Technical Evaluation and Impact Assessment of Automated Driving

Felix Fahrenkrog, Christian Rösener, Adrian Zlocki and Lutz Eckstein

Abstract Currently different research activities on automated driving are conducted around the globe. The European flagship research project on automated driving functions is the research project "AdaptIVe" (Automated Driving Applications and Technologies for Intelligent Vehicles). Besides the development of automated driving functions, the project deals with general research on legal aspects, human factors and evaluation. The evaluation and impact assessment of automated driving functions faces different challenges considering the complexity of the technology. In this context, this paper describes the evaluation approaches that are taken in the project for the technical evaluation and impact assessment.

Keywords AdaptIVe · Evaluation · Technical assessment · Impact assessment

1 Introduction

Automated driving has been a vision since the early 20th century. A first step towards vehicle automation was the introduction of cruise control in the Chrysler Imperial as the so called "auto pilot" in 1958 [1]. A first step towards this vision was the introduction of ADAS (advanced driver assistance systems) in the last

F. Fahrenkrog (✉) · C. Rösener · L. Eckstein
Institut für Kraftfahrzeuge, RWTH Aachen University,
Steinbachstraße 7, Aachen, Germany
e-mail: fahrenkrog@ika.rwth-aachen.de

C. Rösener
e-mail: roesener@ika.rwth-aachen.de

L. Eckstein
e-mail: eckstein@ika.rwth-aachen.de

A. Zlocki
fka Forschungsgesellschaft Kraftfahrwesen mbH Aachen,
Steinbachstraße 7, Aachen, Germany
e-mail: zlocki@fka.de

G. Meyer and S. Beiker (eds.), *Road Vehicle Automation 3*,
Lecture Notes in Mobility, DOI 10.1007/978-3-319-40503-2_19

decade of the 20th century. Following the successful introduction of ADAS research on higher automated driving functions has been ongoing since many years. These functions were intensively investigated and demonstrated during the DARPA Challenges [2, 3] as well as activities of Google and their so called Google self-driving cars [4] in the US and in Europe by the Berta Benz Drive [5] and the GCDC [6].

This chosen path is continued by the European research project AdaptIVe, which develops new automated driving applications in order to promote safer and more efficient driving [7]. Within AdaptIVe automated driving functions are developed for different speed ranges (parking, urban and highway driving) and at different automation levels [8] ranging from level 2 up to level 4 functions.

Due to its nature as a research project, AdaptIVe focuses less on the market introduction of automated driving, but rather on contribution to fundamental research questions of automated driving. These are the legal framework for automated driving functions [9], the human interaction with these functions [10] and the evaluation. For the latter one the objective is to set up a general evaluation framework that can be applied for different types of automated driving functions including a safety impact analysis and an environmental impact analysis.

2 Evaluation Methodology

The initial point for the definition of an evaluation framework for automated driving functions in AdaptIVe was a review of existing evaluation approaches for ADAS and automated driving functions.

During the function development typically a continuous and iterative technical evaluation is conducted. The main objective of this evaluation is to check, whether the pre-defined requirements are fulfilled by the functions and whether the defined performance is reached (e.g. SARTRE [11], HAVEit [12] or KONVOI [13]). This type of evaluation can be seen in the sense of a verification process, which is defined according to [14] as the evaluation of whether or not a product, service, or system complies with a regulation, requirement, specification, or imposed condition. It is often a company- or project-internal process.

On the other hand a technical evaluation can be conducted more in the sense of a validation, which is defined as "the assurance that a product, service, or system meets the needs of the customer and other identified stakeholders" [14]. Here, the function is assessed against certain pre-defined evaluation criteria. In this context important evaluation criteria are the acceptance and, in particular for a research project, the impact on traffic. This has been investigated in detail for different ADAS functions (e.g. TRACE [15], interactIVe [16], eIMPACT [17]), but not for automated driving functions. A general evaluation methodology for this evaluation stage has been introduced by the PReVAL [18] project. This approach considers three evaluation areas (technical, user-related and safety impact assessment).

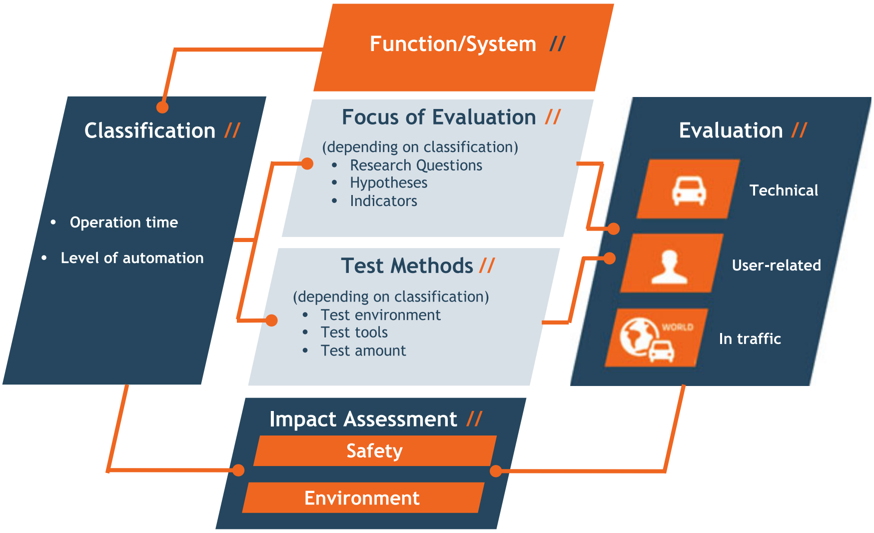

Fig. 1 Evaluation framework in AdaptIVe

The AdaptIVe evaluation framework (see Fig. 1) takes up the PReVAL approach of three different evaluation areas and extents it by a fourth evaluation area that is analysing the changes in the traffic behaviour of the surrounding traffic. The following evaluation areas are considered in AdaptIVe: technical assessment, user-related assessment, in-traffic assessment and impact assessment (safety and environmental effects). For AdaptIVe the focus is mainly on the validation of the developed functions, however also the verification is considered in the technical assessment. Within the AdaptIVe evaluation framework for each evaluation area the tools and indicators that should be applied are described in detail [19]. In general similar tools are used as for the evaluation of ADAS (field test, test in controlled environment, simulator test and simulation). However, the applied tools need to be adapted to the tested automated driving function. Therefore, different automated driving functions have been reviewed. The results of this review is that next to the automation level the operation time must to be considered when selecting appropriate evaluation tools [6].

Regarding the operation time functions have to be distinguished:

- Functions that operate only for a short period of time (seconds up to few minutes). Typical examples are automated parking functions and the minimum risk manoeuvre function. These functions are called *event-based* operating functions.
- Functions that once they are active can be operated over a longer period of time (minutes up to hours). A typical example for this type of functions is a highway pilot or a motorway automation function. These functions are called *continuous operating* functions.

In the following chapters the evaluation approaches for the technical assessment and impact assessment are presented in detail.

3 Technical Assessment

The focus within the technical assessment is on the performance of automated driving functions. A major challenge in setting up the evaluation framework is to limit the test effort in order to guarantee an efficient evaluation while ensuring that all important aspects are covered to ensure an as comprehensive as possible evaluation. Since the operation regime of automated driving functions is covering nearly the whole driving process, a large number and variety of driving situations needs to be taken into account [20]. It is obvious that detailed analyses of all driving situations it is not feasible due to the limited resources, as shown in Fig. 2.

3.1 Technical Assessment for Event-Based Operating Functions

For the assessment of event-based operating functions an approach similar to the assessment of today's ADAS functions is used (e.g. PReVAL [18], interactIVe [21]). According to this approach the first step the assessment scope by means of the research questions is defined. Based on these questions, related hypotheses are

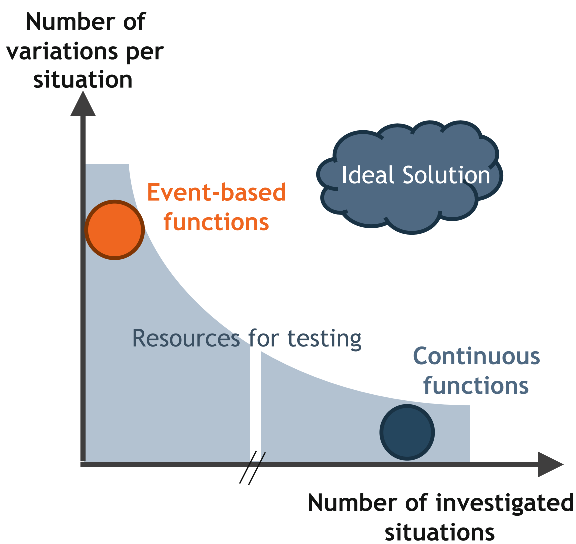

Fig. 2 Test approach for the technical assessment within AdaptIVe

defined to be analysed. Therefore adequate performance indicators are chosen for the analysis.

Once the definition of the assessment scope is defined the relevant test cases are selected. The basis for the test cases is typically the function's use cases respectively other situations that are relevant like e.g. accident scenarios. The actual testing is the second last step of this approach. The tests are typically conducted in a controlled field—mainly test tracks or in closed test garage for parking scenarios. During the testing the parameter set of the test case (e.g. initial velocities or relative distances) is varied. Ideally, each test—a test case with certain parameter set—is repeated several times in order to ensure statistically valid results. The analysis of the test data is the last step of the process. This step includes the calculation of derived measures as well as indicators. Based on the indicators the pre-defined hypotheses are analysed.

3.2 Technical Assessment for Continuously Operating Functions

For the continuously operating functions the focus is slightly different in the technical assessment. Since these functions cover different driving situations also for the assessment, a wider scope in the assessment is required. For the continuously operating functions the focus is less on the performance in a certain driving situation but rather on overall performance during the whole driving process.

Therefore, it is not useful to define certain single test cases. Instead a holistic assessment approach that covers as many different driving situations as possible is needed. Such an assessment approach is the (small) field test on public roads, in which the function must be able to handle different driving situations. The drawbacks of the field test approach are both, the rather uncontrolled test set-up and the relatively high effort for a field test in general. Hence, the extent of the field test needs to be limited to a feasible amount.

Analogue to event-based functions, the assessment approach of continuously operating functions starts with the definition of the research questions and hypotheses. This includes the definition of adequate performance indicators and criteria for the assessment of the hypotheses as well. In order to investigate the performance over the whole driving process adequate indicators are required. For this purpose the basic requirements for automated driving functions need to be considered, which are:

- The function must drive the vehicle in safe manner,
- The function must be able to operate in mixed traffic conditions,
- The function should not affect the other traffic in a negative way.

These basic requirements imply that the automated driving functions need to operate within the range of normal driving behaviour and should at least be as safe as non-automated driving.

Thus, the baseline for the assessment should be the human driver respectively his/her behaviour. Since the driving behaviour of each human driver is different, it can only be described in distributions. These distributions of driver behaviour need to be obtained before the actual assessment is performed. For this purpose two approaches are used in AdaptIVe:

- Data of previous field test projects will be used (e.g. filed operational tests (FOT) like the euroFOT project [22]), since it provides information on the driving behaviour of many different drivers.
- Each test route is driven several times before the test with and without the function in order to consider specific characteristics of the test region.

Next to the distribution of normal human driving also legal boundaries must be considered for the assessment (e.g. speed limits, restriction on passing) respectively their violation.

In the next step the actual tests are prepared. This includes the definition of the test route and test length. To limit the test effort to a feasible extend, the test route has to be chosen in a way that all relevant driving situations will occur with a certain probability. Therefore, the required test length respectively duration needs to be estimated a priori based on the number of expected relevant driving situations that occur while driving in public traffic. For this estimation, the data and the knowledge gained in previous field operational tests is used.

The actual tests are split into two steps. The first step is the pre-tests, in which the basic functionality is checked as well as the accuracy of the used sensors is analysed. These tests are conducted similar to the event-based functions tests on a test track. If the tests have been finished and the function is operating properly, the main test on public roads will be conducted. During the tests the test route will be driven several times with and without the function under assessment. After the tests, the data is evaluated analogue to the process described for the event-based function. The analysis will be conducted on the level of the whole test trip as well as on the level of certain relevant driving situations.

In order to enable a holistic technical assessment of the automated driving function in all relevant driving situations, the previously described methodology foresees a situation-based assessment. During the tests on public roads different driving situations occur stochastically. For identification and classification of these driving situations, detection-rules-based [23] or machine-learning-based approaches [24] can be used. These approaches are used to identify the following situations: Free driving, car following, lane change (left/right), cut-in of other vehicle (left/right) and approaching object.

In order to ensure that all relevant driving situations are occurring during the test, the length of the small field test is estimated by means of data from previous FOT, such as euroFOT [22]. The reference data of the field operational test is clustered in

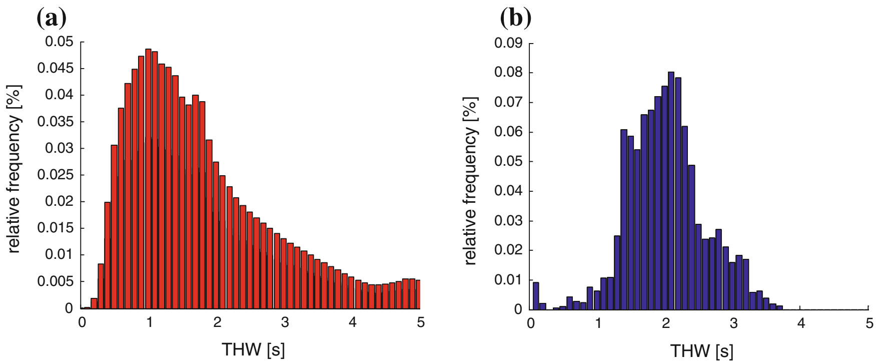

Fig. 3 Distribution of indicator time headway within the scenario "car following" for reference/euroFOT (*left*) and for the exemplary test of an automated function (*right*)

relevant driving situations by using a situation space approach. Afterwards, the distribution of spatial frequencies of all relevant driving situations is calculated. In accordance to [20] a Poisson distribution for the occurrence of driving situations is assumed and the minimal test length for the occurrence of at least k = 5 driving situations is calculated. While most driving situations are detected by using data obtained from field operational tests, driving situations influenced by the infrastructure, e.g. "enter motorway" are obtained by using statistics regarding the infrastructure of motorways in the test region.

After classification of the relevant driving situations the predefined hypotheses can be evaluated. For this purpose, it will be checked, whether the distribution of the performance indicator of the tested automated driving function deviates from the distribution of the reference data. The distributions of the indicator time headway (THW) within "car following" situations are shown for reference/euroFOT (left) and with an automated driving function for the exemplary test (right) in Fig. 3.

Driving with the automated driving function shifts the distribution of the performance indicator THW to become compared to the driving without the automated driving function. The median value of the THW increases from 1.5 s in the reference data to 2.0 s in the tests with the function. The analysis of the consequences of this effect with respect to the traffic in terms of efficiency and safety is part of the impact assessment.

4 Safety Impact Assessment

In the safety impact assessment it should be analyzed, which benefits in terms of avoided accidents or injuries automated driving functions can provide. Different approaches (field of application analysis, accident re-simulation, and field test) were developed and applied for ADAS in the past [25]. However, for the impact

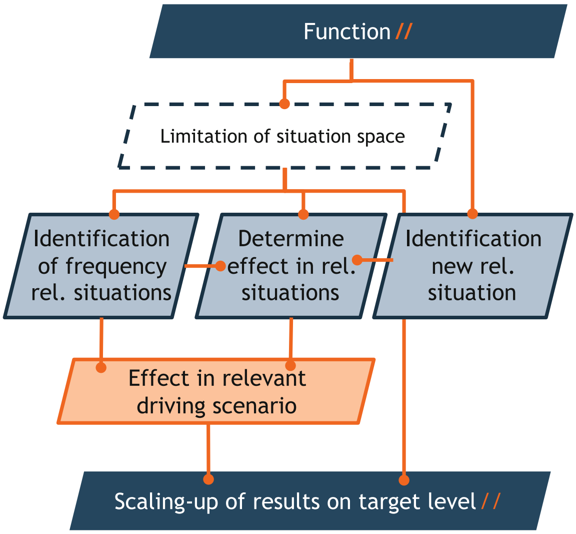

Fig. 4 Safety impact assessment approach in AdaptIVe

assessment of automated driving function different challenges compared to the impact assessment for ADAS must be considered:

- Today's accident data do not consider collisions of automated vehicles,
- Automated driving function operate already before a critical situation occurs, which makes the re-simulation of accidents more difficult
- Interaction with other road users (automated/non-automated) in mixed traffic condition has only rarely been studied so far.

In order to overcome these challenges an approach for the safety impact assessment has been defined that bases on [26] and consists of three main steps (see Fig. 4):

1. *Identify relevant situations*
   By means of the first step all relevant (critical) situations should be identified as well as their frequencies. Therefore, by means of microscopic traffic simulation the traffic flow at different road section is simulated. The microscopic traffic simulations allow to vary the penetration rates and to identify changes in the frequencies of driving situations depending on the penetration rate.
2. *Investigate the relevant situation in detail*
   In this step the effect of functions is investigated. The applied method is similar to the re-simulation. This means that certain driving situations are simulated with and without the function under study. The effect of the function is determined by comparison of both results. The analyzed driving situations include accidents as well as critical driving situation. In this step also the results of the other assessment (technical, user-related, in-traffic assessment) are considered.

3. *Identification of new situations*
   Automated driving function will cause also (critical) situations that are not part of today's traffic. Examples are the "transition of control" or "minimum risk maneuver" situation. These situations are investigated in the third step by means of risk analysis

5 Summary and Outlook

The paper describes the evaluation approach for the technical and safety impact assessment that is taken in the European research project AdaptIVe. A major challenge is the diversity of automated driving functions—ranging from automated parking function to automated highway driving functions at different automations levels. Therefore, in the developed evaluation framework the applied evaluation tools are chosen based on the operation time of the function. The developed evaluation framework is applied in AdaptIVe.

The evaluation approach for AdaptIVe does not consider the context of market introduction relevant validation/verification of sufficient safety operation of automated driving functions after the market introduction. To tackle this issue ika has already introduced the circle of relevant situation approach [27, 28]. This approach bases on two basic ideas, which are

- a combination of existing test tools in effective and cost-effective manner,
- re-usage of logged field data to cover by means of simulation the overall situation space.

Thus, the aim is to extend the AdaptIVe framework with the approaches that focus on the safety validation for the field in order to get to a comprehensive evaluation framework for automated driving functions.

Acknowledgments The research leading to these results has received funding from the European Commission Seventh Framework Program under the project AdaptIVe, grant agreement number 610428. The authors would like to thank all partners within AdaptIVe for their cooperation and valuable contribution. In particular the authors want to thank the partners of the Evaluation SP, namely BAST, BMW, CTAG, CRF, TNO and Lund University.

References

1. Rowsome F (1958) What it's like to drive an auto-pilot car. Popular Science Monthly, USA
2. NN Report to Congress, DARPA Prize Authority, DARPA, 2006
3. NN DARPA Urban challenge Media. DARPA urban challenge, http://archive.darpa.mil/grandchallenge/, 2014
4. Markoff J (2010) Google cars drive themselves. The New York Times

5. Ziegler J, Bender P, Schreiber M et al (2014) Making Bertha drive – an autonomous journey on a historic route. IEEE Intell Transp Syst Mag 6(2)
6. Ploeg J, Shladover S, Nijmeijer H et al (2012) Introduction to the special issue on the 2011 grand cooperative driving challenge. IEEE Trans Intell Transp Syst 13(3)
7. Etemad A, Bartels A (2014) A stepwise market introduction of automated driving. In: 21st ITS World Congress, Detroit
8. NN Taxonomy and Definitions for Terms Related to On-Road Automated Motor Vehicles, SAE document J3016, 2014
9. Knapp A (2015) Legal issues addressed in the EU funded AdaptIVe project. In: 22nd ITS World Congress, Bordeaux, 2015
10. Johansson E (2014) Human factors in vehicle automation - activities in the European project AdaptIVe. VRA Webinar
11. Dávila A (2012) Validation test plan and results. SARTRE, Deliverable D4.1 & 4.2
12. Beutner A, Pettersson S, Seglö F et al (2011) Optimized and validated demonstration vehicles. HAVEit, Deliverable D13.1
13. Deutchle S (2006) Das KONVOI-Projekt – Entwicklung und Untersuchung des Einsatzes von elektischen LKW-Konvois auf Autobahnen, 15. Aachner Kolloquium
14. NN Adoption of the Project Management Institute (PMI®) Standard A Guide to the Project Management Body of Knowledge (PMBOK® Guide)–4th edn. IEEE Guide, vol 2011
15. Karabatsou V, Pappas M, Elslande P et al (2006) A-priori evaluation of safety functions effectiveness – Methodologies, Trace, Deliverable D4.1.3
16. Fahrenkrog F, Zlocki A, Dobberstein J et al (2013) Impact assessment of developed application – overall interactive assessment, interactIVe, deliverable D7.5
17. Wilmink I, Janssen W, Jonkers E et al (2008) Impact assessment of intelligent vehicle safety systems. eIMPACT, Deliverable D4
18. Scholliers J, Hendirks F, Aust ML et al (2008) Project final report and recommendation for future assessments. PReVAL project, Deliverable DL IP D12/D16.4
19. Rodarius C, Duflis J, Fahrenkrog F, et al (2015) Test and evaluation plan. AdaptIVe Deliverable D7.1
20. Wachenfeld W, Winner H (2015) Die Freigabe des autonomen Fahrens, Autonomes Fahren, Technische, rechtliche und gesellschaftliche Aspekte. Springer Vieweg
21. Larsson P, Esberg I, van Noort M (2012) Test and evaluation plans. interactIVe Deliverable D7.4
22. Kessler C, Etemad A, Alessandretti G (2012) Final Report, euroFOT, Deliverable D11.3
23. Benmimoun M, Fahrenkrog F, Zlocki A, Eckstein L (2012) Detection and classification of critical events by means of vehicle data. ATZ Worldwide 114(10):60–66
24. Reichel M, Botsch M, Rauschecker R, Siedersberger KH et al (2010) Situation aspect modeling and classification using a scenario based random forest algorithm for convoy merging situations. In: 13th International IEEE Conference on Intelligent Transportation Systems, 2010, pp 360–366
25. Page Y, Fahrenkrog F, Fiorentino A et al (2015) A comprehensive and harmonized method for assessing the effectiveness of advanced driver assistance systems by virtual simulation: the PEARS initiative. In: 24th international technical conference on the enhanced safety of vehicles (ESV)
26. Fahrenkrog F, Zlocki A, Eckstein L (2014) Bewertung aktive Sicherheitssysteme vom Test zur Wirksamkeitsanalyse. Automobiltechnische Zeitschrift 01
27. Eckstein L, Zlocki A (2013) Safety potential of ADAS – combined methods for an effective evaluation, 23rd ESV
28. Zlocki A, Eckstein L, Fahrenkrog F (2015) Evaluation and sign-off methodology for automated vehicle systems based on relevant driving situations. In: 94th Annual TRB Meeting, Washington D.C.