

Challenges for the Validation and Testing of Automated Driving Functions

Halil Beglerovic, Steffen Metzner and Martin Horn

Abstract In this paper, we will explore challenges for the validation and testing of Automated Driving Functions (ADF), which represent one of the major roadblocks for successful integration of emerging technologies into commercial vehicles. We provide an overview of current methodologies used for validation and testing, focusing on the missing parts. Furthermore, we give an insight into promising methodologies, frameworks, and research areas which aim to reduce current testing and validation efforts.

Keywords Challenges · Validation · Testing · Automated Driving Functions (ADF) · Advanced Driving Assistance Systems (ADAS)

1 Introduction

In order to satisfy the increasing demand for safety, reliability, and comfort of commercial vehicles, manufacturers and research groups put great effort in the development of new and sophisticated driving functionalities throughout the decades. This development has undergone several phases and it started with the driver assistance systems which required constant driver control, e.g., Cruise Control (CC), or were active only in certain situations, e.g., Emergency Brake Assist (EBA), Electronic Stability Program (ESP).

H. Beglerovic (✉) · S. Metzner
AVL List GmbH, Hans List Platz 1, 8020 Graz, Austria
e-mail: halil.beglerovic@avl.com

S. Metzner
e-mail: steffen.metzner@avl.com

M. Horn
Institute of Automation and Control, Graz University of Technology, Inffeldgasse 21B/I,
8010 Graz, Austria
e-mail: martin.horn@tugraz.at

With the increase in the processing power of computers, researchers and OEMs were able to include more sophisticated algorithms and sensors into the vehicle, what led to the development of Advanced Driving Assistance Systems (ADAS). The new technologies allowed the development of Adaptive Cruise Control (ACC), Lane Keep Assist (LKA), Lane Change Assist (LCA), traffic signs, pedestrian and vehicle detection, etc. Even though these systems are more complex and sophisticated, the driver still needs to monitor them continuously in order to compensate unforeseen conditions on the road or misbehavior of the algorithm. Nevertheless, the safety benefits of ADAS systems have led governments to make laws which obliges manufactures to include some of them into commercial vehicles (e.g., ABS and ESP), showing that these systems have become an irreplaceable part of the driving experience.

In the recent years, a significant rise in the set of functionalities of Automated Driving Functions (ADF) could be observed, as they are increasingly introduced into vehicles. ADF represent the next technological step as they aim to provide a driving experience where constant monitoring of the system is not needed. In its core, ADF combine several ADAS functions in a comprehensive and complex system. For example, a Highway Pilot ADF is combining ACC, LKA, LCA, traffic sign and vehicle recognition in order to successfully drive the vehicle on a highway. Many manufactures like Tesla, Daimler, Otto, etc., have already demonstrated various capabilities of highway pilots.

Current technology allows limited automated driving on specific scenarios and the aim for the future, as shown in Fig. 1, is to enable full autonomy where human interaction is not needed at all. These systems should be able to perceive and understand the environment in order to act accordingly to all possible situations encountered in real traffic. Such systems would open the door to new types of transportation where fleet of automated vehicles could be shared between people, effectively lowering the overall number of used vehicles. In addition, the mobility of elderly people or people with disabilities would increase drastically.

The main barrier for the commercialization of ADF is the need to test them on a theoretically infinite variety of scenarios, including a huge number of parameters to be varied to reflecting the reality. In order to keep up with the high commercialization demand, rapid developments and advancements in technology, testing, and

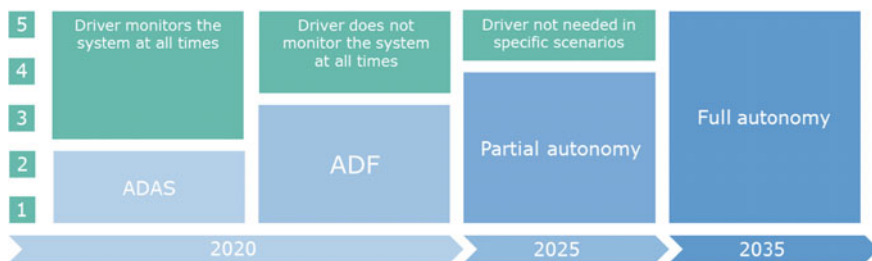


Fig. 1 Development of assistance systems

validation procedures need to lower the current testing effort, and in addition, new innovative methods need to be developed. Watzenig and Horn (2016) state that there is an increasing demand for methodologies and validation procedures that will enable the development of ADF which allow greater safety, traffic flow optimization, reduced emission and enhanced mobility.

2 Challenges for Validation and Testing

The challenges for validation of ADF are encountered on both sides, methods and tools. On the method side, the safety of the intended function (SOTIF) needs to be ensured while remaining economically feasible. On the other hand, the tools side needs to deal with new signal types and resolve the problem of transferring huge amount of data between the control units, which is out of the scope for the classic transfer protocols used in today’s vehicles.

2.1 Complexity of Automated Driving Functions

A general structure of an ADF can be seen in Fig. 2. The task of an ADF is to perceive and understand the environment and take adequate actions depending on the current situation. In order to accomplish this task, ADF use an increasing number of heterogeneous sensor systems and complex algorithms fusing and interpreting the data of the dynamic environment. This sensor inputs are combined with existing knowledge coming either from maps, vehicle-to-vehicle, or vehicle to infrastructure communication. All of these inputs, together with the internal states of the vehicle, e.g., velocity, engine speed, etc., are needed by the decision-making algorithms in order to predict and plan adequate trajectories.

However, the inputs coming from the heterogeneous sensor introduce new data types, e.g., object lists, images, radar data, and current software component verification tools are not able to handle them accordingly. In addition, the programming environments used for development also need to provide an interface to these new

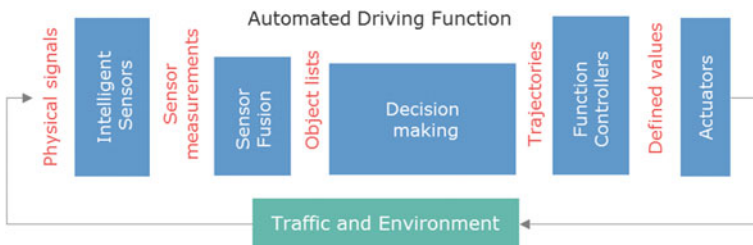


Fig. 2 Structure of an Automated Driving Function

data types. Furthermore, with the advancement of technologies like deep neural networks, dedicated hardware is introduced and developers need to be able to interface both the models and hardware into their existing development environment. Interfacing all these data streams with standard tools during the development phase is still an open issue, as there are still no communication standards.

2.2 Variation of Scenarios and Parameters

Winner et al. (2013) and Wachenfeld et al. (2015) state that an ADF should be driven for more than 240 million kilometers without fatalities to prove that they are, at least, not performing worse than human drivers. Performing such an extensive validation, using classical validation methods on public roads or proving grounds, would not be possible in a sustainable frame. If we take into consideration a typical development cycle, with several iterations of testing and validation, it becomes clear that such extensive test runs are not economically feasible.

On the other hand, simulation has been proven to be a very promising tool for successful development and testing of ADF. The most important advantage is the ability to easily vary the various scenarios and parameters, ensure reproducibility and, in the case of pure software in the loop (SiL), even run test cases in parallel and faster than real-time.

However, even if simulation is used for the scenario and parameter variation, the ADF could theoretically be tested on 10^{12} different test cases within various scenarios, traffic, and environment parameter variations. This type of full factorial testing is not feasible and new advanced methods are needed for appropriate scenario selection and parameter variation.

2.3 Scenario Selection and Test Generation

In order to successfully validate an ADF, an appropriate selection of scenario-subsets from all possible scenario variations, which will lead to a sufficient scenario coverage, is needed. Only a small portion of all available scenarios presents a challenge for the ADF and could potentially lead to faulty behavior. It is important to reduce the number of considered scenarios in order to save time, costs and resources needed for testing and validation.

The main task is the systematic selection of scenarios and corresponding parameters for the variation. A possible approach for the scenario selection is to group all scenarios into different categories, e.g., highway, left turn, right turn, country road, etc., together with different traffic behavior. With this type of grouping methodology it is possible to conduct validation on certain test runs and exclude test runs with similar characteristics from further consideration.

An opposite approach could be applied by identifying those scenarios leading to the most critical behavior of the ADF. Then by knowing these most critical scenarios, which can be taken from existing accident databases, we could expand the test to similar runs and conclude where the critical area stops.

Additionally, a subset of scenarios could be selected depending on specific regulations, imposed by the laws from different countries.

After appropriate scenario selection, an automated test generation procedure is needed. This procedure should be able to generate appropriate test cases taking into consideration the capabilities and level of autonomy of the ADF. Furthermore, no unified metrics, references nor testing criteria exist up to today. The test generation is usually carried out by engineers tailored to meet their specific needs. Better objective evaluation methods are needed which can derive, with some certainty, the overall behavior of the model from just a subset of test cases.

3 Current Methodologies/Technology Overview

Today's ADF functions are developed using environment simulation systems like VTD from Vires, PreScan from Tass or IPG CarMaker (there are much more available). Especially in the beginning of the development process, for both the algorithm and software development, those simulation systems are well accepted. All available environment simulation systems have strengths and weaknesses depending on the origin domain (vehicle dynamics, driver simulator, etc.) and allow manual design and simulation of a wide range of scenarios. Dependent on the complexity of the respective scenario, setup and design of those take a lot of time.

As for usual control units, the functionality and the electronics are integrated and tested on HiL systems. The HiL systems are either used for testing one single control unit or testing several units in compound on an integration HiL. Usually, setup and maintenance need a lot of effort. Variability and the possibility to influence the inputs of the control unit independent from the environment of the vehicle overcome these disadvantages.

After the integration of the functionality into the vehicle, testing and validation is nowadays executed on test tracks and public roads (Benmimoun 2017). Especially on public road, the reproducibility of scenarios is very difficult and can be very time consuming. Although simulation is used for algorithm development extensively, up to now the trust in this methodology is not high enough for testing with prototype vehicles.

In some cases, already ground truth measurement systems are used (Fritsch et al. 2013). Such systems are able to collect the data of the environment with higher precision than the mass productions sensors inside the vehicle. These systems enable the engineers to compare the sensor output of both vehicle sensor and ground truth measurement system, and facilitate high improvements in debugging.

However, reproducing such measured scenarios in simulation environment to debug the source code in detail still means a lot of manual work. In some cases,

limited tool support is available, which converts such measurements in reproducible scenarios to be executed in environment simulation tools.

Forthcoming autonomous driving systems are further extending the testing and simulation requirements of the safety and domain control functions. The vehicle needs to be able to anticipate road hazards in advance, and adjust driving strategies accordingly to increase driver's trust. Thus, there needs to be a way to provide the vehicle with an awareness of the road environment beyond the reach of its onboard sensors.

4 Validation—Global Approach

It is still unclear, how validation, homologation, and certification of automated driving functions above SAE (www.sae.org) automation level 2 shall be done. However, several research projects on national and international level are taking care about this topic. There are few points nearly all experts do agree on:

- The solution uses simulation for a major part of scenario-execution;
- A clever combination of methods and validation environments (SiL, HiL, test-track, public road, etc.) is necessary;
- The number of test cases/scenarios will still be quite big.

One of the research projects dealing with testing and validation optimization is the European ESCCEL research project ENABLE-S3 (www.enable-s3.eu). The aim of the project is to reduce current efforts needed to test and validate an automated system by 50%. The project is divided into two comprehensive parts as shown in Fig. 3. First, the validation methodology deals with novel approaches for scenarios and metrics selection, data acquisition and storage, and test optimization. In order to run the generated test cases, the second part of the project focuses on a reusable validation framework which can support seamlessly various development stages (MIL, HiL, ViL, Proving ground, etc.).

5 Supporting Tools in the Validation Task

The environment simulation tool is the central tool on all development and validation levels involving simulation. Such environment simulation systems are already available from several suppliers and usually include sensor models that can be directly connected to the ADAS/ADF controllers. Dependent on the accuracy of the models, a huge amount of processing power may be required. This includes models involving ray tracing methods necessary for detailed camera or radar models. However, for a lot of cases this accuracy is not necessary or even obstructive. If very accurate sensor models are used and the robustness regarding

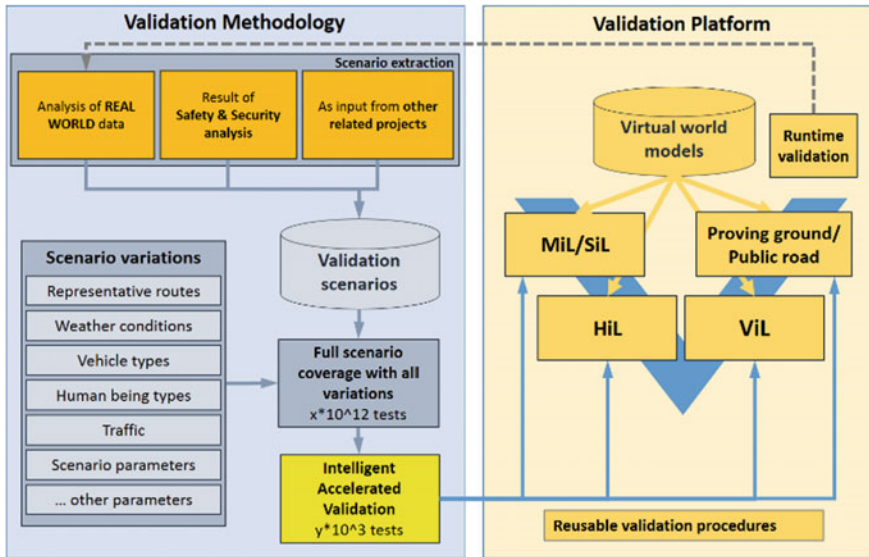


Fig. 3 Modular validation and verification framework of the ENABLE-S3 project (www.enable-s3.eu)

special effects of the sensor shall be tested, it is also necessary to model the environment in the same level of detail to trigger these effects. In such cases, a phenomenological model may be a better choice as the modeling of the scenario is made much easier. This is also valid for other parts of the model, e.g., the vehicle model may be reduced in details, if the test case does not contain stability critical maneuvers. Therefore, simulation platforms need to support the exchange of models dependent on the parameterization of the test case.

Additionally, mixed/real development and validation environments will be part of the validation as they are also in today’s environment. But to enable these new methods of stimulation (e.g., for radar and camera), new data types (object lists and raw/streaming data) are necessary to be used, e.g., vehicle in the loop is a well-established environment for the development of powertrain-related functions. This environment could also be used for ADAS/ADF validation. This raises the complexity of the vehicle tests in the same way as the complexity of the vehicle.

Even if methods are found to reduce the number of test cases, there are still millions of kilometers which need to be driven either virtually or on real roads. In addition, test cases need to be handled in a structured manner and if no real hardware is involved it is possible to execute these test cases in parallel on the cloud. This coordinated distribution of the test cases and collection of the results is also an important requirement to the simulation platform as well.

However, it will not be possible to execute all test cases virtually or with only partial hardware. Some of the test cases or scenarios still need to be executed in real environment or test tracks. The data collected should be as close as possible to the

data collected in simulation to enable usage of the same evaluation metrics. In order to accomplish this metrics transfer, ground truth measurement systems which record the environment with similar sensor as the vehicle but with higher precision, are necessary.

6 Standardization

For development and validation of ADF, both standards for the tooling and for the methodology need to be extended.

On methodology side ISO26262 is a well-accepted and applied functional safety standard for the development and validation of functionalities, realized with electronic control systems. In the concept phase of development, this standard demands, among others, an item definition and a hazard and risk analysis.

Both require the knowledge of the item (which may be available) and its environment, as well as the dependencies of the item to the environment. However, the fact that scenarios and the variation of the environment parameters (daylight, weather, temperature, etc.) are infinite restricts the direct application of ISO26262 on automated functions with automation level 3 and above. However, there needs to be a definition of a finite testing space regardless which method and which tools will be necessary for validation. Finding this testing space is even harder considering that all implementations of the ADF functionalities will have different strengths and weaknesses. A good example is the development of the vehicle dynamics functionality. In all test cases defined by authorities or car-magazines, for e.g., brake distance or stability during a lane change maneuver, the systems from different manufacturers react in a similar manner. But comparing those systems outside of this test space leads to significant differences as the developers are focusing on the defined test cases.

Therefore, the test case space for autonomous driving functionalities needs to be chosen wisely to avoid over- and under-optimized areas. In the worst case, this may lead to a product passing all tests but being still not useful for the customers.

ISO26262 has been published at a time where electronic control units have been already used for many years also for safety critical applications. Hence, it contains a lot of experience gained through the development of those control units. The same is to be expected for publication of the standard covering automated driving.

On the tools side the same problems as on control unit site occur as all bus systems, especially the CAN-bus, are optimized for signal-based controls-communication. This kind of communication will still play a big role in the future, but with new functions considering the objects of the environment, also object-based communication becomes necessary. For small objects containing few data, the CAN-bus may still be used, although the CAN-drivers need to be adapted. But the CAN-communication description and well known DBC-file format cannot be applied. Here it is necessary to establish new formats describing the exchanged objects and bus systems able to handle this kind of communication properly. This adaption needs to be applied to simulation, measurement and evaluation tools and on all related file types.

7 Conclusion

Currently, the main approaches for validating the ADF are based on proving ground or public road test drives. The challenges that arise from testing on proving ground or public road are that they are very expensive, time consuming, requiring huge effort and are hard to reproduce. Based on these challenges, simulation offers a solution as we are able to achieve high reproducibility with low effort. However, the question of how to select proper scenarios and parameter variations which will cover the infinite set of variation in a comprehensive manner is still an open issue. New methodologies and frameworks, able to automatically generate adequate test cases from input requirements or use ground truth data are necessary. In addition, standards dealing with new data types and new communication protocols need to be established. Furthermore, unified metrics, references and testing are needed for the overall standardization of testing procedures for ADF.

Acknowledgements The project leading to this application has received funding from the European Union's Horizon 2020 research and innovation programme under the Marie Skłodowska-Curie grant agreement No. 675999.

References

- Benmimoun M (2017) Effective evaluation of automated driving systems. SAE technical paper, 28 Mar 2017
- Fritsch J, Kuhl T, Geiger A (2013) A new performance measure and evaluation benchmark for road detection algorithms. Intelligent Transportation Syst (ITSC)
- Wachenfeld W, Winner H (2015) Die freigabe des autonomen fahrens. In: autonomes fahren. Springer, Berlin, Heidelberg
- Watzenig D, Horn M (eds) (2016) Automated driving: safer and more efficient future driving. Springer
- Winner H, Wachenfeld W (2013) Absicherung automatischen Fahrens, 6. FAS-Tagung München, Munich
- www.sae.org—society of automotive engineers
- www.enable-s3.eu—European initiative to enable validation for highly auto-mated safe and secure systems