S-A	Sr Ai	32-1-Lat
35		**32x-Lat**	S-B	Sr Bd	32-2-Lat
36		**32x-Lat**	S-A-B	Sr Ai Bd	32-3-Lat
37		**32x-Lat**	S-B-A	Sr Bd Ai	32-4-Lat
38	decrease longitudinal acceleration + increase lateral acceleration	**33x-Long**	B-S	Bi Sl	33-1-Long
39		**33x-Long**	A-B-S	Ad Bi Sl	33-2-Long
40		**33x-Long**	B-A-S	Bi Ad Sl	33-3-Long
41		**33x-Long**	B-S-A	Bi Sl Ad	33-4-Long
42		**33x-Lat**	S-B	Sl Bi	33-1-Lat
43		**33x-Lat**	A-S-B	Ad Sl Bi	33-2-Lat
44		**33x-Lat**	S-A-B	Sl Ad Bi	33-3-Lat
45		**33x-Lat**	S-B-A	Sl Bi Ad	33-4-Lat
46	decrease longitudinal acceleration + decrease lateral acceleration	**34x-Long**	B-S	Bi Sr	34-1-Long
47		**34x-Long**	A-B-S	Ad Bi Sr	34-2-Long
48		**34x-Long**	B-A-S	Bi Ad Sr	34-3-Long
49		**34x-Long**	B-S-A	Bi Sr Ad	34-4-Long
50		**34x-Lat**	S-B	Sr Bi	34-1-Lat
51		**34x-Lat**	A-S-B	Ad Sr Bi	34-2-Lat
52		**34x-Lat**	S-A-B	Sr Ad Bi	34-3-Lat
53		**34x-Lat**	S-B-A	Sr Bi Ad	34-4-Lat
54	undefined longitudinal acceleration + increase lateral acceleration	3xx	A-B-S	Ai Bi Sl	3x-1-Long
55		3xx	B-A-S	Bi Ai Sl	3x-2-Long
56		3xx	A-S-B	Ai Sl Bi	3x-3-Long
57		3xx	B-S-A	Bi Sl Ai	3x-4-Long
58		3xx	A-B-S	Ad Bd Sl	3x-5-Long
59		3xx	B-A-S	Bd Ad Sl	3x-6-Long
60		3xx	B-S-A	Bd Sl Ad	3x-7-Long
61		3xx	S-A-B	Sl Ai Bi	3x-1-Lat
62		3xx	S-B-A	Sl Bi Ai	3x-2-Lat
63		3xx	S-A-B	Sl Ad Bd	3x-3-Lat
64		3xx	S-B-A	Sl Bd Ad	3x-4-Lat
65		3xx	A-S-B	Ad Sl Bd	3x-5-Lat

Legend
A: Accelerator
B: Brake pedal
S: Steering
i: increase
d: decrease
l: left, r: right
Long: longitudinal
Lat: lateral

Table 6.
Low-level RTYPE classification (3 of 3)

No.	Effect	Mid-level code	Control unit usage sequence	Control unit action	Low-level code
	RTYPE 3xx: combined reaction (3 of 3)				**3xx**
66	undefined longitudinal acceleration + decrease lateral acceleration	3xx	A-B-S	Ai Bi Sr	3x-1-Long
67		3xx	B-A-S	Bi Ai Sr	3x-2-Long
68		3xx	A-S-B	Ai Sr Bi	3x-3-Long
69		3xx	B-S-A	Bi Sr Ai	3x-4-Long
70		3xx	A-B-S	Ad Bd Sr	3x-5-Long
71		3xx	B-A-S	Bd Ad Sr	3x-6-Long
72		3xx	B-S-A	Bd Sr Ad	3x-7-Long
73		3xx	S-A-B	Sr Ai Bi	3x-1-Lat
74		3xx	S-B-A	Sr Bi Ai	3x-2-Lat
75		3xx	S-A-B	Sr Ad Bd	3x-3-Lat
76		3xx	S-B-A	Sr Bd Ad	3x-4-Lat
77		3xx	A-S-B	Ad Sr Bd	3x-5-Lat
	RTYPE 4xx: no reaction				**4xx**
78	no reaction	**40x**	n.a.	n.a.	40-0
79		**40x**	A	Ad	40-1

Legend
A: Accelerator
B: Brake pedal
S: Steering
i: increase
d: decrease
l: left, r: right
Long: longitudinal
Lat: lateral

Table 7.
Parameters and dependencies of model variables RTYPE, RT and RINT

	Reaction type (RTYPE)	**Reaction time (RT)**	**Reaction intensity (RINT)**
Dependencies	Scenario dependent (TTCP, PL)	Scenario dependent (TTCP), reaction type dependent (RTYPE)	Reaction time dependent (RT)
Parameter	Probability of occurrence of a RTYPE	Control unit specific reaction time (mean value, standard deviation)	Control unit specific and algorithm specific parameters

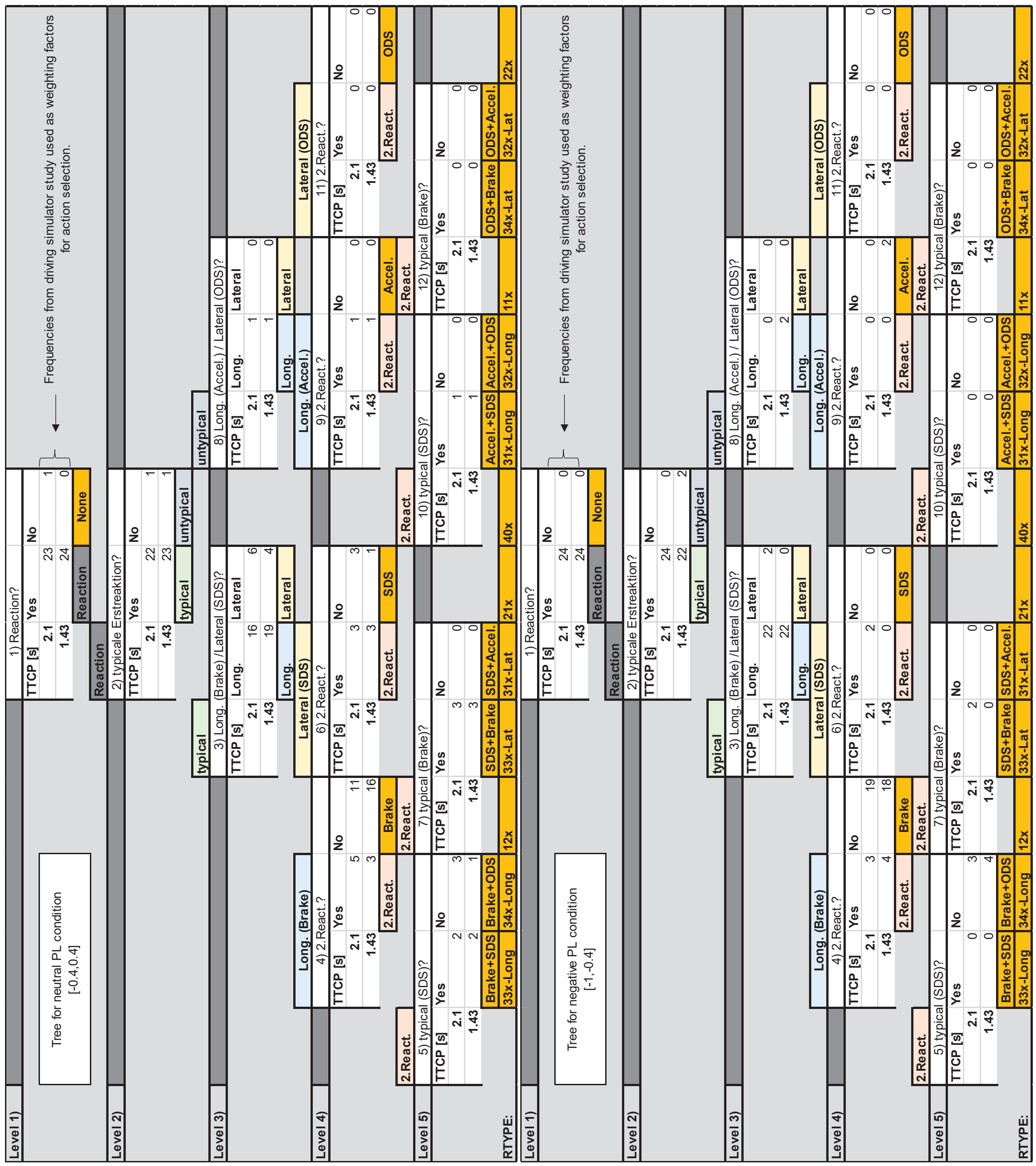


Figure 12. Parametrized decision trees for the priority level range of [-0.4,0.4] and [-1,-0.4]

Table 8.
Parameter values of the RTYPE-specific reaction times

ØTTCP [s]	**RTYPE**	**Frequency of occurrence in study**	**RT [s]**					
			Accelerator pedal		**Brake pedal**		**Steering wheel**	
			MW	**Std.**	**MW**	**Std.**	**MW**	**Std.**
1.42	**11x**	2	0.642	0.153	-	-	-	-
1.43	**12x**	34	-	-	0.826	0.223	-	-
1.45	**21x**	1	-	-	-	-	1.267	0.000
1.42	**31x-Long**	1	0.633	0.000	-	-	0.917	0.000
1.42	**33x-Long**	2	-	-	0.717	0.047	1.025	0.153
1.43	**33x-Lat**	3	-	-	0.917	0.202	0.850	0.188
1.43	**34x-Long**	5	-	-	0.757	0.158	1.123	0.119
2.10	**12x**	30	-	-	0.896	0.240	-	-
2.11	**21x**	3	-	-	-	-	1.628	0.208
2.10	**31x-Long**	1	1.433	0.000	-	-	1.833	0.000
2.09	**33x-Long**	2	-	-	0.950	0.236	1.967	0.613
2.11	**33x-Lat**	5	-	-	1.437	0.140	1.083	0.216
2.10	**34x-Long**	6	-	-	0.783	0.211	1.189	0.323

Table 9.
Key values for the subdivision of reactions into reaction intensity groups

Control unit reaction:	**Operate accelerator pedal**		**Operate brake pedal**		**Steer left**		**Steer right**	
Key value:	Max. accelerator position [0-1]		Max. brake pedal position [0-1]		Max. steering wheel angle [°]		Max. steering wheel angle [°]	
RINT group	From	To	From	To	From	To	From	To
1 (very low)	0	0.2	0	0.2	0	24	0	24
2 (low)	0.2	0.4	0.2	0.4	24	48	24	48
3 (medium)	0.4	0.6	0.4	0.6	48	72	48	72
4 (high)	0.6	0.8	0.6	0.8	72	96	72	96
5 (very high)	0.8	1	0.8	1	96	120	96	120

Table 10.
Aggregated control unit reactions and frequency of occurrence in study data

Control unit reaction	Origin of the parameters for the reaction intensity specific algorithms						
	Aggregated RTYPEs	Frequency of the RINT group in study data					
			RINT group				
		∑	1	2	3	4	5
Operate accelerator pedal	11x, 31x-Long, 32x-Long, 31x-Lat, 32x-Lat	4	0	0	0	0	4
Operate brake pedal	12x, 33x-Long, 33x-Long, 33x-Lat, 34x-Lat	87	1	1	5	10	70
Steer left	21x, 31x-Lat, 33x-Lat, 31x-Long, 33x-Long	18	4	7	1	1	5
Steer right	22x, 32x-Lat, 34x-Lat, 32x-Long, 34x-Long	11	2	5	2	2	0

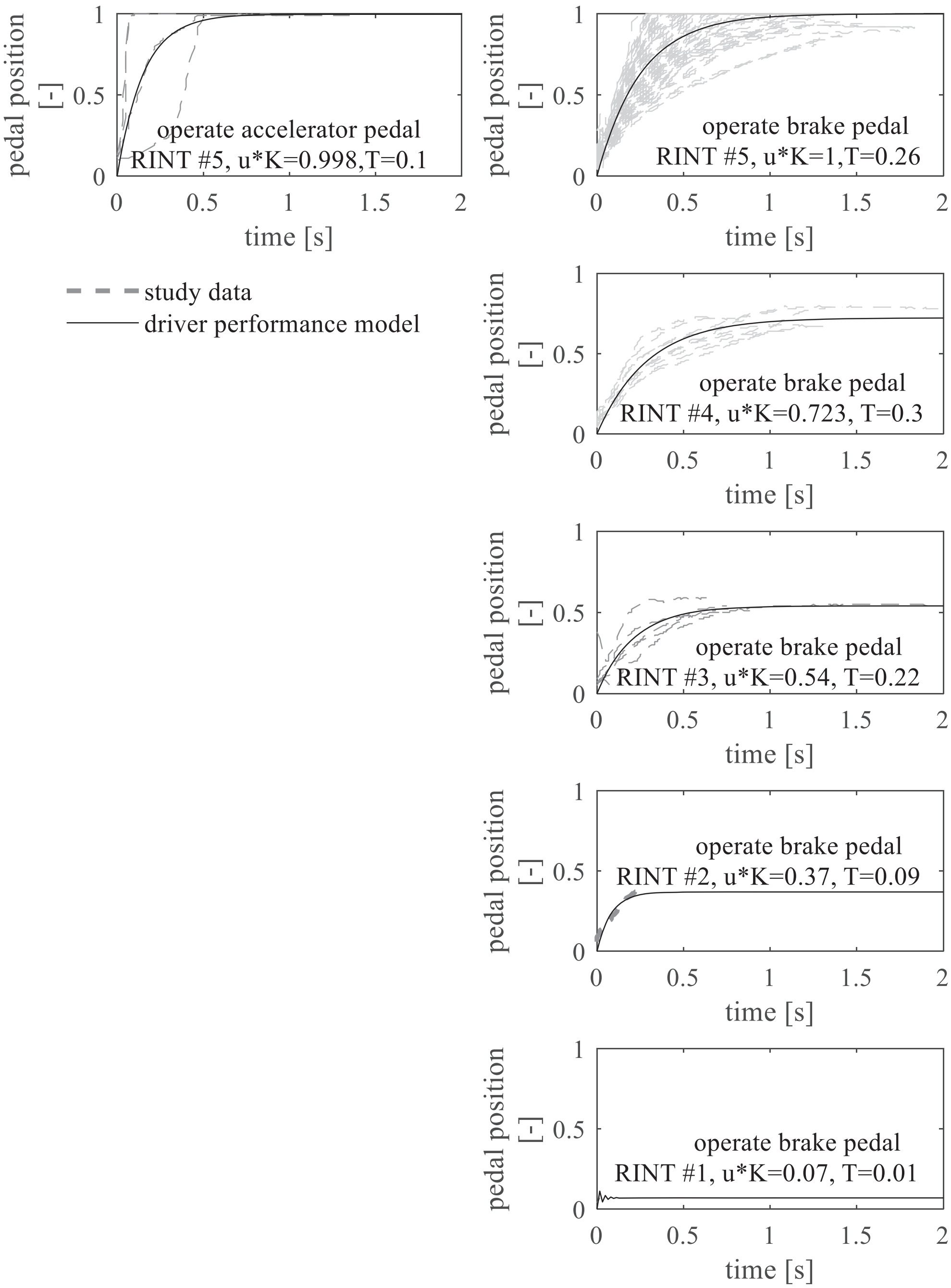


Figure 13. Reaction intensity specific parameters of the pedal control algorithm of the driver performance model

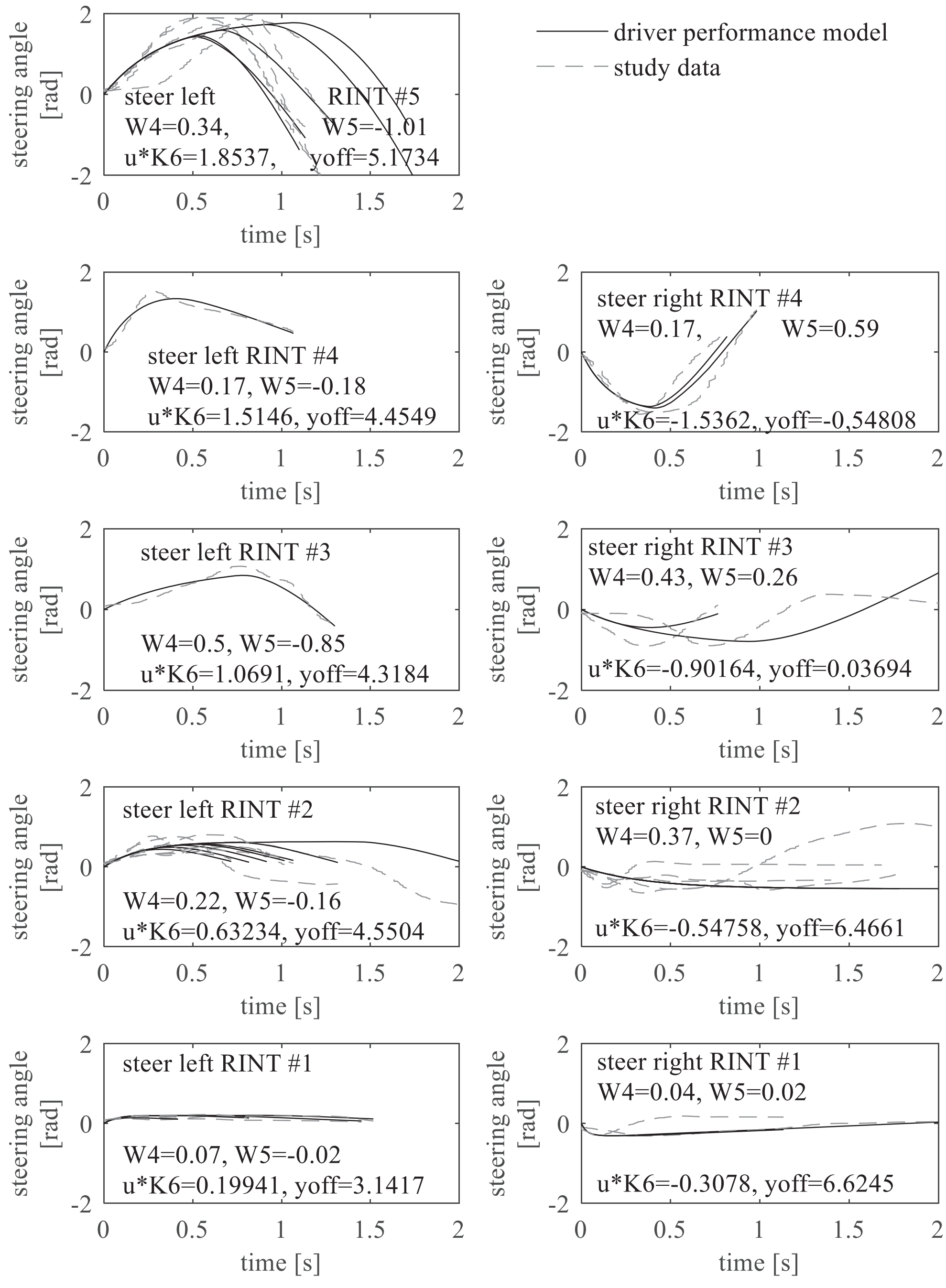


Figure 14. Reaction intensity specific parameters of the steering algorithm of the driver performance model

```
{
  "$schema": "http://json-schema.org/draft/2019-09/schema",
  "title": "Parameter file",
  "required": ["scenario", "RTYP"],
  "properties": {
    "scenario": {
      "description": "C1"
    },
    "RTYP": {
      "12x": {
        "comment": "acceleration: long neg, lat none",
        "classification": "typical",
        "RT": {
          "long": {
            "independent_var": {
              "comment": "'name' needs to be attribute of driver, e.g. 'ttcp' (driver.ttcp)",
              "name": "ttcp",
              "val": [1.43,2.10]
            },
            "mean_val": [0.826,0.896],
            "std": [0.223,0.240],
            "dist": "normal",
            "device": "brake",
            "rt_transfer": -0.2
          }
        },
        "RINT": {
          "long": {
            "independent_var": {
              "comment": "'name' needs to be attribute of driver, e.g. 'ttcp' (driver.ttcp)",
              "name": "rt_long",
              "val":[0,0.016667,0.033333,0.05,0.066667,0.083333,0.1,0.11667,...]
            },
            "branches": ["very_low","low","mid","high","very_high"],
            "weights_very_low": [0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,...],
            "weights_low": [0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,,...],
            "weights_mid": [0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0.,...],
            "weights_high": [0,0,0,0,0,0,9.9285e-05,9.9374e-05,9.9453e-05,0.00019902,0.00029872,...],
            "weights_very_high": [1,1,1,1,1,1,0.9999,0.9999,0.9999,0.9998,0.9997,0.9996,0.9994,...]
          }
        },
        "inputs": {
          "comment": "low level RTYP",
          "long": {
            "low": {
              "comment": "Accelerator lift and low intensity brake apply",
              "A": {
                "comment": "define input here",
                "accelerator_open_loop_target": 0,
                "accelerator_open_loop_gain": 1,
                "accelerator_open_loop_timeconstant": 0.1
              },
              "B": {
                "comment": "define input here",
                "brake_pedal_open_loop_target": 0.37,
                "brake_pedal_open_loop_duration": 10,
                "brake_pedal_open_loop_gain": 1,
                "brake_pedal_open_loop_timeconstant": 0.09,
                "tau_control": "False"
              }
            },
            "mid": {
              "comment": "Accelerator lift and medium intensity brake apply",
              ...
            },
            ...
          }
        }
      },
      ...
    }
  }
}
```

Figure 15. Exemplary implementation of decision tree in model based on study results

ADAPTING APPROVAL REGULATIONS TO ACCOMMODATE AUTOMATED VEHICLES

Matthias Seidl, Dr Mervyn Edwards, Dr Kostas Kourantidis, Jack Radcliffe, Tadas Andriejauskas, Dr Tim Barlow
Transport Research Laboratory (TRL Limited)
United Kingdom

Tony Soper
UTAC
United Kingdom

Paper Number 23-0174

ABSTRACT

The UK government are committed to bringing forward legislation to allow the safe and secure deployment of self-driving vehicles, as set out in the recent policy paper Connected & Automated Mobility 2025: Realising the benefits of self-driving vehicles in the UK. As part of the Connected and Automated Vehicles Process for Assuring Safety and Security (CAVPASS) programme, TRL was commissioned to propose approaches to vehicle classification, and suitable technical requirements for aspects not related to the Automated Driving System (ADS). These included crashworthiness, occupant protection, protection of Vulnerable Road Users (VRUs), and the lighting, braking and steering systems.

The initial focus of this work was on Low-Speed Automated Vehicles (LSAVs). The work involved selection and adaptation of existing pre- and post-deployment regulation to enable it to be applied to LSAVs. A main part was the adaptation of the technical regulations for M- and N-category vehicles, laid down in Great Britain's Road Vehicles (Approval) Regulations 2020 (SI 2020 No. 818), which implements retained Regulation (EU) 2018/858.

The study proposed the introduction of two new vehicle categories (for LSAVs with and without occupants, respectively) to allow approval of designs not compatible with the M- and N-category definitions, such as passenger shuttles with six seats and space for standing passengers, or goods vehicles without any seats. Of 132 technical items collated from the existing body of regulations, 65 were found to be generally applicable for LSAVs with occupants, and 53 for LSAVs without occupants. Technical clarifications for regulations were developed relating to references to the driver or driver's seat, controls, warnings and tell-tales and relating to bi-directional vehicles in general. The study further found that a general permission to carry standing passengers in light vehicles could present unreasonable risks to occupants in braking manoeuvres or collisions, but that it could be safe in some Operational Design Domains (ODDs). A concept was proposed which offers manufacturers a choice between two Crashworthiness Approval Levels (CALs). The less demanding CAL allows standing passengers but restricts the subsequent ODD of the vehicles. VRU impact protection was a high priority due to the expected operation in areas with a high density of pedestrians and cyclists. However, LSAV aspects such as their typically flat-fronted shape cause issues for the application of the current regulation, so modifications were proposed.

Low-speed vehicles are not in widespread use today, which means no directly relevant real-world collision data was available to base safety decisions on. The guiding principle applied in this study was to provide 'at least equivalent safety', i.e. to offer safety levels relating to non-ADS aspects, which, based on the limited data available and expert judgement, are comparable to or better than those of current vehicles used in similar scenarios.

This study proposes a novel approach to link approval regulations to the vehicle's ODD and a set of technical requirements for non-ADS-related aspects of passenger- and goods-carrying LSAVs, which could help enable the approval of new vehicle concepts. The proposals have been presented to the United Kingdom Department for Transport for consideration.

INTRODUCTION

The United Kingdom (UK) government are committed to bringing forward legislation to allow the safe and secure deployment of self-driving vehicles, as set out in the recent policy paper Connected & Automated Mobility 2025: Realising the benefits of self-driving vehicles in the UK [1]. As part of the Connected and Automated Vehicles Process for Assuring Safety and Security (CAVPASS[1]) programme, TRL was commissioned to propose approaches to vehicle classification and suitable technical requirements for aspects not related to the Automated Driving System (ADS). These included crashworthiness, occupant protection, protection of Vulnerable Road Users (VRUs), and the lighting, braking and steering systems. The focus of the work presented was on Low-Speed Automated Vehicles (LSAVs), i.e. fully electric vehicles with a maximum speed not greater than 20 miles per hour (approx. 32 kilometres per hour) to be used in mixed traffic on roads with a speed limit not higher than 30 miles per hour (approx. 48 kilometres per hour). This initial focus on low-speed vehicles was selected because they have been suggested as a potential early use case for self-driving technology by industry and, while not in widespread use today, there is an emerging class of speed-limited automated vehicles being developed. Table 1 provides additional details of the use cases and vehicle designs in scope. Further, it should be noted that LSAVs can be bi-directional, i.e. capable of travelling forwards and backwards at close to their maximum speed.

Table 1.
Scope of use cases and vehicle designs considered

Characteristic	Scope
Body shape	To include novel vehicle designs which do not conform with legacy design conventions such as windscreens, long bonnets, driver controls, etc.
Purpose	Carriage of goods or passengers (seated, standing or mixed)
Powertrain	Fully electric
Maximum speed	20 mph
Maximum mass (gross vehicle weight)	5,000 kg for passenger-carrying vehicles 3,500 kg for goods vehicles
Operating environment	Roads with a speed limit up to 30 mph with mixed traffic (including VRUs), or Dedicated roadways (which may or may not have segregation barriers) Areas which may include high density of pedestrians Operating on a fixed route or within a fixed geographical area

This paper presents the work performed to develop a suitable set of pre-deployment regulations for LSAVs, drawing from the existing body of UN and EU type-approval regulations. Note that further work was performed to adapt post-deployment regulations, i.e. construction and use regulations, which is not reported in this paper.

METHODS

Vehicle Classification

Technical requirements for type-approval are organised by vehicle category. Which category a vehicle belongs to is determined by certain key design aspects, such as vehicle mass, number of seating positions, or primary purpose (passenger or goods transport). A market review identified the design characteristics of vehicles, either already on the market or announced by manufacturers, which would fall within the scope of the study. These designs and plausible iterations of them were compared with existing vehicle category definitions to identify if and how the vehicles would integrate with the current system: Great Britain's type-approval categories for passenger vehicles (M category) and goods vehicles (N category) are laid down in Retained Regulation (EU) 2018/858 [2]. Relevant aspects are contained in Article 3 (Definitions), Article 4 (Vehicle categories) and Annex I (General definitions, criteria for vehicle categorisation, types of vehicle and types of bodywork). Potentially relevant sub-categories, based on the LSAV mass-limited scope, are M_1 (passenger cars), M_2 (minibuses) and N_1 (vans). Note that L-category definitions, laid down in Retained Regulation (EU) 168/2013 [3], were also considered but deemed unsuitable due to the limited number of occupants and low maximum mass allowed under L-category definitions.

[1] https://www.gov.uk/guidance/connected-and-automated-vehicles-process-for-assuring-safety-and-security-cavpass

Developing a Suitable Set of Technical Requirements for LSAVs

In order to identify the most suitable basis for LSAV legislation, technical items in existing Great Britain and EU approval schemes were collated. It was found that the requirements for unlimited series approvals of M/N-category vehicles under Retained Regulation (EU) 2018/858 [2] were the most appropriate basis because they provided the most comprehensive list of items as a starting point. In addition, items contained in Regulation (EU) 2019/2144 [4] were added to the base list to arrive at a complete set of technical items for consideration, even though it should be noted that, at the time of writing, a requirement for compulsory fitment of those items has not been incorporated into Great Britain type-approval. Separate risk analysis identified four additional items relevant for LSAVs, which were added to the list: software updates, maximum vehicle speed limitation, manual operation at very low speeds, and static vehicle stability. An overview of the approach is presented in Figure 1.

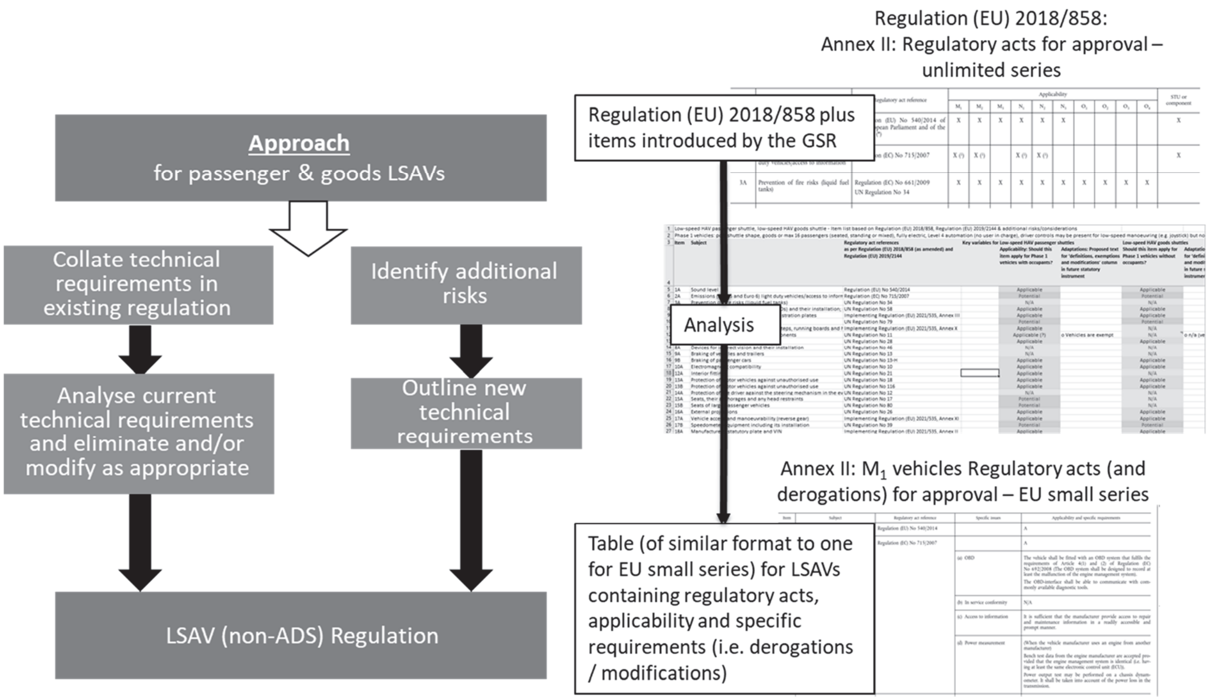


Figure 1. Overview approach for developing a suitable set of technical requirements for LSAVs.

In a first step, this list was reduced to items relevant for LSAVs, i.e. low-speed vehicles without a driver. Proposals for technical requirements and test procedures were then developed, as far as possible by selection of suitable existing regulations. The guiding principle when deciding what technical regulations to apply was to provide ‘at least equivalent safety’, i.e. offering levels of protection comparable to those of current vehicles used in similar scenarios, while ensuring that requirements were proportionate to the low-speed use case and were not over-burdensome for smaller manufacturers. For most items, various technical regulations existed to choose from, including those for light or heavy M/N-category vehicles or those for L-category vehicles. In a second step, new outline regulations were developed for the items ‘manual operation at very low speeds’ and ‘static vehicle stability’ because none could be identified in the existing body of regulations. For both steps, key safety and security areas that affected multiple technical items were considered collectively in order to develop a coherent and well-balanced approach for these requirements. These areas were: occupant restraint and crashworthiness, vulnerable road user protection, vehicle stability, and protection against unauthorised use and security. This paper presents only the results of the first two, more complex areas.

In a last step, the regulations selected for application were analysed for incompatibilities with LSAVs and appropriate modifications of the text or procedures were developed. Incompatibilities were identified using a keyword search to highlight aspects that presume the presence of a driver or a vehicle that has one main driving direction. The keywords included: ocular, front, rear, driver, passenger, seat, push, press, pull, button, control, mass, actuate, activate, switch, foremost, R-point, H-point, direction, forward, rearward, behind, fore, and aft. This search was backed up by a full manual analysis of the regulations which highlighted additional incompatibilities, e.g. with regard to applying the pedestrian impact test procedure to flat-fronted LSAV designs.

RESULTS

Vehicle Classification

Many envisaged LSAV designs have no conventional vehicle equivalents on-road, for example passenger shuttles with six seating positions and space for standing passengers, or goods vehicles without any seating positions. Barriers to classifying these vehicles in the existing system were identified within each relevant sub-category definition (M_1, M_2 and N_1). The most common barrier consisted of references to a driver's seating position, which is generally not present in LSAVs. Other barriers, relevant for certain LSAV designs, included:

- Standing passengers: The M_1 vehicle category definition excludes standing passengers. To permit standing passengers, an LSAV would need to be classified as an M_2 category vehicle. Whether a vehicle is classified as an M_1 or M_2 category is dependent on how many seating positions it offers: up to 9 seats: M_1 category; 10 or more seats: M_2 category. To carry standing passengers, vehicles therefore need at least 10 seating positions in the existing classification system.
- Low-speed vehicles: The definition of a motor vehicle with respect to M and N categories includes a prerequisite that the vehicle has a maximum design speed of more than 25 kilometres per hour (ca. 15.5 miles per hour). Thus, LSAVs with a maximum speed of 15.5 miles per hour or less would fall outside the scope of the categories.

Consequently, to accommodate LSAVs, the existing categories need to be amended or the classification system extended. Following careful consideration, it was recommended that new categories should be introduced because modifications of existing M and N categories could have too far-reaching impacts. New categories will also open a route to assign more suitable technical requirements specific to LSAVs, because some technical requirements associated with current M and N categories do not apply well, e.g. crashworthiness requirements are for more severe collisions than are likely to be experienced by an LSAV.

The most dominant factor determining how many and which technical requirements are needed was found to be the presence of occupants, because a substantial number of requirements are related to occupant protection. This aspect was deemed important enough to warrant the introduction of two separate LSAV categories: Passenger LSAV (for vehicles with occupant positions) and Goods LSAV (vehicles without occupant positions). Category definitions were proposed as detailed in Table 2.

Table 2.
Proposed definitions for new vehicle categories including a top-level definition of LSAV applicable for both

<table>
<tr><td>Top-level definition</td><td colspan="2">'LSAV' means any power-driven vehicle that is designed and constructed to be moved by its own means and to be driven, in normal operation, only by an automated driving system without a driver, that has at least four wheels, is complete, completed or incomplete, has a maximum design speed of at least 10 mph but not more than 20 mph.</td></tr>
<tr><td>Vehicle category definitions</td><td>'Passenger LSAV' means an LSAV intended to be used on public roads, with at least one occupant position (seated and/or standing), and having a maximum mass not exceeding 5,000 kg.</td><td>'Goods LSAV' means an LSAV intended to be used on public roads, without occupant positions, and having a maximum mass not exceeding 3,500 kg.</td></tr>
</table>

Selection of Technical Regulations for Applicable Items

The full list of 132 regulatory items collated for consideration is provided in Appendix A. The table details the items' proposed applicability for Passenger and Goods LSAVs, respectively. Out of the total 132 items considered, 65 were proposed to be applied for type-approval of Passenger LSAVs and 54 for Goods LSAVs. The dominant reasons found to exclude items in the applicability analysis were that they: were specific to heavier vehicles, vehicles with capability to tow trailers or vehicles propelled by a combustion engine or hydrogen drivetrain; were intended to support a driver (e.g. mirrors, windscreen wipers) or to protect vehicle occupants (and thus not relevant for Goods LSAVs); or because they were related to the dynamic-driving task and therefore not within the scope of this study.

The technical regulations recommended for application following the ‘at least equivalent safety’ principle are detailed in the same table. These are mostly UN and EU regulations for M- and N-category vehicles. For Item 115, maximum vehicle speed limitation, an L-category regulation was proposed to be used as the available M/N category regulations were deemed unsuitable: Maximum vehicle speed is not a relevant categorisation criterion in the M/N-category framework and is therefore not regulated or routinely tested for type-approval. Tests according to UN R68[2] may be performed at the request of the manufacturer or, in some cases, the technical service, but commonly the maximum vehicle speed is declared by the manufacturer in the Information Document without an approval test. UN R68 is a technically robust regulation to determine the achievable maximum speed, but it does not contain technical stipulations as to how a reliable and tamper-proof speed limitation shall be achieved. UN R89 applies to top-speed limiters for heavy vehicles (M_2, M_3, N_2, N_3). It does not contain specific provisions for electric vehicles and the tolerances permitted in the regulation are too wide for low-speed vehicles.

For certain aspects identified as additional risks for LSAVs, no suitable regulatory text was available in the existing body of regulations to draw from. For these items, new regulatory proposals were developed:

- Item 58: Pedestrian protection
- Item 116: Manual operation at very low speeds
- Item 117: Static vehicle stability

Modification of Technical Regulations for Self-driving Vehicles
Where existing UN or EU regulations were recommended for application, in many cases modifications were required to make them suitable for LSAVs: Some parts of regulations could be waived or alleviated due to the specific use cases (low-speed, goods vehicles without occupants); other parts needed to be modified because of novel vehicle shapes or the absence of a driver, driver’s seat and conventional vehicle controls. Common themes identified that caused incompatibility of regulatory text with self-driving vehicles included:

- References to:
    - driver’s seating position
    - driver controls (steering wheel, pedals, buttons, adjustment controls, etc.)
    - instrument panel
    - actions performed by the driver (e.g. full application of brake control, setting/unsetting of protective devices)
    - driver side/passenger side
    - front/rear, forward-/rear-facing (conflict with bi-directional vehicles)
- Information presented to the driver (tell-tales, indicator or warning lights, acoustic or haptic warnings); system reaction to failures not defined
- Reliance on muscular energy (e.g. braking)
- No consideration of bi-directional vehicles
- Occupant protection for goods vehicles (without occupants)
- Feasibility of type-approval tests in absence of driver controls

[2] Shorthand for UN Regulation No 68, applied equivalently for other UN Regulations

Modifications to the existing regulatory texts were proposed to detail the application of each regulation for LSAVs. An example overview of the outcomes of the analysis and the proposed modifications to the regulation for braking systems is presented in Table 3.

Table 3.
Analysis and modification of technical regulations for LSAVs, example results for Item 9B, braking

<table>
<tr><th colspan="2">UN Regulation No 13-H.01, consolidated to Supplement 1</th></tr>
<tr><td>Content summary:
Requirements relating to the characteristics and performance of braking system for passenger cars and light goods vehicles to ensure safe design and minimum performance levels</td><td>Applicability for self-driving vehicles:
• Passenger: Applicable
• Goods: Applicable</td></tr>
<tr><td colspan="2">Applicability and suitability of technical requirements/test procedures:
• Aspects related to driver controls and tell-tales not applicable for self-driving vehicles
• Functional safety and complex electronics – required, but propose to allow combined assessment of systems (e.g. steering / braking) for self-driving vehicles because of interdependencies
• ‘Full-stroke actuation’ and ‘full application’ of brake controls, which is referred to throughout the regulation, not possible in the absence of controls and requires clarification
• Alternative energy sources required to replace muscular operation
• Pneumatic braking systems not covered in regulation</td></tr>
<tr><td colspan="2">Requirements to cover additional risks/considerations for self-driving vehicles:
• For LSAVs equipped with compressed air braking systems: UN Regulation No 13, Paragraph 5.1.3., Annex 6, Annex 7 and Annex 10</td></tr>
<tr><td colspan="2">Definitions, modifications, exemptions and additions:
• The safety of complex electronic systems with regard to braking, Paragraph 5.1.3. and Annex 8, shall not be assessed as part of the brake systems approval but instead as part of an overall assessment of the ADS.
• All muscular generated performances in conventional vehicles (e.g. secondary brake) shall be replaced by alternative energy sources.
• Paragraph 5.2.2.8. (two completely independent energy reserves) shall continue to apply even though the service brake system is not ‘controlled by the driver’ as stipulated in the regulation.
• In the event of any failure or defect that would result in a red warning signal according to the regulatory requirements an electronic signal shall be sent to the ADS.
• References to ‘full-stroke actuation’ and ‘full application’ of brake controls shall be interpreted as ‘automatically commanded maximum braking demand’ by the ADS; references to ‘release’ of the brake controls shall be interpreted as ‘applying no automatically commanded braking demand’.
• Vehicles equipped with compressed-air braking systems shall fulfil, in addition to the requirements of UN Regulation No 13-H as modified above, the requirements for compressed-air braking systems set out in UN Regulation No 13 (11 series of amendments), Paragraph 5.1.3., Annex 6, Annex 7 and Annex 10.</td></tr>
</table>

Where equivalent aspects requiring modification were identified in multiple regulations, the authors found that these were best addressed in a single place in a uniform way. Item 0, cross-cutting prescriptions, was created to capture modifications intended to apply to all regulations concerned. The elements proposed to cover within this item include:

- Driver controls:
    - A driver’s seating position or driver controls shall not be required and thus requirements relating to the fitment of driver controls, their geometric positioning (e.g. achieve braking action from the driving seat) or their labelling, and requirements governing maximum permitted control application forces are not applicable.

- References to activation or operation of a system or device by the driver, driver's demands (e.g. driver's braking demand) and driver's intention shall be interpreted as activation or operation by the ADS, ADS's demands and ADS's intention. The moment when a control begins to be actuated shall be interpreted as the moment when the automatically commanded signal (e.g. brake or steering demand signal) is issued.
- Vehicles shall be so designed to allow the technical service to perform the required type-approval tests in the absence of conventional driver controls, for example by allowing temporary manual control for the technical service or by offering an automated test mode performed by the ADS which is fully representative of the performance characteristics in normal mode. In production vehicles intended for road use, these test modes or functions shall not be available.

- Requirements relating to information presented to the driver or other vehicle occupants, e.g. in the form of tell-tales, indicator or warning lights or acoustic warnings, shall not apply, unless stated otherwise in the context of specific regulations. Instead, relevant information on failures, defects or emergency situations shall, as appropriate to control the situation, be acted upon appropriately by the vehicle (e.g. initiate minimum risk manoeuvre) or made available to a remote control-centre. Other information, e.g. about the activation status of lights or on-board systems, may be made available to a remote control-centre (optional) or not be provided.
- 'Driver side' of a vehicle shall be understood as the right side with respect to the direction of travel; 'passenger side' shall be understood as the left side with respect to the direction of travel.
- Bi-directional vehicles, defined as vehicles that can travel in either direction, forwards or backwards, at a speed of 80 percent or more of the vehicles maximum speed, shall meet the applicable requirements in both driving directions unless stated otherwise in the context of specific regulations. Where the requirements represent different levels of safety, the more stringent requirements shall apply. Tests shall therefore also be performed in both directions or, at the discretion of the technical service, in the direction determined to represent the worst case.

Operational Design Domain-based Occupant Restraint and Crashworthiness Requirements
The level of occupant protection and crashworthiness appropriate for a Passenger LSAV will change depending on its Operational Design Domains (ODD) – an aspect that is not catered for in the current type-approval system for M- and N-category vehicles where a single set of requirements applies to allow unrestricted (ODD-independent) operation.

A main factor determining appropriate requirements is the risk of injury from collisions which could cause large accelerations and/or occupant compartment intrusion. The ADS is responsible for primary safety for collision avoidance, but the action of other road users may still result in collisions which require mitigation through secondary safety measures. Thus, on the assumption that the vehicle drives well (e.g. in a manner which adheres to the traffic rules, avoids being the cause of a collision), the main types of collisions were anticipated to be 'not at fault' ones, i.e. other vehicles colliding into the LSAV. The typical collisions envisaged on which to base protection needs are a vehicle (most likely a car) travelling at 30 miles per hour impacting the side, front or rear of the LSAV.

The injury risk will be low in ODDs where the collision risk and/or consequence is reduced. Examples include separated lanes, where the collision risk will be reduced compared to open roads because less traffic is encountered, and 20 miles per hour zones or business/university campuses where the collision consequences will be reduced because of the low traffic speed. Conversely, in ODDs such as roads with 30 miles per hour speed limits on which mixed traffic is present, the injury risk will be much higher because the presence of many other vehicles increases the collision risk and because they are likely to be travelling faster at 30 miles per hour or potentially above, which also increases the collision consequences. Different levels of protection will be required for these different operating environments.

When considering whether or not standing or unrestrained passengers should be permitted in passenger LSAVs, the presence of VRUs in the ODD should be taken into account: VRUs' trajectories are more difficult to predict than motor vehicles' and their presence is thus expected to result in more automatic emergency braking manoeuvres by the LSAV which could cause occupants to fall and/or be thrown forward. Standing/unrestrained passengers could be permitted only in environments with low/no prevalence of VRUs or a very low speed restriction could be imposed on vehicles with such passengers in areas with VRUs.

Table 4.
Proposed occupant restraint and crashworthiness requirements for LSAV approval to CAL Reduced and CAL Standard

<table>
<tr><th></th><th>CAL Reduced</th><th>CAL Standard</th></tr>
<tr><td>Basis</td><td>Category M_2, Class A requirements for operating environments with a low risk of collision and/or low consequences</td><td>Category M_1 requirements for all operating environments, including those with a higher risk of collision and/or higher consequences</td></tr>
<tr><td>Occupant restraint</td><td>Standees:
Permitted
Requirements include handrails/handholds as per UN R107
Seated:
Side-facing seats permitted
Requirements include guard or at least 2-point safety belt for exposed seats only</td><td>Standees:
Not permitted
Seated:
Side-facing seats not permitted
Front and rear impact: Requirements include 3-point safety belts and head restraints for all seating positions (including rear-facing seats). Fit impact friendly interior for parts likely to be hit by passengers; relevant regulations: UN R14, UN R16, UN R17 and UN R21
Side impact: Require side impact protection, e.g. meet dummy injury assessment reference values in side impact test; relevant regulations: UN R95</td></tr>
<tr><td>Crash-worthiness</td><td>No requirements</td><td>Front impact:
Aim: Ensure structural integrity for electrical power train protection and to limit compartment intrusion to allow occupant restraint to function correctly
Test: UN R95 type test at 50 km/h to the front of the vehicle and impose requirements for electrical power train safety as per the regulation, and structural performance through an assessment of intrusion
Rear impact:
Aim: Ensure structural integrity for electrical power train protection and to limit compartment intrusion to allow occupant restraint to function correctly
Test: UN R95 type test at 50 km/h to the rear of the vehicle and impose requirements for electrical powertrain safety as per the regulation and structural performance through an assessment of intrusion as per UN R32
Side impact:
Aim: Ensure structural integrity for electrical power train protection together with associated padding / airbags for occupant protection
Test: UN R95 test with electrical power train and occupant protection requirements as per the regulation</td></tr>
</table>

With this in mind and based on the principle of achieving a safety level at least equivalent to current non-automated vehicles, it is proposed to introduce two ODD-based Crashworthiness Approval Levels (CALs) for passenger LSAVs with outline requirements as detailed in Table 4. It is envisaged that the manufacturer will be able to choose which CAL to approve the LSAV to, depending upon the intended use. An LSAV approved to CAL Reduced shall be limited to ODDs where the risk of collision and/or consequences within the domain are appropriately low (considering the risk of impacting/being impacted by other vehicles and the risk of emergency braking manoeuvres endangering standing/unrestrained occupants). An LSAV approved to CAL Standard on the other hand shall be permitted to operate in all environments within the overall scope of this study, including the aforementioned low-risk ODDs.

Adapting Requirements for VRU Protection

LSAVs are expected to operate frequently in areas with a high density of VRUs which makes VRU protection a high priority. Given high exposure to VRUs, the risk of harm is greater when operating in these environments. Therefore, it is proposed that secondary safety countermeasures are required. The relevant technical areas to consider were: external projections (to reduce the risk of injury to a person hit by sharp edges and protrusions on the vehicle body work), VRU protection (to reduce risk of injury from impacts) and frontal protection systems (to prevent additional injury risk from these structures). The outline requirements proposed for these areas are summarised in Table 5.

Table 5.
Proposed VRU protection requirements for LSAV approval

	Proposed requirements
External projections	Apply requirements of UN R26 with newly added requirements for ADS sensors, which are not currently covered within the regulation. If vehicle features are present that are regulated only in UN R61, the Technical Service may consider the relevant prescriptions of UN R61 in their assessment. An exemption for certain features can be given by the Approval Authority where the special purpose of a vehicle makes it impossible to fully comply. In general, the Technical Service should pay specific attention to features likely to cause leg injuries because pedestrian leg impact testing will not be performed.
Vulnerable road user protection	For all vehicles with a maximum speed > 16 km/h, submit documentation to demonstrate that frontal areas of the vehicle that are likely to be hit by a VRU's head have safety levels in line with the principles of UN R127. The documentation will be assessed by the Technical Service. Assessed areas should include windscreen / window areas if likely to be hit, i.e. wrap-around distance of 800 mm to 2500 mm or a height above ground of 2000 mm for vehicles with close to vertical front shapes. Test areas to be defined using UN R127 procedures or, if not appropriate, an equivalent type of method. Safety levels may be demonstrated with headform tests or, for limited areas, with logical argumentation to the satisfaction of the Technical Service/Approval Authority. Headform test parameters may be adjusted where appropriate, e.g. impact angle for different vehicle front shapes, or impact speed to account for a vehicle's maximum speed being < 40 km/h, the nominal pedestrian impact speed used as the basis for the development of UN R127. For child and adult headform tests the Head Injury Criterion (HIC) recorded shall meet the requirement to not exceed 1000 over all of the test area. Note: This is different to UN R127 requirements, which are < 1000 over 2/3 of test area and < 1700 over remaining area, because it is anticipated that a reduced impact test speed will make it technically feasible to meet this requirement in the entire area.
Frontal protection systems	Fitment of frontal protection systems not permitted.

For VRU protection it should be noted that UN R127 cannot be applied because many LSAVs are anticipated to have a close to vertical front shape (as opposed to a long bonnet type shape of typical passenger cars) which makes the current pedestrian safety protocol, including leg impact, very difficult to perform. Additionally, the maximum LSAV speed is lower than the pedestrian impact represented in the current protocol. Therefore, adapted procedures focussing on head protection have been proposed.

Frontal protection systems are defined as a separate structure or structures, such as a bull bar, or a supplementary bumper which, in addition to the original-equipment bumper, is intended to protect the external surface of the vehicle from damage in the event of a collision with an object, with the exception of structures having a mass of less than 0.5 kilograms, intended to protect only the vehicle's lights. It is not envisaged that frontal protection systems will be needed in the environments LSAVs are expected to operate in. To avoid an elevated injury risk to VRUs from these typically stiff structures, no frontal protection systems should be permitted on LSAVs.

DISCUSSION AND LIMITATIONS

This paper proposes to create new vehicle categories for LSAVs because many envisaged designs do not have current, non-automated equivalents. In the future, dual mode vehicles, i.e. those that can self-drive in specified ODDs and be operated by a driver in other operating environments, are likely to be built by adapting current vehicles. Unlike LSAVs, these will likely fit into current vehicle categories well and hence for these vehicles, the better way forward may be to create automated vehicle sub-categories under existing vehicle categories.

Analysis of the regulatory body for M- and N-category vehicles found that more than half of the existing technical items were superfluous for LSAVs. The dominant reasons for eliminating items were not related to the low maximum design speed but to the absence of a driver or occupants in general, so would also hold for other self-driving vehicles.

Low-speed vehicles are not in widespread use today, which means no directly relevant real-world collision data were available to base safety decisions on. The guiding principle applied in this study was to provide 'at least equivalent safety', i.e. to offer safety levels relating to non-ADS aspects, which, based on the limited data available and expert judgement, are comparable or better than those of current vehicles used in similar scenarios. The technical requirements were mostly based on M_1- and N_1-category regulations mainly because of their mass similarity to LSAVs. Reduced occupant protection requirements (CAL Reduced) were drawn from M_2-category regulations because the vehicles share similar use cases.

The concept of linking crashworthiness and occupant protection requirements at type approval to the future ODD of a vehicle by offering different CALs is a novel concept which does not exist in current type-approval frameworks. It remains to be seen if this concept proves practical in real-world application, i.e. if sufficient low-risk environments can be identified to warrant the design and production of restricted-use vehicles. If successful, it might allow the production of more cost-effective and mass-reduced vehicles for low-risk environments, and also enable the carriage of standing passengers in small vehicles to enable frequent and fast passenger changes. The risk of vehicle ODDs being expanded post-approval by software updates while the CAL, determined by the mechanical design of the vehicle, remains unchanged should be considered. This might require mitigation by appropriate post-deployment or licencing legislation.

The head impact requirements proposed for VRU protection should be regarded as a first step to ensuring secondary VRU safety for flat-fronted vehicles, which could otherwise not be assessed at all. In the medium to longer term, these should be updated with an appropriate regulation for a full assessment of VRU impact protection for LSAVs when lessons learnt from in-use safety monitoring can be taken into consideration.

CONCLUSIONS

This independent study developed proposals for a comprehensive set of technical requirements to allow the safe deployment of LSAVs with a focus on non-ADS aspects, such as crashworthiness, occupant and VRU protection, and the lighting, braking and steering systems. The requirements were arranged in relation to two new vehicle categories with the ability to carry passengers being the main differentiator. For Passenger LSAVs, a novel approach was proposed to link approval regulations to the vehicle's ODD, which could help enable the approval of new vehicle concepts, such as small vehicles with standing passengers. The proposals have been presented to the UK Department for Transport for consideration.

ACKNOWLEDGMENTS

The authors wish to acknowledge the UK's Department for Transport for funding this study. This paper was prepared by TRL Limited, and the information presented does not necessarily reflect the views of the Department.

REFERENCES

[1] HM Government. (2022). *Connected & Automated Mobility 2025: Realising the benefits of self-driving vehicles in the UK.* Available at: https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/1099173/cam-2025-realising-benefits-self-driving-vehicles.pdf

[2] *Retained Regulation (EU) 2018/858 of the European Parliament and of the Council of 30 May 2018 on the approval and market surveillance of motor vehicles and their trailers, and of systems, components and separate technical units intended for such vehicles.* Available at: https://www.legislation.gov.uk/eur/2018/858/contents

[3] *Retained Regulation (EU) 168/2013 of the European Parliament and of the Council of 15 January 2013 on the approval and market surveillance of two- or three-wheel vehicles and quadricycles.* Available at: https://www.legislation.gov.uk/eur/2013/168/contents

[4] *Regulation (EU) 2019/2144 of the European Parliament and of the Council of 27 November 2019 on type-approval requirements for motor vehicles and their trailers, and systems, components and separate technical units intended for such vehicles, as regards their general safety and the protection of vehicle occupants and vulnerable road users*, as last amended by Commission Delegated Regulation (EU) 2022/1398. Available at: https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX%3A02019R2144-20220905

[5] *Commission Delegated Regulation (EU) 2022/2236 of 20 June 2022 amending Annexes I, II, IV and V to Regulation (EU) 2018/858 of the European Parliament and of the Council as regards the technical requirements for vehicles produced in unlimited series, vehicles produced in small series, fully automated vehicles produced in small series and special purpose vehicles, and as regards software update.* Available at: https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=uriserv:OJ.L_.2022.296.01.0001.01.ENG

APPENDIX A

Table 6.
List of applicable and non-applicable regulatory items for Passenger and Goods LSAVs with proposed technical regulation for requirements and test procedures; item numbers: left column as per Retained Regulation (EU) 2018/858, Annex II, Part I [2] and consecutive numbers for additional items; right column provides corresponding updated item numbers as per Commission Delegated Regulation (EU) 2022/2236 [5] for reference

Item number		Subject	Regulation	Passenger LSAV	Goods LSAV
0	n/a	Cross-cutting prescriptions	New regulatory text	Applicable	Applicable
1A	G1, B14	Sound level	UN Regulation No 138	Applicable	Applicable
2A	G2–G12, H1	Emissions (Euro 5 and Euro 6) light duty vehicles/access to information	Regulation (EC) No 715/2007 (light duty vehicle emissions)	Applicable	Applicable
3A	A14	Prevention of fire risks (liquid fuel tanks)	n/a	Not applicable	Not applicable
3B	A12	Rear underrun protective devices (RUPDs) and their installation; rear underrun protection (RUP)	UN Regulation No 58	Applicable	Applicable
4A	F1	Space for mounting and fixing rear registration plates	Commission Regulation (EU) 1003/2010	Applicable	Applicable
5A	C1	Steering equipment	UN Regulation No 79	Applicable	Applicable
6A	F4	Vehicle access and manoeuvrability (steps, running boards and handholds)	Commission Regulation (EU) 130/2012, Annex II	Applicable	Not applicable
6B	F3	Door latches and door retention components	UN Regulation No 11	Applicable	Not applicable
7A	D1	Audible warning devices and signals	UN Regulation No 28	Applicable	Applicable
8A	B13	Devices for indirect vision and their installation	n/a	Not applicable	Not applicable
9A	C4	Braking of vehicles and trailers	n/a	Not applicable	Not applicable
9B	C4	Braking of passenger cars	UN Regulation No 13-H	Applicable	Applicable
10A	D2	Electromagnetic compatibility	UN Regulation No 10	Applicable	Applicable
12A	A1	Interior fittings	UN Regulation No 21	Applicable	Not applicable
13A	D3	Protection of motor vehicles against unauthorised use	UN Regulation No 18	Applicable	Applicable
13B	D3	Protection of motor vehicles against unauthorised use	UN Regulation No 116	Applicable	Applicable
14A	A22	Protection of the driver against the steering mechanism in the event of impact	n/a	Not applicable	Not applicable
15A	A2	Seats, their anchorages and any head restraints	UN Regulation No 17	Applicable	Not applicable

Item number		Subject	Regulation	Passenger LSAV	Goods LSAV
15B	A3	Seats of large passenger vehicles	n/a	Not applicable	Not applicable
16A	F5	External projections	UN Regulation No 26	Applicable	Applicable
17A	F2	Vehicle access and manoeuvrability (reverse gear)	Commission Regulation (EU) 130/2012, Annex III	Applicable	Applicable
17B	D5, D6	Speedometer equipment including its installation	n/a	Not applicable	Not applicable
18A	F7	Manufacturer's statutory plate and VIN	Commission Regulation (EU) 19/2011	Applicable	Applicable
19A	A4	Safety-belt anchorages	UN Regulation No 14	Applicable	Not applicable
20A	D15	Installation of lighting and light-signalling devices on vehicles	UN Regulation No 48	Applicable	Applicable
21A	D13	Retro-reflecting devices for power- driven vehicles and their trailers	UN Regulation No 3	Applicable	Applicable
22A	D11	Front and rear position lamps, stop- lamps and end-outline marker lamps for motor vehicles and their trailers	UN Regulation No 7	Applicable	Applicable
22B	D11	Daytime running lamps for power- driven vehicles	UN Regulation No 87	Applicable	Applicable
22C	D11	Side-marker lamps for motor vehicles and their trailers	UN Regulation No 91	Applicable	Applicable
23A	D11	Direction indicators for power-driven vehicles and their trailers	UN Regulation No 6	Applicable	Applicable
24A	D11	Illumination of rear-registration plates of power-driven vehicles and their trailers	UN Regulation No 4	Applicable	Applicable
25A	D12	Power-driven vehicle's sealed-beam headlamps (SB) emitting an European asymmetrical passing beam or a driving beam or both	UN Regulation No 31	Applicable	Applicable
25B	D14	Filament lamps for use in approved lamp units of power-driven vehicles and their trailers	UN Regulation No 37	Applicable	Applicable
25C	D12	Motor vehicle headlamps equipped with gas-discharge light sources	UN Regulation No 98	Applicable	Applicable
25D	D14	Gas-discharge light sources for use in approved gas-discharge lamp units of power-driven vehicles	UN Regulation No 99	Applicable	Applicable
25E	D12	Motor vehicle headlamps emitting an asymmetrical passing beam or a driving beam or both and equipped with filament lamps and/or LED modules	UN Regulation No 112	Applicable	Applicable
25F	D12	Adaptive front-lighting systems (AFS) for motor vehicles	UN Regulation No 123	Applicable	Applicable
26A	D11	Power-driven vehicle front fog lamps	UN Regulation No 19	Applicable	Applicable
27A	F8	Towing device	Commission Regulation (EU) 1005/2010	Applicable	Applicable
28A	D11	Rear fog lamps for power-driven vehicles and their trailers	UN Regulation No 38	Applicable	Applicable

Item number		Subject	Regulation	Passenger LSAV	Goods LSAV
29A	D11	Reversing lights for power-driven vehicles and their trailers	UN Regulation No 23	Applicable	Applicable
30A	D11	Parking lamps for power-driven vehicles	UN Regulation No 77	Applicable	Applicable
31A	A5, A9	Safety-belts, restraint systems, child restraint systems and ISOFIX child restraint systems	UN Regulation No 16	Applicable	Not applicable
32A	B8	Forward field of vision	n/a	Not applicable	Not applicable
33A	D9	Location and identification of hand controls, tell-tales and indicators	n/a	Not applicable	Not applicable
34A	B11	Windscreen defrosting and demisting systems	n/a	Not applicable	Not applicable
35A	B12	Windscreen wiper and washer systems	n/a	Not applicable	Not applicable
36A	D10	Heating systems	UN Regulation No 122	Applicable	Applicable
37A	F9	Wheel guards	Commission Regulation (EU) 1009/2010	Applicable	Applicable
38A	A2	Head restraints (headrests), whether or not incorporated in vehicle seats	UN Regulation No 25	Applicable	Not applicable
41A	G2–G12, H1	Emissions (Euro VI) heavy duty vehicles/access to information	n/a	Not applicable	Not applicable
42A	A13	Lateral protection of goods vehicles	n/a	Not applicable	Not applicable
43A	F10	Spray suppression systems	n/a	Not applicable	Not applicable
44A	F11	Masses and dimensions (M1)	Commission Regulation (EU) 1230/2012	Applicable	Applicable
45A	B10	Safety glazing materials and their installation on vehicles	UN Regulation No 43	Applicable	Applicable
46A	C15	Installation of tyres	UN Regulation No 142	Applicable	Applicable
46B	C10	Pneumatic tyres for motor vehicles and their trailers (Class C1)	UN Regulation No 30	Applicable	Applicable
46C	C10	Pneumatic tyres for commercial vehicles and their trailers (Classes C2 and C3)	UN Regulation No 54	Applicable	Applicable
46D	C10	Tyre rolling sound emissions, adhesion on wet surfaces and rolling resistance (Classes C1, C2 and C3)	UN Regulation No 117	Applicable	Applicable
46E	C11	Temporary-use spare unit, run-flat tyres/system	UN Regulation No 64	Applicable	Applicable
47A	D7	Speed limitation of vehicles	n/a	Not applicable	Not applicable
48A	F11	Masses and dimensions (non-M1)	n/a	Not applicable	Not applicable
49A	F6	Commercial vehicles with regard to their external projections	n/a	Not applicable	Not applicable
50A	F12	Mechanical coupling components of combinations of vehicles	n/a	Not applicable	Not applicable

Item number		Subject	Regulation	Passenger LSAV	Goods LSAV
50B	F12	Close-coupling device (CCD); fitting of an approved type of CCD	n/a	Not applicable	Not applicable
51A	F16	Burning behaviour of materials used in the interior construction of certain categories of motor vehicles	n/a	Not applicable	Not applicable
52A	F14	M2 and M3 vehicles Regulation (EC)	UN Regulation No 107	Applicable	Not applicable
52B	F15	Strength of the superstructure of large passenger vehicles	n/a	Not applicable	Not applicable
53A	A20	Protection of occupants in the event of a frontal collision	n/a	Not applicable	Not applicable
54A	A25	Protection of occupants in the event of lateral, frontal or rear collision	UN Regulation No 95	Applicable	Not applicable
56A	F13	Vehicles for the carriage of dangerous goods	n/a	Not applicable	Not applicable
57A	A11	Front underrun protective devices (FUPDs) and their installation; front underrun protection (FUP)	n/a	Not applicable	Not applicable
58	B1	Pedestrian protection	New regulatory text	Applicable	Applicable
59	G13	Recyclability	UN Regulation No 133	Applicable	Applicable
61	G14	Air-conditioning systems	Directive 2006/40/EC	Applicable	Applicable
62	A17	Hydrogen system	n/a	Not applicable	Not applicable
63	n/a	General Safety	n/a	Not applicable	Not applicable
64	D18	Gear shift indicators	n/a	Not applicable	Not applicable
65	C8	Advanced emergency braking system (heavy vehicles)	n/a	Not applicable	Not applicable
66	C2	Lane departure warning system (heavy vehicles)	n/a	Not applicable	Not applicable
67	A15	Specific components for liquefied petroleum gases (LPG) and their installation on motor vehicles	n/a	Not applicable	Not applicable
68	D3	Vehicle alarm systems (VAS)	UN Regulation No 97	Applicable	Applicable
69	A19	Electric safety	UN Regulation No 100	Applicable	Applicable
70	A16	Specific components for CNG and their installation on motor vehicles	n/a	Not applicable	Not applicable
71	A24	Cab strength	n/a	Not applicable	Not applicable
72	A28	eCall system	n/a	Not applicable	Not applicable
73	A7	Partitioning systems	n/a	Not applicable	Not applicable
74	A10	Enhanced child restraint systems	n/a	Not applicable	Not applicable
75	A18	Hydrogen system material qualification	n/a	Not applicable	Not applicable

Item number		Subject	Regulation	Passenger LSAV	Goods LSAV
76	A21	Frontal full-width impact	n/a	Not applicable	Not applicable
77	A23	Replacement airbag	n/a	Not applicable	Not applicable
78	A26	Pole side impact	n/a	Not applicable	Not applicable
79	A27	Rear impact	n/a	Not applicable	Not applicable
80	B2	Enlarged head impact zone	n/a	Not applicable	Not applicable
81	B4	Advanced emergency braking for pedestrians and cyclists ahead (AEBS PC)	n/a	Not applicable	Not applicable
82	B5	Pedestrian and cyclist collision warning (MOIS)	n/a	Not applicable	Not applicable
83	B6	Blind spot information system (BSIS)	n/a	Not applicable	Not applicable
84	B7	Reversing safety	n/a	Not applicable	Not applicable
85	B9	Heavy duty vehicles direct vision	n/a	Not applicable	Not applicable
86	C3	Emergency lane keeping	n/a	Not applicable	Not applicable
87	C9	Advanced emergency braking on light duty vehicles (AEBS)	n/a	Not applicable	Not applicable
88	C12	Retreaded tyres	n/a	Not applicable	Not applicable
89	C13, C14	Tyre pressure monitoring	n/a	Not applicable	Not applicable
90	C16	Replacement wheels	n/a	Not applicable	Not applicable
91	D4	Protection of vehicle against cyberattacks	UN Regulation No 155	Applicable	Applicable
92	D8	Intelligent speed assistance	n/a	Not applicable	Not applicable
93	D16	Emergency stop signal	n/a	Not applicable	Not applicable
94	D17	Headlamp cleaners	UN Regulation No 45	Applicable	Applicable
95	E1	Alcohol interlock installation facilitation	n/a	Not applicable	Not applicable
96	E2	Driver drowsiness and attention warning	n/a	Not applicable	Not applicable
97	E3	Advanced driver distraction warning	n/a	Not applicable	Not applicable
98	E4	Driver availability monitoring system	n/a	Not applicable	Not applicable
99	E5	Event data recorder	n/a	Not applicable	Not applicable
100	E6	Systems to replace driver’s control	n/a	Not applicable	Not applicable
101	E7	Systems to provide the vehicle with information on state of vehicle and surrounding area	n/a	Not applicable	Not applicable

Item number		Subject	Regulation	Passenger LSAV	Goods LSAV
102	E8	Platooning	n/a	Not applicable	Not applicable
103	E9	Systems to provide safety information to other road users	n/a	Not applicable	Not applicable
104	A6	Safety-belt reminders	UN Regulation No 16	Applicable	Not applicable
105	A17	Hydrogen safety	n/a	Not applicable	Not applicable
106	B3	Frontal protection system	n/a	Not applicable	Not applicable
107	C5	Replacement braking parts	n/a	Not applicable	Not applicable
108	C6	Brake Assist	n/a	Not applicable	Not applicable
109	C7	Stability control	n/a	Not applicable	Not applicable
110	D12	Cornering lamps	UN Regulation No 119	Applicable	Applicable
111	D13	Retro-reflective markings (heavy and long vehicles)	n/a	Not applicable	Not applicable
112	D14	LED light sources	UN Regulation No 128	Applicable	Applicable
113	A8	Child restraint anchorages	UN Regulation No 145	Applicable	Not applicable
114	H2	Software update and software updates management system	UN Regulation No 156	Applicable	Applicable
115	n/a	Maximum vehicle speed limitation	Commission Delegated Regulation (EU) No 3/2014, Annex XVIII	Applicable	Applicable
116	n/a	Manual operation at very low speeds	New regulatory text	Applicable	Applicable
117	n/a	Static vehicle stability	New regulatory text	Applicable	Applicable

SAFETY METRICS ASSESSMENT USING LOGGED VEHICLE TRAJECTORY DATA

Xintao Yan
Department of Civil and Environmental Engineering, University of Michigan, Ann Arbor

Shuo Feng
David J. LeBlanc
Carol Flannagan
Henry X. Liu
University of Michigan Transportation Research Institute

Alrik L. Svenson
National Highway Traffic Safety Administration
United States

Paper Number 23-0224

ABSTRACT

Assuring safe operation remains as one of the main challenges for developing and deploying automated driving systems (ADS). Real-time safety risk metrics may play important roles in calculating a risk measure of driving situations. Although a number of safety metrics have been proposed previously, it is difficult to compare different safety metrics and assess their performance because different behavioral assumptions underly for each. In this paper, a method to assess the behavior of safety risk metrics by determining the subject vehicle (SV) situational safety using logged vehicle trajectory data is proposed. Specifically, it is examined whether the SV is in a collision unavoidable situation at each moment, given the near-future trajectories of all surrounding principal other vehicles (POVs) recorded in the dataset after this moment. The main benefit of using logged vehicle trajectory data is the elimination of behavior prediction errors caused by model assumptions and approximations. This establishes a ground truth for crash outcomes independent of the risk metrics. Using the proposed methodology, the performance of different real-time safety metrics can be evaluated using simulated and/or real-world vehicle trajectories. The proposed methodology also has the potential to be applied in scenarios with vulnerable road users (VRU) interactions. In the case study, three real-time safety metrics are considered: time-to-collision (TTC), the PEGASUS Criticality Metric (PCM), and the Model Predictive Instantaneous Safety Metric (MPrISM). The results can help practitioners to better understand the characteristics and applicability of different safety metrics for different situations. The evaluation results can also help researchers improve and refine existing safety metrics.

INTRODUCTION

The overall idea of real-time safety metrics is to evaluate the situational safety risk based on the predicted future trajectories of both the subject vehicle (SV) and surrounding principal other vehicles (POVs). Utilizing vehicle state information (e.g., position, speed, heading, etc.) at or up to the evaluation moment, different behavior assumptions can be made to infer future trajectories, and these will lead to distinct safety metrics outcomes. An example is shown in Figure 1, where the blue vehicle represents the SV, and the green vehicle is the POV. The arrows represent different potential actions that SV and POV might take in the future. If the POV does not sense the SV in its blind spot and changes lanes in front of the SV, and the SV does not slow down or make an evasive lane change, then the resulting situation potentially can be dangerous. Alternatively, if the POV remains in its current lane, then the current moment in time is safe. This example demonstrates that different behavior assumptions can lead to different safety metrics results.

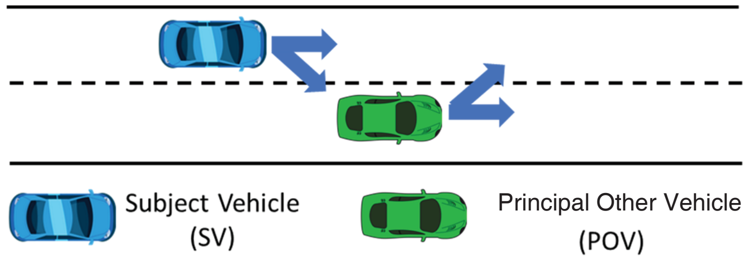


Figure 1. An illustration that different behavior assumptions can lead to different safety metrics results

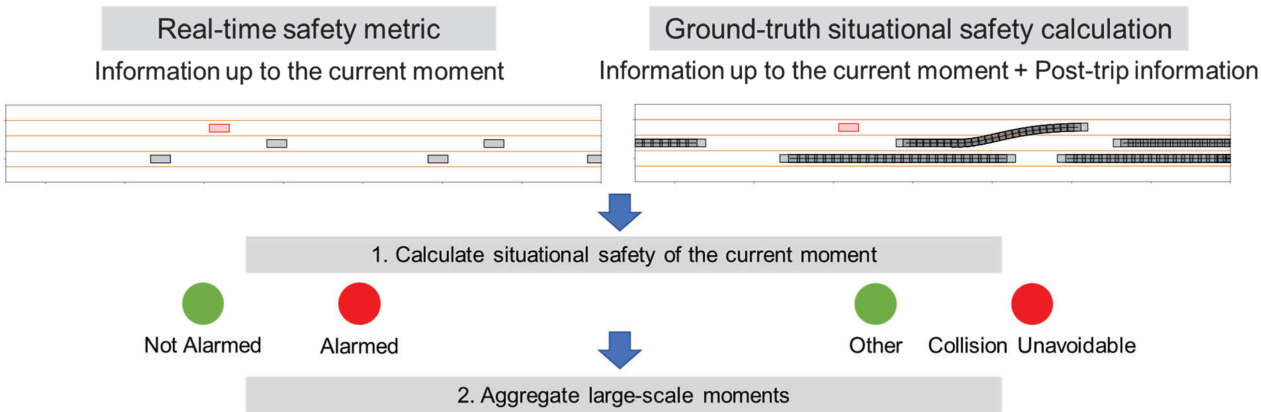
Over the past few decades, a number of safety metrics were proposed in the literature. One of the most commonly used safety metrics is Time-to-Collision (TTC) [1]. It assumes both the SV and its leading vehicle will maintain their current velocities. Based on TTC, Modified Time-to-Collision (MTTC) [2], and Worst Time-to-Collision (WTTC) [3], etc., were also proposed, attempting to capture different behavior assumptions including the acceleration and deceleration processes of both SV and POV. Junietz et al. proposed a Criticality Metric [4], which adopts similar behavior assumptions as in TTC regarding POVs behavior but assumes the SV will take evasive action to avoid the collision. Based on these assumptions, a model predictive control (MPC) method was applied to examine whether the current situation is at risk for the SV. Since it has been developed as part of the PEGASUS project, it is abbreviated here as PCM (PEGASUS Criticality Metric). The Model Predictive Instantaneous Safety Metric (MPrISM) was proposed in [5]. It also applies the MPC-based approach to generate a high-dimensional model predictive TTC (MPrTTC) that can consider both longitudinal and lateral possible risks. It assumes there exists one challenging POV, which will take the worst-case behavior trying to challenge the SV, and the SV will take the best response trying to avoid the collision. To account for the stochastic nature of human drivers, several probabilistic-based metrics have been proposed [6-8]. The collision probability is usually estimated to evaluate situational safety. In order to calculate the collision probability, multiple predicted future trajectories are generated to estimate the collision expectation using different methods, e.g., Markov chain [6], Monte Carlo simulation [7], and Bayesian network [8], etc. The potential stochastic behaviors of the SV and POVs are considered in these methods so they can better capture interactions between vehicles. However, the computational burden of probabilistic methods is heavier which might limit their real-time application.

The key challenge of evaluating a safety metric's performance is the lack of ground-truth measurement of risk at each moment in time because the future behavior of POVs is uncertain. Most existing evaluation methods rely on expert knowledge or heuristics, which suffers from subjective biases. The objective of this research is to propose an optimization-based method to utilize the post-trip information from vehicle logged trajectory data to evaluate and compare the performance of different safety metrics. In this way, a fair comparison of different safety metrics can be made, as a safety metric should alarm in advance to the collision-unavoidable moment. When trajectory data from a large number of trips are available, different metrics' statistical performance can be systematically evaluated and compared.

METHOD

In this study, a systematic evaluation framework has been applied to assess the performance of real-time safety metrics. The method can leverage post-trip information to calculate the situational risk objectively, as illustrated in Figure 2. For a real-time safety metric, information at or up to the current moment can be used to evaluate the SV's (shown by the red rectangle in the figure) situational safety. But when the trip is played back to determine the situational safety of the SV at any moment during the trip, trajectories of POVs are given, and no behavior assumptions are needed. By incorporating this post-trip POV trajectory information, the prediction errors of POV behaviors caused by assumptions and approximations of real-time safety metrics can be eliminated. Note that the post-trip information is obtained from logged trajectories, which can be either simulated trajectories or real-world recorded trajectories. Therefore, the safety metric output and the ground-truth can be compared to evaluate its effectiveness. Specifically, collision-unavoidable moments can be identified when evasive maneuver(s) to avoid the crash do not exist. They are objective without relying on expert knowledge or heuristically defined rules. The goal of a safety metric is to alarm no later than the collision-unavoidable moment. To identify the collision-unavoidable moment, an optimization problem can be formulated to examine whether there exists an evasive trajectory for the

SV, given the near-future trajectories of all POVs after this moment. It is assumed the POV behaviors will not be influenced by the SV actions within a short look-ahead horizon, so they will follow the observed trajectories. Therefore, the current moment is collision unavoidable if the SV has no possible evasive trajectory. The detailed formulation of the proposed algorithm can be found in [9].



Confusion Matrix		Real-time Safety Metric: Alarmed (Positive)	Real-time Safety Metric: Not alarmed (Negative)
Ground-truth situational safety	Should alarm (Positive)	True Positive (TP)	False Negative (FN)
	Should not alarm (Negative)	False Positive (FP)	True Negative (TN)

Figure 2. Illustration figure of the proposed evaluation framework

Note that each observed trip is a realization of the real-world stochastic driving environment. Therefore, in the second step, a large number of trips need to be aggregated to construct the confusion matrix. The trajectory dataset is expected to reproduce the real-world driving environment and cover both normal and safety-critical driving situations in order to investigate the safety metric performance. Note that not only collision-unavoidable moments but also some moments before the earliest collision-unavoidable time need to be included in the "should alarm" category to examine the safety metric predictive capability. Based on the confusion matrix obtained from large-scale trajectories, the safety metric performance can be systematically evaluated. The evaluation results can not only be used to compare different metrics, but also to help build a feedback improvement process to iteratively tune, refine, and improve the safety metric performance.

CASE STUDIES

Introduction of three evaluated real-time safety metrics

Three real-time safety metrics, including TTC, PCM, and MPrISM, were selected to demonstrate the performance of the evaluation framework using simulated trajectories. TTC is one of the commonly used safety metrics to evaluate vehicle situational safety. PCM and MPrISM are state-of-the-art safety metrics proposed to evaluate ADS safety. Detailed information on the three metrics is introduced in the following paragraphs.

The first selected safety metric is TTC, which is defined by

$$TTC = \frac{x_{i-1} - x_i - l}{v_i - v_{i-1}}, \#(1)$$

Where x_{i-1} and x_i denote the longitudinal position of the leading and following vehicles, respectively. v_{i-1} and v_i denote their corresponding longitudinal speed. l denotes the length of the leading vehicle. In this study, the TTC is considered alarmed if it is smaller than 1 second.

The second selected safety metric is PCM. It formulates the situational safety evaluation problem into a SV trajectory planning task. An optimization problem is developed to find the best SV trajectory as defined by an objective function with four weighted terms. The decision variables of the optimization problem are the SV action sequences. The authors assume POVs will follow their current speeds and headings during the prediction horizon and the objective is to minimize the SV "criticality" within the prediction horizon. The optimization formulation is as follows

$$\min_{\boldsymbol{u}=[u(0) \ \dots \ u(N-1)]} \sum_{k=1}^{N-1} \left(w_x R_x(k) + w_y R_y^2(k) + w_{ax} \frac{a_x^2(k)}{(\mu_{max} g)^2} + w_{ay} \frac{a_y^2(k)}{(\mu_{max} g)^2} \right), \#(2)$$

$$s.t. \quad s(k+1) = \boldsymbol{A}(k) s(k) + \boldsymbol{B}(k) u(k), \forall t = 0, \dots, N-1, \#(3)$$

$$c_r(k) \leq y(k) \leq c_l(k), \#(4)$$

$$x(k) \leq c_f(k), \#(5)$$

$$\boldsymbol{G} u(k) \leq \boldsymbol{h}. \#(6)$$

The objective function, i.e., the criticality, is composed of four parts, which include the longitudinal margin $R_x(k)$, lateral margin $R_y(k)$, longitudinal acceleration $a_x(k)$, and lateral acceleration $a_y(k)$. These terms are weighted using four pre-defined parameters w_x, w_y, w_{ax}, and w_{ay}. The decision variable is the SV action at the k-th moment $u(k)$ which includes the longitudinal and lateral accelerations $a_x(k)$ and $a_y(k)$. N denotes the number of look-ahead steps. μ_{max} denotes the road friction coefficient. The constraints of the optimization problem include the SV vehicle dynamics (Equation 3), where $s(k)$ denotes the SV state at the k-th timestep. Eqs. 4-5 denote the SV safety constraints with surrounding POVs during the prediction horizon. c_l, c_r, and c_f denote the left, right, and front boundaries, respectively, which are calculated based on the predicted trajectories of surrounding POVs. The last set of constraints (Equation 6) characterizes the admissible action space (i.e., the Kamm's circle as illustrated in Figure 3) of the SV. By solving this optimization problem at each time step, the safety metric can be obtained using the maximum estimated acceleration during the prediction horizon. If the maximum acceleration exceeds certain pre-defined thresholds (i.e., the maximum acceleration of the SV), then this moment is considered dangerous and the PCM is alarmed. The model parameters have been set according to the original paper, where $w_x = 1, w_y = 1, w_{ax} = 0.1, w_{ay} = 1$, time resolution 0.1s, look-ahead steps N =20 [4]. The PCM is considered alarmed if the maximum expected acceleration is greater than 8 m/s^2.

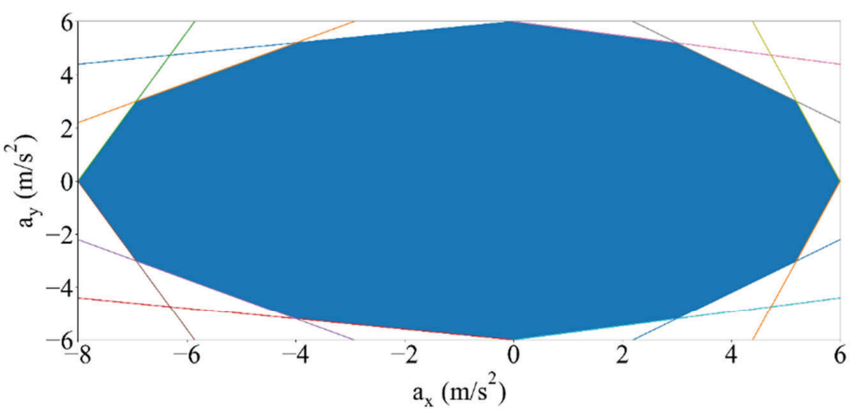


Figure 3. Vehicle admission action space (i.e., Kamm's circle)

The third selected metric is MPrISM. To evaluate the situational safety, it considers the pairwise interaction between the SV and each surrounding POV. It assumes the POV will take the worst-case behavior trying to collide with the SV. At the same time, it assumes the SV will take the best response trying to avoid the collision. Therefore, for each moment, to find the riskiest POV action and the best SV response, a min-max calculation is made. This is done every moment with the SV state (including vehicle position, speed, and heading) s_0, and the POV state, s_i. For each look-ahead step $N = 1, \ldots, T$, the Euclidean distance between the SV and the POV at the end of the prediction horizon N can be obtained by solving the following minimax problem:

$$\min_{\boldsymbol{u_i}} \max_{\boldsymbol{u_0}} \quad \sqrt{\left(x_i(N) - x_0(N)\right)^2 + \left(y_i(N) - y_0(N)\right)^2}, \#(7)$$

$$s.t. \quad s_0(k+1) = \boldsymbol{A_0}(k) s_0(k) + \boldsymbol{B_0}(k) u_0(k), \forall t = 0, \ldots, N-1, \#(8)$$

$$s_i(k+1) = \boldsymbol{A_i}(k) s_i(k) + \boldsymbol{B_i}(k) u_i(k), \#(9)$$

$$\boldsymbol{G_0} u_0(k) \leq \boldsymbol{h_0}, \#(10)$$

$$\boldsymbol{G_i} u_i(k) \leq \boldsymbol{h}_i, \#(11)$$

where x_0, y_0 and x_i, y_i represent the longitudinal and lateral positions of the SV and POV. $\boldsymbol{u_0}$ and $\boldsymbol{u_i}$ denote action sequences of the SV and the POV, respectively. For the constraints, Equation 8 and Equation 10 denote the system dynamics and Kamm's circle constraints of the SV, respectively. Equation 9 and Equation 11 denote the corresponding constraints for the POV. The predicted time-to-collision equals to $N\Delta$ (Δ is the time resolution) if the distance between SV and POV is smaller than a pre-defined collision threshold C at the end prediction horizon $N\Delta$. After iterating over all POVs, the least time-to-collision value across all POVs can be obtained and denoted as the model predictive time-to-collision (MPrTTC), which is used as the safety metric output. The model parameters have been set according to the original paper, where time resolution $\Delta = 0.1s$, look-ahead steps $N = 10$, and collision threshold $C = 4m$ [5]. The MPrISM is considered alarmed if the MPrTTC is smaller than 1 second.

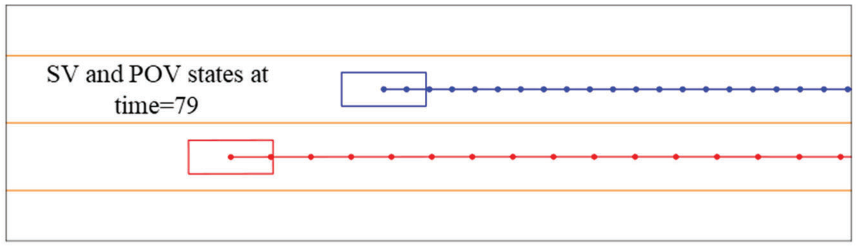
Simulation results

In this study, the performance of the proposed method is demonstrated using three simulated scenarios: SV overtaking POV, POV cutting in front of the SV, and POV and SV both moving into the same lane. The scenarios are generated using simulation models. In these scenarios, the real-time safety metrics may fail, for example, generating false alarms, or misidentifying potentially dangerous situations. The results show that the proposed method can successfully identify potential weaknesses caused by metric assumptions, approximations, and parameters.

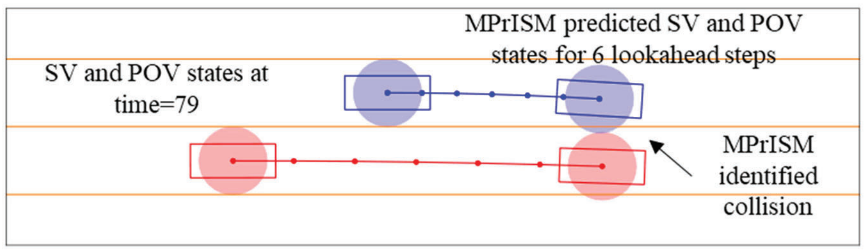
Scenario 1: SV overtakes POV

The first scenario is a normal overtaking case, where the SV is overtaking the POV on the adjacent lane and there is no crash happening in this scenario. The logged SV trajectory (shown by red curve and box) and POV trajectory (shown by blue curve and box) starting from the timestep 79 are shown in Figure 4a. The real-time safety metric results and the ground truth calculated with logged POV trajectory are shown in Figure 5. The red color moments in the ground truth denote the collision unavoidable moments. The red color in the safety metric results indicates that it alarmed at that moment. The MPrISM produces false alarms in this scenario.

As shown by the ground truth, it can be seen that moment 79 is not dangerous, and all metrics except for the MPrISM do not alarm. The MPrISM predicted trajectories of SV and POV are shown in Figure 4b. The shaded area denotes the assumed vehicle geometry by MPrISM. The MPrISM predicts that the POV will steer into the SV's lane since it makes worst-case behavior assumption regarding POV behavior. Therefore, a collision is predicted to occur in 0.6 seconds even if the SV takes the best response to avoid the collision. According to the collision definition in MPrISM, a collision is considered to happen when the Euclidean distance between two vehicles (considered as mass points with a certain radius) is smaller than the collision threshold. From the results as shown in Figure 4b, it is apparent that the predicted SV and POV circles overlap, so a collision is predicted by MPrISM. However, the two vehicles are not colliding with each other if considering their real geometry as shown by the two rectangles. Therefore, a false positive case is produced by the MPrISM due to both the worst-case behavior assumption and the single-circle approximation. Note that there is a trade-off for the MPrISM collision threshold. Using a larger threshold will cause more false alarm cases as shown in this scenario, whereas using a smaller threshold will cause false-negative cases (i.e., miss dangerous situations) as shown in Scenario 3.



(a)



(b)

Figure 4. (a) the logged SV (Red) and POV (Blue) trajectories starting from timestep 79, (b) the MPrISM predicted SV (Red) and POV (Blue) trajectories starting from timestep 79 of Scenario 1. The shaded area denotes the assumed vehicle geometry by the safety metric.

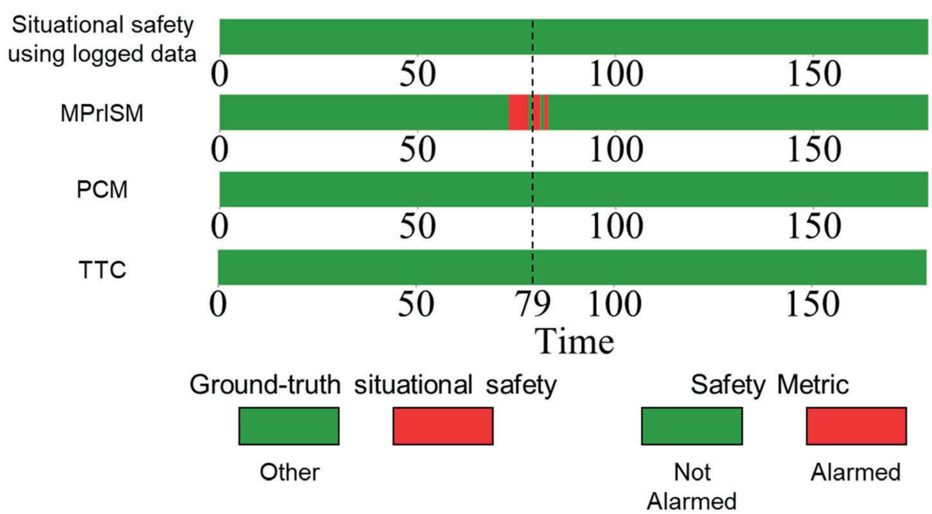
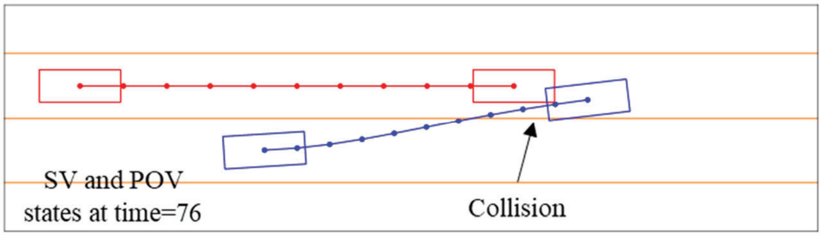


Figure 5. The ground-truth situational safety calculated using post-trip logged trajectory data and real-time safety metrics results of Scenario 1.

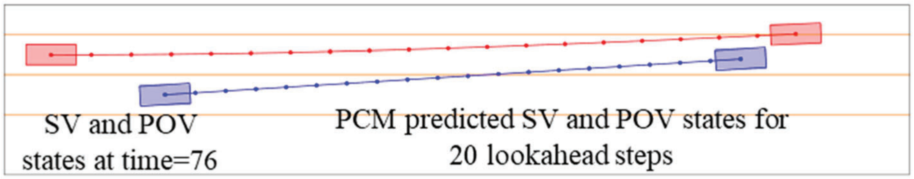
Scenario 2: POV cut-in SV

In the second scenario, the POV cuts in front of the SV, and a crash happens during the process. The logged trajectories starting from timestep 76 are shown in Figure 6a. The post-trip calculation and safety metrics results are shown in Figure 7. The PCM and TTC produce false-negative cases in this scenario.

From post-trip calculation results, starting from moment 76, the state of the SV is collision unavoidable, given the actual POV actions after that moment. The MPrISM successfully alarms starting from the moment 76. It does not generate false positives as in the previous case, since the worst-case behavior assumption regarding the POV can accurately characterize the aggressive cut in behavior. However, both PCM and TTC did not activate at the moment and generated false negatives. The PCM predicted SV and POV trajectories are shown in Figure 6b. The PCM assumes POV will maintain its current speed and heading within the prediction horizon. As a result, the SV can avoid the crash by turning to the adjacent lane. However, in the real situation, the POV is cutting in more aggressively compared with PCM behavior assumptions. Therefore, the PCM fails to activate at this collision unavoidable moment. For the TTC, it fails to consider the cutting in POV since they are not in the same lane at the current moment. Thus, the SV does not consider the POV as its leading vehicle and therefore produces false-negative cases.



(a)



(b)

Figure 6. (a) the logged SV (Red) and POV (Blue) trajectories starting from timestep 76, (b) the PCM predicted SV (Red) and POV (Blue) trajectories starting from timestep 76 of Scenario 2. The shaded area denotes the assumed vehicle geometry by the safety metric.

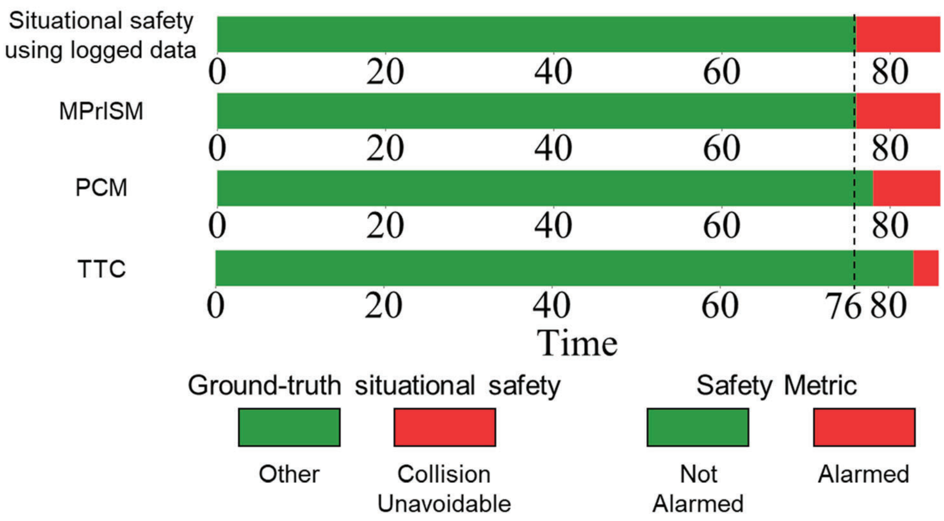
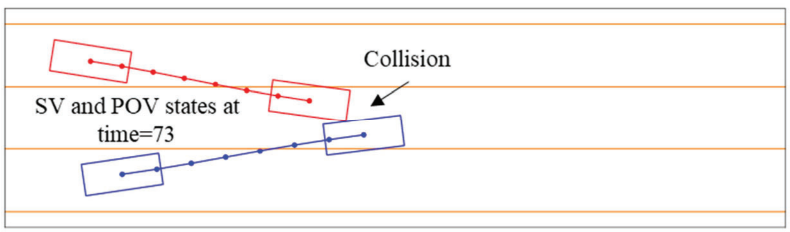


Figure 7. The ground-truth situational safety calculated using post-trip logged trajectory data and real-time safety metrics results of Scenario 2.

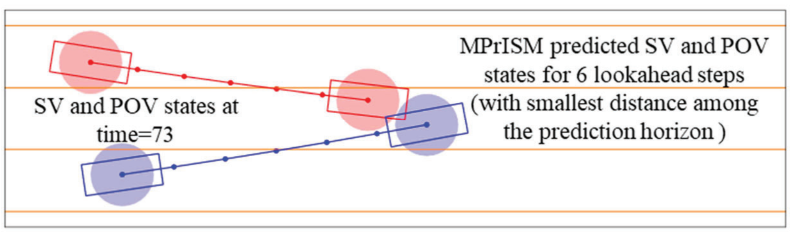
Scenario 3: POV and SV both move into the same lane

In the third scenario, the POV and SV are lane-changing to the same target lane and a crash happens during this process. All three metrics produce false-negative cases in this scenario. The logged trajectories starting from timestep 73 are shown in Figure 8a. The post-trip calculation and safety metrics results are shown in Figure 9.

From post-trip calculation results, starting from moment 73, the SV is collision unavoidable. However, all three metrics fail to alarm before the moment and cause false negatives. To analyze the MPrISM result, the MPrISM predicted trajectories of SV and POV are shown in Figure 8b. Within the 10 looking-ahead steps of MPrISM, the minimum predicted distance between the SV and POV occurs at the 6 look-ahead steps. As shown in Figure 8b, the two vehicles have already collided with each other as shown by the rectangles. However, due to the single circle approximation adopted by the MPrISM, it fails to identify the collision. One method to improve accuracy is to use multiple circles to represent the SV and POV. The benefit of using three circles rather than one circle is that it can more accurately capture the vehicle geometry and determine whether a collision happens or not. For the PCM, the reason why it fails to detect the crash in this scenario is similar to that discussed in Scenario 2. For the TTC, the reason why it fails to detect the crash in this scenario is that the leading vehicle velocity is higher than the following vehicle velocity.



(a)



(b)

Figure 8. (a) the logged SV (Red) and POV (Blue) trajectories starting from timestep 73, (b) the MPrISM predicted SV (Red) and POV (Blue) trajectories starting from timestep 73 of Scenario 3. The shaded area denotes the assumed vehicle geometry by the safety metric.

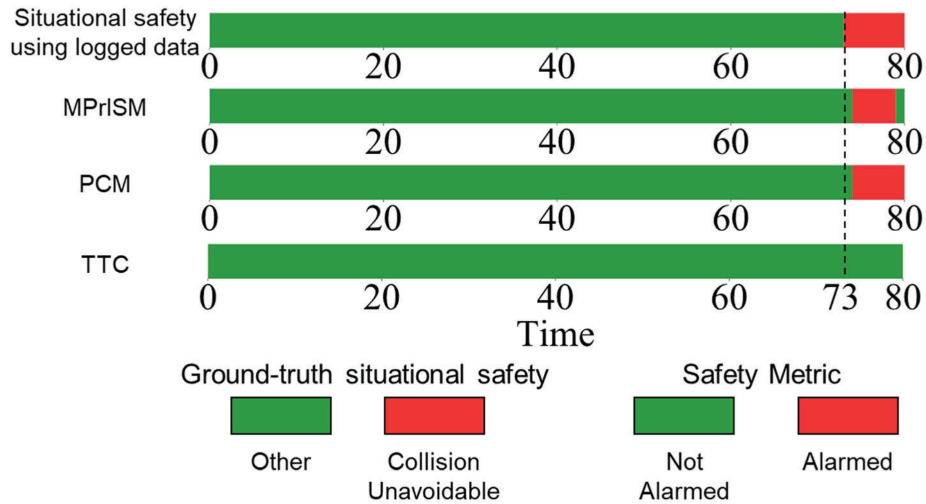


Figure 9. The ground-truth situational safety calculated using post-trip logged trajectory data and real-time safety metrics results of Scenario 3.

CONCLUSIONS

In this study, an assessment framework of real-time safety metrics performance is developed, utilizing post-trip information from vehicle logged trajectory data. Specifically, the logged POVs trajectories are utilized to determine whether the SV is in a collision unavoidable situation, which is an objective reflection of the situational safety, and this can serve as the ground truth of safety metrics. The goal of a safety metric is to alarm no later than the collision unavoidable moment to indicate the situational safety of the SV, but also without over-alarming (i.e., without too many false-positives). By analyzing the safety metric outputs and ground truth, one can identify the safety metrics' failed scenarios and investigate their potential weaknesses caused by model assumptions, approximations, parameters, etc. These results can help developers improve the performance of existing safety metrics. It can also help researchers and practitioners better understand the characteristics of different safety metrics, which may result in a choice of more appropriate metrics for different application purposes. This study has shown the benefit of a large number of logged trajectories for conducting a systematic analysis of safety metric performance. These can be used to compare the safety metric results with the calculated ground truth in order to obtain their statistical performance evaluation. It is recommended that logged trajectories cover diverse scenarios and locations to examine safety metrics performance under different circumstances. In this study, simulated trajectories were used to showcase the evaluation results. In the future, with the development of sensing technologies and data acquisition systems, real-world trajectories can be recorded and used to enrich and complement the simulated datasets.

STUDY LIMITATIONS

- In this study, it is assumed that perception systems are perfect in identifying POVs and their relevant attributes. Metric performance implications of uncertainties and inaccuracies that may be associated with a perception system's performance are not explored in this study.
- The metrics are assessed using three relevant but simple scenarios involving a single POV in each case. There are other complexities (e.g., computational) associated with these metrics when calculating a crash risk measure in real-world scenarios where multiple safety relevant objects may operate within the future time horizon of a metric's consideration. These are not explored in this paper.
- The crash risk metrics studied in this paper make assumptions on the bounds of how POVs may behave. In this study, those parameters are set as recommended by their authors with citation. However, metric performance is very sensitive to these settings. The use of the cited parameters does not imply agreement on these as reasonable or acceptable settings.

REFERENCES

[1] Hayward, J. C. (1972). Near miss determination through use of a scale of danger.

[2] Ozbay, K., Yang, H., Bartin, B., & Mudigonda, S. (2008). Derivation and validation of new simulation-based surrogate safety measure. Transportation research record, 2083(1), 105-113.

[3] Wachenfeld, W., Junietz, P., Wenzel, R., & Winner, H. (2016, June). The worst-time-to-collision metric for situation identification. In 2016 IEEE Intelligent Vehicles Symposium (IV) (pp. 729-734). IEEE.

[4] Junietz, P., Bonakdar, F., Klamann, B., & Winner, H. (2018, November). Criticality metric for the safety validation of automated driving using model predictive trajectory optimization. In 2018 21st International Conference on Intelligent Transportation Systems (ITSC) (pp. 60-65). IEEE.

[5] Weng, B., Rao, S. J., Deosthale, E., Schnelle, S., & Barickman, F. (2020, June). Model predictive instantaneous safety metric for evaluation of automated driving systems. In 2020 IEEE Intelligent Vehicles Symposium (IV) (pp. 1899-1906). IEEE.

[6] Althoff, M., Stursberg, O., & Buss, M. (2009). Model-based probabilistic collision detection in autonomous driving. IEEE Transactions on Intelligent Transportation Systems, 10(2), 299-310.

[7] Althoff, M., & Mergel, A. (2011). Comparison of Markov chain abstraction and Monte Carlo simulation for the safety assessment of autonomous cars. IEEE Transactions on Intelligent Transportation Systems, 12(4), 1237-1247.

[8] Schreier, M., Willert, V., & Adamy, J. (2016). An integrated approach to maneuver-based trajectory prediction and criticality assessment in arbitrary road environments. IEEE Transactions on Intelligent Transportation Systems, 17(10), 2751-2766.

[9] Yan, X., Feng, S., LeBlanc, J. D., Flannagan, C., & Liu, X. H. (2022) Evaluation of automated driving system safety metrics with logged vehicle trajectory data. Under Review.

Assurance Through Safety Cases—There’s a Claim For That

By Nat Beuse and Chan Lieu
Aurora Innovations, Inc.

Aurora Submission for ESV 2023

Executive Summary

Designing, developing, testing, and deploying an Automated Driving System (ADS) for use on public roadways in the United States is challenging for a variety of reasons, including for the ADS developer in defining and describing their approach for ensuring the safety of their vehicles. An autonomous vehicle is subject to National Highway Traffic Safety Administration (NHTSA) motor vehicle safety requirements, despite there being no defined Federal Motor Vehicle Safety Standards (FMVSS) that govern ADS performance requirements. The operation itself may be subject to other federal safety, state, and local laws and regulations depending on the type of operation (e.g., commercial motor vehicle or passenger service operation) and operating location. In addition, there is federal voluntary guidance containing priority safety design elements and a growing number of relevant industry-developed consensus standards and best practices available to an ADS developer to consider and incorporate in the design of their ADS. In navigating these various regulatory frameworks, standards, and best practices, the ADS developer is still ultimately responsible for defining and ensuring safety for their own vehicles. A safety case-based approach is a valuable way to provide such assurance.

A safety case is a structured argument, supported by evidence, intended to justify that a system is acceptably safe for a specific application in a specific operating environment. While this approach is not entirely new – safety cases have been incorporated into other safety-critical industries – safety cases for the development of autonomous vehicles are novel.

A safety case-based approach creates value through both flexibility and a high degree of rigor, if applied correctly. It is flexible because it provides the ADS developer with the latitude to determine what claim to make, and it is rigorous because there must be evidence to substantiate it. For example, there are now several publicly available voluntary industry standards and guidance spanning many important topics related to the development and safe operation of an ADS. These topics include functional safety, behavioral safety, and safety assurance for machine learning systems.[1] The emergence of these standards provide varying perspectives that ADS developers should consider and how an ADS developer implements these standards can be the basis of a safety case claim related to adhering to industry standards.

This paper will present Aurora's experience and lessons learned in developing and implementing its own Safety Case Framework. This includes discussion regarding how Aurora integrates existing industry standards into the ADS development process, while also

[1] https://safeautonomy.blogspot.com/2022/04/maturity-levels-for-autonomous-vehicle.html
https://www.eetimes.com/ul-4600-draft-puts-safety-onus-on-av-hopefuls/

building on them by incorporating into the development process vehicle product engineering requirements, enterprise wide processes, and operational elements (such as a Safety Management System). A safety case-based approach is important to ensure that the integration of many new, overlapping standards is managed correctly. And ultimately, a safety case-based approach provides transparency and insight into safety assurance.

What is a Safety Case?

A safety case is a structured argument, supported by evidence, to justify that a system is acceptably safe for a specific application in a specific operating environment.[2] A structured argument includes a specific claim – in Aurora's case, that our self-driving vehicles are acceptably safe to operate on public roads – that is then divided into subclaims. These subclaims may be further broken down into additional subclaims that ultimately result in evidence to substantiate the claim.

Safety cases are not a new concept – they have been widely used in various other safety-critical industries, such as oil and gas exploration, aviation, rail, medical devices, and nuclear energy.[3] (See Fig. 1 for other examples and timelines)

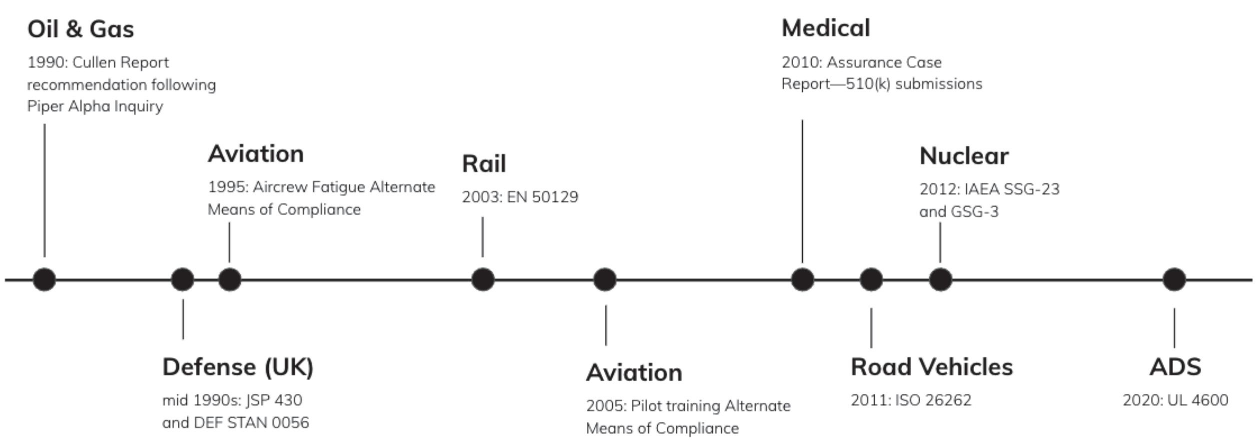


Fig. 1 Safety cases in other industries

Aurora has adopted a safety case-based approach because we believe it is the most logical and efficient manner to show and explain how Aurora determines that our self-driving vehicles are acceptably safe to operate on public roads. The heart of our safety case is a

[2] Defence Standard 00-56 Issue 4 (Part 1): Safety Management Requirements for Defence Systems. UK Ministry of Defence.
[3] See Cullen, W. Douglas. *The Public Inquiry into the Piper Alpha Disaster*. London, UK, 1990.

structured argument, supported by evidence, to demonstrate the claim for why our vehicles are acceptably safe for operation on public roads.

No single piece of evidence captures the totality of safety. There are complex interactions and relationships between the many elements that go into developing an ADS and it can be difficult to track and trace these interactions. We can provide various pieces of evidence, but without the appropriate context it would be difficult to understand why and how these pieces of evidence are relevant.

Ultimately, evidence without a claim is simply trivia and, conversely, a claim without evidence is baseless. A safety case-based approach brings these two essential concepts together in a logical manner to effectively show the work that we have undertaken to determine our vehicles are acceptably safe to operate on public roads.

Safety Case Framework vs. Safety Cases

Aurora's Safety Case Framework is designed to be adaptable to different vehicle platforms and operational contexts, and is the superset that captures all the claims that Aurora argues are necessary to safely deploy our ADS. It provides a logical mechanism to describe, characterize, and justify that the ADS is acceptably safe to operate. The structured argument of the Safety Case Framework outlines ***what*** the completed safety argument must achieve without providing prescriptive requirements on ***how*** it must be achieved (see Fig. 2). This abstraction is intentional and provides flexibility for different engineering teams to define and develop the evidence to substantiate the claim. Aurora's safety cases are derived from this framework and each is simply the compilation of the evidence addressing the relevant claims.

Aurora's Safety Case Framework is built upon five principles that describe our approach to developing our self-driving technology – Proficient, Fail-Safe, Continuously Improving, Resilient, and Trustworthy. The Principles are the broad categories that guide and describe our goals for developing safe autonomous vehicles, including:

- **Proficient**: The vehicle is acceptably safe during normal driving. Essentially, everything is working as intended.

- **Fail-Safe**: The autonomous vehicle is acceptably safe when there is a fault or failure. We design our vehicles in such a way that, if some component fails (like if a sensor is damaged or a tire blows out), the vehicle should behave in a manner that does not endanger its passengers or other road users.

- **Continuously Improving**: Aurora is committed to continuously improving. We are constantly learning and striving to identify, evaluate, and resolve anomalies that could affect the safety of the vehicle.

- **Resilient**: Our vehicles are acceptably safe in the case of reasonably foreseeable misuse and unavoidable events. For example, our cybersecurity-related claims mostly reside under this principle.

- **Trustworthy**: The public can have confidence in not only Aurora's autonomous vehicles, but our entire company – that we not only design, build, and test our self-driving vehicles in a dependable manner, but also that we have a safety and organizational culture in place to quickly address and resolve issues.

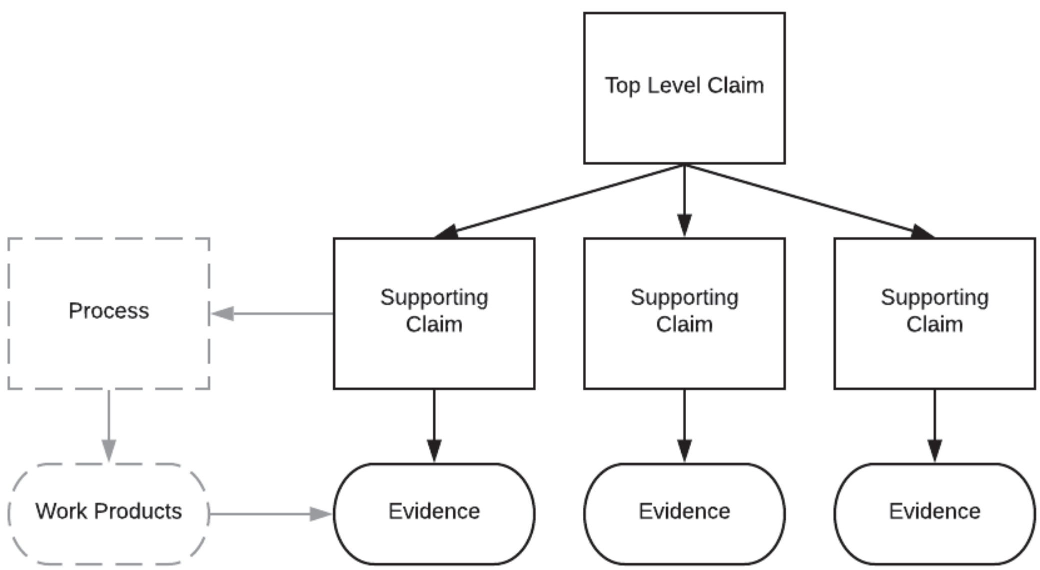


Fig. 2 Top down claim structure

Safety Case Tailoring

We use the Safety Case Framework to create a tailored safety case, taking care to define its specific context and application in each instance. A tailored safety case is a subset of the overall framework that includes the claims and subclaims appropriate for a specific use case, such as autonomy operations with a vehicle operator (VO). Aurora has multiple tailored safety cases, the primary two of which are centered around—[4]

1) operating in autonomy on public roads with VOs onboard (VO Road), and

[4] Aurora also has other safety cases based on specific hardware configurations (e.g. sensors and computing hardware) on various vehicle platforms (e.g. Class 8 tractor or passenger vehicle).

2) operating in autonomy with no VOs onboard (NVO Commercial).

Since the VO Road safety case is designed for ADS on-road operations with a VO present in the vehicle, the claims in this tailored safety case emphasize VO hiring, training, and certification and recertification, as well as driver monitoring systems. However, claims related to VO hiring, training, and certification and recertification become irrelevant when we progress to the NVO Commercial safety case. As a result, those VO related claims are no longer applicable, or are "tailored out." At the same time, we "tailor in" other claims, such as the ADS's ability to execute a minimum risk maneuver to achieve a minimum risk condition when a fault is detected, for the NVO Commercial safety case since there is no longer a VO in the vehicle to intervene.

This tailoring of the framework is unique and depends on the assumptions, claims, and arguments we make in each safety case. As a result, while the Safety Case Framework may be the superset, the tailoring activity is essential to guide evidence development. Aurora already maintains VO safety cases and expects to maintain multiple safety cases for NVO and different vehicle platforms (e.g., Class 8 trucks vs. passenger vehicles) so that all vehicle configurations we operate on public roads have an associated safety case.

Types of Claims

Aurora's Safety Case Framework claims can be broadly categorized into three major buckets: product, process, and operational safety. Product claims are related to the engineering of the ADS, and the evidence to support those claims will change over time as we iterate and improve the ADS's capabilities. Process claims are related to our process for how we develop the ADS and the evidence supporting those claims tend to be less frequently updated. Since most companies will look to improve their development processes, we are not implying that the evidence for this will never change, but it will do so less frequently. Finally, operational claims are related to how the ADS is operated on public roads.

Aurora's Safety Case Framework is more than a set of product requirements by encapsulating contributions from the entire company. This means that, in addition to the engineering involved with developing the ADS, we are incorporating non-engineering contributions from other portions of the company, such as government relations or legal, that are critical to deploying an ADS. By aligning cross-functional contributions when developing and deploying the ADS, we have created a safety culture throughout the company in which everyone at Aurora plays a role in safety, and our safety case helps recognize all of those contributions in a manner that requirements alone don't capture.

Incorporating Standards & Best Practices

The National Highway Safety Administration (NHTSA) has promulgated a number of FMVSS that define performance standards for automotive safety requirements that apply to all motor vehicles, including autonomous vehicles. These standards include clear test procedures, parameters, and pass/fail criteria that motor vehicle manufacturers meet when designing their vehicles. While NHTSA has not yet published FMVSS specifically related to ADS performance requirements, many ADS developers leverage voluntary industry developed standards and best practices as they design, develop, and deploy their ADS.

The industry, through standards development organizations, has produced a number of useful voluntary standards and best practices for ADS developers to consider and adopt when designing their ADS. These voluntary standards and best practices cover a variety of safety relevant topics, including system and operational safety and ADS verification and validation, operational design domain, and cybersecurity. Notable standards and best practices ADS developers can consider include ISO 26262 Road vehicles – Functional Safety [5] and ISO 21448 Road vehicles – Safety of The Intended Function (SOTIF)[6] regarding system safety and a number of Automated Vehicle Safety Consortium (AVSC) best practices, ranging from ODD definitions, to first responder interactions, to metrics and methods for assessing safety outcomes (see Table 1).

[5] ISO 20262 Road vehicles – Functional safety, 2018. https://www.iso.org/standard/68383.html
[6] ISO 21448 Road vehicles – Safety of The Intended Function, 2019. https://www.iso.org/standard/70939.html

AVSC ID	Title
AVSC0006202103	Metrics and Methods for Assessing Safety Outcomes of Automated Driving System (ADS)
AVSC00009202208	Interactions Between ADS-DVs and Vulnerable Road Users (VRUs)
AVSC0007202107	Information Report for Adapting a Safety Management System (SMS) for Automated Driving System (ADS) SAE Level 4 and 5 Testing and Evaluation
AVSC0005202012	First Responder Interactions with Fleet-Managed Automated Driving System-Dedicated Vehicles
AVSC00008202111	Evaluation of Behavioral Competencies for Automated Driving System Dedicated Vehicles (ADS-DV)
AVSC0004202009	Data Collection for Automated Driving System Dedicated Vehicles to Support Event Analysis
AVSC00002202004	Describing an Operational Design Domain: Conceptual Framework and Lexicon
AVSC00001201911	Safety operator selection, training, and oversight procedures for automated vehicles under test

Table 1 AVSC best practices

ADS developers each have their own processes to evaluate whether and how to adapt and conform with these various standards and best practices. That type of work is captured in Aurora's Safety Case Framework in multiple different places. For example, under one Aurora claim –"G5 The Self-Driving Enterprise is trustworthy" – there is a subclaim related to how prevailing industry best practices and standards are reviewed and adherence documented (G5.1.1.1.3). Evidence to support this subclaim might include a process of how Aurora periodically surveys the publication of new industry standards or best practices. It might also include a library of identified standards and the documentation on how Aurora is conforming to these standards. We would also expect to also see evidence in the G1 Proficient principle related to how these standards or best practices are traced to requirements. Tying all these different pieces of evidence together is what substantiates the claim.

Safety Case Challenges

Maintaining flexibility

Aurora chose to be less prescriptive in developing the safety argumentation and the claim language in its Safety Case Framework because flexibility is key to our G3 Continuously Improving principle. This flexibility enables broader applicability and longevity before having to revisit and update the overall framework. For example, in our claim regarding incorporating industry standards and best practices described above, the claim language itself states that "prevailing industry best practices and standards are reviewed and adherence documented, on a continual basis." It purposely does not specify *which* industry best practices and standards because the identification and selection of those standards is part of the point of the claim. This flexibility enables us to retire a best practice or standard should it become outdated or deprecated, without having to redefine the argument. Furthermore, prescriptive language has the risk of introducing rigidity and risk the argument becoming invalid.

Claim ownership

When it comes to project management of a Safety Case Framework claim, it is important to identify a single claim owner who will be designated as responsible for ensuring that the evidence to support the claim is completed. While this may be straightforward in a hierarchical structure, it becomes more difficult to manage with cross functional teams of an organization. Cross functional teams bring together resources from multiple parts of the organization to jointly work together (e.g., incorporating resources from product development, software, and hardware). However, the challenge now becomes which functional organization would be responsible for overseeing work on the single claim. Ultimately, as with many development efforts, coordination and communication is key to ensuring that all collaborators understand the task and are aligned on the deliverables.

Adapting UL 4600

UL 4600 is intended to help ensure that an acceptably thorough consideration of safety for an autonomous product has been performed during the design process and will continue to be done throughout the system life cycle. It does so by emphasizing repeatable assessment of the thoroughness of a safety case.[7] It is also the first standard to address the entirety of safety assurance in the design, development, and deployment of automated vehicles.

[7] UL Standard for Safety for Evaluation of Autonomous Products, UL 4600, 2nd Edition, March, 2022.

While UL 4600 was comprehensive in many aspects of autonomous vehicle development, there were several gaps that were deliberately out of scope by the authors, most notably the road testing of prototype vehicles that include human operators responsible for supervising the autonomous systems.[8] Since Aurora is currently testing vehicles in development with VO supervision of the ADS, it was imperative to incorporate those safety arguments into our overall Safety Case Framework. By developing those additional claims, we are then able to control when they come into scope.

Communicating a safety case

Since safety cases are new for the automotive industry, it is necessary to educate industry stakeholders, regulators, and others on what safety cases are, how they are constructed, and how they can be interpreted. This initial effort is necessary because safety cases can be complex and difficult to comprehend, often requiring context and explanation in order to fully grasp the safety argument. Beyond this initial education, there is also the challenge of bringing all the different pieces of evidence together in order to tell that story. A comprehensive report that breaks down the safety case argument, details the claims, and puts the various pieces of evidence into the necessary context, without forcing the reader to wade through potentially hundreds or thousands of pieces of evidence, would be the most effective way to summarize that work.

Conclusion

We've built our Safety Case Framework and each of its five supporting principles to guide responsible ADS development. Each principle is supported by multiple claims and will be substantiated by hundreds of pieces of evidence. Only by validating our system with hardened evidence through this process can we build confidence in the Aurora Driver's ability to safely operate on public roads without a human driver.

We believe this is a powerful tool that not only can guide a company as it develops ADS technology, but can also be useful in telling a coherent safety story. At the same time, it would be inadvisable to mandate a safety case because doing so runs the risk of turning this introspective activity into a check-the-box exercise whereby the motivation would be rooted in compliance and whose value would be further diminished via enforcement action. We believe every company building this transformational technology should openly share their safety case. At Aurora, our safety case shows that we are doing more than just committing to safety in principle, we are putting safety into practice.

---

[8] *Id.* section 2.1.2