

Controlling Risk for Highly Automated Transportation Systems Operating in Complex Open Environments

A white paper of the

SafeTRANS Closing the Gap Initiative¹

Inhalt

Recommendations	4
Summary	5
1. Motivation and Industrial Relevance	7
1.1 Motivation.....	7
1.2 Industrial Relevance	9
1.3 Structure of White Paper	10
1.4 Disclaimer.....	11
2. Overall Approach.....	14
3. Quality Metrics and Quality Guarantees	17
3.1. Determining Relevance of Environmental Objects.....	18
3.2. Bounding Uncertainty for Components in the Perception Chain	20
3.2.1. Absence of Adversarial Conditions	20
3.2.2. Accuracy of Perception for Individual Component.....	21
3.2.3. Emergence and Propagation of Uncertainty	21
3.3. Quantifying Perception Uncertainty Through Virtual Testing	22
3.3.1. Sensor and Environment Models	22
3.3.2. Modeling Sensor Fusion and Classifier Components.....	24
3.4. Quantifying the Overall Accuracy of the Perception Chain	25
4. Sensor Characterization and Sensor Modelling	26
4.1 Environment Simulation	26
4.1.1 Use Cases	26
4.1.2 Sensor modalities.....	26
4.1.3 Object and material properties	27
4.2 Radar	28

¹ For a complete list of authors and their affiliations, see list at the end of the document

4.2.1	Principles of Operation	29
4.2.2	Artifacts and Effects.....	29
4.2.3	Assessment of Radar Sensors	30
4.2.4	Modeling of Radar Sensors.....	31
4.2.5	Model verification/ validation	31
4.3	Lidar	34
4.3.1	Lidar Technologies	34
4.3.2	Principles of Operation	35
4.3.3	Characteristics.....	36
4.3.4	Artifacts.....	36
4.3.5	Modeling of Lidar Sensors	37
4.3.6	Lidar Sensor Modeling	37
4.3.7	Model Validation	39
4.4	Camera.....	40
4.4.1	Principles of Operation	40
4.4.2	Artifacts.....	40
4.4.3	Characterization of Video cameras.....	43
4.4.4	Models of video cameras.....	43
4.4.5	Model validation	43
4.5	Digital Maps	46
4.5.1	Principles of Operation	46
4.5.2	Artifacts.....	46
4.5.3	Environment Simulation based on Digital Maps	48
4.5.4	Research needs.....	48
5.	Sensor Fusion and Classification	49
5.1	Introduction	49
5.2	State of the Art.....	49
5.3	Safe Maneuver Execution in the presence of possibly incomplete of incorrect beliefs.....	52
5.4	A reference architecture of the perception chain supporting uncertainty propagation for relevant objects.....	56
5.5	Bounding Uncertainty for AI-based Classifier components.....	59
5.6	Propagating Uncertainty guarantees in sensor fusion.....	61
6.	Credible Co-Simulation and Model Composition	67
7.	Verification and Validation Methods and Processes.....	73
7.1	Supporting Continuous Risk Management through Traceable Verification and Validation Processes	74
7.2	Validation and Verification of ADS Perception Systems	75
7.3	Tool-based Continuous Software Development.....	77
7.4	Test Environment Credibility and Test Case Allocation	79

7.5	Challenges in Decompositional Verification and Validation.....	81
7.6	Future Directions and Research Needs.....	81
8.	Architectural Requirements	83
8.1	Related Work	83
8.2	Research Questions and Possible Ways Forward	84
9.	Putting it all together: Deriving Safety Assurance Cases for highly automated systems exploiting digital twins	86
9.1	Motivation / Need for assurance cases	86
9.2	State-of-the-art assurance case frameworks.....	86
9.3	Research needs	89
	References.....	92
	Authors.....	104
	Annex 1: References in Tables 1, 2, and 3 of Section 5.2 from [FJG+2020]	106
	Annex 2: Glossary	110

Recommendations

The SafeTRANS Expert group has in this presented White Paper assessed the potentials for reaching a key cornerstone towards safety and acceptability for highly automated vehicles, in providing a holistic approach towards bounding uncertainty in the perception chain.

To reach the strategic objective of the German Government to become a leading nation in innovation for automated and connected driving, we propose the following recommendations:

Research Funding

- This milestone can only be achieved by integrating expertise in sensor technologies, AI, digital twins, V&V, safety methods and processes with industrial experts in building highly automated vehicles in one R&D strategy.
- A seamless extension of, and continuous integration into, the achievements from preceding large R&D formats is another pre-requisite.
- Cross-sectorial benefits can be exploited by integrating multiple application domains (road vehicles: Cars, robo-taxis, trucks, rail vehicles: Trains, Offroad vehicles: Vessels, tractors)

Reference suites of standards for testing ADAS/AD and SDV

- To assure high quality of the perception chain, we consider it mandatory to establish open suites of standards documenting the current understanding of all types of objects in the environment of AVs which must be recognized by AV.
- This suites must cover the complete range of environment models such as demanded for sensor components, up to and including those objects which must be observed in world models of the AV
- This test suite must include agreed models of dynamics of all traffic participants
- Processes must be put in place to regularly update this test suite based on incidents and accidents related to mis-classifications and mis-assessments during operation of AVs and subsequently integrated into the operating systems of software-defined vehicles.

Sharing models and data

- The investments required for achieving high-quality sensor models demand the establishment of instruments allowing the sharing of data and enabling interoperability while protecting individual IP and ensuring compliance to anti-trust regulations.
- It is suggested, that the GAIA-X Initiative is involved in providing solutions to support such data exchange platforms and interoperability of models.

Regulations

- Type certification must be understood as a continuous process allowing for over-the-air updates of software components of the perception chain mandated from learning from in-field incidents and accidents.
- The development of open standards and global harmonization led by German industry are key to success.

Summary

This paper provides an approach for controlling the level of risk when operating highly automated transportation systems like cars, trains and similar. Such systems replace human perception and decision-making by employing highly sophisticated solutions based on electronics, IT, and AI. Such systems have demonstrated the potential for building highly automated vehicles, but, as of today, encounter challenges in correctly understanding the extremely complex open contexts, into which such vehicles could be deployed. Its key focus is on **bounding the risks stemming from uncertainty in the perception of the environment**.

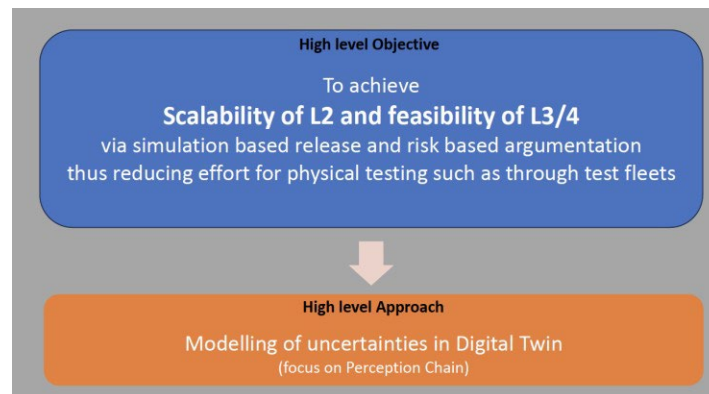


Figure 1 High Level Objectives and Approach of the White Paper

The approach presented in this paper is building on and elaborating major results achieved in previous projects of the VDA strategic initiative on highly autonomous driving (Pegasus, VVM, SetLevel, KI-Absicherung), the Joint Undertaking ECSEL (Enables-S3), the Vivid (Vivaldi und DIVP) Consortium and the safe.trAIIn Project. We collectively refer to these as *background projects*.

To bound the risks stemming from uncertainty in the perception chain, we follow the high-level flow depicted in Figure 2 below.

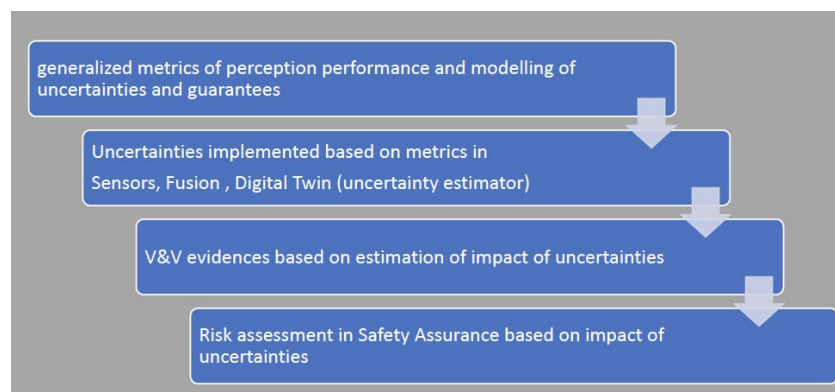


Figure 2 High Level Flow

Specifically, we aim to synergistically combine advanced approaches in:

1. Building and validating "highly accurate" sensor models in field trials, including modelling of artefacts causing distortion of perceptions.
2. Building and validating "highly accurate" environment models of complex open world contexts, including environment conditions influencing distortion of perceptions.
3. Building and validating "high accuracy" digital twins of the perception chain to establish (situation dependent) statistical guarantees for bounding the level of uncertainty in the

- perception of the environment, taking into account dynamic reconfiguration and degradation.
4. Developing metrics which take safety relevance explicitly into account in giving precise definition to the required degree of accuracy for items 1-3 above, providing the basis for safety assurance cases.
 5. System architectures, which allow to dynamically tune the degree of precision of perception
 - a. by determining the criticality of objects of the ego system's environment
 - b. based on prediction of the short-term evolution of this environment.
 - c. The system architecture combines this top-down attention focus with a bottom-up propagation of the current level of uncertainty
 - d. allowing for optimal resource usage of dynamically integrating sensors and classifiers and sensor fusion components
 - e. so as to compensate weaknesses of sensors operating in currently distorting environments by high quality measurements of sensors operating in favorable environmental conditions
 - f. This allows to inductively derive formal assurance guarantees in achieving the required degree of precision needed for safety case arguments.
 6. On-line monitoring of all conditions critically influencing the achievable degree of precision, such as conditions causing distortions of sensors and compliance to ODDs
 7. System architectures offering safe degradation when such monitors signal risks in distortion of perception, or lack of compliance to ODDs, or alarms raised by health-state monitoring of system components.

1. Motivation and Industrial Relevance

1.1 Motivation

Automated vehicles are one of the most promising solutions for the greatest challenges of modern mobility: emission reduction, effective time management and comfort for modern citizens. This paper presents an approach to provide (probabilistic) guarantees on the maximum level of uncertainty in the perception chain of highly automated transportation systems, building on significant results achieved in previous projects of the VDA strategic initiative on highly autonomous driving (Pegasus, VVM, SetLevel, KI-Absicherung), the Joint Undertaking ECSEL (Enables-S3), the Vivid (Vivaldi und DIVP) Consortium and the safe.trAIIn Project. These projects have set the stage to enable verification and validation methods for highly automated transportation systems² operating in complex environments.

The *VVM project*³ has established in combination with the SetLevel project a holistic approach for building safety assurance cases for highly automated vehicles. Figure 3 below highlights the key addressed concepts for a methodology for establishing safety of level 4/5 vehicles.

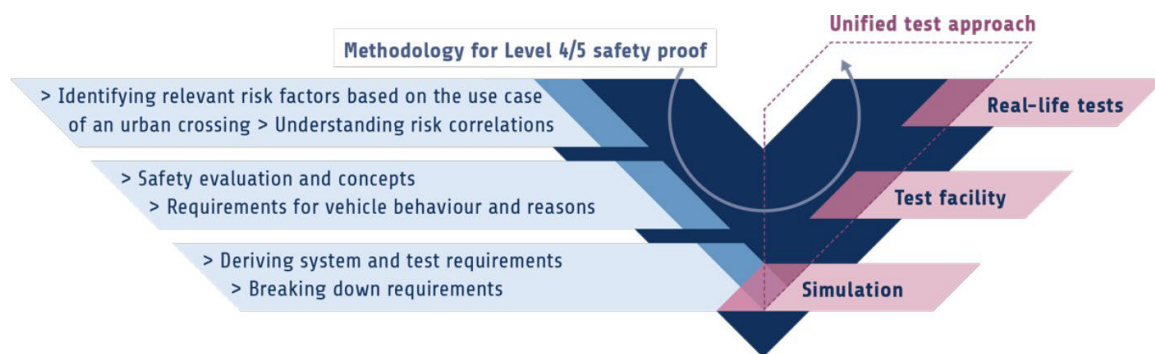


Figure 3 Key research areas addressed by the VVM project

It developed

- A Methodology for efficient control of the test area
- A Validation methodology across all system levels
- An approach for the generation of a continuous testing sequence across all test platforms from simulation to real world driving

The *Set Level Project*⁴ has been supporting V&V of highly automated system by providing a credible simulation platform and supportive processes methods and tools supporting the key industry trend towards digitalization and virtualization in product development and therefore virtualized validation and release.

The Vivaldi project⁴ has been working on the assessment of virtual validation methods for autonomous driving functions with a focus on sensors. Vivaldi has teamed up with the Japanese DIVP (Driving Intelligent Validation Platform)⁵ project in order to synergize and collaborate under the label VIVID (German Japan Joint Virtual Validation Methodology for Intelligent Driving Systems)⁶. The key objective of both projects is the design and implementation of a virtual validation tool chain, reaching from SiL- (Software-in-the-Loop) to OTA/ViL-methods (Over-The-Air, Vehicle-in-the-Loop), connecting software-based traffic and sensor simulations with propagation modelling and over-the-air hardware-in-the-loop testing in virtual environments.

² SAE levels 2 and higher in the automotive domain, see [SAE]

³ <https://www.vvm-projekt.de/en/concept>

⁴ <https://setlevel.de/en>

The European research project *ENABLE-S3 (European Initiative to Enable Validation for Highly Automated Safe and Secure Systems)*⁵, funded through the ECSEL JU programme has established a comprehensive platform for the cost-effective validation and verification of autonomous and highly automated vehicles, trains, tractors, ships, aircraft, satellites and medical examination equipment. ENABLE-S3 has been following a use-case driven approach: requirements for the project are coming from industrial use cases within the 6 industrial domains, and each technical solution is required by a specific use case.

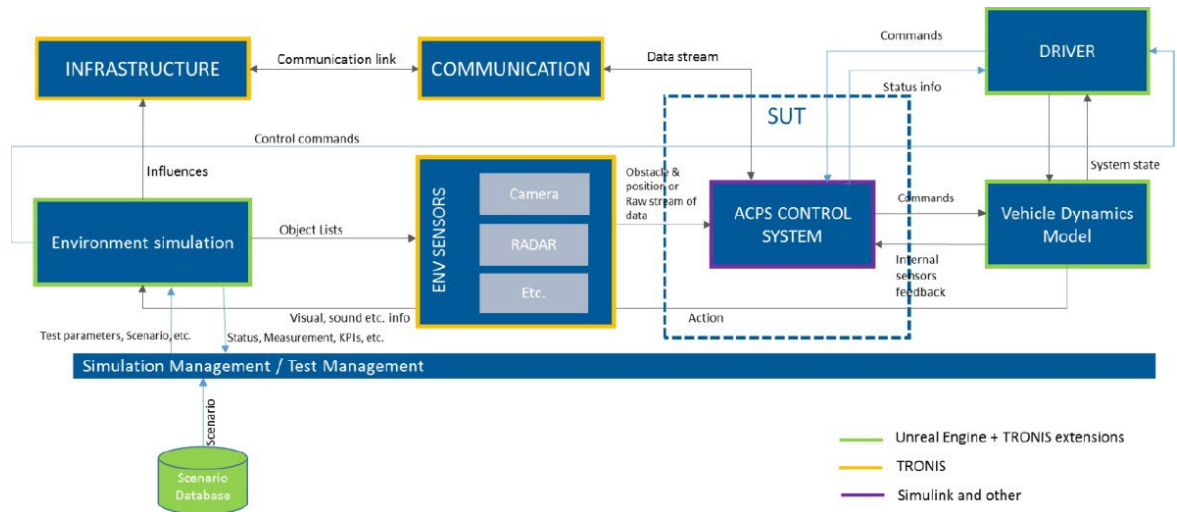


Figure 4 A sample instantiation of the EnableS3 integration platform building on Tronis

The basic architecture developed within ENABLE-S3 is nowadays used by leading companies, e.g., customers of AVL, to design and test alternative solutions for automated driving and its multifaced demands within the development process.

In the railway industry solutions for completely driverless and unattended operation of trains have been successfully established on the market and in operations. Until now, however, these systems have been operating exclusively in controlled and closed environments, such as subway tunnels. The safe.trAIIn project, funded by the BMWK and running since January 2022, is focusing on applying this technology for use in regional trains. Such trains operate in more open environments in which it is necessary, in particular, to reliably recognize obstructions – such as people on the lines as well as fallen trees or mudslides on the tracks, etc.

The project goals are to perform integrated development of testing standards and of methods for using AI to automate rail transportation and to use example applications to verify the suitability of test standards. Focal points here will be on AI-based methods for driverless regional trains, approval-relevant validation of the product safety of the AI components, as well as testing processes and testing methods.

While certain cross project synergies have been achieved through partners involved in all projects (such as the PerCOLLECT & CEPRA methods on hazard analysis of sensor systems [LRS+2021] and derivation of resulting requirements for hazard containment in testing sensor systems), *this white paper brings together for the first time the joint expertise of the above projects to close the remaining gaps towards safety assurance of highly automated transportation systems*. Key challenges remain in their compromised maneuverability and performance in bad weather conditions, such as rain, snow, dust storms, or fog, which can compromise vision and range measurements (degradation of the visibility distance). In such conditions, the performance of most current active and passive sensors is significantly compromised, which in turn can lead to erroneous and even misleading outputs of the perception chain. The current achievements still suffer from **significant limitations in realism of**

⁵ <https://cordis.europa.eu/project/id/692455/results>

models and simulations in the perception chain, such as anomalies based on noise, EMI, bad lighting (low sun angle, specular reflections), poor target resolution, and issues in dealing with shadows, foreign objects on road, reflections, glare, worn or occluded signs and markings, the contextual diversity of traffic interaction, etc. Even if we were able to find models which faithfully represent all objects of the physical system which is modelled, such models would be way too complex to allow large-scale cloud-based system testing. This White Paper performs a gap analysis: what is missing to achieve safety assurance cases for highly autonomous systems meeting standards such as ISO 26262 and ISO 21448 under affordability constraint maximize virtually generated evidences to reduce physical testing to demonstration of faithfulness of virtual models, and proposes research directions to close these gaps. It proposes a divide and conquer approach in generating bounds on uncertainties, in propagating quality guarantees bottom up along the different stages of the perception chain, thus paving the way for reducing the complexity of models. Jointly, the proposed research will allow to **provide (probabilistic) guarantees on the maximum level of uncertainty in the perception chain of highly automated transportation systems**. These guarantees will subsume **guarantees on level of confidence in existence and classification of objects** in the world model determined by perception chain, as well as **guarantees on precision of measurements for all physical attributes of all objects in the world model**. Jointly, these will provide an approach to construct **assurance cases for the safe execution of maneuvers of highly automated transportation systems within given ODD**, being based on **guarantees** of the *quality of perception of all entities in the world model* provided by the perception chain. In identifying what properties and characteristic of which element of the perception chain can be established on what abstraction level, we strive for reducing the complexity of models integrated in digital twins to a level amenable for large scale virtual driving.

1.2 Industrial Relevance

Sophisticated sensor systems are now an integral part of consumer vehicles, and 92% of new cars sold in the US include some Advanced Driver Assistance Systems (ADAS) such as autonomous emergency braking systems, adaptive cruise control, park assist, and lane departure warning, which are defined as partial automation (Level 2 autonomy) under Society of Automotive Engineers [SAE2018]. The Global Advanced Driver Assistance Systems Market size was valued at USD 20.74 billion in 2020 and is estimated to grow USD 48.37 billion by 2028, at a CAGR of approximately 11.6% between 2021 and 2028, according to a recent research study published by Zion Market Research⁶. The first Level 3 features have recently been approved for use in commercial vehicles, e.g., both Honda and Mercedes have received regulatory approval for Advanced Lane Keeping System (ALKS) [UN2021], with the Drive Pilot of Mercedes expected to be available in the market by the end of this year in Germany within its S-Class. According to a market study by Fortune Business Insights⁷, the increasing demand for road safety is driving the growth of the Advanced Driver Assistance Systems. The perception system forms an integral part of such ADAS systems. It is clearly safety critical at Level 3 as, for example, failure to correctly identify lane markings could lead an ALKS-equipped vehicle to stray from the current lane. While aiming to support construction of safety assurance cases for Level 3 and Level 4 systems, the approach of the project to heavily rely on highly accurate digital twins, thus allowing to reduce the amount of physical field testing, is immediately relevant for Level 2 applications as well because of the expected reduction of V&V costs. Additionally, even for Level 2 systems, reducing the risk of misperceptions⁸ through the proposed quality assurance measures is of immediate market relevance. Achieving a significant part of assuring the quality of perception through digital twins is thus not only an enabler for achieving the necessary coverage levels required for L3/L4 safety assurance cases, but

⁶ <https://www.globenewswire.com/en/news-release/2023/05/29/2677668/0/en/Latest-Global-Advanced-Driver-Assistance-Systems-ADAS-Market-Size-Share-Worth-USD-48-37-Billion-by-2028-at-a-11-6-CAGR-Zion-Market-Research-Industrial-Trends-Report-Analysis-Player.html>, published May 29 2023

⁷ <https://www.fortunebusinessinsights.com/industry-reports/adas-market-101897>

⁸ potentially causing accidents, e.g., rear-end collisions caused by unwarranted emergency braking

also contributes to reducing costs for V&V of L2 and L3 systems. Moreover, given the regional differences in regulatory approaches, it strengthens the competitiveness of German automotive industries in being able to achieve a large part of the learning curve towards highly automated driving not only by test fleets but by driving in high-accuracy digital models of the environment and the perception system. Achieving this level of precision in such digital twins demands a concerted effort of industry and research, integrating top expertise in modelling and validation and verification across all levels of the perception chain. Integrating these results with the overall system verification approaches including as well trajectory planning and maneuver execution as provided by the VVM and SetLevel projects thus closes the gap towards achieving safety assurance cases for highly automated vehicles.

In automated trains there are already existing commercial offerings for fully automated trains, such as the Nuremberg Metro which started operations in 2008⁹, which however operates in a highly constrained Operational Design Domain. Other existing market offerings include assistance functions for lower Grades of Automation (GoA). However, like in the automotive domain, an unsolved challenge is the availability of robust perception systems that enable highly-automated operation in unconstrained environments. Nevertheless, due to legal regulations and the underlying trackside infrastructure (e.g., train protection systems) the kind of objects which have to be reliably detected differ from those required in the automotive domain. For instance, it can be assumed that no persons and no other trains occupy the trackway, while in contrast all obstacles which could lead to a derailment of the train have to be reliably detected. Moreover, as such critical incidents are highly rare events, it is unrealistic to generate sufficient amounts real-world training data, e.g., for the training of AI based perception systems, only by live recordings of fleet data. Therefore, suitable approaches for generating synthetic data are needed.



Figure 5 Automation in the Railway Domain

1.3 Structure of White Paper

⁹ <https://www.railway-technology.com/projects/neuremburgautobahn/>

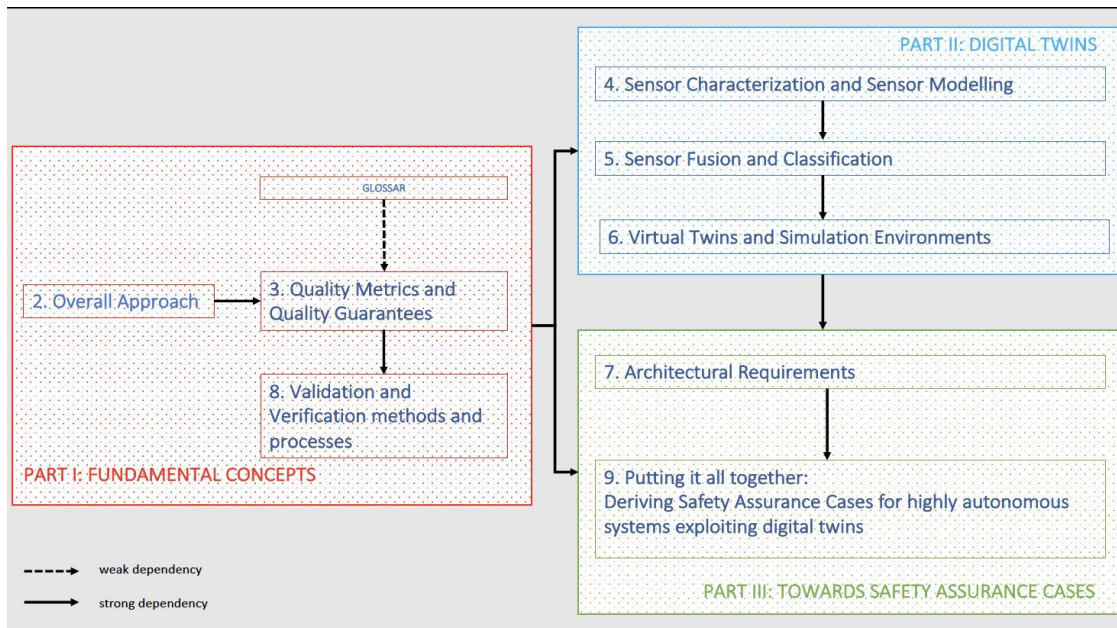


Figure 6 Structure of the White Paper

This paper is organized as follows.

Part I *Fundamental Concepts* contains three sections. We describe the overall approach towards bounding uncertainties in Section 2. Section 3 discusses the type of Quality Metrics and Quality Guarantees required to provide a rigid formal argumentation bases for bounding the level of risks stemming from uncertainties. The Verification and Validation methods required to demonstrate these are described in Section 8, while requiring customization for their specific use-cases described in Parts II and Part III.

Part II focusses on the *Digital Twin*. Section 4 discusses, for all relevant classes of sensors, the construction of highly accurate sensor models and sensor environment models, and refers to approaches developed in Section 8 for verifying the type of quality guarantees described in Section 2. Section 5 discusses how the various techniques for sensor fusion used in industry can be enriched in order to propagate and improve quality guarantees from sensors by sensor fusion, and demonstrating these with instantiations of verification and validation methods described in Section 8. Section 6 discusses methods to fuse such quality guarantees to achieve high confidence world models of the environment of the system as an interface to trajectory planning, and discusses requirements on a faithful “credible” open simulation framework for digital twins allowing to incorporate the sensor models of Section 4 and sensor fusion components, such that simulations executed in the digital twin meet quality requirements as defined in Section 3 on the level of accuracy in matching the physical closed loop systems of the perception chain and the environment.

Part III *Towards Safety Assurance Cases* discusses in its section 7 architectural requirements to support bounding the uncertainty in the perception chain. The key Section 9 shows how the combined evidences gained from field testing and virtual testing in the highly accurate digital twin can be combined using the V&V techniques of Section 8 into a Safety Assurance Cases for highly automated vehicles.

1.4 Disclaimer

This paper focusses on risks coming from the uncertainty of the perception of the environment of the ego-system operating in complex environments. While bounding this risk is a necessary precondition for assuring safety of highly automated vehicles, it is by no means a sufficient condition. Specifically,

in this paper we abstract from all of the following sources all impacting overall safety, by assuming a number of idealizations for system components, as outlined below.

1. The execution platform in the deployed system differs from the execution platform used for the digital twin
2. HW/SW failures
3. Systematic causes of risks in the prediction engine
4. Incomplete characterization of environment
5. Faults of maneuver decision layer or Faults of maneuver execution layer
6. We now list for each of these the idealization/abstraction from these aspects used in this paper.

Risk source 1: The execution platform in the deployed system differs from the execution platform used for the digital twin

We assume an idealized execution platform meeting the so-called synchrony hypothesis underlying the synchronous programming paradigm (see e.g., [BER2004]), which essentially states that the underlying hardware is so powerful, that all timeliness constraints and causality constraints of all tasks are met. See e.g. [TPB+2008] for an example of how real architectures can be built to achieve this. In general, it requires a separate verification effort to demonstrate that all non-functional requirements such as regarding real-time requirements and causality constraints of a given application are actually met in the target architecture.

Risk source 2: HW/SW failures

For traditional functional safety applications, the approach and risk management of programmable systems are well defined within the framework standards that organize the design, development and deployment of a safety system within a given operational domain/sector. To achieve functional safety in the presence of HW/SW failures in safety-critical applications, rigorous processes are mandated to ensure necessary levels of fault tolerance, redundancy or reliability needed to achieve safety goals. Systems must be designed, verified, validated, built, and operated in a way that minimizes the risk of harm or more formally achieves a demonstratable tolerable risk, see e.g. (ISO/IEC 610508 Series, and ISO 26262). We rely on such processes in assuming in this paper, that all components of the execution platform are free from HW/SW failures.

Risk source 3: Systematic causes of risks in the prediction engine

In this paper we propose a methodology for designing the perception chain of the ego-system to constantly compute high accuracy models of the environment of the ego-systems called *Lagebild*, with guaranteed bounds on the uncertainty on state and physical attributes of all relevant objects in the environment of the ego system (see the following Chapter 2 for a first summary of the overall approach used to achieve this, and the notion of “relevance”). Given such a *Lagebild* at time t , it is the task of the prediction engine to assess the potential future evolution of the state and physical characteristics of all relevant objects in the environment of the ego-system in order to determine next possible maneuvers of the ego-vehicle in order to meet the current set of goals of the ego-system.

Even if we assume, that classification of such objects and their state meets such quality criteria, there are inherent causes for mispredictions of the future evolution of such systems: models of the dynamics of such systems are by necessity only of statistical nature, hence any individual relevant object occurring in the *Lagebild* might deviate within the distributions coming with such models. Moreover, such models are not able to reflect internal invisible mode-changes of such a system; the perception chain must be designed to perceive all cues of such systems possibly triggering a mode-change, hence leading to a potentially radically different dynamic model for the anticipated behavior of this system. The identification of all such cues for all types of mode-changes of the behavior of environment systems is out of the scope of this paper. In this paper we assume, that such an analysis has been carried out, and that dynamic models of objects in the environment are parametrized to cater for mode-changes whenever such cues have been perceived. We assume that all such relevant cues are part of the annotation of relevant objects in the *Lagebild*, and thus also labeled with the degree of confidence with which these cues have been identified.

Risk source 4: Incomplete characterization of environment

This paper assumes, that there is an agreed current state of understanding about what objects in the environment of the ego-system are potentially relevant for safe operation of the ego-system.

This assumption is extremely difficult to meet.

Taking the automotive domain as an example, this assumption goes far beyond the current level of pre-standardization which is currently developed as part of the projects of the *VDA Leitinitiative for autonomous driving*, since it must among others encompass all aspects of the environment impacting reflection and absorption properties of active sensors.

The overall underlying challenge of operating in an open context has been addressed by the system-safety community with by now well-established heuristics, learning from previous incidents and accidents, and turning this knowledge into guiding questions for system design, demanding systems engineers to study for any aspect of the environment found to be causally relevant for an accident or incident, whether appropriate measures have been taken to observe such aspects.

In this paper, we assume, that there is a documented and eventually standardized state of the art of what entities in the environment of the ego-system must be observable, and assume, that deployed systems are designed to not only observe all such aspects of the environment, but also allow updates in the field to accommodate for any changes in standards as to what must be observable.

Risk source 5: Faults of maneuver decision layer or Faults of maneuver execution layer

In this paper we completely abstract from the implementation of the maneuver decision layer and the maneuver execution layer of the ego system and assume these to be free of faults.

2. Overall Approach

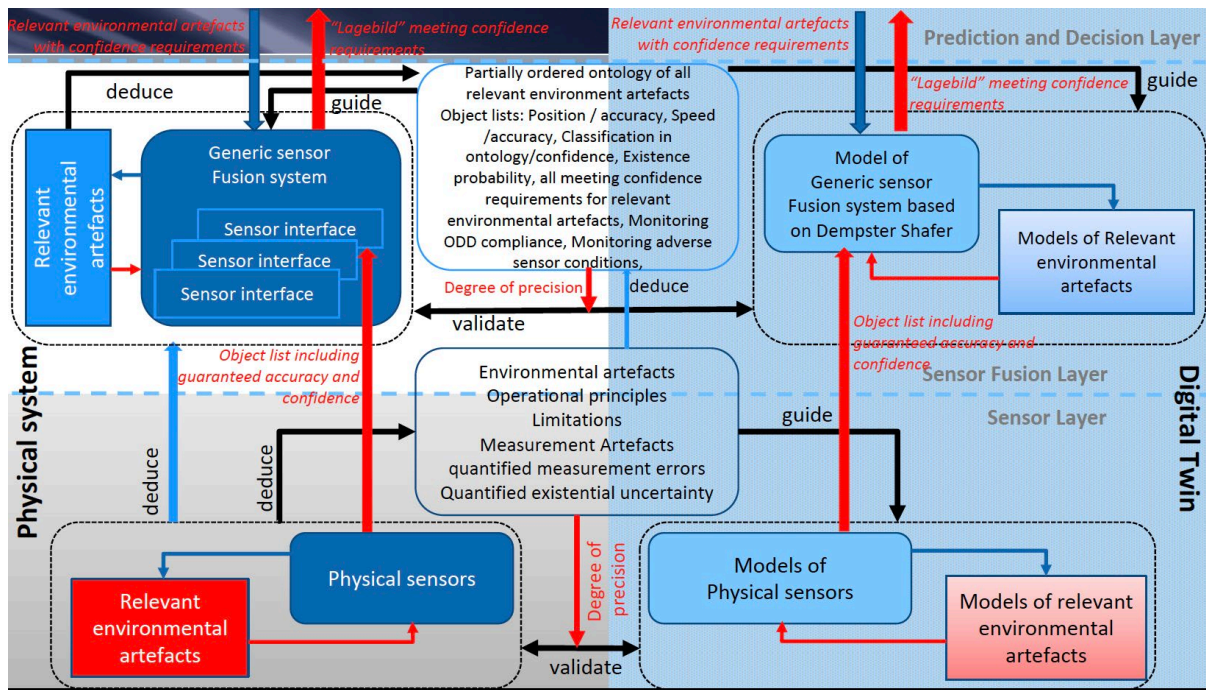


Figure 7

Bounding the risks stemming from uncertainty in the perception of complex open environments of highly autonomous transportation systems relies on a layered and segmented structure, consisting of the real physical system and environment and its digital twin, both reaching from the component level of the physical sensors to the level of sensor fusion, and both considering relevant environmental artefacts, eventually leading to a world model ("Lagebild") meeting given confidence requirements. This approach rests on a notion of what we call "sufficiently perfect components" of the perception chain, which allow to inductively derive such guaranteed bounds on the maximum level of uncertainty, provided the system is currently operating in sensor-specific specified environmental conditions, and in well-defined operational design domains (ODD).

The induction basis will be provided by "sufficiently perfect sensors". Such sensors come with a characterization of their behavior under adverse environmental conditions known to be detrimental to such guarantees in an intrinsic sensor-specific way, such as type and intensity levels of precipitation (e.g., fog for lidar, rain or snowfall for radar), illumination (e.g., straylight or backlight for camera and lidar, poor reflectivity of objects for lidar), or multipath propagation (e.g., reflection, refraction, or diffraction for all sensor modalities).

We will use real-field tests to generate scenario-based reference data for virtual sensor models demonstrating not only sufficiently precise processing of raw data in non-adverse conditions, but which are also additionally able to demonstrate the same degradation effects as real sensors. This will allow us to quantitatively assess the level of uncertainty not only based on field measurements, but using large test sets of highly reproducible, adjustable and scalable environmental conditions in digital twins of the sensors and their environment, continuously validated against the reference data. In addition to these intrinsic uncertainties resulting from the type and performance of the sensor and its integration in the perception chain, we will also characterize extrinsic uncertainties from environmental factors which are partly controllable by the ego system, such as analyzing the level of precision and uncertainty of dynamic attributes like position, velocity, direction, as well as current vibration levels etc. We also include "sufficiently perfect digital maps" with precise information of geometrical arrangements and material properties of the environment suitable for each sensor

modality as anchoring components in the perception chain.

For each type of output of a sensor, the probabilistic guarantees given under non-adverse conditions and known levels of controllable disturbances of the ego vehicle will include for each individual object in a given dynamic situation of the representation at that sensor output

- a quantification of existential uncertainty,
- a quantification of the confidence categorization of the type of artefact, if applicable, and
- a quantification of the distributions of imprecision of physical attributes of such artefacts observable by this particular sensor system at this particular interface.

We collectively refer to these guarantees as *guarantees for bounding uncertainty*.

We propagate such guarantees along the perception chain by requiring what we call “*sufficiently perfect sensor fusion components*” and “*sufficiently perfect classifier components*”. This idea of propagating uncertainties across the perception chain was first mentioned in [STE2016]¹⁰ and recently in [PPF2023]. It was implicitly also anchored in [MDM2010], in that sensor components were required to be enriched with quality attributes as a basis for Dempster-Shafer based sensor fusion. The meta-requirements for such components demand that each such component comes again with a characterization of adverse environmental conditions and allowed ODDs. For example, a classification component can only be required to provide guarantees for bounding uncertainty if the actual environment of the ego system is matching the characteristics in the data used for training classifier components with respect to types of objects, distribution of objects, and (if applicable) dynamic properties of objects in the analyzed sensor stream. [SMH+2022] demonstrates an approach bounding uncertainty in neural networks. For sensor fusion components, adverse conditions will be derived dynamically. Specifically, all inputs must be decorated by a characterization of those sensor-specific relevant environmental conditions under which they were collected, time stamps of raw data used, the component type delivering this input, and the adverse conditions of this component type. As proposed in [MDM2010], we will use approaches akin to Dempster-Shafer adapted to the guarantees for bounding uncertainty, to compute the maximal probabilistic guarantees for uncertainty for output streams of such components, and determine the adverse conditions by conjoining adverse conditions of components providing input streams with high relevance in strengthening guarantees for bounding uncertainty. At the highest level of the perception chain, sensor fusion components provide what is often called a “*world model*”. This is comprising all artefacts in the environment of the ego system relevant for maneuver decisions, with guarantees for bounded uncertainty. Such guarantees can be further strengthened by exchanging such world models with neighboring systems or information provided by infrastructure components. Jointly, we can thus derive for each stage in the perception chain the joint (as various objects may induce the same driving decision and individual uncertainties can thus be amortized) level of uncertainty for relevant environmental entities¹¹ for maneuver decisions of the ego system, such as those provided from the prediction and decision layer.

We combine this approach with what is referred to in the literature as attention driven perception, which relies on a quantitative assessment of the dynamic evolution of the ego-systems environment to determine, which objects and entities and which of their physical attributes are critical for guaranteeing safe evolution of the dynamics of the ego-vehicle in the currently perceived world-model. This attention focus allows to dynamically configure components of the perception chain maximizing quality of perception for these critically relevant environment observations, while at the same time optimizing resource usage supporting the computation of such guarantees for bounding uncertainty, as initially proposed in [HMG+2023].

¹⁰ We note, however, that the approach of Stellet does not attempt to model the effect of adverse conditions (see e.g., p.27: „Due to the high complexity of the interplay between lighting conditions, object surface and algorithm, this effect is not modelled explicitly but subsumed by a fixed percentage of valid measurements.”)

¹¹ See e.g., [HMG+2023] for a definition of „relevant“

To formally establish both the quality of models and verify the guarantees for bounding uncertainty, we will build on the results of the background projects to extend these methods to the type of quality guarantees required by this approach.

In consequence, this overall approach is further detailed according to the following building elements:

- Quality metrics and quality guarantees,
- Sensor characterization and sensor modelling,
- Sensor fusion and classification,
- Virtual twins and simulation environments
- Architectural requirements
- Validation and verification methods and processes, and the
- Derivation of safety assurance cases.

3. Quality Metrics and Quality Guarantees

Highly automated vehicles (HAV) promise safer and more efficient mobility solutions. However, their successful deployment hinges on their ability to navigate and make decisions in the face of complex, dynamic, and often unpredictable real-world environments. When operating in these open contexts, HAVs must contend with many challenges, especially perception uncertainty. Any incomplete or incorrect information can lead to potentially hazardous outcomes. As such, understanding and managing perception uncertainty is paramount to ensuring the overall safety and reliability of HAVs.

This chapter investigates the representation of perception uncertainty within sensor models, environment models, and digital twins of the complete perception chains, to define quality metrics and quality guarantees as a formal foundation for bounding the risks posed by perception uncertainties. In general, uncertainty can be categorized as follows (see [\[KD2009\]](#)):

- **Epistemic Uncertainty:** This type of uncertainty arises from incomplete knowledge or information gaps. It is often reducible through further research, data collection, or improved understanding. Epistemic uncertainty is associated with known unknowns and can be reduced with more data or improved models.
- **Aleatory Uncertainty:** Also known as stochastic uncertainty, aleatory uncertainty is inherent randomness or variability in a system. It is often irreducible and represents the inherent unpredictability of certain events or processes.

In assessing the safety of HAVs, it is important to recognize that the evaluation cannot be confined solely to the perception chain. Hazards arising from imperfect perception only manifest when we consider the entire ego vehicle within its dynamic environment. These hazards often exhibit intricate interdependencies with the vehicle's planning algorithms and interactions with surrounding traffic. In the forthcoming discussion, we will introduce a mathematical framework designed to bound uncertainty comprehensively. This framework serves as a robust foundation upon which both the planner and vehicle control systems can operate, ensuring the safe navigation of HAVs in complex real-world scenarios.

In the following, we first outline a mathematical setting for quantifying perception uncertainty that considers the various influencing factors impacting perception quality. As a starting point, we formalize a meta requirement for the quality of the perception chain in a probabilistic linear-time temporal logic, with observables defined by valuations of all attributes of all instances of classes within the electronic horizon of the ego, over a typed first-order signature induced by the types of attributes in the ontology. Ideally, for each point in time, the ground truth of all relevant objects (note that relevance is a dynamic property and cannot be fixed at design time) in ego's electronic horizon, in particular position and speed of surrounding traffic participants, road- and weather conditions, coincide exactly with the ego's beliefs about these objects.

We must relax this unachievable ideal by considering standard measurement errors, and allowing classifications to be vague, as long as they are correct with respect to the ordering relation in the ontology (i. e., the ground truth classification is a specialization of the believed classification). While misclassifications and misperceptions will occur, we want these to be bounded by the societally accepted level of risk¹². We assume for each relevant object a a metric d_a that measures the distance between ground truth and the beliefs ego has about a . Additionally, we assume a safe level of measurement tolerance¹³ δ_a . We want these to be consistent almost always for a minimal time period Δ sufficient to take safe maneuver decisions for the HAV. Comparing this to the acceptable level of misperception (derived from the overall risk of the HAV operation) P_M leads us to the following formal

¹² It should be noted here, that perception uncertainty has a strong impact on the overall safety of HAVs. However, it is not feasible to assess the risk to passengers or other traffic participant without considering downstream functions such as the planner or existing safety mechanisms. For that reason, we will avoid using the term risk throughout this chapter and rather argue about the probability of misperception.

¹³ For realization of this performance metrics from ISO PAS 8800 can be used. See also [Chapter 9](#).

requirement on the confidence of the perception chain of the HAV vehicle ego.

$$\begin{aligned} \text{Safe_perception}(\text{ego}) &= \forall a \in \text{observables}(TS(\text{ego})) : \\ \square &\left(\text{relevant}(a) \Rightarrow P \left(\neg \left(\square_{\leq \Delta} (d_a(a, \text{belief}(a)) \leq \delta_a) \right) \right) < P_M \right) \end{aligned}$$

For now, we state that a perception is said to be *precise* if for all relevant objects in the proximity of the *ego* vehicle holds that the probability of misclassification or misperception of these relevant objects is bound by a certain threshold P_M for a minimal time horizon Δ . Here, one might think that a misclassification or misperception can be indicated by an object-dependent metric d_a exceeding an also object-dependent tolerance level δ_a .

The approach to meet this meta requirement laid out in this white paper relies on constructing the necessary building blocks (sensor models, ontologies, interface formats, etc.) to achieve the following requirements:

1. It must be possible to state which environmental objects are relevant at a given point in time, based on the current traffic situation as well as a set of planned maneuvers (see [section 3.1](#)).
2. It must be possible to bound the contributions of each component in the perception chain to the overall uncertainty (see [section 3.2](#)).
 - a. We must be able to specify for each component in the perception the conditions under which this component can be “trusted”) and complement this by continuous monitoring of trust conditions (i. e., absence of adversarial conditions).
 - b. We must be able to quantify the accuracy of perception for each component in the perception chain.
 - c. We must be able to describe the emergence and propagation of uncertainty through the system.
3. It must be possible to quantify the accuracy of the perception of a single component via virtual testing (see [section 3.3](#)).
 - a. We must be able to quantify the degree of precision of sensor models and environment models in closed-loop simulations.
 - b. We must be able to quantify the impact of misclassifications or misperception of every relevant object for each component on the overall uncertainty.
4. It must be possible to compose models from requirement 3 into a full digital twin of the perception chain and the relevant environment in order to quantify the overall perception uncertainty (see [section 3.4](#)).

In the following sections, we will discuss for each of the building blocks which approaches, methods, and tools from the literature can be used to guarantee the respective properties.

3.1. Determining Relevance of Environmental Objects

A way to measure the **relevance** of an element of a given scenario to the *ego*, requires an **ontology** (or, at least, a taxonomy) containing classes all safety-relevant objects. Such an ontology has to be created at design time, starting early in the design process. For example, SOTIF triggering conditions¹⁴, which lead to hazardous behavior of the *ego*, can be identified and incorporated into the ontology. Then, all subsequent methods have a foundation to access such elements. For example, the system can initiate certain safety mechanisms during run-time if it is confronted with such triggering conditions. A sketch of how to iteratively construct a “good enough” ontology based on the outputs of safety engineering procedures in the design and development phases is outlined by Stierand et al. [[SWH2023](#)]. Westhofen et al. [[WNB2022](#)] have sketched in more detail how in one concrete step – the

¹⁴ An example for such a triggering condition could be the misclassification of a person riding an e-scooter as a pedestrian provoking the *ego* to assume a completely different range of motion and potentially selecting hazardous driving maneuvers.

identification of abstract classes of hazards – an ontology can be iteratively completed ([[WNB2022](#)], Section IV-A).

The evaluation of object relevance in dynamic environments can be based on the following parameters:

- distance to the object
- relative velocity
- typical criticality metrics (e. g., Time-To-Collision (TTC))
- objects classification (e. g., vulnerable road users (VRUs))
- predicted behavior (e. g., crossing the ego's path)
- contextual situation (e. g., pedestrians close to a crosswalk)

Very naively, measuring safety-relevancy of objects in a given scenario can be done in a standard way, e.g., using criticality metrics [[WNK2023](#)]. If any object exceeds a certain threshold given a set of metrics applicable to the current scenario, this object is clearly relevant. For example, anything with a constant-velocity model $TTC < 6$ seconds might be considered relevant in car-following scenarios. Obviously, such an approach has limits, as relevancy estimation actually requires prediction of all possible futures. If there exists a future with a sufficiently large probability of realization in which a certain object is impacting the safety of the ego, this object can be understood as relevant. Therefore, more work is required especially in the area of probabilistic criticality metrics under worst-case (and not average-case) assumptions.

A related approach has been established in the literature: Mori et al. [[MSP2023](#)] define a relevant object in a given scenario as any object that limits “the set of safe actions available to the ego under consideration of all uncertainties”. However, this approach requires a formal definition of the set of safe actions of the ego, which Mori et al. showed only exemplarily for a rather simple highway use case. The complexity of such a definition grows by orders of magnitudes when advancing to urban environments. Moreover, the approach requires valid and over-approximating prediction models for all other dynamic objects, which is, again, highly involved for complex ODDs. However, their evaluation for highway contexts is encouraging for expanding such a definition of relevance to more complex contexts. Finally, these ideas consider only dynamic elements for “relevancy”. However, a lot of traffic infrastructure, such as traffic signs or markings, are relevant for the ego's behavior as well. Thus, it has to be researched whether the definition of relevancy from Mori et al. [[MSP2023](#)] can be extended to include static objects, and how such a formal relevancy estimation for static objects may look like. A possible solution is a counterfactual (thus, causal) argumentation (“Would the element not have been perceived, would the likelihood of the ego exhibiting hazardous behaviors be increased?”).

Rakow [[RAK2023](#)] suggests a game-theoretic characterization of relevance. The central idea is to capture relevance as “What do you need to know in order to achieve your goals?”. Relevance is reduced to the question of whether there exists a strategy for an autonomous system S so that it is successful when following this strategy, given certain knowledge (K), observations (O) and resources for encoding its beliefs (B). The approach regards a tuple (K,O,B) to be relevant, when it is sufficient for S to be successful and the tuple (K,O,B) minimal. The notion hence not only captures whether the observed objects are factually relevant for achieving the system's goals (whether the road is slippery is relevant if the S wants to brake or do a turn) but rather whether the system needs to perceive/know certain aspects (whether the road is slippery is not relevant to S if S wants to brake or do a turn and S knows that is a speed limit requesting low speed anyway).

Damm et al. [[DFH+2019a](#)] investigate conflicts between traffic participants arising dynamically in traffic scenarios where participants act on local and incomplete perception to fulfill their goals. Determining such a conflicting situation provides valuable input into identifying which objects are relevant.

Damm et al. [[DFH+2019b](#)] propose to require the planning component to explicitly list the objects in the environment that are relevant for ensuring the safety of planned maneuvers. To this end, the planning component will evaluate which maneuvers would be suitable to achieve the current objectives of the ego-system based on dynamic models of the objects identified in the current situation

picture. It will then compute weakest preconditions on objects and their classifications that must be guaranteed by the perception chain for such maneuvers to be safe.

Storms et al. [SMP2023] show that it is possible to validate analytic relevance criteria by analyzing the effect on a motion prediction component. The motion predictor leverages a deep neural network (DNN) as proxy for human behavior.

3.2. Bounding Uncertainty for Components in the Perception Chain

For each component in the perception chain (sensors, fusion components, classifiers), probabilistic guarantees given under non-adverse conditions and known levels of controllable disturbances of the ego vehicle have to be considered. These guarantees include for each type of object:

- a quantification of existential uncertainty and cardinality
- a quantification of the confidence categorization of the type of object, if applicable,
- a quantification of the maximal degree of imprecision of physical attributes of such objects observable by this particular sensor system or classifier component at this particular interface.

We collectively refer to these guarantees as *guarantees for bounding uncertainty*.

The issue here is threefold:

1. We need to identify whether the perception component is in its **applicable domain** (otherwise, the perception component must not be trusted at all).
2. If we can be sure that the perception component is within its domain, an estimate of its probability of **error** (i.e., rating its performance) has to be provided.
3. We need to access the specific details on the **emergence of uncertainty and its propagation** through the perception chain.

These three issues as covered in the following sub-section, 3.2.1 covers the absence of adversarial conditions, 3.2.2 addresses perception accuracy and 3.2.3 deals with emergence and propagation of uncertainty.

3.2.1. Absence of Adversarial Conditions

Anomaly detection is a useful tool for identifying whether the current context contains a rare event, is out of distribution (i. e. not matching the distributions of objects used for training and testing), or actually demonstrates a domain shift where new objects have become statistically relevant in the considered ODD since the definition of the operational context. In all such cases, the perception component may behave unexpectedly. Various approaches for anomaly detection exist depending on the technology used, such as cameras, radar, or object-level sensors [BNZ2022]. One example of an implementation of an anomaly detection is to construct a confidence score, which estimates the accuracy of the perception output for a given input. This can be done using the probability scores provided by a classifying neural network or assessing the variance in the output classes. Even more involved approaches exist, such as, training a reconstruction module which is then used to reconstruct an image from a semantic segmentation output. If the reconstruction is vastly different from the input to the semantic segmentation, the input is likely an anomaly and the segmentation component may not be trusted. These approaches are all technology-dependent, for instance, detecting anomalies in lidar-data since point clouds are sparser during rain. It is evident that this area is under active research and more work is required to handle the detected anomaly (e. g., a low confidence score) downstream (see [Section 3.2.3](#)). An approach how this can be achieved is discussed in [Chapter 5](#).

3.2.2. Accuracy of Perception for Individual Component

In the *Safe_perception(ego)* formula above, the term *da* represents the uncertainties introduced by individual components in the perception chain. This metric measures the distance between the ground truth and the *belief(ego)* established by the perception chain. It is important to note that we are not aiming for the greatest possible approximation of the belief to the ground truth. It is only required that the deviation is smaller than the given maximum deviation δ_a . With suitable abstraction in the construction of component and environment models, considerable complexity can be saved while it is still guaranteeing the necessary high-level properties.

In addition, the respective quality metrics must consider that uncertainties do not only depend on the sensor type (radar vs. lidar vs. camera). Other uncertainties arise from the respective realization of the sensor (e. g., wavelength of the radar may have an impact on its susceptibility to weather conditions) as well as from factors such as placement of the sensor on the ego vehicle. Representation of uncertainties in component models and the impact on quality metrics will be further discussed in Section 3.3.

Salay et al. [SCK+2021] give a method to represent risk-aware performance metrics which could be relevant for this. The main idea is to incorporate the probability that the perception failure leads to a crash into the performance metric, thus excluding performance degradations that have no impact on the overall system safety (e. g., a systematic performance degradation in an irrelevant area behind the vehicle).

3.2.3. Emergence and Propagation of Uncertainty

Risk propagation poses a serious challenge, as it requires, per definition, a valid representation of how system components hierarchically contribute to the overall risk. A possibility of such a representation is the use of a safety argumentation that follows a decomposition approach along the system's architecture.

Salay et al. [SCK+2021] propose exactly this: to expand a safety argumentation in a Goal Structuring Notation to incorporate risk propagation from the individual perception components. The core idea is to use deductive reasoning, starting from a system-level safety goal of bounding the collision risk due to misperceptions. This risk is decomposed into the risk induced by a finite set of "hazardous misperception patterns" (i. e., a system-level hazard that can be caused by a misperception pattern). Subsequently, the bound of the risk induced by an individual hazardous misperception pattern is again calculated by decomposition, similar to that of a system-level hazard, into what roughly equals exposure (how likely is the hazardous scenario to occur), controllability (how likely is a misperception pattern in this scenario), and severity (how likely is a crash given this scenario and misperception pattern). The first can be taken from the main safety case. For the latter, the authors propose a method of simulation combined with fault injection to obtain an estimate on the crash rate. Therefore and similar to the claims made in this white paper, a valid digital twin of the ego vehicle, including the perception chain, with a high degree of detail is required. Finally, the "controllability" component is decomposed further by collecting conditions under which this misperception pattern is triggered disproportionately often, for which again exposure of the condition and controllability of the misperception pattern given this condition are separately estimated. In the end, this approach shows how a propagation from perception component performance ratings to the overall system level can be done, based on a standard decomposition approach. Note that this argument hinges on the completeness and correctness of the identified hazardous scenarios and thus requires methods to validly derive such scenarios even for open and complex contexts. Moreover, it demands further research in how to estimate the single quantities required as evidences: among others, the crash risk given a misperception, the occurrence of hazardous scenarios, or performance metrics for perception components. Since those quantities can often not be estimated in real-world or proving ground settings due to their rareness, highly detailed simulation models (digital twins) in valid simulation

environments are needed.

Fränzle [FHD+2023] argues that machine learning-based systems are more susceptible to noise overlaying percepts than humans. Assessing the functional relevance of percepts is challenging due to the time dimension and prediction horizon, which depend not only on observable objects but also on inner states. Determining the necessary degree of perceptual precision is a complex issue. The impact of (mis-)percepts on safety varies significantly. This approach considers the guard conditions of actions specified in propositional logic. The formulas refer to perception atoms. To justify that a HAV is allowed to perform maneuvers, the guards must satisfy certain rates (false positive/true positive) indicating that the classifier is believed to perform well. Given the Boolean structure of the guards, it is possible to construct an optimization problem for certain types of guards. This can improve confidence in perception for certain types of actions.

3.3. Quantifying Perception Uncertainty Through Virtual Testing

Quality metrics that enable quantifying uncertainty for the different components in the perception chain should reflect environmental conditions, the capability to reproduce physical operation-based failure modes, and the overall completeness of sensor measurements. For instance, sensor models must be assessed in conjunction with corresponding environment models. This includes the ability to simulate sensor failure modes caused by physical phenomena (e.g., double images in lidar due to motion distortion, as illustrated in Chapter 4, Figure 12) and account for corrective mechanisms.

Sensor-specific metrics should be tailored to each type of sensor and consider their placement on the vehicle, necessitating 3D modeling. The sensor's position affects its field of view, influencing the disparity between perceived and actual conditions.

Sensor fusion layers depend on these "imperfect" sensor outputs to filter and identify accurate data. Their metrics must address sensor-specific signals, like intensity clouds for radar or point clouds for lidar, and discuss the completeness of sensor measurements.

3.3.1. Sensor and Environment Models

The separation between the environment simulation models (including moving and stationary objects, lanes, traffic rules, weather conditions, etc.) and the sensor simulation models cannot be maintained when aiming for detailed perception sensor performance modeling. For instance, the shape and materials of surrounding objects are crucial for calculating sensor data, such as lidar point clouds. Therefore, the signal processing model must consider the reflection calculation as its input. To validate perception sensor simulation, it is necessary to validate the materials and geometries beforehand.

The initial stage for a valid perception sensor simulation is to have control over the 3D objects and the digital twin of the environment. This is achieved through the use of the evolving ASAM standard OpenMATERIAL, which is highly supported by this initiative.

Another often ignored aspect in model validation is the necessary diligence when collecting the measurement data that is later used for model validation. This means that with all collected sensor data, all reference data about the environment has to be collected that is later used for re-simulation of the measurement to produce the simulation data for validation. When these reference data (object positions, weather conditions, etc.) are collected, the used measurement devices itself are not perfect either and therefore introduce epistemic and aleatory uncertainties into the validation. Consequently, they must be propagated through simulation adequately, resulting in multiple slightly different simulations per validation sample. Introduced by Roy and Balch [RB2012], this results in multiple data

from simulation that can be combined as so-called p-boxes¹⁵, reflecting their origin and their statistical properties.

In addition, it is important to consider the specific nature of perception sensor simulation, which does not predict a single value for given parameters, but mimics sensor behavior and performance over time, including noisy output. This highlights the need for metrics that distinguish between model bias and model scattering error when validating sensor models, as described by Rosenberger [ROS2022]. In this case, model bias refers to the estimated mean error between the model's output and the actual sensor data, which itself may have a bias when compared to the ground truth value. According to Rosenberger [ROS2022], the applied metrics must account for epistemic and aleatory uncertainties from reference data by being applicable to p-boxes, and they must also provide values for model bias and model scatter error. Therefore, the Double Validation Metric is considered a strong candidate for comparing synthetic and real perception sensor data. In addition to the described benefits, this method provides results in the unit of the measured quantity. Elster et al. [ESP2023] have demonstrated its applicability for radar detections as well.

As already described by Viehof [2018], only sample-wise validation is possible. This means that after all results are determined for all these samples, the interpolation of the errors within the parameter space of the later targeted application domain of the models has to be performed, resulting in model error predictions and the respective uncertainties of these predictions. How to interpolate within the application domain is seen as one of the research questions within the near future.

Sensor models based on effects:

Modeling Recent studies have made significant contributions to the understanding and implementation of sensor models, particularly in terms of how they relate to specific effects and conditions, such as triggering conditions or failure modes for different sensor types.

Adee et al. [AGL2023] developed a methodology utilizing Bayesian networks to model the performance limitations and triggering conditions specific to sensor systems, as demonstrated in a LiDAR-based perception case study. This approach allows for a more nuanced understanding and prediction of sensor behavior under various conditions.

Following a similar direction, Reckenzaun et al. [RWR+2024] have presented a comprehensive best practice guide for validating driving functions, including the perception system, within a virtual environment. The authors show how requirements for the perception systems can be derived from an ODD and how they have to be implemented. They provide detailed insights into characterizing and classifying the effects and properties modeled, as well as guidance on validation processes at different vehicle and system abstraction levels. Their approach illustrates the use of a validated toolchain to ensure the effectiveness and accuracy of the simulation models.

Additionally, Linnhoff et al. [LRS+2021] have introduced a methodological framework to identify relevant effects that need representation in simulation models. They focus on analyzing sensor effects using a tree structure to delineate the relationships between various effects and employ an FMEA-like analysis for understanding the cause-effect chains. The tree structure is documented as part of a GitHub project (see <https://percollect.github.io/lidarLimbs/>).

Sensor models based on measurement data:

Aust et al. [AHD+2023a] present a methodical approach to deriving sensor model requirements based on radar measurement data at the detection level. They employ a variety of metrics, such as point-cloud-to-point-cloud distance normalized by maximum detection number, intersection over union, total variation distance for histogram comparison, and Hellinger distance, also for histogram comparison. The bounds for the sensor model are established from these metric results, supplemented by several experimental replicates. This approach allows for a precise and comprehensive understanding of sensor performance and its limitations.

¹⁵ P-Box

Another work by Elster et al. [ESP2023] presents an approach to derive requirements from radar measurement data, focusing on the radar cuboid and detection level. They utilize the Double Validation Metric to analyze various static scenarios, examining the impact of different environmental conditions such as asphalt surfaces, vegetation, rain, and other objects. The research highlights the significance of reproducibility and repeatability in measurements. These factors are essential for validating sensor models and ensuring their reliability as basis for deriving quality guarantees.

3D object detection and mapping:

Mapping each sensor's measurements onto a 3D grid requires quantifying the disparity between the sensor measurements and the ground truth for each grid cell. This assessment is essential for determining the accuracy and reliability of the sensor data in representing the real world. Fernandez (2015) emphasized the importance of this process.

Mori et al. [MSP2023] present a comparison between machine-based 3D object detection capabilities and human perceptual abilities. The authors argue that for a machine's 3D object detection system to be effective, it should match human capabilities in various respects, given that humans can drive effectively based on their perception. This benchmarking approach focuses on human perception and sets a clear goal for the development of automated vehicle systems. It ensures that these systems are designed to meet or exceed human-level perception.

3.3.2. Modeling Sensor Fusion and Classifier Components

Fusion components play a key role in integrating data from different sensor types to correct perception errors and derive advanced information such as object tracking. Quality metrics for these components must therefore cover a wide range of effects to ensure comprehensive evaluation.

Each layer within the system architecture, especially the sensor layer, is responsible for accurately representing the current level of perception uncertainty. This detailed representation is critical to the effective operation of the fusion layer.

The reliability of the fusion layer is highly dependent on a thorough understanding of the sensor characteristics, including type, positioning, performance parameters, and how environmental conditions impact sensor performance. Chapter 5 reviews different fusion concepts and AI architectures used in classifiers, and discusses phenomena that affect perception uncertainty, referencing work such as Munz et al. [MDM2010].

Reducing uncertainty through sensor data fusion, such as combining lidar and camera data, hinges on the statistical independence of disturbance sources. Hence, the quality metrics must carefully consider and account for potential confounding factors.

Classifier components, typically based on machine learning (ML), require specific considerations for their quality metrics. They need to account for:

- The appropriateness of the selected ML architecture and its training process.
- The accuracy, correctness, and completeness of training and test datasets.
- Representation of rare events in both training and testing phases.

To reduce and limit the uncertainty inherent in classification processes, several strategies have been found to be successful. When defining quality metrics and guarantees targeting classifier components the following approaches should be taken into consideration:

- Implementing a strategy where inputs are processed multiple times through classifiers, cross-referencing results with existing world knowledge to minimize misclassification.
- Identifying objects that contradict known physics (e.g., erratic class changes or unrealistic physical location shifts) to flag potential perception errors.
- Utilizing standardized and comprehensive ontologies to detect inconsistencies in measurements.
- Leveraging high precision maps and preexisting knowledge to resolve conflicting perceptions.

Confidence in classifiers is maximized when the conditions leading to errors are well understood. This includes characterizing favorable conditions for AI components, quantifying the disparity between current measurements and ground truth, and monitoring for adversarial conditions. Damm et al. [DFH+2019b] propose a method for increasing confidence in the perception chain by bounding [perception](#) uncertainty.

Furthermore, Simon Burton and Benjamin Herd [BH2023] discuss several sources of uncertainty of in ML-based classification. Their work proposes a range of methods and measures aimed at limiting misclassification and enhancing the overall performance and safety assurance of ML models.

3.4. Quantifying the Overall Accuracy of the Perception Chain

Assuming that the models of all sensors relevant for the construction of the world model, the environment, the fusion components and the classifiers (covered in Chapters 4 and 5) are good enough to reproduce all relevant sources of uncertainty, the overall accuracy of perception can be assessed based on a digital twin of the perception chain. Neurohr et al. [NKM+2023] demonstrate this by comparing the distributions of the generated observation of the digital/virtual twins and the ground truth process.

Additional deviations between the digital twins of the perception chain and the ground truth, such as those resulting from different characteristics of the execution platform, are out of scope in this context (see [Section 1.4](#)).

The integration of component guarantees can be based on safety cases as proposed by Salay [SCK+2021]. In these, system-level arguments are linked to unit-level arguments, allowing the investigation of misperceptions as a function of context and the assessment of their impact. The safety case template focuses on hazardous misperception patterns that separate the analysis of HAV dynamics from perception. This approach aims to identify patterns of misperception that can lead to hazardous situations, allowing for targeted mitigation strategies. Risk-aware performance metrics are defined that compute a measure of the misperceptions generated by a component. Unlike generic performance metrics (e.g., recall, mAP, AuPR), which consider any deviation from ground truth as bad, these metrics consider only deviations that are hazardous.

4. Sensor Characterization and Sensor Modelling

In this chapter, the environment simulation and an overview of radar, lidar and video camera models is given. The chapter outlines a model overview from the literature, presents validation initiatives, and highlights existing research gaps. Subsequently, digital maps are described, along with their application in the same format. All topics listed here share a commonality in lacking a methodology that defines the requisite precision of the models with respect to the specific use case.

4.1 Environment Simulation

The use of environment simulation plays a crucial role in the development and safety validation of automated vehicles, particularly in the area of sensor simulation for cameras, lidar and radar. These models are central to virtually recreating realistic scenarios and testing the performance of vehicle systems under different conditions. Cameras, lidar and radar capture the environment and provide data for navigation and vehicle control. Extensive testing and simulation is required to ensure that vehicles can operate safely in all possible conditions. This is where environment simulation comes into play.

4.1.1 Use Cases

Environment simulation represents the vehicle's physical environment in a virtual world. They include roads, buildings, other vehicles, pedestrians and all potential obstacles. Sensor simulation models are used to generate data similar to what a vehicle's sensors would collect in the real world in different conditions. This allows engineers to simulate different scenarios, including challenging weather conditions, traffic congestion and unpredictable events. Environment simulation allows accurate and repeatable evaluation of the performance of sensors and vehicle systems. Errors can be detected and corrected early, increasing the safety and public acceptance of automated vehicles.

The adaptation of the environment model to the sensor model and its interfaces is of paramount importance. In the case of object-based ideal models, for instance, positions and corresponding classifications of objects in the physical environment of the vehicle are sufficient. However, for complex models like ray tracing-based methods for sensor simulation, intricate 3D geometries are required, complete with material properties. This allows the virtualization of the effects of reflection, transmission, and absorption at surfaces struck by electromagnetic radiation in the virtual world.

4.1.2 Sensor modalities

In addition to the objects in the vehicle environment, the properties of the atmosphere also play a significant role in sensor performance. Radar is less affected by rain than time-of-flight lidar. Camera sensors are limited in their performance due to changes in brightness (e.g., tunnel entry/exit) or day and night situations. Consequently, such environmental conditions must also be part of a valid environmental model.

As shown in Figure 8, the different wavelength ranges of the various sensor modalities are affected differently. Radar is in the mm range, lidar is in the infrared range between 1550 to 850 nm, and the visible light for the classic camera is between 750 and 300 nm. Also shown in the figure is the attenuation of the output due to various weather effects.

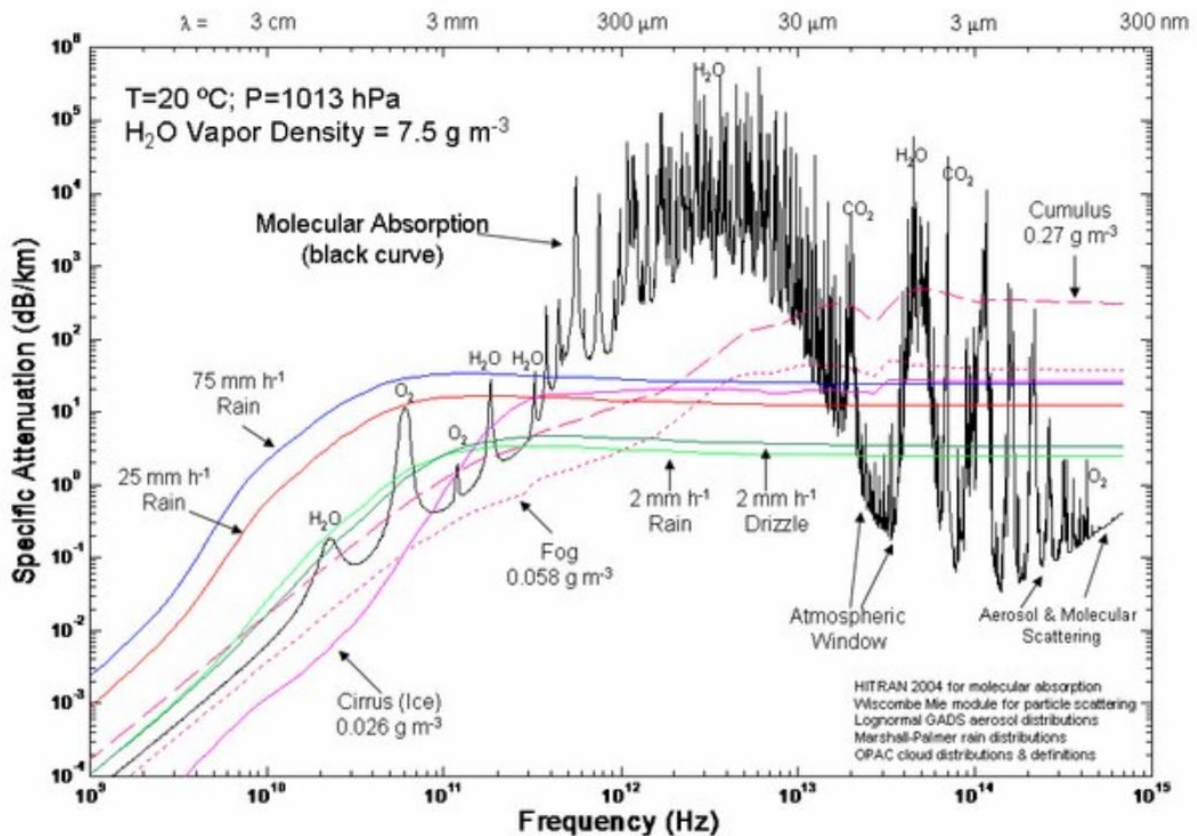


Figure 8: Specific attenuation over frequency for radar, lidar and camera and different environmental conditions ¹⁶

4.1.3 Object and material properties

The fundamental three-dimensional structure of objects is theoretically dissociated from the sensor modality. However, considerations such as the granularity of 3D-object meshes hold significant importance. This is because various sensor resolutions impose distinct prerequisites on the creation of a valid environment model. Nevertheless, it is noteworthy that the current state of scientific knowledge lacks comprehensive investigations addressing this aspect in detail.

Furthermore, distinct challenges manifest themselves, particularly with regard to material properties. The variations in wavelength ranges result in disparities in the phenomena of reflection, transmission, and absorption. In this context, a significant reliance on factors such as surface roughness, surface temperatures, material thicknesses, and surface moisture is anticipated. One exemplary environment model implementation is shown in Figure 9. The 3D geometry is based on a glTF-file with an addition of different extensions. Thereby the material file can be defined by means of the OpenMATERIAL format, which is just started as an official ASAM standardization project.

¹⁶ [FBK+2008]

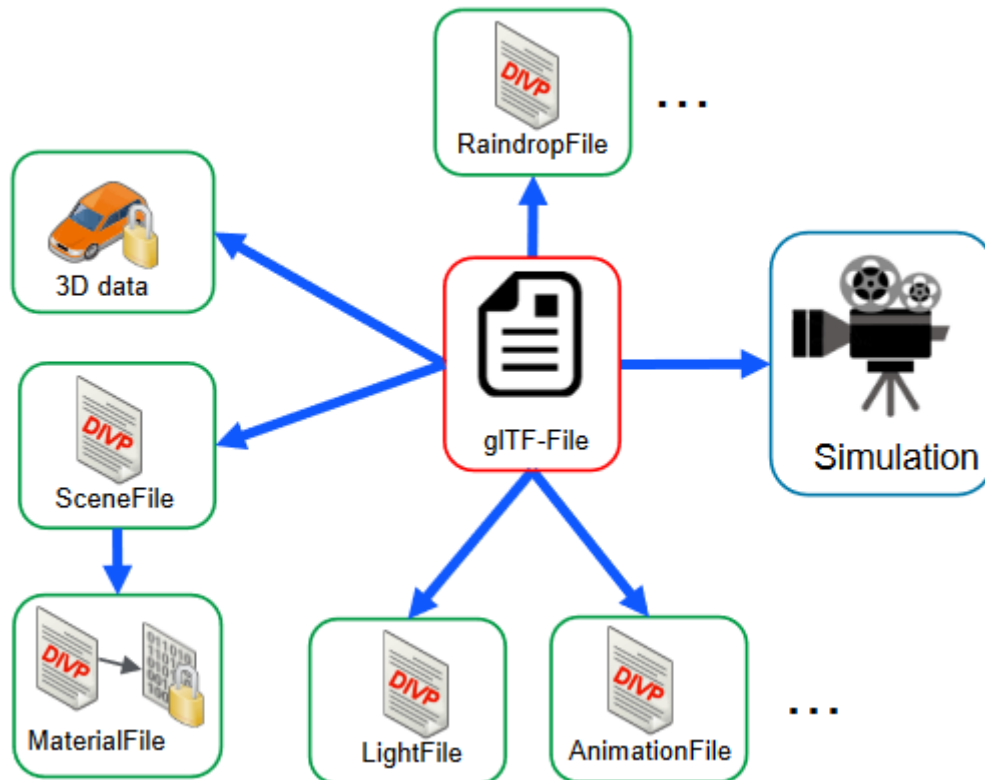


Figure 9: Exemplary material properties implementation based on the Japanese DIVP project¹⁷

In conclusion, the following open research questions arise from the current state of the art:

1. In which quality does an environment model have to be provided at different distances from the sensor (level of detail and mesh granularity of e.g. 3D models)?
2. How can an environment model and its quality be validated to ensure credible simulation?
3. Which number of materials and which properties (physical properties like permeability, permittivity, temperature, roughness) are necessary for a valid environment model?
4. How can the material assignment and 3D geometry generation be accelerated to enable industrial virtualization of scenes?

4.2 Radar

In the rapidly advancing realm of automotive technology, Radio Detecting and Ranging (radar) sensors have emerged as pivotal components in enabling a safer and more efficient driving experience. These sensors utilize electromagnetic waves to detect and track objects in the vehicle's vicinity, providing essential data for advanced driver assistance systems (ADAS) and autonomous driving functionalities. One of the key concepts in radar technology is the frequency modulated continuous wave (FMCW) chirp sequence radar, which offers enhanced accuracy and reliability. In the context of automotive applications, radar can be categorized into short- and long-range radar. Short-range radar sensors use the 24 GHz so-called K-band and long-range radars use the 77 GHz so-called W-band.

New developments in the radar field deal with Synthetic Aperture Radar applications, bistatic arrays around the vehicle and 150 GHz frequency ranges.

¹⁷ Hideo Inoue, Matthias Hein: DIVP material initiatives for OpenMaterial through VIVID collaboration, 21th March 2023, ASAM Technical Seminar 2023, <https://www.asam.net/index.php?eID=dumpFile&t=f&f=6860&token=b1e6378e2411d5b68cc6e3cfd615c6a7c9c582d4>, accessed on 4th October 2023

4.2.1 Principles of Operation

Radar sensors emit electromagnetic waves, typically in the microwave frequency range, which travel through the environment and bounce off objects in their path. By analyzing the returning signals in comparison to the emitted signal, the sensor can determine by FFT algorithms the distance, speed, and estimates the azimuth and elevation angle of these objects. This process allows radar sensors to detect obstacles, pedestrians, vehicles, and other potential hazardous and non-hazardous objects around the vehicle. In Figure 10, a simplified radar processing chain is visualized with the different sensor output interfaces. In the current state of the art, object lists are normally used to realize environment detection in combination with other sensors. In the future, less processed data will gain importance due to machine learning approaches like detections, which are already defined in ISO 23150 as a sensor interface, or the radar cuboid.

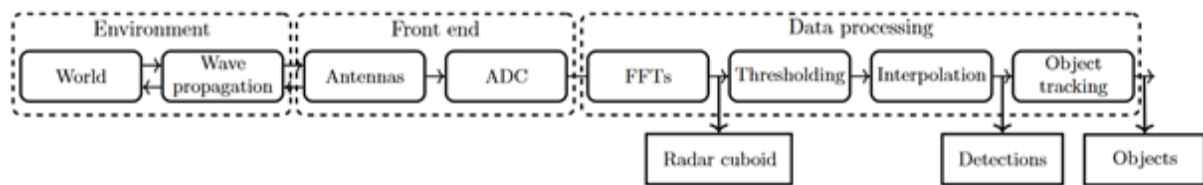


Figure 10: An abstracted radar processing chain with elements visualized as blocks with rounded corners. The group within the processing chain is visualized as dashed rounded blocks and the sensor interfaces are marked as edged blocks.¹⁸

4.2.2 Artifacts and Effects

The terms “artifacts” or “effects” describe characteristics in radar measurement which lead to a deviation in the output interface with a varying degree of significance. Holder¹⁹ defines an artifact as “noticeable deviation from ground truth in the sensor readings that is inherent in the sensor measurement principle and its system design”. Linnhoff et al. define an effect as “the deviation from the originally existing information about the environment in the signal or data”.²⁰ Both definitions have in common that the deviation from an undisturbed measurement in comparison to the resulting measurement is an artifact respectively effect. Due to some main effects in radar measurements, a list of the strengths and weaknesses can be derived:

Strengths of Radar Sensors:

1. **Versatile Weather Resistance:** Radar sensors can reliably operate in adverse weather conditions such as rain, snow, or fog, as the electromagnetic waves they use can penetrate these obstacles.
2. **Extended Range:** Radar sensors can detect objects at significant distances, which is crucial for responding to potential hazards or obstacles well in advance.
3. **Insensitive to Lighting Conditions:** Unlike optical sensors, radar sensors are not dependent on daylight or darkness, providing consistent performance around the clock.
4. **Object Detection Regardless of Color and Shape:** Radar sensors are based on electromagnetic reflection, allowing them to detect objects regardless of their color or shape.
5. **Detection of Multiple Objects:** Many radar sensors can simultaneously track and identify multiple objects within their detection range, advantageous for complex traffic scenarios.
6. **Speed Measurement:** By applying the Doppler effect, a radar sensor can directly measure not only the distance but also the speed of moving objects.
7. **Influence of Multipath Propagation:** In multipath propagation, radar waves can be reflected

¹⁸ [ERH+2023]

¹⁹ [HOL2023]

²⁰ [LRS+2021]

from the road surface and from objects, making obscured objects visible.

Weaknesses of Radar Sensors:

1. **Reduced Spatial Resolution:** Compared to optical sensors, radar sensors offer lower capability to distinguish objects with similar distances and relative radial velocities from one another like parking cars.
2. **Limited Detail Recognition:** Automotive radar sensors are not able to provide detailed information about the shape, size, or type/class of detected objects.
3. **Limited Accuracy at Short Distances:** Due to the limited number of radar oscillations within a short distance, distance resolution can be compromised in close proximity.
4. **Challenges in Classification:** Radar sensors might struggle to differentiate between different types of objects since they rely on electromagnetic reflection and do not consider visual features.
5. **Interference from Other Sensors:** In densely populated areas or near other vehicles, radar sensors can be affected by interference from other radar systems.
6. **Ghost targets from Clutter and Multipath Propagation:** Due to signal processing based on dynamic thresholds, radar sensors are prone to clutter caused by the environment. In addition, the material properties of road objects result in multipath propagation of the electromagnetic wave, which can create ghost objects or detect objects in front of other vehicles.
7. **Angle and relative velocity ambiguities:** Due to the aperture size and radar signal processing, ambiguities occur in the angle and velocity measurements, causing detections and objects to be misinterpreted

More detailed artifacts as well as a hierarchical structuring with information about the interaction of different effects can be found on Github at PerCOLLECT RadarRami.²¹

4.2.3 Assessment of Radar Sensors

To analyze different effects in radar measurements in the context of model validation, different experiment designs arise in literature. Ngo gives an overview about radar model validation experiments separated into the validation scope, the corresponding author and the validation method.²² Additionally, the publications in Table 1 address radar sensor model validation.

Table 1: Radar model validation experiments by author and the corresponding validation method

Authors	Validation method
Aust et al. ²³	Qualitative, one dynamic scenario
Aust et al. ²⁴	Quantitative, one dynamic scenario
Elster et al. ²⁵	Qualitative, one dynamic scenario with multiple objects
Elster et al. ²⁶	Quantitative, one static scenario
Elster et al. ²⁷	Quantitative, six static scenarios
Magosi et al. ²⁸	Quantitative, one dynamic scenario

From the sources listed, it appears that there is currently no uniform methodology for creating

²¹ Github, "RadarRami", <https://percollect.github.io/RadarRami/> (accessed on 03.09.2023)

²² [NGO2023]

²³ [AHD+2023b]

²⁴ [AHD+2023a]

²⁵ [EHR2023]

²⁶ [ERH+2023]

²⁷ [ESP2023]

²⁸ [MWT+2022]

experiments for radar sensors. In particular, reference data and their accuracies are underrepresented in the model evaluation. Furthermore, within radar sensors, various effects exist as presented in 1.2.2, each of which necessitates individual validation to substantiate the model's capabilities with corresponding thresholds and metrics. Therefore, another gap in the state of the art is present:

1. Which experiments are necessary for which effects and thereby ensure that effect isolation is possible and quantifiable?
2. How can the uncertainties introduced by reference sensors be accounted for in the validation process, and what are the necessary reference sensor accuracies associated with the specific perception sensor in this regard?

4.2.4 Modeling of Radar Sensors

The simulation of radar sensors in virtual environments is of great interest in the ADAS community. Besides the different effects and their corresponding modeling approaches, the categorization as well as the application to different use cases is not yet standardized. Magosi et al. summarize in the literature known categorizations and available modeling approaches in the context of automotive applications.²⁹ Thereby, they introduce their own categorization scheme, which is shown in Figure 11, as well as an explanation about the differences in the modeling approaches and the used simulation techniques.

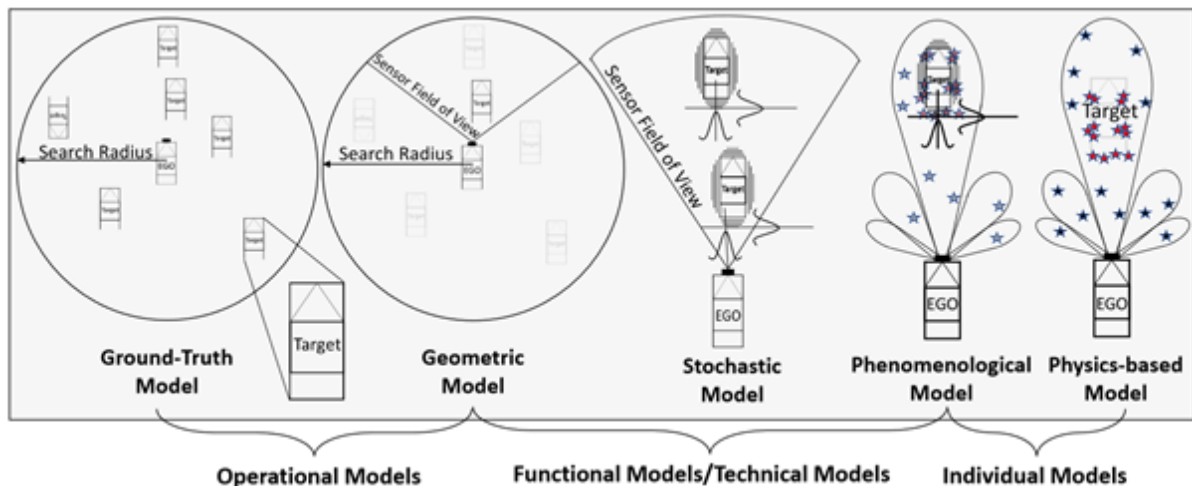


Figure 11: Categorization of modeling approaches defined by Magosi et al.

Based on the mentioned survey, still open research questions arise, which have to be addressed in the future:

1. How exactly do artifacts need to be implemented in the radar model and how can their influences be quantified in means of testable requirements?
2. Which effects have to be observed on which output in which quality and how can this be verified?
3. What effects result from the interaction of multiple radar sensors, be it sensor systems on the ego vehicle or sensors from other road users?

4.2.5 Model verification/ validation

Model verification and validation is a very extensive field and is strongly influenced by the respective sensor technology (radar, lidar, camera, ultrasonic) and the targeted data interfaces. The generated synthetic data from the model has not only to follow the physical laws but also needs to be comparable with the original sensor measurement data for any given scenario and sensor variant. The potential

²⁹ [MLR+2022]

data interfaces (based on radar) are described in the following Figure 12.

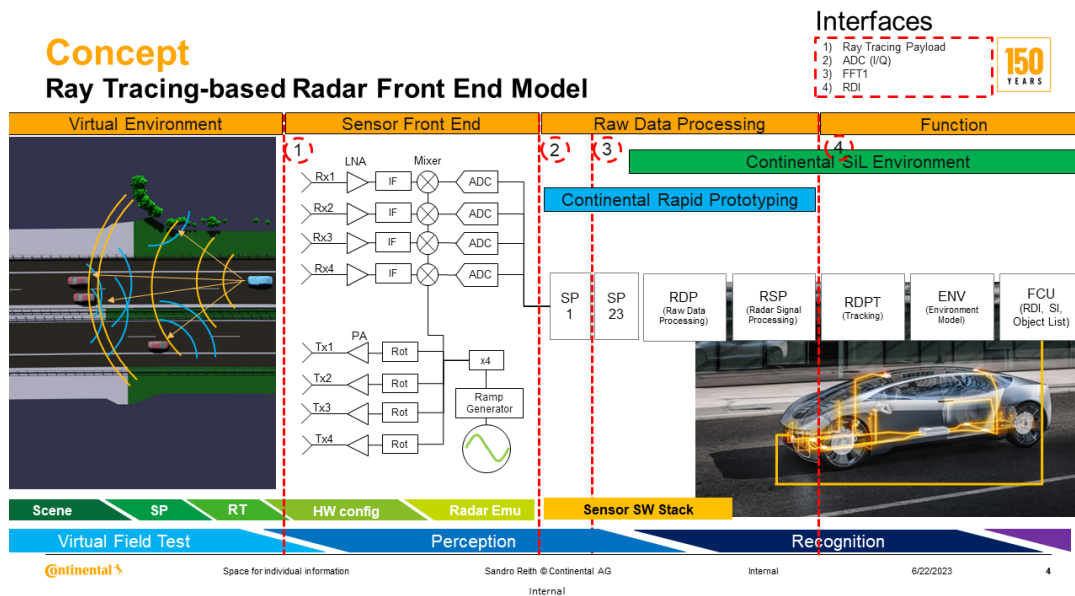


Figure 12: Potential interfaces and processing steps of a virtual (radar-) sensor.

The complexity of the verification and validation processes increases in correlation with the number and impact of various effects when data is utilized within the radar sensor at earlier stages. The number of dimensions and the dependencies to specific hardware effects increasing rapidly. It is recommended to use the interface 4 (s. Figure 12), on the Radar Detection Image (RDI - ISO23150) this interface is already standardized and also includes some processing parts of the original input data (I/Q - data [Interface 1]) of the sensor. Therefore, the complexity is already reduced and the handling of the data is easier.

For validating this data, a generic method and metric is required. In the VIVALDI project the chosen metric is and Area Validation Metric (AVM). In Continental a submetric of the AVM was chosen, the Mahalanobis Distance. This metric judges the quality of the virtual processed data points compared to the real measurements based on the Key Performance Indicators (KPI). The Mahalanobis Distance is suitable as it accounts for correlation between datasets. It is used frequently with large datasets with manifold correlations. In Figure 13 following observation can be made for and Car-to-Pedestrian Longitudinal Adult scenario from the Euro NCAP catalog.

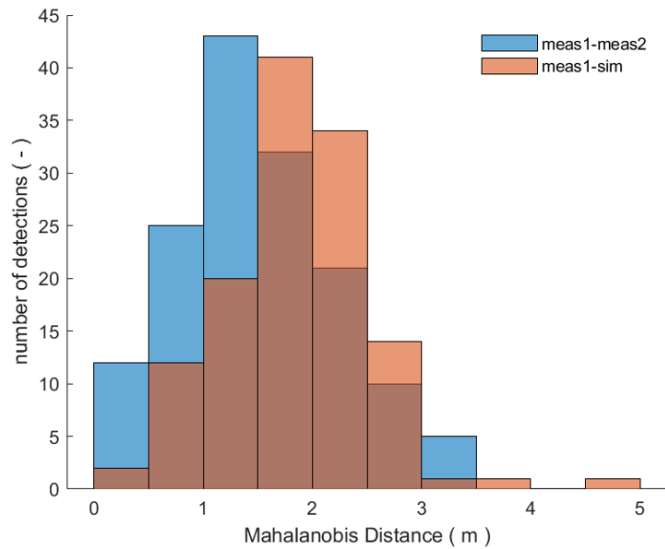


Figure 13: Mahalanobis Distance for a CPLA with ego velocity $v_{ego}=30$ km/h and pedestrian velocity $v_p = 5$ km/h.

Good initial overlap between measurement and simulation data. The spread of the distance is in both cases very similar. Some differences occur due to walking and driving uncertainties of real people compared to simulated scenarios. Most of these effects are from the differences in the ground truth data.

Based on the presented exemplary evaluation final research questions arise:

1. Which metric is suited for which interface in the radar processing chain to prove validity of effects and have additional statistical methods be provided?
2. How can simulation models utilize interpolation and extrapolation to extend their validity within a specific parameter space within the ODD?

4.3 Lidar

Lidar sensors, as shown in Figure 14, have gained significant attention over the past few years for their use in advanced driver-assistance system (ADAS) applications because they provide outstanding angular resolution and high ranging accuracy compared to radar [[BIL2022](#)].



Figure 14: Lidar sensor on a test vehicle

4.3.1 Lidar Technologies

To make it simple, each automotive lidar sensor consists of three subsystems: transceiver with transmitter (laser) and receiver (detector), beam steering and the required control/processing. Commercially available automotive lidar sensors can be classified according to their beam steering subsystem into two categories, non-scanning, and scanning sensors, based on their beam steering technology (Figure 15).

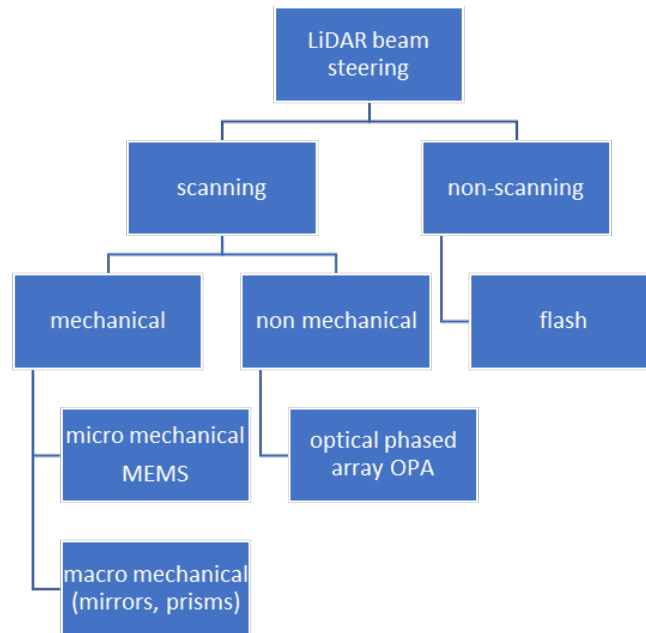


Figure 15: Lidar beam steering technologies for automotive applications

Flash lidar sensors are non-scanning types of lidars. They illuminate their entire field of view (FoV) at once by a laser source and do not contain any mechanical moving parts to steer the beam [RCG2021]. Non-scanning lidar sensors can measure up to 50 m and are used for forward collision warning (FCW) and blind spot detection (BSD) [THA2016]. Scanning lidar sensors steer the laser beam in the FoV by using mechanical moving parts to obtain the complete 3D view of the vehicle’s surroundings in a specific frame rate [RB2019]. Moreover, a scanning lidar sensor focuses its laser beam in a particular area for one shot. Therefore, they can measure objects up to 200 m with a typical horizontal and vertical resolution of 0.1 deg, depending on the frame rate [RB2019]. That is why they are used for lane departure warnings (LDW), simultaneous localization and mapping (SLAM), FCW, and BSD [RB2019]. Specifically, MEMS-based lidar sensors are getting more attention for automotive applications because they are small, lightweight, and power efficient [HB2014]. Due to this relevance, MEMS-based lidar sensors are considered in the following. Furthermore, MEMS-based lidar sensors are also used for agricultural, archaeological surveys, and crowd analytics [BLI2023].

4.3.2 Principles of Operation

MEMS-based lidar sensors consist of a laser and detector module (LDM) and a beam deflection unit (MEMS-based mirrors) for beam steering, as shown in Figure 16. The laser source emits laser pulses, and the beam deflection unit deflects the beam in different directions to obtain a holistic imaging of the environment. The photo detector receives the laser pulse partly reflected from the target’s surface. The sensor acquires the round-trip delay time (RTDT) τ that the laser light takes to hit an object and return to the detector. With τ , the range R can be calculated as [HPK+2022]:

$$R = \frac{c \cdot \tau}{2}$$

where c is the speed of light and the RTDT is denoted by τ .

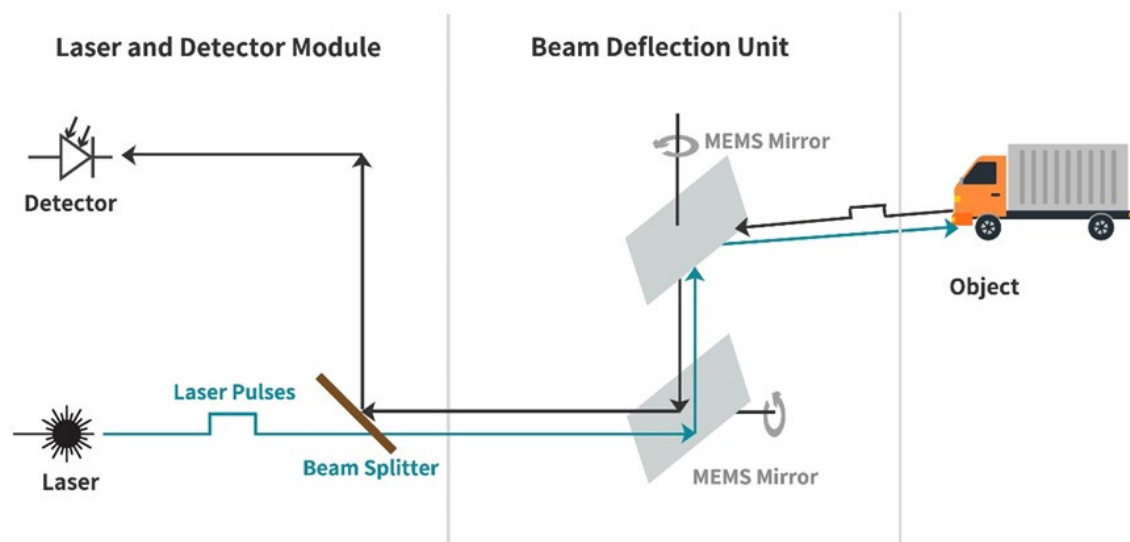


Figure 16: Block diagram of MEMS lidar sensor, source: adapted with permission from [PET2022a].

4.3.3 Characteristics

Table 2 shows lidar sensor-specific parameters essential for the application. In addition, exemplary values of a MEMS-based lidar sensor are assigned to the parameters.

Table 2: Parameters of a MEMS-based lidar sensor [BLI2022]

Parameter	Exemplary Values
Typical application range	1.5 m ~ 75 m
Range resolution	< 1 cm
Range accuracy	< 2 cm
Maximum number of scanlines	400
FoV (H x V)	70 ° x 30 °
Angular resolution	0.4 ° ~ 1 ° (user configurable)
Frame rate	1 Hz ~ 50 Hz (user configurable)
Laser wavelength	905 nm

4.3.4 Artifacts

Like other environmental perception sensors, including RADAR, camera, and ultrasonic sensors, lidar sensors also have strengths and weaknesses.

Strengths of lidar sensors:

- **Range:** Lidar sensors can detect objects at long distances, which is important for responding to potential hazards or obstacles well in advance.
- **Detection of Multiple Objects:** Lidar sensors can simultaneously track and identify multiple objects within their detection range, advantageous for complex traffic scenarios.
- **Object detection in Darkness:** Unlike camera sensors, a lidar sensor's performance is not affected at night.

- **Resolution & Accuracy:** Lidar generates instantaneous, dense measurements and can be accurate to a centimeter.
- **3D Mapping:** 3D point clouds of the environment generated by the lidar sensors can be used to produce 3D maps to interpret the environment.

Weakness of lidar sensors:

- **Costly:** Current lidar sensors available for automotive applications are very costly in comparison to RADAR and camera sensors.
- **Sensitivity to Environmental Conditions:** The performance can decrease significantly in rain, fog, and snow. In addition, the lidar sensor's performance also degrades in direct sunlight.
- **Motion Distortion:** The fast relative motion between the scanning lidar sensor and objects can lead to a distortion of the point cloud.
- **No Color information:** Unlike Camera sensors, lidar sensors don't provide the color information of the detected objects.

4.3.5 Modeling of Lidar Sensors

Automotive LiDAR sensor models are typically divided into ideal, phenomenological, and physical models depending on their modeling approach and covered effects [AHH+2020].

Ideal sensor models, also known as “ground truth” (Ground truth provides the simulated objects' actual values, dimensions, position, velocities, orientation, and bounding box) sensor models, use the object list provided by the simulation framework in the world coordinate system as an input. The term ground truth is borrowed from remote sensing, where it refers to location information for data calibration [NKL+2021]. These models' output is a filtered object list for the sensor specific FoV [HHB+2015]. An ideal lidar sensor model does not consider any sensor-related imperfections except the FoV and object occlusion. Therefore, these models have low complexity, require less computation time, and can test the highly automated driving (HAD) function operation in the early stage of development. It should be noted that ideal models, which are described in the literature, are mostly generic, and they can fulfill the requirements of different environment perception sensor types, including lidar, radar, and camera [SN2018]. The ASAM Open Simulation Interface (OSI) provides the `osi3::GroundTruth` interface for such sensor models.

Phenomenological lidar sensor models use the object list as an input and apply weather conditions, false alarms (positive/negative), detection probabilities, and sensor-related effects, including the FoV and a limited detection range. This type of sensor models outputs either detections (point clouds) or object lists [HHR+2015].

Physical sensor models are based on the physical aspects and can be numerically complex. Hence, they usually require a lot of computational power and, thus, might not be real-time capable. The subsequent models use the rendering techniques provided by the simulation framework as input and generate the detections (point clouds) as an output containing distance, intensity, and timestamp. Several rendering techniques generate synthetic lidar sensor detections; ray tracing, ray casting, rasterization (z-buffers), and ray path [SLB2017].

4.3.6 Lidar Sensor Modeling

Figure 17 depicts the toolchain and the signal processing steps of the proposed lidar model. The sensor model considers the scan pattern and complete signal processing steps of Blickfeld Cube 1. As mentioned earlier in Section 1, the model itself is built as an OSMP FMU and uses the virtual environment of IPG CarMaker. It provides the ray tracing framework with a bidirectional reflectance distribution function (BRDF) that considers the direction of the incident ray θ , material surface, and color properties [IPS2021]. The lidar FMU model uses the ray tracing module of IPG CarMaker. The material properties of the simulated objects, angle-dependent spectral reflectance $R_\lambda(\theta)$, and reflection types, including diffuse, specular, retroreflective, and transmissive, are specified in the

material library of IPG CarMaker.

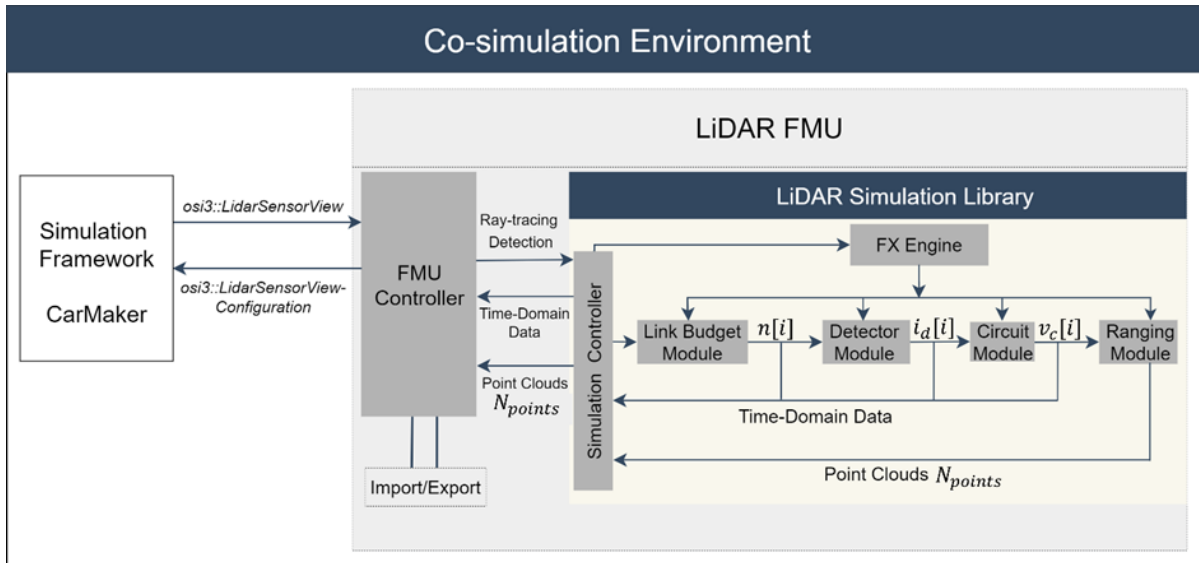


Figure 17: Co-simulation framework of the lidar FMU model.

The FMU controller passes the required input configuration to the simulation framework via `osi3::LidarSensorViewConfiguration`. The simulation tool verifies the input configuration and provides the ray tracing detections via `osi3::LidarSensorView::reflection` interface time delay τ and relative power $P_{rel}(R)$ [ASA2022].

Afterward, the FMU controller calls the lidar simulation library and passes the ray tracing data for further processing. The central component of the simulation library is the simulation controller. It is used as the primary interface component to provide interactions with the library, for instance, configuring the simulation pipeline, inserting ray tracing data, executing the simulation's steps, and retrieving the results.

The next block in the pipeline is the link budget module, which calculates the photons over time. The task of the detector module is to capture these photons' arrivals and convert them into an electrical current signal $i_d[i]$. In the proposed lidar model, a silicon photomultipliers (SiPM) as a detector [FSB+2022] is implemented. Still, it can also support avalanche photodiode (APD) and single-photon avalanche diode (SPAD) detector models.

The third block in the pipeline is the circuit module. Its task is to amplify and convert the detector's photo current signal $i_d[i]$ to a voltage signal $v_c[i]$ that is processed by the ranging module.

The last part of the toolchain is the ranging module, which determines the range and intensity of the target based on the $v_c[i]$ received from the analog circuit for every reflected scan point. Finally, the effect engine (FX engine) is a series of interfaces that interacts with environmental or sensor-related effects and the blocks of the simulation pipeline. These interactions can involve, for example, the consideration of thermal noise in electrical components, signal attenuation due to weather phenomena, and backscattering. It should be noted that this paper only considers the environmental condition sunlight effect.

This section will cover a detailed description of scan patterns and lidar simulation library components.

A detailed description of the individual modules and covered effects of the sensor model can be found in [HPK+2022, HCP+2023, HPK+2023].

4.3.7 Model Validation

The lidar sensor model developed by Kempten University of Applied Sciences and the Blickfeld GmbH in the VIVALDI project [HPK+2022] is validated on the time domain, point cloud, and object recognition level. The sensor model contains the complete signal processing toolchain of the Blickfeld Cube 1 lidar sensor and the sensor-specific imperfections, including optical losses, inherent detector effects, effects generated by the electrical amplification, and noise produced by sunlight. In addition, the effect of rain and fog is also modeled in it. A detailed description of the sensor model validation can be found in [HPK+2022]. In order to validate the model at all three levels, both static and dynamic tests were performed.

Static Lab Tests:

The static tests are performed in the lab. Therefore, a 10% diffuse reflective Lambertian plate is placed at different distances in front of the sensor.

To verify the model on the time domain level, only single-point scattering is considered from the surface of the target.

To validate the sensor models at point cloud level, three metrics are defined. A detailed reasoning for choosing these metrics can be found in [HPK+2022].

- Number of received points N_{points} from the surface of the simulated and real objects of interest (OOI).
- Comparison between the mean intensity I_{mean} values of received reflections from the surface of the simulated and real targets.
- Distance error d_{error} of point clouds obtained from the real and virtual objects should not be more than the range accuracy of the real sensor, which is 2 cm in this case.

Dynamic Proving Ground Tests:

Dynamic test drives were conducted at the Jtown proving ground. Therefore, it should be noted that the daylight intensity is recorded and modeled in the simulation environment too. The simulated and real test drive results have been compared frame by frame to validate the environment and sensor modeling for the defined metrics.

To validate the sensor model at the object recognition level we trained a state-of-the-art deep learning based PointPillars network [LVC+2017] for object detection using simulated lidar data. The model has been tested with real and simulated data of the vehicle target. The average orientation similarity (AOS) metric [GLU2012] has been used to find the correlation between the object's 3D ground truth orientation and the object's estimated 3D orientation by the object detection algorithm.

A detailed description of the scenarios used for the validation as well as an overview of the results obtained can be found at [HCP+2023, HPK+2023].

4.4 Camera

The camera sensor is expected to be one of the major components of an advanced driver assistance system (ADAS) or automated driving system. Environment recognition by a camera sensor is essential since the camera provides a reliable full scene understanding and furthermore it gives information's lidar, radar and ultrasonic sensors are not capable to provide like detection of traffic-lights, -signs and emergency signals. Cameras can furthermore provide visual information to the driver about the car surrounding e.g., during parking scenarios. The number and characteristics of camera systems in a car can vary, depending on the use case. In general, we distinguish between near range cameras which usually have a wide field of view (i.e., fisheye optics) for rear or surround view applications and front view cameras with a high range and small field of view (tele optics) for i.e., emergency braking functions. The most significant differences are the field of view and the resolution.

4.4.1 Principles of Operation

Camera sensors are optical and passive sensors, which means that the incoming light is focused by the camera lens system and hits the electronic image sensor-unit which is located at the focus point of the lens.

The light spectrum “visible” to a camera is mostly bounded below by material properties and bounded above by an IR-cut-off-filter which yields a spectrum-range of approx. 380nm-780nm. The sensor-unit consists of light-sensitive photodiodes (pixels) arranged in a two-dimensional array. The number of collected photons is then converted to the corresponding number of electrons transferred into a digital number by an A/D-Converter.

Cameras function through a combination of various components (i.e. lenses, filter, lens-housings, and physical principles, these include i.e. exposure control, which regulates the amount of light reaching the sensor.

In automotive cameras sampling happens at discrete locations (pixel), at discrete time (capture time), with discrete spectral weighting. In the process from analog to digital signals the algorithms of the signal processing reduce these sampling- and noise-effects, before detection (perception) algorithms are applied to detect objects, like pedestrians. In Figure 18 a simplified video processing chain is visualized.

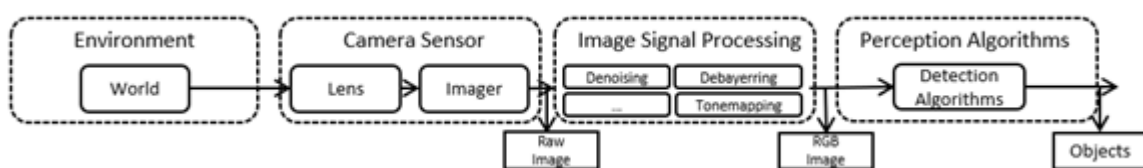


Figure 18: A simplified camera processing chain with elements visualized as blocks with rounded corners. The group within the processing chain is visualized as dashed rounded blocks and the sensor interfaces are marked as edged blocks

Summarized, automotive camera systems must deal with unconstrained environments, i.e., a wide range of weather, illumination, and temperature conditions. The process of environmental visual data acquisition is the result of a complex effect chain, which starts from a light source and ends with the final image stored in memory. In this information transfer chain, the signal suffers from a variety of intermediate disturbances, thus degradation of the signal quality will always take place to some extent. It is important that the system is designed so that enough relevant information about the world is still preserved in the chain.

4.4.2 Artifacts

Camera sensors have different strengths compared to other sensor modalities as well as weaknesses. The characteristic (more precisely the degree of aberrations and distortions) of a camera sensor

depends on the design requirements and the implemented design.

Therefore, the following strengths and weaknesses are formulated in a quite general manner:

Strengths of camera sensors:

1. **Reliable Spatial Resolution:** Optical sensors have a high solid angle resolution and the angle determination can be carried out horizontally and vertically. It offers the capability to distinguish objects with similar distances and relative radial velocities from one another like parking cars.
2. **Detail Recognition:** Camera sensors are capable of providing detailed information about the shape, size, color, or type/class of detected objects. Especially traffic signs and lights and emergency signals are detectable and crucial information about the traffic light color, or the speed limit can be provided.
3. **Detection of Multiple Objects:** Camera sensors can simultaneously track and identify multiple objects within their detection range, advantageous for complex traffic scenarios.

Weaknesses of Camera Sensors:

1. **Sensitive to Lighting Conditions:** Without active illumination, cameras have detection limitations in low illuminated scenes, i.e. darkness, as well as low contrast scenes, i.e. blinded by the sun.
2. **Limitations at adverse Weather Conditions:** Camera sensors cannot reliably operate in adverse weather conditions such as rain, snow, or fog since the contrast is limited here.
3. **Not accurate in Speed Measurements:** Camera speed estimation is not measured, but derived from subsequent position measurements and therefore is inaccurate e.g. compared to radar.
4. **Detection Range is limited:** Compared to other sensors, the detection range of cameras is in general shorter.
5. **Modulated Light Challenges:** If parameters of the camera (i.e. exposure time) and modulated LED-light sources do not fit, the LED source could appear to be OFF in the camera (i.e. traffic signs and signals, headlights).

In sum, cameras have advantage over lidar and radar concerning resolution, color detection and detail detections, but perform worse than radar and lidar in terms of weather robustness, 3D shape detection and the measurement of relative speed. Furthermore, they are limited in darkness or if straylight happens.

The most relevant camera properties to the advanced driver assistance systems were systematically derived and confirmed also in expert interviews: the so-called lens impairments, like distortion, vignetting, and blur, and the image impairments, like noise and rolling shutter. There is a vast literature on this topic and therefore reference is made to the white paper - IEEE P2020 Automotive Imaging, IEEE, 2018 (https://www.image-engineering.de/content/library/white_paper/P2020_white_paper.pdf).

Furthermore, detailed artifacts as well as a hierarchical structuring with information about the interaction of different effects can be found on Github at PerCOLLECT CameraCope <https://percollect.github.io/CameraCope/>.

Ultimately, the uncertainty of a camera's object recognition determines not only the camera properties and their phenomena but also the algorithms used. What is ultimately relevant for vehicle guidance is how accurately the objects can be reproduced in a vehicle movement and under different environmental conditions. Figure 19 shows an example of the measurement of the road curvature of two current vehicles (SUV and sedan) with two different camera systems under different weather conditions. The measurements were taken over several weeks as part of the SensIndex project, funded by the Bavarian state as part of a large-scale test subject study. The advantage there was that the study was carried out on digitized ground truth routes and the weather conditions were very different. The sequence of the test drives was clearly defined and could always be reproduced. This revealed major differences in sensor behavior. The comparability of several journeys is a key challenge due to different speeds. The analysis based on the digital ground truth maps enables the measurement data to be synchronized locally. The location-based analysis enables an objective comparison of different sensors.

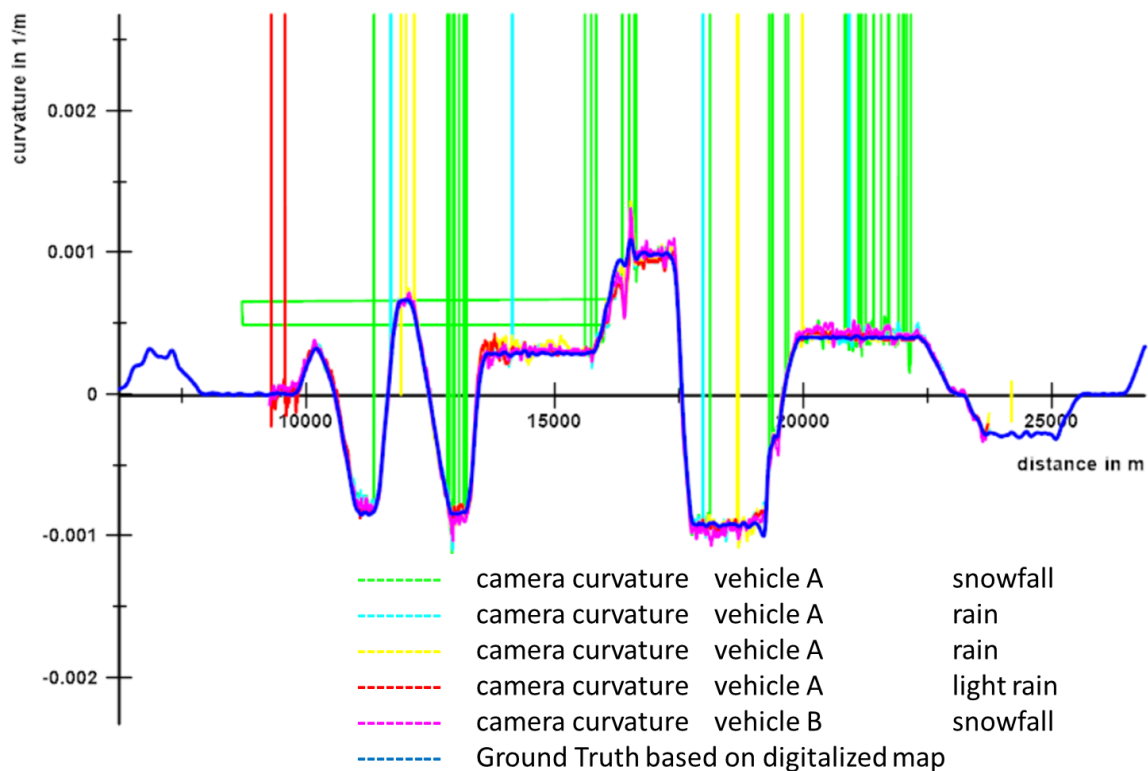


Figure 19: Camera performance characteristics based on Ground Truth under different weather conditions

Figure 19 shows the camera lane curvature compared to ground truth data in different weather conditions. As ground truth data, digital maps of the B19 Kempten - Immenstadt were generated as a look-up table with an absolute accuracy of $>5\text{cm}$ [SHL+2018;KHM+2019] in a curved reference grid as curve reference objects (CRO) [KHM+2019]. The contents of the digital maps include lane markings, marking widths, lane gradients, bends, crash barriers, guide posts and signs. The camera failures (drop-offs) due to environmental conditions such as rain, snow, sun and driving situations such as drop shadows under bridges/short tunnels and the accuracy of the curvature reproduction under different conditions can be recognized very well. Local curvature errors of up to 50% could be detected, which limit stable and good track guidance [HFS+2019]. This allows, among other things, the performance differences between the various cameras to be illustrated. As part of SensIndex, a method was developed at Kempten University of Applied Sciences to characterize sensor performance under real conditions and compare it with a reference condition (degradation). It was also noticeable that the sensor performance interacts with the vehicle and its movements. The same camera, for example in an SUV or sports car, can exhibit different performance. Figure 19 illustrates the large gap between the desired sensor behavior and the actual behavior of cameras, which can lead to reduced availability and reduced functional quality.

Research Needs:

This motivates the development of camera technology and fusion technologies in particular (chapter 5). Fusion with digital maps, for example, can help to bridge drop-offs and minimize curvature errors. Fusion with other sensors such as Lidar can also help to identify distance-to-line more robustly. Overall, this can significantly increase system availability and functional quality. For vehicle and system development, this results in the great need to be able to virtually simulate the sensor fusion based on scenarios in the early phase. This applies at vehicle level (system), at the level of the environment

simulation (sub-system) and at sensor level (component level). This requires highly accurate and detailed digital ground truth maps for the road, road infrastructure and the surrounding area in the simulation environment (see [chapter 4.1.3.](#) and [4.5.](#)). This requires highly accurate and detailed digital ground maps for the road, the road infrastructure and the environment in the simulation environment. These are not currently available in the necessary and desired level of detail, accuracy, time and cost for practical use. In addition, the uncertainties are not known.

4.4.3 Characterization of Video cameras

A detailed presentation of the camera impairments and their modeling is given in the doctoral thesis “Automotive Camera Modeling and Integration with Standardized Interfaces”, Saad, K., 2019. The relevant impairments are understood and their models can be abstracted mathematically, mainly with matrix-vector products. These matrix-vector multiplications also determine the run time behavior of the models.

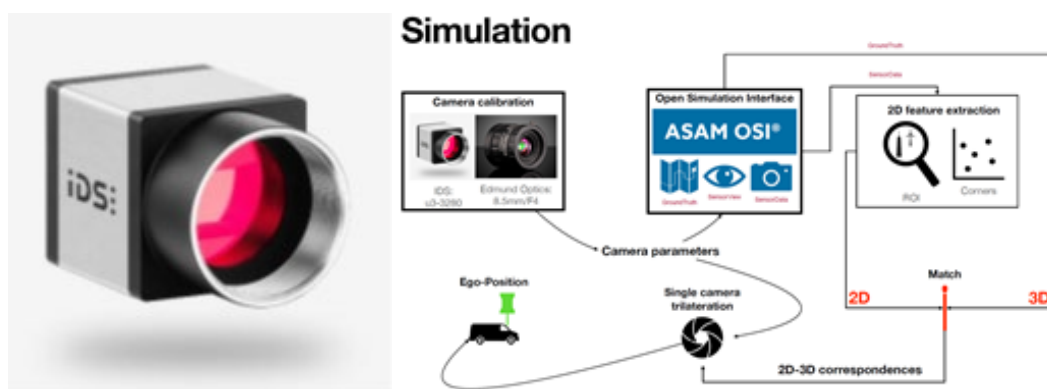


Figure 20: The used camera IDS3280-CP (right) and an architecture of the corresponding algorithm (left)

4.4.4 Models of video cameras

In literature, e.g. [DS2019], video camera models are classified into three different categories ideal, phenomenological and physical sensor models. Standardization for sensor models is still missing. On the one hand, the models for video cameras must consider the relevant impairments with regard to the usability of the processing for ADAS as well as the resources storage space and computing time. Another aspect that should not be neglected is standardized models that ensure (comparable) use in different simulation environments.

Some prototypical implementations of camera sensor models based on OSI and FMU were already done in the publicly funded projects PEGASUS and SetLevel.

4.4.5 Model validation

There are essentially two different methods to validate the models: either it is proven that the errors of the corresponding exposure values of the real and simulated images for the relevant class of possible images for the driving function to be verified differ negligibly, or it is proven that for the further processing in the fusion and afterward in the driving function the relevant image information, such as features or object lists, lead to the same reactions of the driving function. This is generally not easy and therefore individual effects are tried to be studied in detail, which is not easy because images with real cameras always have all the effects attached. Replicating the behavior of a camera depends primarily on the modeled ray tracer. However, a good introduction to the literature for the verification and validation of camera behavior models can be found in the papers [ED2022], [ESN2022a] and [ESN2022b]. In [GMS+2021] the focus is laid on efficient and realistic perception sensor models.

However, the impact of the camera behavior is studied in the VIVALDI project in the case study Campus University Sciences Kempten (TE building). The camera is put on the position like in the case study 3 with Ground Truth [598504.25 5285509.12 701.374] in UTM. This time the camera faces towards the building TE.



Figure 21 Position of the camera (red point) and the Campus University of Applied Sciences Kempten TE building on the left side (left figure). Measuring the camera height (middle figure) and taking images of the building TE in the campus courtyard (right figure)

Features of the SensorData image of the real camera are compared with the SensorData image of the simulated camera with camera behavior model impairments.

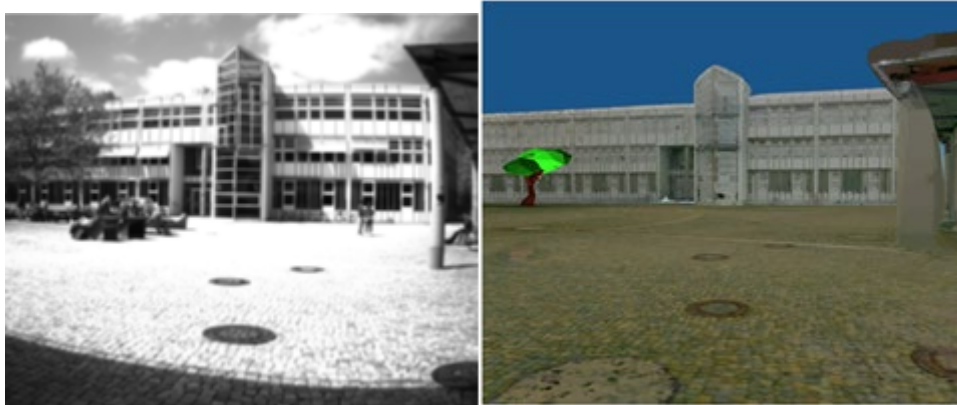


Figure 22: Sensor Data of the real camera of the building TE (left). SensorView of the simulated environment of the building TE (right).

The SensorView image of the environment is based on a lidar scan of the campus. The SensorData of the image is generated from the SensorView of the camera. Finally, some features in the SensorData image are automatically identified.

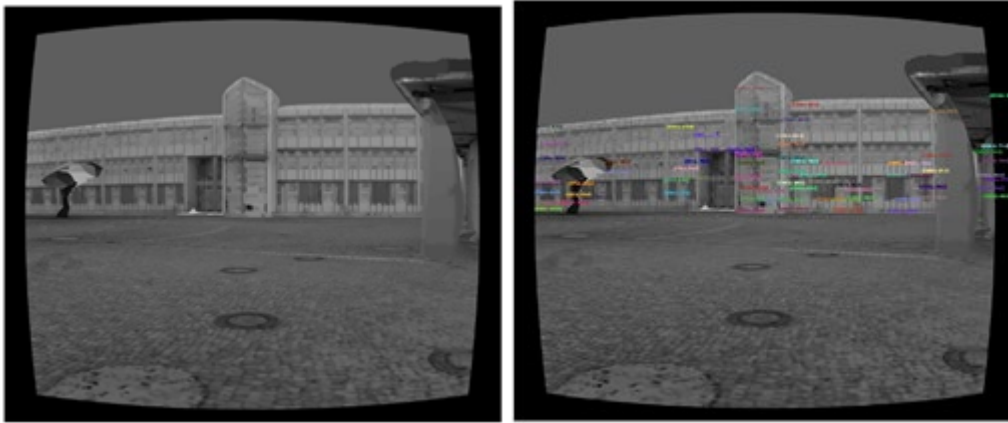


Figure 23: SensorData of the simulated environment of the building TE (left). Features of the SensorData of the simulated environment of the building TE (right).

The camera position is then calculated with the corresponding Ground Truth data of the city of Kempten database. Finally, four positions can be compared: the Ground Truth position, the estimation on the basis of the IDS3280 image, the estimation on the SensorView image and the SensorData image of the simulated environment are calculated. Then, the Euclidean Distance between GroundTruth on the one side and IDS3280 image, SensorView image and the SensorData image on the other hand are calculated for each coordinate direction.

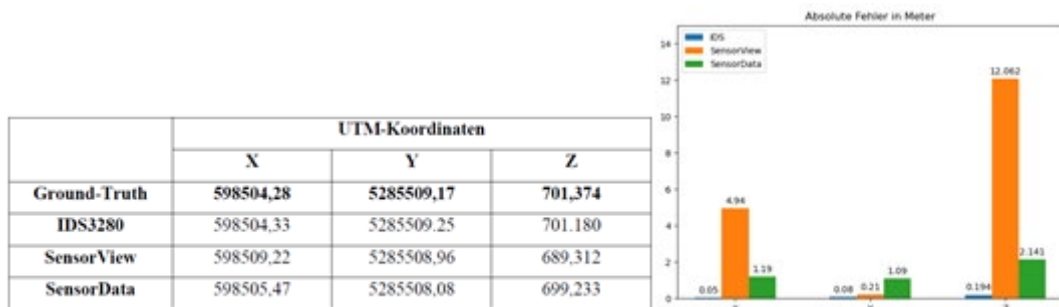


Figure 24: Measured and calculated coordinates of the camera position (left). Differences in the Euclidian distance in x-, y- and z-direction (right).

In the very first application of the localization method, the impact of the camera behavior model reduces the error in x from 4.94 m to 1.19 m and in z from 12,062 m to 2.141 m whereas the error in y increases from 0.21 m to 1.09 m. The error in respect to the taken picture is comparable small with in x: 0.05 m, in y: 0.08 m and z: 0.194 m. These results need to be analyzed in more details.

On the bases of these results, the research goal of the project VIVALDI “how accurate is accurate enough” of 5 cm seems possible. However, there are still analyzes to be made in the future. The results are promising and further case studies are therefore recommended.

4.5 Digital Maps

As additional virtual sensor information, digital maps offer immense potential in the perception of the environment and in the precise localization of vehicles and traffic objects. Sensor fusion with digital maps enables a preview with regard to the static road and road infrastructure and can help to bridge sensor drop-offs and minimize errors. Overall, this can significantly increase system availability and functional quality. In particular, static objects such as roadways, road markings, footpaths and cycle paths, signs, traffic lights, buildings and vegetation and much more can be labeled and stored here in geo-referenced form. If an environment sensor detects certain static objects on the map, the vehicle can use these landmarks to locate itself and other road users. However, this requires high-resolution maps (HD maps) and, above all, accurate maps in order to reliably locate your own vehicle and other road users. Only then can the scenario be interpreted reliably. For example, objects must not be incorrectly located in the adjacent or even opposite lane. It is therefore very important to be aware of the uncertainty of this source of information and to take it into account when recognizing and interpreting the environment.

4.5.1 Principles of Operation

Digital maps, also called digital cartography, is the technology of measuring, creating and using maps for different applications, in this context for ADAS/AD functions. The primary challenges of digital maps is the accurate measuring and creation, including the accurate representations of a particular geo-referenced areas and objects such as roadways, road markings, footpaths and cycle paths, signs, traffic lights, buildings etc. Digital maps are usually measured at great expense using measuring vehicles or satellites based on various measuring principles such as camera, lidar or radar and are usually part of the geographical information system (GIS). There are a wide variety of digital maps such as topographical maps, building maps from the cartographic office, road maps, hiking maps etc. for many different applications. These are also based on positioning using GPS (Global Positioning System). From this, digital maps are generated for a wide variety of target applications, which are based on very different requirements such as content, resolution, and accuracy.

Digital maps are offered by various suppliers for automotive applications and have been used in navigation systems for a long time. For example, the online geodata service HERE was acquired in 2015 by the three German car manufacturers Audi, BMW and Mercedes-Benz Group (previously Nokia, Windows). HERE Technologies is now one of the world's leading location data and technology platforms and also offers digital HD maps as live data and addresses applications in the context of ADAS/AD.

4.5.2 Artifacts

Digital maps have different strengths compared to other sensor modalities as well as weaknesses. The characteristic of a digital map depends on the application requirements and the measurement and labeling technologies. When measuring digital maps, the basic problem is incorrect position measurement, e.g., using satellite-based GPS or other global positioning systems. Digital maps are usually measured using special measuring vehicles and device. Differential GPS supports IMUs (Inertial measurement Units) to measure the vehicle motion (3x translation, 3x rotation). This allows the transformation of sensor data such as point clouds from a lidar to be georeferenced. The problem here is obvious. The position measurements are subject to significant errors due to limited satellite coverage such as bridges and urban canyons. While with good satellite coverage an accuracy of +/- 2 cm can be achieved, the error increases to 30 - 130 cm even with moderate satellite coverage. The poor satellite coverage leads additional to position drifts for IMUs. Absolute and relative errors are passed on in the measurement sequence, which can lead to significant overall errors. The transformation of the sensor data is equally affected. This leads to significant uncertainties in the digital maps. Multi-path

propagation of the GPS, sensor latency times, synchronization errors of the sensors and much more also influence the absolute and relative error of the map. This can lead to simple offset errors, but also to distortion errors over longer distances. All digital cards for the different applications (chapter 1.5.1.) are more or less affected by these errors and are therefore difficult to bring together.

As part of a research project, reference measurements of landmarks in urban areas at numerous intersections were carried out for validation. Due to the development, very different and challenging measurement conditions arose. Figure 18 shows a GeoTIFF of Durlacher Allee in Karlsruhe. Static reference measurements of landmarks such as traffic lights and signs are drawn here. Here we can see clear position deviations and errors in the GeoTIFF. GeoTiffs also have corresponding offset and distortion errors in their measurement chain. These are approx. 30 - 100 cm in this sample.

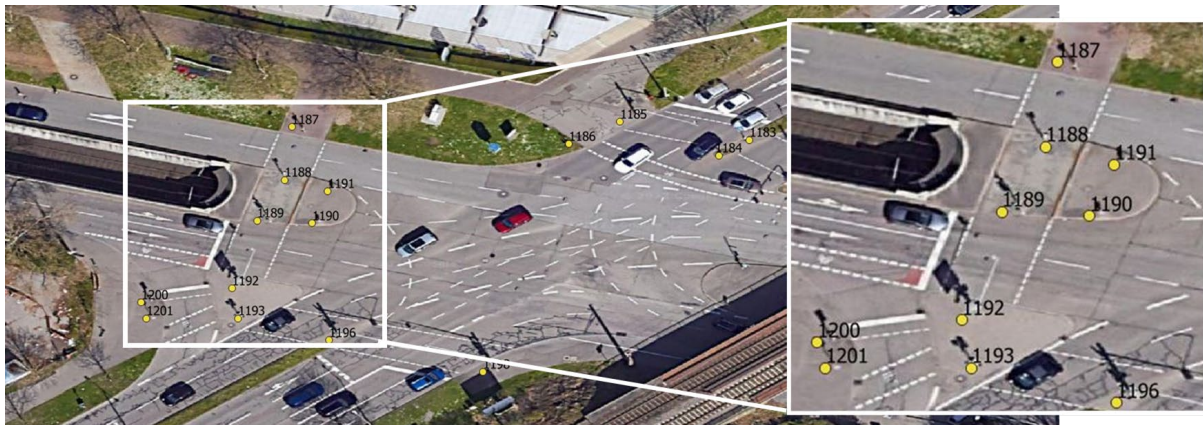


Figure 25: Precise static reference measurements of landmarks as ground truth in comparison to geo-referenced images (GeoTIFF)

Therefore, the following strengths and weaknesses are formulated in a quite general manner:

Strengths of digital maps:

1. **Absolute reference:** Maps represent the absolute static reference that can be checked at any time and to which other measurements can be based.
2. **Defined Resolution:** Digital maps usually have a clearly defined resolution/grid. This helps with location.
3. **Defined Object range:** Digital maps enable a clearly defined range of objects.

Weaknesses of digital maps:

1. **Error sensitivity:** Digital maps are prone to errors due to error propagation.
2. **Accuracy difficult to prove:** Digital maps are very difficult to check for accuracy and uncertainty due to the measurement problems. There are no suitable procedures here.
3. **Update capability:** Digital maps are difficult to keep up to date. Structural changes cause them to lose validity or accuracy.
4. **Fixed Resolution:** The fixed resolution of digital maps for different applications makes multiple use difficult.

If there are errors in the digital map, it is difficult to accurately match measured scenarios such as positions and movements of various traffic objects.



Figure 26: Matching of digital map (static ground truth) with measured scenario (dynamic ground truth)

4.5.3 Environment Simulation based on Digital Maps

Digital maps can be converted into an environmental simulation. There are various standards for this, such as OpenDrive or OpenCRG (Curved Reference Grid), but also numerous proprietary formats from various simulation environments such as IPG CarMaker Road5 etc. All formats have certain restrictions and were originally developed for other applications. For example, OpenDRIVE was developed as a road format for building routes in driving simulators. Although extensive routes and environments can be created in OpenDRIVE, many parameters such as track width, road markings, road inclination and curvature are only valid for entire sequences and can lead to limitations in accuracy. OpenCRG was developed as surface models for the tire-road system to model road excitations. OpenDRIVE cannot display road networks today. What remains is that the environmental models for sensor simulation are incomplete regarding material properties and standardized 3D assets.

4.5.4 Research needs

The limitations of digital maps are obvious. On the one hand, these are created for very different applications and associated requirements. Resolutions, contents and absolute accuracies vary considerably. On the other hand, the maps are not suitable as an environment model in the sense of your ground truth. This requires special, very complex track/route measurements and their transfer into the simulation. Such projects, with purely geometric representation, usually take more than 12 months, even for 20-30 km. The material properties from Chapter 1.3.1 are not even taken into account, here. This creates enormous barriers for the development.

The following research questions arise here:

- Which reference/ground truth can be used to determine and correct the uncertainties in measurements and digital maps?
- How can different sources of digital maps re-used and be brought together as precisely as possible to form a ground truth environment model?
- How can the accuracy of the environment model be controlled and optimized?
- How can the process of creating an environment model be automated and significantly accelerated from > 12 months up to few days?
- How can you create 3D objects (meshing) from measured point clouds as automatically as possible?
- How can you easily and quickly add the material properties that are important for the sensor models to the environmental models, if possible in an automated way?
- What can a standardized data and exchange format for the objects and environment look like?

5. Sensor Fusion and Classification

5.1 Introduction

This Chapter addresses the fundamental challenge of bounding the uncertainty of perception of automated vehicles. To this end we propose to inductively derive what we call “*guarantees for bounding uncertainty*” along the structure of the perception chain. This calls for major research in providing characterizations of all elements of the perception chain to provide local quality guarantees of perception, which justifies to designate such components as “*sufficiently perfect components of the perception chain*”. This seemingly contradictory term is used to on one side acknowledge the fact, that e.g., all sensors suffer from severe degradation in adverse weather conditions, but on the other hand, they come with a characterization of such adverse conditions, and provide, possibly weak, quality guarantees in non-adverse conditions. The key role of sensor fusion is to alleviate weak or even lacking quality guarantees of one sensor by stronger guarantees for bounding uncertainty of other sensors. While there is a rich literature and strong body of industrial experience for building sensor fusion systems, we propose the following research challenges:

1. How to provide quality guarantees for all types of sensors, in spite of
 - a. Intrinsic uncertainties, resulting from the structure of the sensor and its integration in the signal chains, such as distortions, disturbance sensitivity
 - b. Extrinsic uncertainties that come from the external conditions in which the sensor is embedded during operation, such as accuracy of attributes such as velocity, locality, ... which are key for assessing confidence and accuracy of measurements
 - c. Uncertainties stemming from employed measurement techniques, such as multipath phenomena, reflection phenomena, occlusion, ...
2. How to provide quality guarantees for AI-based classification components
3. How to derive strongest quality guarantees when fusing such information
4. How to dynamically adjust the structure of the perception allowing to exploit additional fusion steps until uncertainty is bounded to a level allowing safe execution of selected maneuvers.

This chapter is structured as follows. Section 5.2. summarizes the state of the art in sensor fusion, based on the survey [Fay et al 2020]. Section 5.3 addresses the fundamental challenge of deriving strategies for safe manoeuvre execution in the presence of possibly partial or imprecise or even incorrect beliefs about the environment of the ego system, and derives from this a meta requirement on the quality of perception of the perception chain. Section 5.4. proposes a functional architecture for dynamically adjusting the quality of perception to the quality requirements of what we call relevant objects in the environment of the ego system, if this is at all possible. Together with the monitoring components on adverse conditions anchored in this architecture, this allows to determine if the required quality level has been achieved, or else degrade to emergency rescue manoeuvres. Section 5.5. proposes research directions for providing quality guarantees for AI-based classification components. Section 5.6 proposes a mechanism for propagating quality assurances through sensor fusion components based on Dempster Shafer Theory, following the recommendation of [Mu et al 2010].

5.2 State of the Art

It is well known and state of industrial practice to use sensor fusion to mitigate weakness or limitations of sensors for tasks which are fundamentally required in SAE levels 2-5 [SAE2021] and in general in highly automated systems. We refer to the excellent survey [FJG+2020] for a discussion of the state of the art. The following tables taken from [FJG2020] summarizes key use cases for sensor fusion as well as classical as well as AI-based sensor fusion algorithms.

Study	AV Application	Fused Sensors	Limitations without Fusion	Fusion Advantages
[34–36]	Pedestrian Detection	Vision and LiDAR	Sensitive to illumination quality; Night vision difficulties by vision camera only Low resolution of LiDAR 3D scene reconstruction when used alone.	Ability to measure depth and range, with less computational power; Improvements in extreme weather conditions (fog and rain)
[37–42]	Pedestrian Detection	Vision and Infrared	Night vision difficulties with vision camera only; Thermal cameras lose fine details of objects due to their limited resolution.	Robustness to lighting effects and nighttime detection; Infrared camera provides distinct silhouettes of objects; Ability to operate in bad weather conditions.
[43–46]	Road Detection	Vision and LiDAR	Illumination and lighting conditions; High computational cost for vision depth measurements; Limited resolution and range measurements by LiDAR; Sparse and unorganized point cloud LiDAR data	Road scene geometry measurements (depth) while maintaining rich color information; Calibration of scattered LiDAR point cloud with the image
[47]	Road Detection	Vision and Polarization camera	Sensitive to lighting conditions; Lack of color information	Polarized images enhance scene understanding, especially with reflective surfaces.
[48–50]	Vehicle Detection Lane Detection	Vision and Radar	Low resolution of radar. Camera needs special lenses, arrangements, and heavy computation to measure distance.	Measure distance accurately; Performs well in bad weather conditions; Camera is well suited for lane detection applications
Study	AV Application	Fused Sensors	Limitations without Fusion	Fusion Advantages
[51]	Visual Odometry	2D Laser scanner and Vision	2D scanners can miss detection of objects in complex environments; 2D images are insufficient for capturing all the features of the 3D world.	Fusion of vision and 2D scanners can replace the need for 3D LiDAR, and hence reduce price and computation load.
[52,53]	SLAM	Vision and Inertial Measurement Unit	Illumination and lighting conditions by the camera; Camera suffers blur due to fast movement; Drifting error for IMU	Improved accuracy with less computational load; Robustness against vision noise, and corrective for IMU drifts.
[54]	Navigation	GPS and INS	GPS outage in denied and canyon areas; Drift in INS readings	Continuous navigation; Correction in INS readings
[32,55]	Ego Positioning	Map, vision, GPS, INS	GPS outage; INS drifts; HD map accuracy; Visibility of road markings	Accurate lateral positioning through road marking detection and HD map matching.

Table 1 Summary of AV applications, limitations of sensors, and advantages of sensor fusion, [FJH+2020]. For references listed in this table, please refer to Annex 1.

Table 2 below summarizes “classical” algorithms for sensor fusion, as contrasted with deep-learning based approaches surveyed in their Table 3.

Algorithm	Characteristics	Advantages	Disadvantages	Applications Areas	Level of Fusion
Statistical Methods	Utilized to enhance data imputation using a statistical model to model the sensory information [64,67]	Can handle unknown correlations; Tolerant [68,69]	Limited to linear estimators; Computation complexity is high [65]	Estimation	Low [70]
Probabilistic Methods	Based on probability representation for the sensory information [64]	Uncertainty in the provided information is handled. handles nonlinear systems (particle filter, UKF, ...)	Requires prior knowledge of systems model and data	Estimation/ Classification	Low→Medium [70]
Knowledge-based Theory Methods	Utilizes computational intelligence approaches inspired by human intelligence mechanisms. [71]	Handles Uncertainty and imprecision; Ability to handle complex nonlinear systems [72]	Depends on the expertise knowledge and extraction of knowledge	Classification/ Decision	Medium→High [70]
Evidence Reasoning Methods	Depends on the Dempster combination mechanism to implement the model [71]	Uncertainty degree is assigned to the provided information. Identification of conflicting situation. Modeling of complex assumption	High computation complexity. Require assumption of evidence.	Decision	High [70]
Interval Analysis theory	Shares the operating space in intervals [73]. Constraint satisfaction problem [74,75]	Guaranty integrity. Ability to handle complex nonlinear systems	Discretization of the operating space. High computation complexity.	Estimation	Low

Table 2 A comparison between traditional sensor fusion algorithms, their advantages, disadvantages, applications, and fusion level, taken from [FJG+2020]. For references listed in this table, please refer to Annex 1.

DL Algorithm	Description	Applications
Convolutional Neural Network (CNN)	A feedforward network with convolution layers and pooling layers. CNN is very powerful in finding the relationship among image pixels.	Computer Vision [82–84]; Speech Recognition [85]
Recurrent Neural Network (RNN)	A class of feedback networks that uses previous output samples to predict new data sample. RNN deals with sequential data; both the input and output can be a sequence of data.	Image Caption [86]; Data Forecasting [87]; Natural Language Processing [88]
Deep Belief Net (DBN)	Multilayer generative energy-based model with a visible input layer and multiple hidden layers. DBN assigns probabilistic values to its model parameters.	Collaborative Filtering [89]; Handwritten Character Recognition [90]; acoustic modeling [91]
Autoencoders (AE)	A class of neural network that tends to learn the representation of data in an unsupervised manner. AE consists of an encoder and decoder, and it can be trained through minimizing the differences between the input and output.	Dimensionality Reduction[92]; Image Retrieval [93]; Data Denoising [94]
Transformer	A class of neural networks that consider contextual relationships of the input data using the self-attention mechanism. Transformers originate from sequence processing and were recently adopted by computer vision applications.	Computer Vision; Natural Language Processing; Audio-Visual Speech Recognition

Table 3 A summary of deep learning algorithms, their main properties, and applications, taken from [FJG+2020] and extended. For references listed in this table, see Annex 1

5.3 Safe Maneuver Execution in the presence of possibly incomplete of incorrect beliefs

The selection of manoeuvres and, in general, selection of strategies of an HAV is inherently restricted by the imperfections of the “Lagebild” computed by the perception chain. This Lagebild, often also referred to as the world model of the HAV, is bound to deviate from what in the AI community is called ground truth. Problems causing the withdrawal of the license to run robotaxis in the city of San Francisco³⁰, or as leading to the recall of more than 2 million Tesla’s³¹ and as evidenced by the related accident analysis carried out by the NHTSA give ample evidence of situations where the perception chain may miss or incorrectly classify objects in the environment of the ego system. To stress this difference, we referred to in [DHS+2024] to the Lagebild as *Descriptive Beliefs* of the ego system, which in general will differ from the ground truth of the state of the environment of the HAV. This entails, that there is a constant need for monitoring the plausibility of the computed Lagebild, the need for enabling belief revision if new perceptions are inconsistent with previous observation, as well as the need to regularly adapt components of the perception chain by learning from mis-classifications and mis-observation, and adapting sensor- and/or classifier characterizations by addressing the causes of misinterpretation or inaccuracies, requiring over-the-air update capabilities.

We present in this section a meta-requirement on the perception chain, which, when established by the components of the perception chain through propagating robustness guarantees, enforces that what the ego car believes to be true about its environment, and the actual ground truth, *rarely* differ for all aspects of the environment which are *relevant* for ensuring the safety of the ego vehicle. The vague term “rarely” is given a precise meaning in the formal definition of this meta-requirement on the quality of the perception chain given below. The term “relevant” is also given a more precise meaning below; intuitively, an object in the environment of HAV is *relevant* if missing or mis-qualifying this object or its attributes could lead to accidents. Our meta-requirement on the perception chain demands that the perception of “relevant” environmental objects errs with rate at most r . This rate of misperception can then be used in the system hazards analysis to quantify the resulting risks of such misperceptions, taking into account the hazard severity.

Within the scope of this document, we only sketch the underlying formal semantics necessary for giving meaning to this formal requirement specification. We refer the reader to [DHS+2024] for a formal definition of the underlying transition system semantics.

Our mathematical model uses, as e.g., [GAL2015], labelled occupancy grids for fusion of sensor data from radar, lidar, video, etc, and as interface to learning algorithms-based components for classifying objects in the environment of the ego-vehicle according to a partially ordered ontology. We assume that each object in this ontology comes with class definitions characterizing both static and – if applicable – dynamic aspects, such as ODD dependent models for typical traffic behaviour (e.g., characterizing variations in lateral and longitudinal acceleration of vehicles in a neighbouring highway lane, or of pedestrians in an urban pedestrian crossing). Such knowledge is exploited in prediction engines for online prediction of the evolving traffic. These also give formal meaning to the vague term of “relevant” environmental objects: the prediction engine provides feedback regarding criticality of queries for individual fields in the occupancy grid: errors arising from misperceptions of objects only count, if they relate to such criticality.

For each sensor our approach requires the capability to identify harsh environmental conditions

³⁰ ORDER OF SUSPENSION of October 24, 2023: The Autonomous Vehicle Testing Permit-Driverless Vehicles issued to Cruise LLC is hereby suspended immediately for violations pursuant to California Vehicle Code 38750 (d)(3). and California Code of Regulations (CCR), Title 13, Division I, Chapter I, Article 3.7, Section 227.42 (b)(5) and (c).

³¹ *New York CNN* — (13.12.2023)

Tesla is recalling nearly all 2 million of its cars on US roads to limit the use of its Autopilot feature following a two-year probe by US safety regulators of roughly 1,000 crashes in which the feature was engaged.

(where no sufficiently tight bounds for risks of misperceptions can be given), and exploits this information in mechanisms for sensor fusion. Finally, we propose a “safety net” reducing likelihood of misperceptions, in declaring “blindness” for perceptions, where neither the evidence for existence of an object nor the evidence for its absence is sufficiently strong: such declaration of “blindness”, if sustained over several cycles for relevant objects in the environment of the HAV, should automatically induce minimal risk manoeuvres (as does any detected usage of the HAV outside the allowed ODD), or require takeover of mission management by an outside agent.

We assume as given a partially ordered ontology containing all relevant environmental artifacts for HAV. Initiatives towards identification of such an ontology are currently part of a number of R&D projects on HAV, including PEGASUS and VVM. The ordering relation reflects the degree of precision of classification of objects, with \perp (read: bottom) representing inconsistency, and \top (read: top) denoting no knowledge. The ontology (c.f. e.g. [SCH+2021]) contains different kinds of artifacts, e.g., relating to weather conditions (rain, snow, ...), road configurations (x-lane highway, T-type intersection, ...), road conditions (dry, icy, ...), roadside infrastructure (traffic signs, traffic lights, ...), traffic participants (car, truck, pedestrians, animals, ...), and surroundings (trees, buildings, ...). Objects in different categories form separate sublattices, which are turned into a complete lattice by adding a new top element, meaning that the type of this object is completely unknown.

With each object o in the ontology, we assume as given a specification of its HAV-related aspects through a class definition $cl(o)$. For artifacts of type road configurations, this includes a specification of all geometric aspects including slope, number of lanes, width of lanes, etc. Road configurations are built from segments of a parametric length. An operational design domain (ODD) is defined by constraints on types of road-configurations, and constraints on prevailing environmental conditions and road conditions.

For each object veh of type *vehicle*, attributes of $cl(veh)$ include the type of an instance of class road configuration, on which the vehicle is currently located, as well as its position. Moreover, each vehicle maintains its beliefs about its environment in appropriately typed attributes. For example, a car, *ego*, driving on a country road may believe the road surface is dry, that there is an obstacle in 250 m distance ahead blocking the lane, and that some vehicle of unknown type is approaching on the opposite lane. Importantly, $cl(veh)$ contains as well a characterization of ODD-dependent dynamics of veh . We assume that these are given by probabilistic hybrid automata (PHA, see e.g., [SPR2000] and extensions thereof such as [O22]), where mode changes are triggered based on believed changes of road configurations, weather conditions, road conditions, and observations of surrounding traffic and roadside infrastructure,

A key point to be exploited is that behavioural models are increasingly unconstrained along the generalization hierarchy: e.g., for class *vehicle*, the associated probabilistic hybrid automaton is constructed from those of the next level of specialization by introducing a new start state, branching non-deterministically into the entry states of the PHA of cars, two-wheelers, trucks, emergency vehicles, etc. Similarly, we assume such models for pedestrians, animals, obstacles etc. Based on this, we can define a mathematical model of traffic evolution from the perspective of a given ego-vehicle:

We define the electronic horizon of the ego vehicle to be the best-case range of perception of its on-board sensory system, and for simplicity of exposition assume this to be given by a rectangle aligned to the current pose of the ego vehicle centred at the geometric centre of the ego vehicle. The mathematical model is an infinite-state transition system, whose state space is constructed as follows: It contains for each point in time t all instances of road segments extending from the instance on which the ego car is currently positioned, completely covering its electronic horizon. For each of these road segments, it contains for time t position and speed of all vehicles on the road segment, as well as road- and weather- conditions, positions of pedestrians, animals, obstacles, etc.

The (dense time) transition relation of the mathematical model is determined by following a driving strategy based on ego’s beliefs, where the environmental constraints are instantiated randomly according to the ODD and the dynamic models associated with other traffic participants are

considered.

We refrain for space reasons from giving the formal definitions (see [DHS+2024]), and refer to this transition system by $TS(ego)$. A state of $TS(ego)$ is given by a valuation of all its observables. A run of $TS(ego)$ defines for each observable its evolution over time, thus including trajectories of all vehicles, pedestrians, animals, obstacles within ego's electronic horizon, including the evolution of the beliefs of the ego vehicle, and the evolution of its perceived road segments. We call $RUNS(TS(ego))$ the set of all runs of $TS(ego)$.

We can now formalize the meta requirement for the quality of the perception chain, up to the (to be defined) defined notion of *relevance*, in a probabilistic linear time temporal logic, with observables defined by valuations of all attributes of all instances of classes within the electronic horizon of ego, over a typed first-order signature induced by the types of attributes in the ontology. Ideally, for each point in time, the ground truth of all relevant objects in ego's electronic horizon, in particular position and speed of surrounding traffic participants, road- and weather conditions, coincide exactly with ego's beliefs about these objects. We must relax this unachievable ideal by considering standard measurement errors, and allowing classifications to be vague, as long as they are correct with respect to the ordering relation in the ontology, i.e., the ground truth classification is a specialization of the believed classification. While mis-classifications and misperceptions will occur, we want these to be bounded by a given probability r . We assume for each object o existence of a metric d_o to measure the distance between ground truth and beliefs of o , and an instant-dependant safe level of discrepancy between ground-truth and beliefs of o δ_o . We assume a maximal time period Δ sufficient to take safe manoeuvre decisions for the ego vehicle. We want the discrepancy between beliefs and ground truth to be bounded by δ_o , up to a probability of at most r that this requirement is violated. This leads to the following formal requirement on the confidence of perception:

$$Safe-perception(ego) = \forall o \in observables(TS(ego)) :$$

$$\begin{aligned} & \Box(relevant(o) \Rightarrow \\ & P(\neg(\Box_{\leq \Delta} d_o(o, belief(o))) \leq \delta_o) < r) \end{aligned}$$

Equation 1 Meta requirement for quality of the perception chain

There is a recursive dependency of the notion of *relevance* in the above formula on the capability of the perception chain to meet this meta requirement: assuming that the world model (or: "*Lagebild*") computed by the perception chain meets this meta-requirement, then this will be input to the *prediction engine* to compute the likely *evolution* of the *Lagebild* over time, in order to determine the next maneuver m to be executed by the ego system in order to meet its goals, notably including safety. This prediction highly depends on the correct classification of the objects in the ego-system's environment, as these determine the associated model of their expected dynamics in the current scenario. The prediction engine will compute guards $pre(m)$ over the signature determined by the ontology for maneuver m to be safe, annotated by requirements on level of confidence on existence of objects and classification of objects as well as quantification of guaranteed maximal measurement uncertainty, for each relevant maneuver m in the current *Lagebild*, across the ODD. A perception of an object o in the environment is *relevant at time t*, only if it occurs in the precondition of one of the maneuvers proposed by the prediction engine at time t.

We propose a uniform quality measure for different type of sensor systems by assessing their capability to detect all relevant objects in what is often called the occupancy grid enclosing the ego vehicle. We take a suitable 3d extension of the electronic horizon of the ego vehicle, assume a sufficiently finely grained, not necessarily homogeneous partitioning of this 3d space, and refer to this as *occupancy grid(ego)*. We view sensors as labelling partitions of the occupancy grid, providing information about whether a given sensor has identified some object in this grid partition, depending on sensor type filled with additional information such as speed (for radar), temperature (for infrared), gray-scale

distribution of pixels for video cameras, typically as confidence levels, e.g. regarding the likelihood of some object being located in this grid partition.

We provide a formal quality requirement on sensors by adapting the safe perception requirement in Equation 1. Specifically, for a given position pos of sensor s on vehicle ego , let us denote by $visible(s, pos)$ the coherent subspace of the occupancy grid that is observable by sensor s when mounted at this position. Then for any relevant object o , and any property $label(o)$ discernable by s in this position, we would want the sensor s to almost always, up to some bounded risk r_s , correctly label the grid field with $label(o)$ at time t , when o is at a visible grid partition in ground truth, as determined by the state of $TS(ego)$ ³² at that point in time:

$$\begin{aligned} & Safe_labelling(s, pos) \equiv \\ & \forall o \in observables(TS(ego)) \forall p \in occupancy_grid(ego) : \Box (relevant(o) \wedge o \text{ is} \\ & \text{ at } p \wedge p \in visible(s, pos) \Rightarrow P(\neg(\Box_{\leq \Delta} d_{o,s}(o, label(o)) \leq \delta_{o,s})) < r_s \end{aligned}$$

Equation 2 Quality requirement on sensor components

In this formula, we interpret the predicate $visible(s, pos)$ dynamically, providing for subspaces of the occupancy grid to be temporarily blocked through artifacts such as other vehicles, or buildings, etc. Moreover, since phenomena such as glare, fog, heavy rain etc. will have strong negative impact on achieving a sufficiently precise distance between belief and ground truth, we propose to characterize for each sensor what we call *adverse conditions*. Only by using advanced sensor systems (such as radar) capable of identifying most ghost objects, and explicating adverse conditions, can we obtain a sufficiently tight bound on the probability of incorrect labelling of the occupancy grid, which then can enter a hazard analysis to determine the induced risks of such a misperception. The downside is the need to then also identify such adverse conditions reliably: In this paper, we integrate such adverse conditions in our ontology, and then use the full power of the perception chain to learn sufficiently reliable classifications of adverse conditions, which can then be integrated as monitors for on-line checking. As discussed in Section 4, we propose to use learning techniques from field observations to characterize adverse conditions for each sensor s and to derive such quality guarantees. We now weaken the formula of Equation 2 by weakening the criteria for sufficient precision of beliefs in that no promises regarding precision are made when adverse conditions ad occur:

$$\begin{aligned} & Safe_labelling(s, pos, ad) \equiv \forall o \in observables(TS(ego)) \\ & \forall p \in occupancy_grid(ego) : \Box (relevant(o) \wedge o \text{ is at } p \wedge \\ & p \in visible(s, pos) \Rightarrow P(\neg(\Box_{\leq \Delta} (d_{o,s}(o, label(o)) \leq \delta_{o,s})) \text{ unless } ad)) < r_s \end{aligned}$$

Equation 3 Quality criteria for sensor components with characterization of adverse conditions

Level 4 and 5 HAV must guarantee by construction what we call *sensor completeness with maximal error probability r* : for all runs of $RUNS(TS(ego))$, there exists a “sufficiently dense” sequence of time instances $(t_j)_{j \in \mathbb{N}}$ s.t. for all t_j and all relevant fields p of the occupancy grid, there must at least be one positioned sensor operating in non-adverse conditions, such that p is visible for this sensor, and the probability of misperception by this sensor is smaller than r , unless there are multiple positioned sensors all operating in non-adverse conditions, which ideally are stochastically independent under such conditions; in this case the products of their error probabilities must be smaller than r . The term “sufficiently dense” depends on the current ODD. We note that this definition is in fact recursive: Relevance depends on the set of actions we are considering possible, which in turn bring us into new situations, where the same applies. Informedness deficits can thus be resolved by more sensors, which

³² See [DHS+2024] for a formal definition of this transition system

in turn lead to more permissible actions and consequentially to more possible follow-up situations increasing the relevance set and calling for informedness, or by confining the number of actions and thereby the relevance set now and in the future.

For each type of output of a sensor, the probabilistic guarantees given under non-adverse conditions and known levels of controllable disturbances of the ego vehicle will include for each type of object of the representation at that sensor output

- a quantification of existential uncertainty,
- a quantification of the confidence in the categorization of the type of object, if applicable,
- a quantification of the maximal degree or the variance of imprecision of physical attributes of such artefacts observable by this particular sensor system at this particular interface.

for environments which meet specified non-adverse conditions and ODD specification.

We require all input data to be time-stamped and all output data to be labeled by the time stamps of all processed input data. Quality statements refer to such time stamps, and thus make guarantees about the perception of the environment of the vehicle at the most recent processed input data. **We collectively refer to these guarantees as *guarantees for bounding uncertainty*.**

5.4 A reference architecture of the perception chain supporting uncertainty propagation for relevant objects

This section proposes a functional reference architecture for the perception chain allowing to provably bound the risk for misperceptions. The following principles guided the proposal of this architecture:

It must be possible to bound the contributions to risks for each component of the perception chain.

This entails the need to specify for each element in the perception the conditions under which this element can be “trusted”. Violations of such conditions must be monitored and –if persistent– must induce declaration of “blindness”, automatically inducing minimal risk manoeuvres or delegation of HAV control to an external agency.

The reference architecture must provide guidance as to the degree of precision of perception required for current manoeuvre decisions.

The reference architecture must support the trade-off between maximal availability and maximal safety.

A key element of our approach is that we dynamically adjust thresholds for declaring blindness, thus allowing optimal trade-offs between availability and safety. To this end, we propose, much as in [HMG+], a dynamically configurable perception chain, allowing to optimize resource usage, but in our approach in order to reach the level of precision in identifying relevant objects in the environment required by the criticality of misperception.

We now elaborate on the functional reference architecture of the perception chain and its components, focusing on their functionality and quality requirements. Figure 27 below shows the overall architecture, depicting three stages of the perception chain: the sensor layer, the sensor fusion layer, and the world model layer. The latter two stages can be passed multiple times because of the dynamic reconfiguration capabilities offered by the interconnection network, such as allowing to pass fused information on to classifier components to reduce uncertainty. We note, that we do not assume any implementation of the Prediction Layer and the Decision Layer. Instead, we are relying only on the following specification of their interface, thus enabling integration of our approach with *any* implementation of the prediction- and decision layer meeting this interface specification:

- The Perception Chain promises to deliver a Lagebild at time $t+1$, where all quality requirements for all relevant objects provided as input to the Perception are met (or else the perception chain declares temporary blindness), provided it receives from the Prediction- and Decision Layers a

list of all relevant objects together with requirements on confidence and precision of measurements at time t ;

- The Prediction- and Decision Layers promise to provide to the Perception Chain a list of all relevant objects together with requirements on confidence and precision of measurements at time $t+1$, as long as they receive a Lagebild where all quality requirements for all relevant objects provided as input to the Perception Chain are met at time t .

Indeed, by propagating down along the perception chain the information, which objects in the environment are relevant, what degree of confidence must be achieved regarding their existence and their classification, and what maximal degree of uncertainty is tolerated in determining physical properties such as distance and velocity, we can reserve computation resources for exactly those sensors, sensor fusion components, and classifier components required to achieve these quality criteria. Each stage determines for such relevant objects the currently achieved degree of uncertainty, allowing the control unit of the interconnection networks to determine required interconnections for further processing fused sensor and classification data until these reach the quality requirements imposed by the prediction engine. We note that this architecture thus supports plausibility checks for AI based components exploiting time-redundancy; moreover, by integrating classifiers covering multiple levels of the ontology, it is also possible to detect inconsistencies in classifiers if the determined classifications violate the ordering constraints of the ontology.

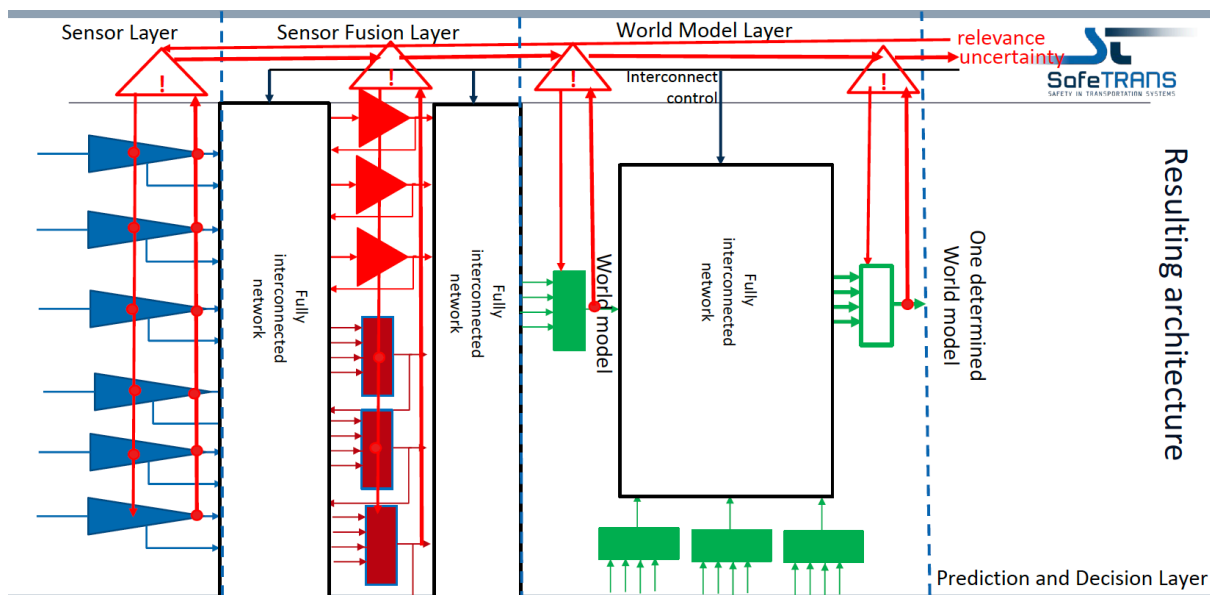


Figure 27 The reference architecture of the perception chain

This architecture is based on a notion of what we call "*sufficiently perfect components*" of the perception chain. These components allow at run time, and thus in a concrete situation, to inductively derive guaranteed bounds on the maximum level of uncertainty, provided the system is currently operating in well-defined ODDs.

The induction basis will be provided by "*sufficiently perfect sensors*". Such sensors are required to come with a characterization of adverse environmental conditions known to be detrimental to their guarantees, such as intensity levels of fog for lidar, strong levels of rain for radar, sunlight reflections on a wet street surface for video, or tunnels for radar. We will use real-field tests to generate virtual sensor models demonstrating not only sufficiently precise processing of raw data in non-adverse conditions, but which are also additionally able to demonstrate the same degradation effects as real sensors. This will allow us to quantitatively assess the level of uncertainty not only based on field measurements, but using large test sets of completely reproducible environmental conditions in digital twins of the environment and the sensor components. We will also characterize environmental factors partly controllable by the ego system, such as analyzing current vibration levels, the level of precision

of positioning information, precision of measurements of the velocity of the ego system, etc. We also include "sufficiently perfect digital maps" as anchoring components in the perception chain.

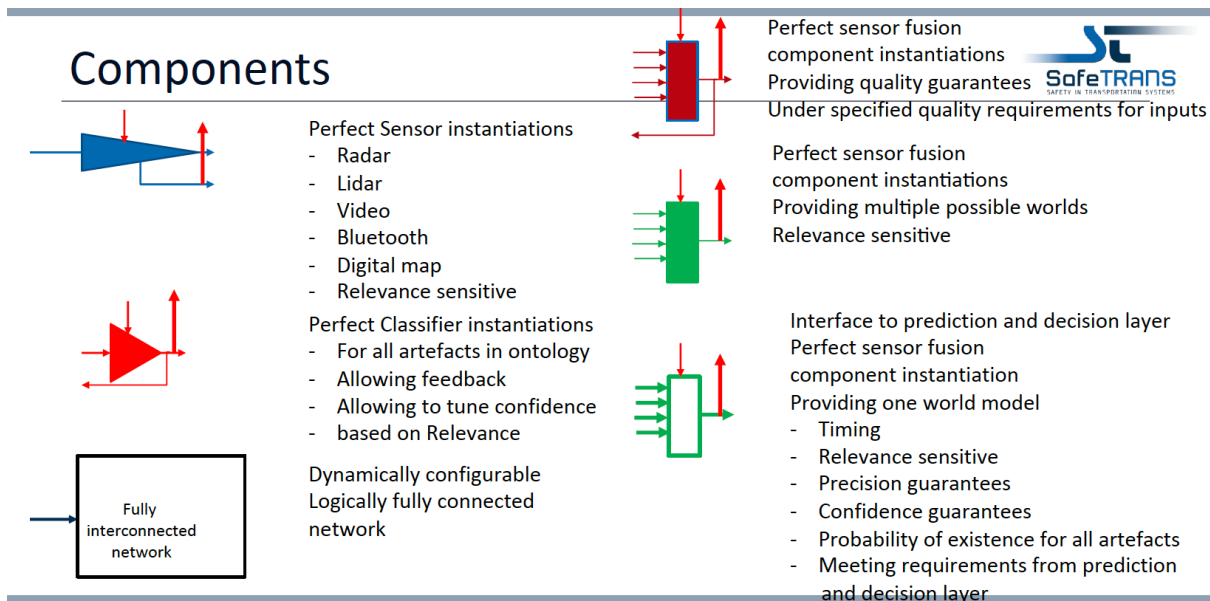
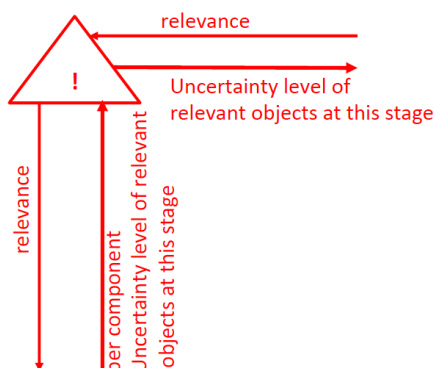


Figure 28 Components of the Reference Architecture of the Perception Chain

We propagate such guarantees along the perception chain by requiring what we call "sufficiently perfect sensor fusion components" and "sufficiently perfect classifier components". The meta-requirements for such components demand that each such component comes with a characterization of adverse environmental conditions and allowed ODDs. E.g., a classification component can only be required to provide guarantees for bounding uncertainty if the actual environment of the ego system is matching the characteristics in the data used for training classifier components with respect to types of objects, distribution of objects, and (if applicable) dynamic of objects in the analyzed video stream. For sensor fusion components, adverse conditions will be derived dynamically. Specifically, all inputs must be decorated by a characterization of environmental conditions under which they were collected, time stamps of raw data used, the component type delivering this input, and the adverse conditions of this component type.

Uncertainty Quantification Components



Per Stage quantification of current uncertainty level for each relevant artefact

- Downward propagation of relevance
- Upward propagation of uncertainty
 - Within stage
 - Between stages

Figure 29 Uncertainty Quantification Components

5.5 Bounding Uncertainty for AI-based Classifier components

We now discuss what could be called “adverse conditions” for AI-based classifier components. This part is based on the paper by [PT2020] performing a generic risk assessment for AI-based classification components based their life cycle. All material in this subsection is quoted directly from [PT2020].

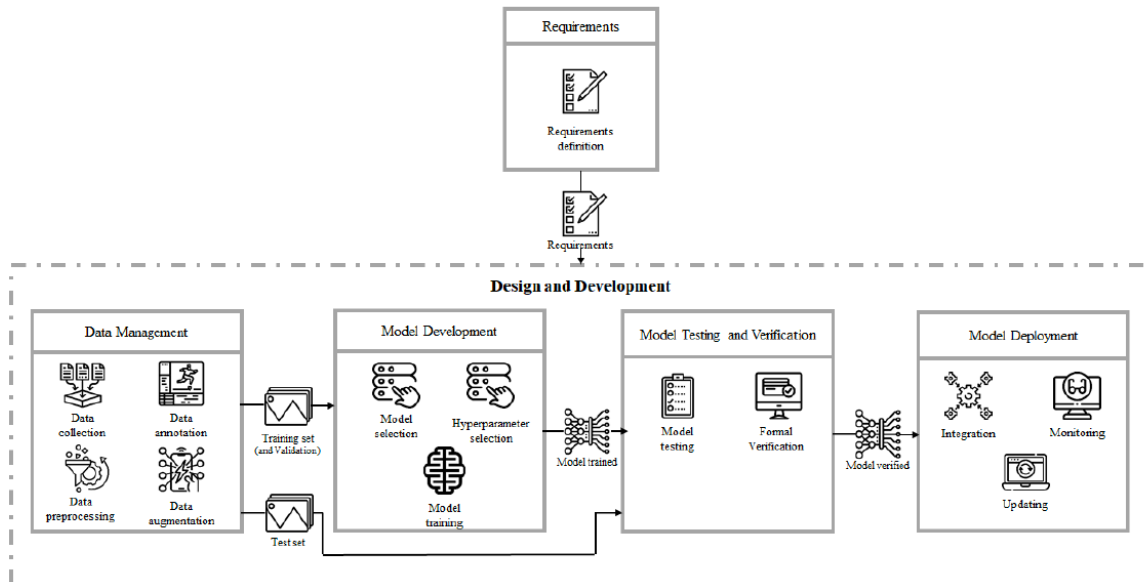


Figure 30 Process Model for AI based components, from [PT2020]

The following list of adverse conditions for AI-based classifiers follow the different phases in the life cycle of AI-based classifier components, and are labelled both by the phase and either as static or dynamic. Static adverse conditions must be eliminated at design time, e.g. by following the guidelines proposed by Simon Burton [BH2023] which are influencing the development of ISO PAS 8800.

- RE1 Incomplete definition of data (static)
- RE2 Incorrect objective function definition (static)
- RE3 Inadequate performance measure (static)
- RE4 Inadequate safe operating values (static)
- DM1 Inadequate distribution:
 - the distribution of the training examples does not adequately represent the probability distribution of (X, Y) (e.g., due to lost or corrupted data) (static)
 - the distribution of input variables can be different between training and the operating environment due to shifting environments during the lifecycle of the system (dynamic)
 - rare examples are absent or under-represented due to their small probability density (both static and dynamic)
 - not representing known statistically relevant examples of adversarial attacks in the training data (both static and dynamic)
 - bias not matching intended operational context (static and dynamic)
- DM2 Insufficient dataset size (static)
- DM3 Irrelevance: the data acquired contains extraneous and irrelevant information (static)
- DM4 Quality deficiencies:
 - data collected (based on sensors but also on human input) are limited in their accuracy (static)
 - during data annotation, quality could be compromised due to the incorporation of incorrect labels or incorrectly annotated area (static)

- the inclusion of non-realistic examples during data augmentation could affect dataset quality by generating data that do not make sense and change the complete meaning in a sample (static and dynamic)

MD1	Model mismatch: the model does not fully cover the requirements (static)
MD2	Model Bias inherently induced by the chosen model (static)
MD3	(Hyper) Parameters mismatch of parameters chosen in modelling (static)
MD4	incorrect error rate (dynamic)
MD5	the probability of failing is intrinsic to an ML model: The system is not able to ensure the complete correctness of an ML module output in the user environment where unexpected input occurs sporadically (dynamic)
MD6	Lack of interpretability (static)
MTV1	Incompleteness: Due to the large input space, it is difficult to test or approximate all possible inputs (the unknown is never tested). This way, the ML model only encounters a finite number of test samples and the actual operational risk is an empirical quantity of the test set. Thus, the operational risk may be much larger than the identifiable actual risk for the test set, not being representative of the real-world operation performance (dynamic)
MTV2	Non-representative distribution of test set (mirroring deficiencies DM1-4) (static and dynamic)
MDY1	Differences in computation platforms: Deploying a model into a device can result in computation limitations and compatibility issues across platforms, requiring adaptations potentially invalidating testing and verification results (static)
MDY2	Operational environment: Differences between the operational environment and data used for model development and testing, can lead to different/new inputs that affect the output produced, (dynamic) e.g. by <ul style="list-style-type: none"> ○ failure of one of the subsystems that provide inputs to the deployed ML model, ○ deliberate actions of an adversary; ○ changes in the underlying processes to which the data are related (changes on the environment or on the way people or other systems behave)
MDY3	Non detection of potentially incorrect outputs (dynamic)
MDY4	New data/Continuous learning. This hazard only considers the case of online learning (i.e., systems that continue to learn parameters and train the model during operation). Despite the fact that the incorporation of new data from the real operation domain suggests improving the model performance, since new data are added to the model training, the dataset distribution could be biased and it is no longer supervised, susceptible to result in lower model performance in scenarios that are no longer as frequent on the new data (e.g., a self-driving vehicle that was trained before operation on an adequate distributed dataset is now operating only at dark scenarios; for this case, the model could start to be optimized for dark conditions and to behave less accurate in the remaining day time scenarios). (dynamic)

Thus, adverse conditions for classifier components are induced by any of the dynamic hazards in the above lists, and results of AI-based classifiers can only be trusted, if both the static hazards have been addressed at design time, and the dynamic hazards are controlled at run-time.

The quality requirements enforcing consistency between the actually observed types and distributions of objects, and those used in training data, entails the need for regular over-the-air updates of AI-based classifier components. Processes must be established which monitor any statistically relevant deviations between actually observed data and data used during training, as well as monitoring for new types of objects requiring updates to the ontology. This need is well understood, as documented by numerous projects addressing the Devops cycles, leading to what is now called the MLOps cycle.

While these process-oriented measures will drastically reduce the risk of misclassification, they must be complemented with AI-centered research on improving accuracy, robustness, explainability, as e.g. discussed in the SafeTRANS Roadmap on Foundations for Safety and Explainability of AI based Safe-Critical Applications³³ and in the recommendations of the [UNA2024] to appear.

5.6 Propagating Uncertainty guarantees in sensor fusion

We will use approaches akin to Dempster-Shafer adapted to the guarantees for bounding uncertainty, to compute the maximal probabilistic guarantees for uncertainty for output streams of such components, and determine the adverse conditions by conjoining adverse conditions of components providing input streams with high relevance in strengthening guarantees for bounding uncertainty.

Let us illustrate on the classification case how we can derive quality attributes for fused data by fusing data of two sensors.

We can derive a composed sensor s as the fusion of two qualified positioned sensors $\langle s1, pos1 \rangle$ and $\langle s2, pos2 \rangle$ as follows:

```

–  $ad(s)$  is the disjunction of  $ad(s1)$  and  $ad(s2)$ 
–  $visible(s) = visible(\langle s1, pos1 \rangle) \cup visible(\langle s2, pos2 \rangle)$ 
– for each  $p \in \text{occupancy grid}(ego)$ :
if  $p \in visible(\langle s1, pos1 \rangle) \cap visible(\langle s2, pos2 \rangle)$  then
  if  $\neg ad(s1) \wedge \neg ad(s2)$  then
    if  $label(s1)(p) \wedge label(s2)(p) \neq false$  then  $label(s)(p) := label(s1)(p) \wedge label(s2)(p)$ 
    else  $label(s)(p) := \perp$ 
      else if  $ad(s1) \wedge \neg ad(s2)$  then  $label(s)(p) := label(s2)$ 
    else if  $\neg ad(s1) \wedge ad(s2)$  then  $label(s)(p) := label(s1)$ 
      else if  $ad(s1) \wedge ad(s2)$  then  $label(s) := \perp$ 
    else if  $p \in visible(\langle s1, pos1 \rangle) \setminus visible(\langle s2, pos2 \rangle)$  then  $label(s) := label(s1)$ 
      else if  $p \in visible(\langle s2, pos2 \rangle) \setminus visible(\langle s1, pos1 \rangle)$  then  $label(s) := label(s2)$ 
  fi

```

Recall that labels are part of an ontology with a lattice structure. The fusion operator \wedge on labels is accordingly based on the classical meet-semilattice.

We illustrate some cases of the above definition:

- Parts of the occupancy grids are labelled free, if all sensors for which this part is visible and which are operating in non-adverse conditions agree on this;
- however, if one sensor operating in non-adverse conditions senses one object, while other sensors operating in non-adverse conditions detect no object, the fusion operator yields \perp .
- In general, any detected inconsistencies are marked \perp .
- If, for example, front radar and front camera, both operating in non-adverse conditions, agree on identifying an object in a field of the occupancy grid, then the fusion of their sensor values at this field of the occupancy grid contains the detected position and speed of the object, as well as a sequence of vertical slices of a series of matrices of gray values of pixels.
- For observations where both sensors are in non-adverse conditions, and both sensors observe p , then the risk of misperception is reduced to $r_{s1} \times r_{s2}$ if the misperceptions are stochastically independent and under operating conditions favorable for both sensors.
- Otherwise the risk for misperception of the fused system for a field visible to both sensors is $\min(r_{s1}, r_{s2})$.

As suggested by Dietmayer et al (see, e.g., [MDM2010], [FHW+2022]), we will generalize this approach using versions of Dempster-Schafer Theory. We assume type classifications based on standardized

³³ See <https://www.safetrans-de.org/en/activities/Roadmapping.php>

partially ordered ontology with top elements static or dynamic and probability of correctness of object type classification. We also assume location information with information of aleatoric uncertainties in precision of measurement and confidence in measurement. Figure 31 below taken from [FHW+2022] illustrates the assumed setting for a concrete instance in an urban driving scenario.

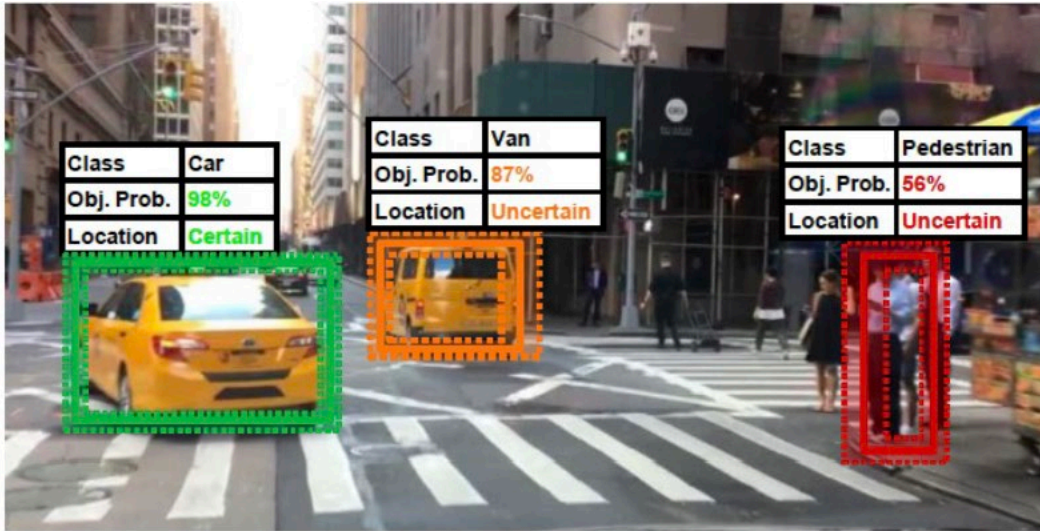


Figure 31 A conceptual illustration of probabilistic object detection in an urban driving scenario (taken from [F et al 22])

As argued in [MDM2010], this calls for a generalized fusion of heterogeneous sensor measurements for multi-target tracking:

- No single sensor system can provide the basis for sufficiently concise perception
- A fusion system which supports different types of sensors, like radar, laser or image-based sensors without adapting the fusion algorithm is desired.
- Such a reusable fusion system would lead to a high reduction of costs and time to market.

This requires a generic characterization of sensors, their generated time series of environment perception, and their confidence in these measurements. A foundation for this approach has been laid out in chapter 4 in characterizing the quality of measurements for different sensor types. Based on this information, the Dempster-Shafer approach allows to derive maximal quality guarantees for fused sensor data.

The key elements of the Dempster-Shafer Theory of Evidence (DST) as developed in [MDM2010] are summarized below:

- The *frame of discernment* Ω is defined as the set of elementary hypotheses

$$a_i: \Omega = \{a_i\}, i = 1, \dots, n.$$

In our multi-sensor fusion system, this should be given as a partially ordered ontology of objects in traffic scenes and their attributes such as relative distance, speeds, azimuth, ...

- A *basic belief assignment (BBA)* m is a mapping from the power set 2^Ω of the frame of discernment to the interval $[0, 1]$ with the following properties:

$$m(\emptyset) = 0,$$

$$\sum_{A \subseteq \Omega} m(A) = 1$$

- The *degree of belief* of a BBA m for a proposition A is defined as

$$Bel_m(A) = \sum_{B \subseteq A, B \neq \emptyset} m(B)$$

- The *degree of plausibility* is defined as

$$Pl_m(A) = \sum_{B \cap A \neq \emptyset} m(B)$$

The plausibility $Pl_m(A)$ is therefore the sum of all probability mass assigned to propositions which are not contradicting A, exploiting partial order.

From a safety perspective, the resulting capability to explicitly characterize the plausibility of both presence and absence of an environmental object A combined with the explicit confession of ignorance is a key strength of this approach. They allow to generalize the simple Boolean fusion operator described above to richer type structures, addressing both the situations where different sensors have contradictory observations, thus detecting inconsistencies, as well as situations where weak plausibility of identification of an environment object A can be compensated by high plausibility of detection of A by a second sensor. Graphically, this fusion is nicely supported by depicting the detection capabilities of a single sensor as shown in Figure 32 below:



Figure 32 Visualization of detection Capabilities of a Single Sensor

Let us illustrate the fusion of such qualified sensor measurements with a simple example, involving fusion of Lidar data and Camera data.

The left side of Figure 33 gives a bird's eye view of the perception of the traffic environment shown on its right side as seen by a Lidar sensor, the right-side highlights detections in this scene from a video camera in magenta, overlaid by projected lidar segments (blue, green, red, yellow).



Figure 33 Detections of one and the same traffic scene as observed by Lidar and Camera

For simplicity, let us assume that we are interested in extracting only three categories from sensor fusion:

- N: No object exists
- O: there is an object at this position, but it is not relevant for the current driving situation

- R: there is an object at this position, and it is relevant for the current driving situation

Let us assume that the video sensor can distinguish between the proposition {N,O} and {R} but not between {N} and {O}, and that the laser scanner module can separate {N} from {O,R} but only very poorly {O} from {R}, as depicted in Figure 34 below:

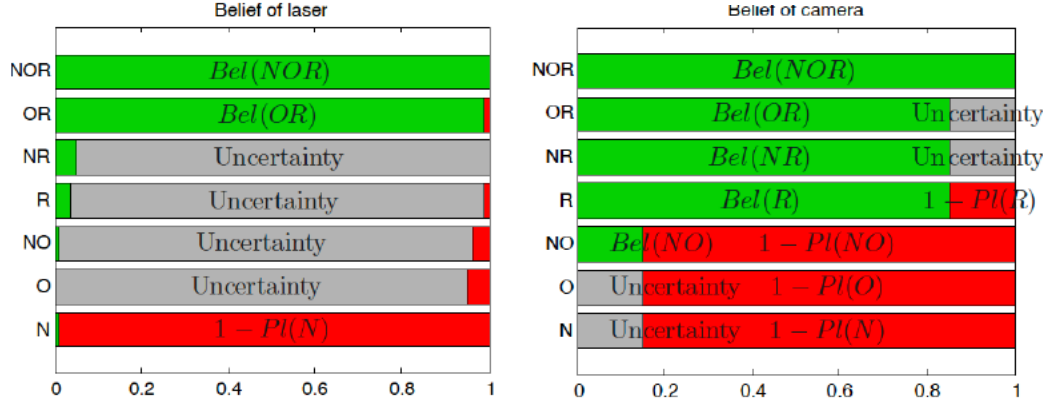


Figure 34 Visualization of Beliefs of Lidar and Camera

This leads to the following set of propositions:

$$2^\Omega = \{\phi, N, O, NO, R, NR, OR, NOR\}.$$

Each of the propositions of 2^Ω can now be assigned a probability.

In the above two sensor example the video camera module would assign the following probabilities to the BBA based on the measurement z with the frame of perception $F^C = \{R\}$:

$$\begin{aligned} m(NO) &= p_{FP}(z) \\ m(R) &= p_{TP}(z) = p_{vehicle}(z) \end{aligned}$$

The BBA of the laser scanner module will be

$$\begin{aligned} m(N) &= p_{FP}(z) \\ m(OR) &= 1 - p_{FP}(z) - p_{vehicle}(z) \\ &= p_{TP}(z) - p_{vehicle}(z), \\ m(R) &= p_{vehicle}(z) \end{aligned}$$

and the corresponding frame of perception $F^L = \{OR\}$.

The fusion of the laser BBA and the video BBA can be done according to the formula

$$m_{1 \oplus 2}(A) = \frac{\sum_{X \cap Y = A} m_1(X) m_2(Y)}{1 - \sum_{X \cap Y = \emptyset} m_1(X) m_2(Y)}$$

For our example, we first consider $m_{1 \oplus 2}(N)$. The nominator is

$$\begin{aligned} Z_1 &= m_1(N) m_2(N) + m_1(N) m_2(NO) + m_1(N) m_2(NR) + m_1(N) m_2(NOR) + m_1(NO) m_2(N) \\ &\quad + m_1(NR) m_2(N) + m_1(NOR) m_2(N) + m_1(NO) m_2(NR) + m_1(NR) m_2(NO) \end{aligned}$$

and the denominator is

$$Z_2 = 1 - m_1(N) m_2(O) - m_1(N) m_2(R) - m_1(O) m_2(N) - m_1(O) m_2(R) - m_1(R) m_2(N) -$$

$$m_1(R)m_2(O) - m_1(N)m_2(OR) - m_1(O)m_2(NR) - m_1(R)m_2(NO) - m_1(OR)m_2(N) - m_1(NR)m_2(O) - m_1(NO)m_2(R),$$

so that

$$m_{1\oplus 2}(N) = \frac{Z_1}{Z_2}$$

Thanks to many BBA assignments being zero, this expression boils down to

$$\begin{aligned} m_{1\oplus 2}(N) &= \frac{m_1(NO)m_2(N)}{1 - m_1(R)m_2(N) - m_1(NO)m_2(R)} \\ &= \frac{p_{1,FP}(z)p_{2,FP}(z)}{1 - p_{1,vehicle}(z)p_{2,FP}(z) - p_{1,FP}(z)p_{2,vehicle}(z)} \end{aligned}$$

Note that the denominator is identical for all BBA fusion arguments. The remaining BBA fusion assignments are

$$\begin{aligned} m_{1\oplus 2}(O) &= \frac{m_1(NO)m_2(OR)}{1 - m_1(R)m_2(N) - m_1(NO)m_2(R)} \\ &= \frac{p_{1,FP}(z)(p_{2,TP}(z) - p_{2,vehicle}(z))}{1 - p_{1,vehicle}(z)p_{2,FP}(z) - p_{1,FP}(z)p_{2,vehicle}(z)} \end{aligned}$$

$$\begin{aligned} m_{1\oplus 2}(R) &= \frac{m_1(R)m_2(R) + m_1(R)m_2(OR)}{1 - m_1(R)m_2(N) - m_1(NO)m_2(R)} \\ &= \frac{p_{1,vehicle}(z)p_{2,vehicle}(z) + p_{1,vehicle}(z)(p_{2,TP}(z) - p_{2,vehicle}(z))}{1 - p_{1,vehicle}(z)p_{2,FP}(z) - p_{1,FP}(z)p_{2,vehicle}(z)} \end{aligned}$$

It holds that $m_{1\oplus 2}(NO) = m_{1\oplus 2}(NR) = m_{1\oplus 2}(OR) = m_{1\oplus 2}(NOR) = 0$, which can be verified by inserting all single BBA assignments or by observing that

$$m_{1\oplus 2}(N) + m_{1\oplus 2}(O) + m_{1\oplus 2}(R) = 1$$

since

$$\begin{aligned} &p_{1,FP}(z)p_{2,FP}(z) + p_{1,FP}(z)(p_{2,TP}(z) - p_{2,vehicle}(z)) + p_{1,vehicle}(z)p_{2,vehicle}(z) \\ &\quad + p_{1,vehicle}(z)(p_{2,TP}(z) - p_{2,vehicle}(z)) \\ &= p_{1,FP}(z)p_{2,FP}(z) + p_{2,TP}(z) - p_{1,FP}(z)p_{2,vehicle}(z) \\ &= 1 - p_{1,vehicle}(z)p_{2,FP}(z) - p_{1,FP}(z)p_{2,vehicle}(z) \end{aligned}$$

due to the relation $p_{1,FP}(z) + p_{1,TP}(z) = p_{2,FP}(z) + p_{2,TP}(z) = 1$.

We refer to [\[MDM2010\]](#) for the formal definition of the fusion operator. While thus a conceptual framework for propagating guarantees of the level of uncertainties exist, we note that Dempster-Shafer works only well for small universes of a handful of propositions, but doesn't scale well due to the necessity of identifying all (i.g. exponentially many) inconsistent combinations. The challenge thus remains to develop algorithms rendering Dempster Shafer reasoning scalable. Also, in many cases, Dempster Shafer would enhance fusing results by rescaling the individual results after eliminating impossible combinations. This, however, also implies that Dempster Shafer breaks stochastic independence and may introduce fallacies when arguments involve it, a topic demanding further

research.

At the highest level of the perception chain, sensor fusion components provide what is often called a “*world model*”. This is comprising all artefacts in the environment of the ego system relevant for maneuver decisions, with guarantees for bounded uncertainty.

A *world model* at time t identifies

- for all relevant objects in the environment of the vehicle a consistent description of their classification with guaranteed confidence levels in existence and classification, as well as their position and relative distance with guaranteed level of precision and additionally, for all dynamic objects, their velocity, with guaranteed levels of precisions;
- all relevant other environmental artefacts, subsuming in particular those relevant for identifying adverse conditions for all components in the perception chain.

Rather than denoting the fusion of high-confidence inconsistent data observed under non-adverse conditions as inconsistent, we propose to view this as a temporary inconsistency which will be resolved by future observations, and thus maintain *multiple* possible world models. To this end, please note that we are generally operating in a setting of monotonic reasoning over possible-world models, attaching likelihoods to possible worlds by means of Dempster-Shafer belief strength. In such a setting, a set of high-confidence inconsistent labeling ought in principle not happen, as it points to wrong assumptions about the individual sensors’ reliability – the empty set of possible worlds has no physical realization and can only be inferred from incorrect assumptions. Nevertheless, we have to be prepared for this case, as such inaccuracies in the assumptions about sensor performance will inevitably arise, as will then the corresponding inferences. The consequences for a monotonic reasoning framework over possible worlds would be devastating, as any possible maneuver would be considered safe over the empty set of possible worlds, i.e., any maneuver be justified. We therefore implement two extra safety measures that are activated when high-confidence inconsistent data arise:

- The respective quantitative inference graph has to be recorded in order to permit root-cause analysis and subsequent adjustment of the quantitative sensor models.
- Rather than deriving the empty world (with high confidence), we consider all the worlds indicated by the different inconsistent sensors and their respective inference chains possible, i.e., we realize a strictly pessimistic sensor fusion in this special case. This is a safety mechanism, as all safety-relevant maneuvers then have to be justified over all the possible worlds, which are bound or at least extremely likely to contain the real one.

As noted above, we don’t expect this state to persist, as subsequent observations will refine the set of possible worlds again, thereby successively removing the pessimistic over-approximation of possible worlds, which however permits to safely survive the sensor (or rather sensor model) failure that manifested in the high-confidence inconsistent data observed.

Vehicles may insert an additional layer in the perception chain allowing for fusing their world model(s) with world models provided by infrastructure or vehicles in the neighborhood. This requires standardized representations of world models used for such interchanges, such as by evolving ontology standards and quality and confidence attributes as defined previously, and time stamps. Using such information for fusing world models must take into account the variable transmission latencies of such messages. Protocols must be established for secure communication and trustworthiness of senders of such messages.

Jointly, we can thus derive for each stage in the perception chain the level of uncertainty for relevant environmental artefacts for maneuver decisions of the ego system, such as those provided from the prediction and decision layer.

6. Credible Co-Simulation and Model Composition

As discussed in the previous chapters, the realm of automated driving functions is characterized by a complex sequence of tasks, encompassing environment perception, sensor fusion, object detection, tracking, prediction, behavior planning, and actuator control. Given the intricacies involved, testing these functions through simulation, especially as part of the safety approval, brings the highest requirements possible when building up the whole simulation out of multiple models. While perception sensor simulation and sensor fusion modeling has been described, the following chapter widens the view towards the overall concept of co-simulation for safety validation. It also proposes a flexible composition of modeled effects to build the models depending on the simulated scenarios and demands to evaluate their fidelity per scenario and how these differently composed models from different sources build a per-scenario valid co-simulation.

To build up a solid safety argumentation during the homologation process, the primary objective of safety validation, whether conducted in a virtual, physical, or mixed environment, is to generate evidence supporting claims. This evidence is crucial for constructing a trustworthy safety argumentation. This process has been researched in depth in the Verification and Validation Methods (VVM) project. The simulation process often involves multiple distinct models from diverse sources in a modular architecture termed co-simulation. This is necessary, when e.g. multiple different sensor manufacturers deliver sensor models for different modalities that come together at an OEMs co-simulation of the vehicle. The combination of multiple models reveals the question of validity as such, even if they are already separately validated for the targeted application domain and scenarios to simulate. To light up the path towards such a valid co-simulation, multiple standards, tools, and the credible simulation process are applied. The challenge as such highlights the necessary interplay between science and industry, the fusion of free and commercial tools and models.

The first challenge is purely technical and involves the integration of diverse models from different sources. Such a modular simulation architecture, as shown in Figure 35 Generic Simulation Architecture, © Persival GmbH, is called co-simulation. It starts with a scenario player that moves the objects' 3D assets within the scene, according to the scenario description. It incorporates models for traffic participants behavior, vehicle dynamics, and many more. The perception sensor models, which have been discussed in chapter 4, produce the synthetic sensor data that is then fused with probabilistic strategies according to the actual predicted capabilities of each sensor, as described in detail in the previous chapter. Subsequently, objects are detected and tracked/predicted, and the vehicle's internal environment model is built. Based on this, the vehicle plans and acts, while both last steps are out of scope of this paper.

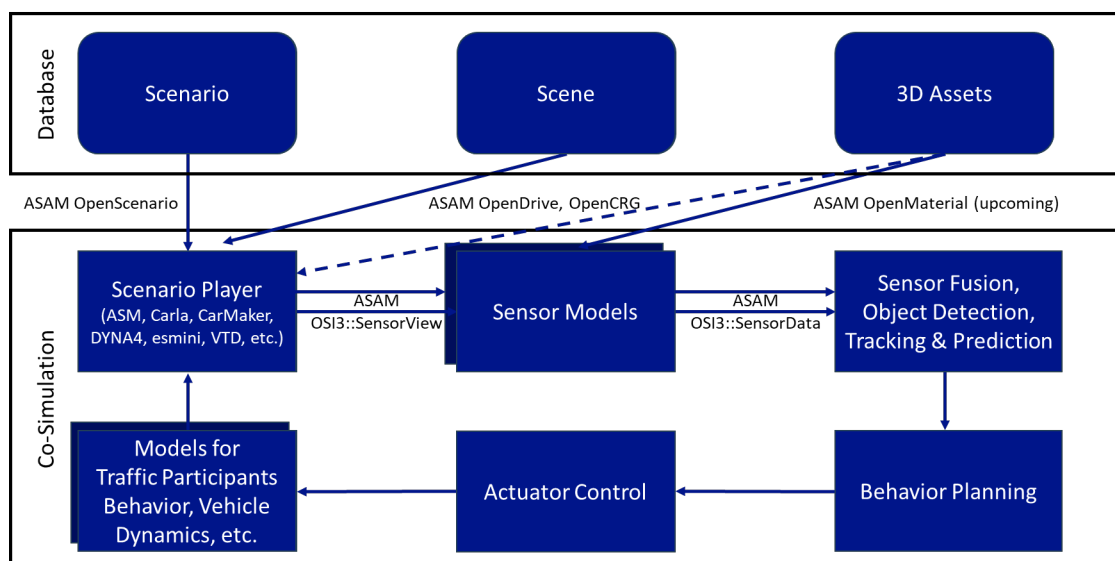


Figure 35 Generic Simulation Architecture, © Persival GmbH

Co-simulation tools³⁴ are needed that act as co-simulation master and therefore take care of the simulation flow, the data exchange between the models, and the timing. These tools facilitate the integration of distinct models, enabling a comprehensive evaluation of complex interactions. The central part of the simulation is the scenario player, which often also acts as co-simulation master. Commercial simulation tools often entail own models for driver and object behavior, vehicle dynamics, environmental conditions, and sensor performance. Most of them also bring a library of static scenes and 3D assets. Still, multiple parts of the whole simulation often come from different sources and modelers, e.g. the sensor models from the actual manufacturers of the sensors, the scenes from real world capturing, etc. The scenarios and scenes, as well as the 3D assets for the static scenes and the movable objects in case of ray tracing and other detailed rendering techniques are taken from one or multiple databases provided by possibly several suppliers. Marketplaces³⁵ are emerging as essential platforms for accessing multiple of the simulation components shown in Figure 35

In previous projects on simulation-based testing like the already mentioned ENABLE-S3, Pegasus, SETLevel, and Vivid, several simulation models have been developed to set up whole co-simulations. These initiatives have paved the way for the establishment of standards that facilitate seamless integration of models from different sources and enable communication between these simulation components. The standards include:

- Modelica FMI: Enabling model exchange and co-simulation through a functional mock-up interface.
- Modelica SSP: Providing a standardized simulation platform for consistent interfaces.
- ASAM OpenSCENARIO: Defining scenarios for testing advanced driver assistance systems (ADAS) and automated driving functions.
- ASAM OpenDRIVE: Describing the road network and contextual aspects.
- ASAM OpenCRG/OpenMATERIAL: Detailing road surface and object material modeling to enhance simulation realism.
- ASAM Open Simulation Interface (OSI): Enabling effective communication between different simulation tools and models.
- ASAM OpenLABEL/OpenTEST/OpenODD: Focusing on labeling, testing, and describing the operational design domain.
- ISO11010-1 and 11010-2 on classification of vehicle dynamics simulation models and perception sensor simulation models

An effective simulation ecosystem thrives on a harmonious amalgamation of free and commercial tools and models. The credibility bestowed upon open-source solutions resonates well within the community, fostering trust through inspectability and collaboration. Interestingly, open-source options might even complement and enhance commercial models. This fusion is pivotal in creating an inclusive environment where innovation is nurtured. The convergence of such diverse simulation models from different vendors underscores the complexity of the simulation task. These varied components coalesce to form the co-simulation architecture, necessitating seamless interoperability and communication. The challenge lies not only in technical integration but also in nurturing a cooperative environment among stakeholders with distinct interests.

The landscape of simulation tools includes both commercial offerings and open-source alternatives. Commercial tools provide pre-built models and assets, streamlining the simulation process. However, open-source tools³⁶ and scenario players³⁷ offer often sufficient functionality while gaining trust through their open-source code. However, comprehensive testing is needed that can be applied on the models within automated testing pipelines, from automated verification of the implementation to continuous comparison against real sensor data, as e.g. demonstrated on OpenMSL on the provided

³⁴ Eclipse Foundation: *OpenMCx* (<https://github.com/eclipse/openmcx>)

³⁵ Automotive Solution Center for Simulation e.V.: *ENVITED Marketplace* (<https://envited.market/>)

³⁶ *CARLA Simulator* (<https://carla.org/>)

³⁷ *Basic OpenSCENARIO player Environment Simulator Minimalistic: esmini* (<https://github.com/esmini/esmini>)

open-source models.³⁸

Smooth transition across testing levels (SiL, HiL, ViL) is a pressing need, as well. Here, standardization comes in handy. For instance, the transfer of OpenDRIVE files, OpenMATERIAL scene assets, and other simulation components should be seamless. This seamless transfer ensures that the testing framework remains coherent as the simulation progresses from virtual environments to real-world scenarios, preventing an unwieldy test bench. Additionally, the simulation runtime on different hardware should be monitored and as soon as hardware gets in the loop, real-time execution speed becomes mandatory. In SiL testing, scalability and cloud infrastructure can be leveraged and speed up development and testing processes. HiL testing involves already embedded systems while others stay simulated as a digital twin. This complex interplay therefore needs a lot of coordination between all involved parties and specification of interfaces. Provided metadata becomes key regarding the traceability in testing applications.

Another often neglected aspect is the need for customization of models depending on the particular testing objectives. This even involves to differently compose the modeled effects per testing scenario. This means e.g. that for some tests that evaluate the general behavior of some sensor fusion aspects are fine with computationally fast, low-fidelity and effect-reduced sensor models, while other tests need high-fidelity sensor models that involve multiple computationally slow effects like multipath propagation on mirroring surfaces with exact intensities computed for each reflection. In other words, there is not only a single radar model involved in simulation-based safety validation, but multiple models of the same real world radar sensor with differently composed effects to suit the respective testing requirements. Consequently, a standardized method for model classification depending on their composition is required to facilitate proper model selection. Luckily, this pressing need is already acknowledged by peer-groups like the IAMTS³⁹ and by standardization activities like the ISO 11010-2 work item proposal⁴⁰ for the classification of perception sensor models. Here, a scheme for the combination of effects is proposed to enable a flexible and modular catalogue-like selection process of the individual required model composition for each test.

Credible Simulation, however, entails more than selecting and coupling model compositions together for a complete simulation of the environment and the objects moving within for each test. According to Liu et al. “the credibility of a model or simulation is an expression of the degree to which one is convinced that a particular model or simulation are suitable for an intended purpose”.⁴¹ The National Aeronautics and Space Administration (NASA) further defines credibility as “the quality to elicit belief or trust in [Modelling & Simulation] M&S results”.⁴² To determine the credibility of a model or simulation, a specific application purpose must be defined.^{Fehler! Textmarke nicht definiert.} According to Liu et al., the factors of validity, correctness, reliability, usability and interoperability must be considered when assessing the credibility of a model. Therefore, not only the composition, but also the fidelity and the level of detail of each modeled effect has to be considered and validated.

As already discussed in the previous chapters, validation of all models with decent metrics and predicting the model error for the targeted usage of the simulation becomes inevitable. This also includes validating the simulated 3D environment and environmental conditions. It highly influences the necessary effort in modeling perception sensor performance but also vehicle dynamics. Replicating real-world conditions with a high degree of accuracy is the often-mentioned objective in simulation, but validation must ensure that the simulation only considers the relevant cause-effect chains and not every possible effect.⁴³ This means to ensure that the models are only composed of the relevant effects

³⁸ asc(s e.V. - ENVITED Open Source Model & Simulation Library (<https://github.com/openMSL/>)

³⁹ International Alliance for Mobility Testing Standardization (IAMTS): A Classification Scheme for Sensors Models with Related Validation Measures and Application Examples for Automated Driving Systems (unpublished)

⁴⁰ [ISOTC22]

⁴¹ [LYW2005]

⁴² [NAS2013]

⁴³ [LRS+2021]

for each simulation run during the overall simulation-based safety validation. As proposed by Linnhoff et al., the Operational Design Domain (ODD) of the current scenario to be simulated results in a specific application are within the parameter space of each model, which in combination with the known sensor to simulate leads to a set of relevant cause-effect chains. This set of relevant effects is the model composition for the concrete simulation task. However, automation is needed at this point, as multiple models are involved, and a large number of scenarios is to be run within a simulation-based homologation. Still, a methodological way of model selection is possible.

As an integral component of the successfully concluded SET Level funding project⁴⁴, innovative methodologies and processes were devised to facilitate the credible selection and interchange of (sensor) models. To achieve this, an initial categorization of simulation models pertaining to perception sensors (such as radar, lidar, and camera) was established. This classification was based on ISO 11010-1 in terms of structure and procedure.

Additional noteworthy processes for assessing credibility include the IAMTS reference process⁴⁵, which ensures the reliable utilization of virtual validation methods. Furthermore, the "Credibility Assessment Framework"⁴⁶ and the "New Assessment/Test Method for Automated Driving (NATM)"⁴⁶ developed by the Validation Method for Automated Driving (VMAD) working group under the United Nations Economic Commission for Europe (UNECE) provide valuable perspectives. The National Aeronautics and Space Administration (NASA) standard STD-7009A also stands out as a relevant benchmark in this context. For the qualification of the used tools the [ISO26262](#) [Tool Confidence Level (TCL) can be utilized.

To conduct an effective credibility assessment, it is essential to model not only the System in Development (SiD) but also the validation system and its dependencies. This approach enables the automatic detection of inconsistencies between the SiD and the validation environments, facilitating assumptions about credibility. In the DFG CRC 1608 (Consistency in the View-Based Development of Cyber-Physical Systems)⁴⁷ methods for modeling both SiD and validation environments, along with the automatic detection of inconsistencies, are being researched.

To compare the credibility of multiple validation configurations or toolchains, considering a calculation of a credibility index appears to be a viable approach. This index could draw inspiration from the methodologies proposed by Liu et al.⁴¹ and Muessig et al.⁴⁸ Given the inherent uncertainties associated with this index, leveraging Bayesian probabilities becomes a valuable consideration. Furthermore, aligning with the principles outlined in the Guide to the Expression of Uncertainty in Measurement (GUM) may provide additional insights.

When modifying components within a simulation toolchain, it may be unclear whether a comprehensive re-evaluation of the entire toolchain is required. The P.E.A.R.S Initiative presents an approach that employs a round-robin methodology. For instance, various simulation tools are applied under identical conditions to evaluate sensitivities systematically.

As introduced by the SET Level project, the so-called Credible Simulation Process (CSP)⁴⁹ covers the mentioned steps from model requirements definition and model meta data as so-called glue-particles, further detailing into the Credible Modeling Process (CMP) and the Model Selection and Exchange Process. It is currently further specified by Prostep IVIP and the SmartSE project⁵⁰, while being adopted by Ahman et al. in the project UPSIM⁵¹ and also further detailed for sensor model testing on the

⁴⁴ <https://setlevel.de/aktuelles>

⁴⁵ [\[IAM2021\]](#)

⁴⁶ [\[UNE2022\]](#)

⁴⁷ DFG CRC 1608: Consistency in the View-Based Development of Cyber-Physical Systems, <https://www.sfb1608.kit.edu/>

⁴⁸ [\[MLW2000\]](#)

⁴⁹ https://gitlab.setlevel.de/open/processes_and_traceability/credible_simulation_process_framework

⁵⁰ <https://www.prostep.org/en/projects/smart-systems-engineering-smartse>

⁵¹ [\[ALF2022\]](#)

OpenMSL platform as the so-called OSMP Test Architecture.⁵²

Furthermore, beyond model validation and consistent metadata for each model, establishing credible simulation involves ensuring a safe and controllable overall development and exchange process. An important aspect in this exchange is of course the intellectual property (IP) protection of the IP that went into building the models of real sensors. So legal aspects play a role and must be handled by the process. Besides, as already mentioned, timing constraints for each individual model or model composition state can't be neglected. Sensor models for example mimic the frequency of a specific sensor in the simulation time while providing simulation results at certain computation times. To plug models together, this must be handled accordingly.

However, building a credible co-simulation even encompasses to ensure that the combination of per-se validated models, each validated for their individual application area, is valid, as well. Implemented and validated models entail implicit and explicit assumptions about their input data and application area for what they are validated for that must be satisfied and therefore checked at first, when combining the models from different sources. Additionally, monitoring the simulation and checks for the currently simulated conditions against each model's validity area must be ensured before and during each simulation.

A holistic view on credible (co-)simulation reveals several challenges, as described. Any models used are to be designed for the specific actual use case. Therefore, multiple different effect compositions are necessary to fit all needs. Some simulations entail bad weather and make very detailed perception sensor simulation necessary. Other scenarios just need very basic sensor performance models with false positive or negative objects that appear or vanish without applying high-fidelity rendering techniques like ray tracing for lidar. The most important question before any scenario is simulated is: Are the selected effects the right ones to compose for this case, are they modeled detailed enough or over-engineered and therefore too performance-costly?

In the trajectory of advancing simulation-based testing of automated driving technologies, collaboration emerges as the cornerstone. The integration of science and industry, the fusion of open-source and commercial models, and the harmonization of diverse simulation components all hinge upon cooperative efforts. Bridging the gap between scientific research and industry expertise is imperative, particularly possible in publicly funded research initiatives. Such cooperative projects establish a nexus where theoretical advancements meet practical application, laying the foundation for credible simulation methodologies. Nevertheless, regarding regulatory institutions, the proof of safety must be the focus of developments and model use.

In a daring vision, credible simulation of sensor performance could be used to determine the actually required minimum capabilities of the wanted real sensor setup to still be sufficient for an automated driving function or for sensor fusion of multiple sensors under different conditions. In other words, the simulation could be used to determine the real sensor performance requirements in the selected ODD and its inherent adverse conditions. Such an approach would be a more problem-centric approach advancing the state of the art, where currently only the available sensors and their performance in different conditions are evaluated with expensive measurement campaigns. However, once the actual requirements are determined, the task of the manufacturers to meet them would become tangible and for the first time there is an achievable goal.

In conclusion, co-simulation with models from different sources is currently a major challenge to solve, as sample validation of each model per-se is not enough. Standardization of data formats and interfaces is paramount. An ODD- and relevance-driven selection process for the individually composed effects for each simulation is needed. Therefore, the road to credible simulation necessitates a joint effort of all parties and stakeholders involved, where challenges are surmounted collectively, and innovation thrives at the intersection of knowledge and practice. In the near future, software-defined vehicles undergo continuous improvement through Over-the-Air (OTA) updates.

⁵² <https://openmsl.github.io/doc/OpenMSL/test-architecture/index.html>

Given the ever-evolving nature of this system in development, there is a constant need for an adapted and enhanced safety validation environment to preserve a certain credibility. Concepts like the Digital-Loop⁵³ emerge as potential approaches to address this ongoing challenge, providing a framework for an iterative simulation-based homologation.

⁵³ Digital-Loop, <https://www.digi-loop.com/>

7. Verification and Validation Methods and Processes

In the context of automated driving, verification and validation takes an important role for assuring safety and preparing a market release. Any automated driving system (ADS) introduces a certain level of risk into a public context which can be mitigated but not eliminated. This level of risk must be assessed by and dealt with by several stakeholders — depending on market context, these stakeholders may be manufacturers, operators, and/or authorities. A safety-by-design paradigm supports an adequate consideration of that risk during the system design.

This chapter elaborates on the assessment of the level of risk by (1) verification, i.e. the “confirmation, through the provision of objective evidence, **that specified requirements have been fulfilled**” [ISO/IEC/IEEE 15288] and (2) validation, i.e. the “confirmation, through the provision of objective evidence, **that the requirements for a specific intended use or application have been fulfilled**” [ISO/IEC/IEEE 15288]. From a safety perspective, one high-level requirement is that the system shall not introduce an unreasonable level of risk in its operational environment. Similarly, the fulfillment of other system-level requirements may be examined by verification and validation activities.

A promising new approach emerging from the Verification & Validation Methods (VVM) project is the safety assurance framework. This proposed framework deconstructs safety argumentation into multiple claims, each substantiated by supporting evidence. [VVM2023] The created evidence is not only dependent on the credibility⁵⁴ of the used validation & verification (V&V) tool chain, but also on leveraging adequate analysis and development processes and a concern management process. The VVM safety assurance framework will be covered in more detail later in this chapter.

The provisioning of supporting evidence for the risk estimation and safety assurance can be, among others, based on credible (co-)simulation environments and digital twins. The credibility of this evidence is contingent not only on the validity of the underlying simulation models and V&V tool chain, as detailed in chapter 6, but also on the issues concerning test coverage and their management through a structured V&V process for automated driving functions, which is being explored in depth in this chapter. Demonstrating satisfaction of a certain requirement across the possible instances of an intended use case or application requires adding some form of coverage of these instances to the application of the validated (co-)simulation environments and digital twins.

Section 7.1 motivates traceable verification and validation processes targeted towards managing the level of risk on the system level. In order to assess the risk contribution from perception chain elements on the system level risk, an adequate decomposition of the actual and accepted risks is necessary. Based on these explanations, section 7.2 elaborates on how the specific evidence for validation and verification can be obtained using various techniques on several levels of abstraction and decomposition.

Testing is essential in order to generate the objective evidence for verification and validation. ADSs are developed in a melting pot of several engineering domains which bring their own approaches to testing to the table. Section 7.3 provides an overview of challenges while applying such approaches from the domains of software engineering and automotive engineering to these systems. In order to generate verification and validation evidences, a multitude of test environments are available, ranging from simulation to real vehicles. Section 7.4 discusses how the test environment interacts with the validity of the V&V evidences and hence influences their credibility. Section 7.5 addresses decomposition and lays out current challenges, especially those relating to the need of representing uncertainty and probabilities during the process.

We summarize the current challenges in the verification and validation of perception systems used in ADS in section 7.6 and provide an overview of the identified needs for further research activities.

⁵⁴ The “quality to elicit belief or trust in [modeling and simulation] results” [NAS2016], [NKL+2021]

7.1 Supporting Continuous Risk Management through Traceable Verification and Validation Processes

Risk Acceptance Criteria

Established automotive safety standards, notably ISO 26262 and ISO 21448, define risk acceptance criteria based on the acceptable level of risk "according to valid societal moral concepts" [ISO 26262-1:2018, 3.176]. It is important to note that the hazards that give rise to these risks arise from the system-level behavior of the ADS-equipped vehicle. Behavior includes the externally observable dynamic states of the vehicle, such as pose and velocity. Identifying hazards based on the behavior significantly improves the efficiency of the functional safety and SOTIF⁵⁵ life cycles [GSB+2020]. This approach can be used to analyze risks at the system level. However, whether or how these system-level risks can be decomposed through the system architecture in order to determine the risk contributions or risk budgets of the perception chain and its components remains an open question. As mentioned in section 5.1, the ability to provide such a derivation could potentially improve the validity of perception quality requirements. Furthermore, if proofs can be made about properties of the perception chain, e.g. using advanced fusion techniques, these proofs could potentially facilitate risk assessment.

Managing Assumptions in the Engineering Process

An important tool for describing the operational context is the operational design domain (ODD). Combined with a structured description of the operational context using a domain model, the operational domain (OD), references can be made to assumptions for the system design. For example, in order to predict the behavior of agents, assumptions must be made about their characteristics (see [section 5.1](#)). It can be argued that such referencing can support rigorous modeling of the risk contribution by the perception chain. These assumptions also influence robustness certificates or guarantees that elements of the perception chain can provide to the rest of the system. Another aspect specific to machine learning (ML)-based components is the provision of a data set for training and validation of ML models. If the data set specification references some ODD description, the underlying assumptions are contained in the data set. Furthermore, out-of-distribution detection techniques may be deployed to detect an ODD exit – in turn requiring to terminate the execution of the dynamic driving task.

Given all the assumptions that underlie the system design, validation arguably involves a critical assessment of whether these assumptions are adequate for the operational context and thus for the intended purpose of the system. If (formal) guarantees are made and rigorously verified, and if the assumptions are validated, this creates a supporting argument for the validity of the system.

In order to provide evidence that an ADS does not introduce unreasonable risk in its operational environment, there are multiple open challenges that have only partially been addressed in the literature and previous research projects.

One key challenge is the definition of the aforementioned risk acceptance criteria that are based on a public consensus on safety expectations of ADSs. For existing systems, such as driver assistance systems in road vehicles, the calibration of automotive safety integrity levels (ASILs) included the consideration of risk acceptance criteria with respect to functional safety according to ISO 26262. For ADSs such an analysis of relevant risk acceptance criteria is yet to be investigated.

Given a sound definition of risk acceptance criteria, design and validation targets need to be derived for the respective system in the aspired operational context. Following a safety-by-design paradigm in order to provide traceability of assumptions and requirements that can be evaluated through verification and validation activities is necessary to draw conclusions regarding the fulfillment of the defined risk acceptance criteria.

⁵⁵ safety of the intended functionality [ISO 21448:2022, 3.25]

Scenario-based Verification and Validation

Traditionally, verification and validation of automotive systems largely relied on distance-based approaches, where achieving a certain mileage with high integration levels is a major cornerstone in eliciting confidence that a requirement is fulfilled. [WW2016] argue that following this approach for validating ADS can end in an “approval trap” due to the complexity of the open context. Instead, scenario-based approaches for the testing, verification and validation of automated functions, such as in the PEGASUS project [PEG2024], have been proposed. In contrast to distance-based approaches, this method structures the operational environments and enhances the effectiveness of V&V activities. Scenarios capture the temporal progression between scenes where a scene describes the scenery, the movable objects and their self-representation. Furthermore, a scenario entails actions and events as well as actors’ goals and values [BM18]. An open challenge remains to utilize scenario-based analyses in the context of risk management. For example, it is an open question how risk acceptance criteria and the derived design and validation targets can be applied to a scenario catalog that is supposed to describe the operational context of an ADS.

A useful tool to analyze and evaluate risks in a scenario-based approach that was proposed in the VVM project [VVM2023] is a solution-independent behavior specification. This behavior specification increases the traceability of verification and validation activities and rigorously provides a set of system-level requirements. Decomposing the system with a capabilities perspective can be an approach to obtain and justify quality requirements towards and determine validation targets for functional components of the ADS. Deriving such requirements towards the perception chain in a traceable manner based on previously established risk acceptance criteria remains a challenge. This decomposition problem is further detailed in section 7.3.

Another noteworthy process stems from the ENABLE-S3 project [ENS2019], aimed at diminishing total validation time by establishing a cross-domain V&V platform. A essential components of this proposed method include defining generic testing architectures which can be used as blueprints. These blueprints are instrumental in guiding users of the method through the development of an effective validation toolchain. Another component of the project are 24 patterns which describe diverse processes like scenario-based testing and support the V&V workflow. The project, with a pronounced focus on practical application, illustrates the concept through 13 representative use cases.

7.2 Validation and Verification of ADS Perception Systems

The testing of complex systems – especially perception systems – must be viewed from different perspectives. On the one hand, the system itself must be tested with sufficiently realistic stimuli. These depend on the test objective and the level of integration. A single sensor can be tested differently than a fusing perception system of different sensor modalities, possibly including localization and high-resolution maps.

On the other hand, the execution of the system in the selected environment must be considered. If models are used, these must be validated in advance for their intended use in order to be credible for their intended use. Requirements for the test systems are again dependent on the test objective. A closed loop simulation for testing the entire system has different requirements than, for example, an open loop simulation based on recorded real images as input.

Finally, the test evaluation for perception systems must also be considered. Common metrics are not designed for the execution of perception tasks in the vehicle context. In order to compare algorithms and select the best one, metrics must be developed that provide information on the quality of perception in the driving context. Object tracking over time and proximity to the ego vehicle can play a role here, for example.

Test Objective and Level of Integration

The decomposition of the overall system creates various subsystems that are developed independently of each other and then integrated. Tests must be carried out at each integration stage

in order to find errors as early as possible. The test objectives change depending on the integration level, as the system under test and its functionality become more extensive. In the case of perception systems, all sensors and their algorithms can initially be tested individually. Individual sensors are then linked and their detections fused. A distinction can be made between early-fusion, late-fusion and early + late-fusion approaches. During fusion, different sensor modalities – e.g. camera plus LiDAR – can be fused. In addition, the same sensors can be fused at different mounting positions on the vehicle, e.g. for a camera belt around the vehicle.

For individual sensors, but especially for the entire perception system, it must be shown during testing that it meets the requirements in defined scenarios. These scenarios need to cover the ODD and the use cases of the ADS adequately, ideally providing some (quantitative) coverage measure. The evidence gathered at each integration level must then be combined in a safety argumentation for the entire perception system. A process must be defined that defines the scenarios and hence test cases to be considered and defines a minimum level of quality for each test case execution. Cross-company standards are essential in order to set a baseline for testing and hence for an important activity to assure road safety.

Test Content

The test content is defined by the data used for the test, i.e. contained in the scenario. This can be, for example, recorded camera images or point clouds from a LiDAR sensor. To ensure an adequate test coverage, the data must be carefully selected. Data recorded in real traffic contains long, redundant sections without any particular events or content relevant to the test. To increase the efficiency of tests, it must be possible to find relevant and specific content in the data. This requires methods for data search and interpretation.

There are various approaches to searching for data. By enriching or tagging the data, additional information can be added to the individual images or sequences. Images and situations can then be searched for in the data using these tags. [RSS2022] Suitable artificial intelligence (AI) methods can also be used to search "directly" in the images. [RLS2023, RPS+2023] show, for example, how AI can be used to perform a semantic search based on natural language. The efficiency of the methods is particularly critical here, as they have to be carried out on very extensive data sets.

Once the relevant test data has been selected, the data set must be analyzed. A systematic approach [PET2022b] must be chosen to avoid biases or data leakage. Biases can occur in various forms, for example by recording data only at certain times of the year or day as well as by the data selection approaches mentioned above. Especially the use of AI for searching for relevant data may lead to a redundant nature of the bias issue. Data leakage describes the problem that the data used for testing is very similar to the training data of a function. In addition, AI can be used to check the extent to which the test data represents the entire data set. [SRL+2022]

An emerging branch of research is also the artificial generation of training data using, for example, generative adversarial neural network architectures [RIG2022], transformer networks, or foundation models (see GAIA-1 / NXTAIM). Another challenge is the provision of consistent test data across different sensor modalities. Approaches that go beyond replaying recorded data are still in their infancy. Both the synthetic generation of realistic raw sensor data for individual LiDAR and radar systems as well as the consistent modification of recorded sensor data for different sensor modalities have not yet been sufficiently researched.

Test Case Execution

Another aspect is the test execution. Here, questions need to be answered about the extent to which a test environment is able to generate valid test results. To this end, the validity of the test environment must be measured and evaluated. Depending on the test environment type, different parts of the system are modeled. The model quality of each individual model must be evaluated depending on the purpose of the model and the test objective, and with respect to credibility. In addition, the interaction of the models and the run time of the test execution must be validated.

Simulations in particular require additional computing time for rendering and environment simulation. Here, it must be validated that these do not have a negative impact on the validity of the test. Furthermore, there are increasingly extensive plausibility checks within the sensors used. These checks must also be sufficiently well stimulated in the simulation. The validity concerns of simulation models for the perception chain are discussed in chapter 6. The test bench types are addressed in section 7.5.

Test Evaluation

A further perspective on the testing of perception systems is a targeted evaluation. A simple calculation of common metrics (e.g. mean intersection over union, mIOU) is not sufficient here. If, for example, the mIOU is calculated for all pedestrians present in the scenario, relevant scenes (pedestrians directly in front of the vehicle) and irrelevant scenes (pedestrians further away and on the sidewalk) are mixed together. In addition, such evaluation methods are not able to infer the triggering conditions for a misperception in the context of SOTIF. Perceptual errors can, for example, be triggered by special properties of surfaces or combinations of foreground and background. These triggering conditions must be extracted from a test in order to be able to increase the degree to which the function performs as intended. This requires structured description formats for scenarios that take into account the different levels of abstraction.

For planning algorithms, there are established description formats such as scenarios and maneuvers for the relevant environment or the content of a test in general, cf. [PEG2024, VVM2023]. The formats for test case specification for testing planning algorithms are unsuitable for testing perception algorithms. There are currently no special description formats tailored to perception – there is a great need for a systematic recording and structured description of the relevant test content.

7.3 Tool-based Continuous Software Development

The industry is facing a set of challenges with respect to continuous integration and testing as a part of the entire software development process when it comes to employing the classical V-model during the development life cycle.

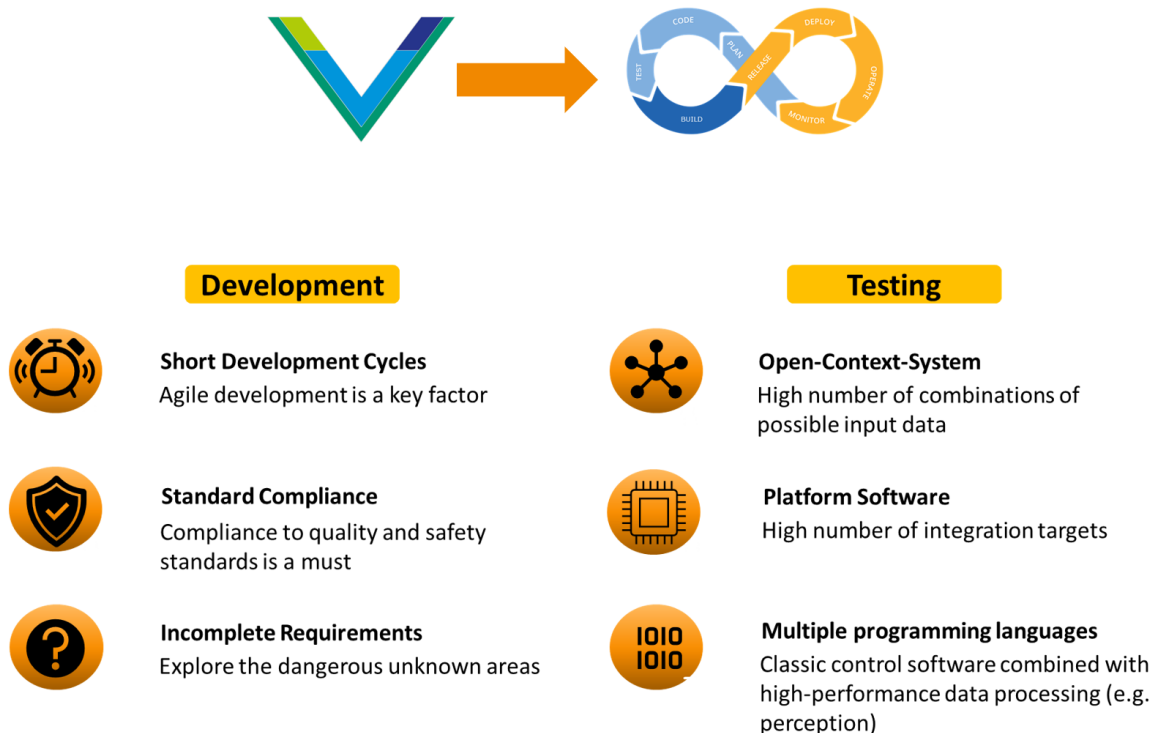


Figure 36 Challenges in applying continuous integration and testing to automotive systems [AVL SFR GmbH]

Figure 36 Figure 36 Challenges in applying continuous integration and testing to automotive systems [AVL SFR GmbH] points to the most relevant challenges in continuous integration and testing. These challenges require a dedicated outline for processes, methods and tools in order to deliver development items within time, cost, and quality bounds and to ensure their compliance to relevant standards for quality and safety. The usage of simulation to integrate and test the ADS in this context is one of the measures required to reach above mentioned challenges. As shown in Figure 37, the continuous life cycle of requirements update, function development, integration, and testing can be executed with high efficiency under use of flexible, but also reliable handling and automation using a simulation environment that includes electronic control units and other electronic subsystems.

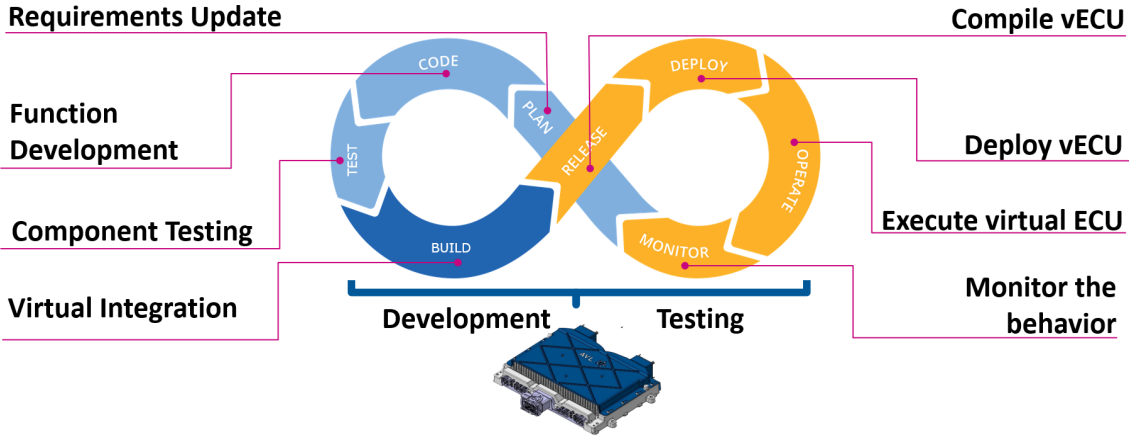


Figure 37 Continuous development and testing using virtual environments (AVL SFR GmbH)

In order to comply with defined processes, sets of enabling or supporting tools are necessary. These development environments have to fulfill specific high-level requirements. First of all, the work flow and execution of integration and testing must represent central aspects of a state-of-the-art model based development. Widely established product development structures require the ability to configure and orchestrate various third-party tools to ensure consistently high-quality work products. Furthermore, efficient continuous integration and testing requires the provision of a seamless connection between the ADS’s architecture and the development and simulation that allows a fully agile continuous integration and continuous testing work flow. A precondition to use the results of testing using virtual environment/simulation as proof for the fulfillment of relevant safety requirement is a credibility assessment as described in chapter 6 including the qualification of the used tools according to ISO 26262.

The Digital Loop showcase⁵⁶ featured a consortium of multiple companies demonstrating a process for addressing the challenges associated with OTA updates for software-defined vehicles [NSB2021]. The primary goal of the project was to expedite the approval process for updates by employing virtual methods and minimizing the reliance on real-world tests. The proposed approach revolves around a continuous integration and continuous deployment framework. Key elements of this process include a high-fidelity virtual simulation environment and automated test execution and evaluation. Several challenges remain, such as into which environments to use for continuous integration and testing and how to ensure their credibility. These issues will be further elaborated in the following section.

56 <https://www.digi-loop.com/>, accessed Jan 14, 2024

7.4 Test Environment Credibility and Test Case Allocation

The increasing complexity of systems-of-systems and the shortened development cycles of cyber-physical systems necessitate a paradigm shift in the application of validation methods involving simulation. Simulation, once primarily employed for analysis and verification, has now become deeply integrated into the product development process, serving as a crucial tool for validation of ADSs. As demonstrated in chapter 6, to validate results obtained through simulation-based test methods, it is essential to establish a framework for conducting credibility argumentation [[UNE2022](#), [VVM2023](#), [DC2022](#), [HS2022](#)]. Credibility depends not only on the quality of correlation between the models and reality of interest, but also on the development process used, the modeling and engineering skills of the actors involved, and the traceability and reliability of the models, parameters, results, and coverage, among others. [[FD2023](#)]

Assessing and Managing Test Environments

Nowadays, there is a common understanding of different test environments such as Model-in-the-Loop (MiL), Software-in-the-Loop (SiL), Hardware-in-the-Loop (HiL), Vehicle-in-the-Loop (ViL) and proving ground. Together with suitable test methods and use cases, these test environments can be found in the ASAM report on testing a software defined vehicle [[ASA2022](#)].

Novel requirements for the V&V of automated vehicles often necessitate the use of a combination of multiple test environments. The once clear separation of those environments becomes therefore increasingly blurred. To meet those new requirements, a modular framework needs to be formed that enables a flexible combination of different testing environments and ensures consistency with the specified objectives and system under test. The current paradigm shift in the use of simulation within the development process is evident through the transformation of simulation into a service. In the creation, execution and evaluation of a simulation, multiple stakeholders may be involved. To accurately assess the level of trust that can be attributed to a result obtained from such a simulation, it is imperative to establish a mutual understanding of the limitations.

Figure 38 illustrates a range of potential test environments applicable to the verification and validation (V&V) process. By examining four cyber-physical test benches, the depiction highlights the integration of multiple environments. For instance, a Hardware-in-the-Loop test bench can be utilized either within a virtual environment with an associated cloud simulation or in conjunction with a Vehicle-in-the-Loop test bench. Conveying the limitations and reproducible effects of test environments poses a challenging task, as many of these aspects (e.g., frequency response, non-linear effects, etc.) are intricate and not easily captured in a simple list of bullet points. Therefore, a novel approach is required. While some initiatives have begun describing the system in development and the associated test environment using model-based systems engineering, as seen in [[WMG+2024](#)], further research is imperative to refine and elaborate on this approach.

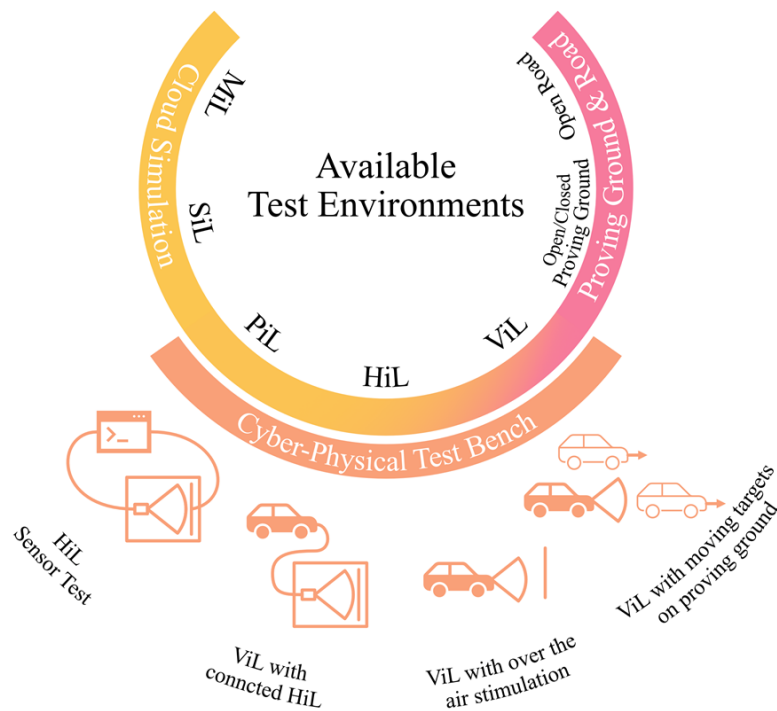


Figure 38 Illustration of diverse test environments (MiL, SiL, HiL, etc.) categorized into cloud simulation, cyber-physical test bench, and proving ground & road. Highlighted are instances of cyber-physical environments which utilize a combination of test environments. [DFE2023]

Choosing and Specifying Suitable Test Environments

The challenge is to find an optimal distribution between the different test environments for specific use cases considering their limitations – a tradeoff between effort to prove credibility, effort for modeling, effort for validation, and process related topics like resources and skills. Based on the risk decomposition on system level, some components might not be relevant in a V&V activity. For instance, when validating a path planning algorithm, a highly accurate representation of the perception chain may not be imperative. Conversely, in the validation of the perception chain, the impact of vehicle dynamics might be negligible. Thus far, there has been no definitive guidance regarding the appropriate level of detail required for various activities within a V&V process. Hence, there is a pressing need to explore and define the optimal level of detail for diverse activities within a V&V process through dedicated research, ensuring more effective and standardized methodologies in the development and validation of ADS.

Figure 39 presents a qualitative comparison of two exemplary test strategies that provide varying levels of validity. Strategy 1 is a common approach, where tests on a test bench and a proving ground are used for approval. In strategy 2, virtual environments are used for screening of scenarios that are critical and need to be re-evaluated in an environment with high credibility. The total effort of both strategies will be different. Hence a formalized process for the selection of environments for the validation is necessary and needs to be researched.

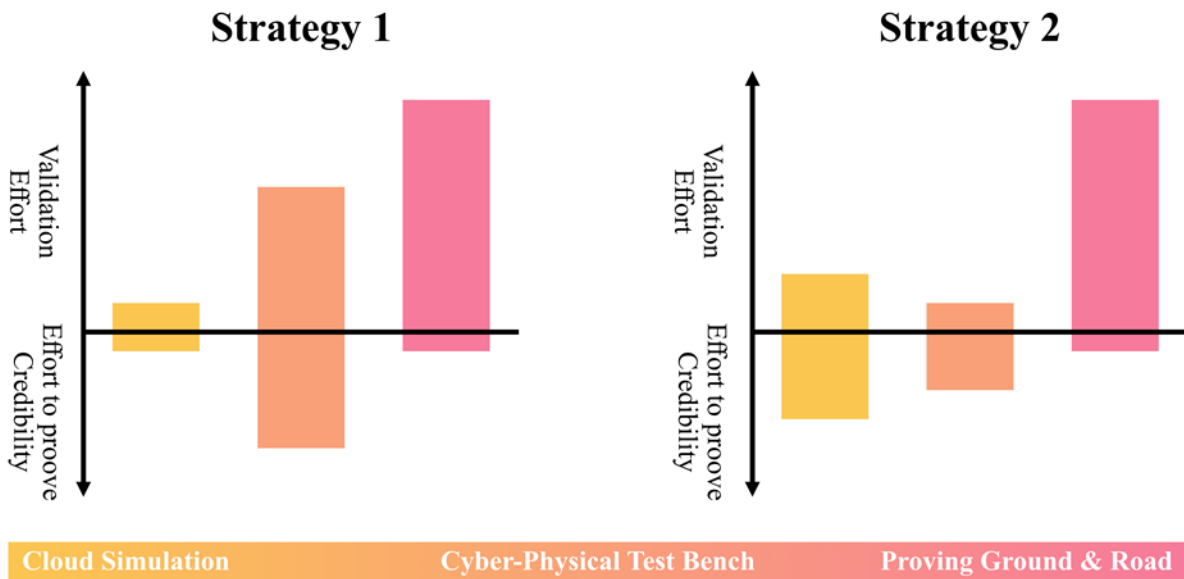


Figure 39 Different exemplary test strategies with associated efforts [DFF2023]

7.5 Challenges in Decompositional Verification and Validation

For component-wise verification and validation of an ADS, as suggested in section 7.1, we need compositional means of system verification and validation, permitting to decompose the overall assurance argument into manageable pieces that focus on sub-functions and architectural components. Compositional V&V builds on two steps, namely verification and validation at component level and the synthesis of an overall validation or verification argument from the component-wise V&V verdicts [BEN2018]. As achieving and verifying 100% absence of undesirable behavior is elusive, especially in the perception chain where a significant positive risk of misperception is inevitable, these compositional methods need to cover quantitative validation and verification up to an agreed quantitative target for component-level requirement satisfaction, with the quantitative requirements at component level being rigorously derived from the quantitative system-level requirements.

Such quantitative V&V shall provide guarantees that the frequency of requirement violations remains below the agreed quantitative target. Unfortunately, applicable compositional methods for quantitative validation and verification are currently missing, as compositionality requires adequate treatment of conditional probabilities rather than absolute probabilities: The frequency of requirement violations in a component will vary situationally and generally depends on operational conditions and context, as will the propagation or masking, respectively, of a component's violating behavior by subsequent components or functions along the ADS' function chain. As a consequence, the overall error rate of the system depends on mutually induced conditions and the interplay of conditional probabilities that these conditions evoke.

It is thus not surprising that the well-established qualitative (i.e. Boolean, generating verdicts of total absence of violating behavior) contract frameworks for compositional analysis and verification (cf. [BEN2018] for an overview) currently have no usable quantitative counterparts. Suggestions like [DCL2011] fail to comprehensively formalize conditional probabilities and consequently are inapt to trace the dependency of guarantees on distributionally varying probabilistic assumptions – an obvious prerequisite for successful compositional quantitative reasoning in verification and validation in general and for simulation- and scenario-based statistical compositional verification in particular.

7.6 Future Directions and Research Needs

As described in sections 7.1 through 7.5 a wide array of advancements has been achieved in previous

research projects. Nevertheless, it has become obvious that many approaches still need further research. Furthermore, several challenges still require the exploration of novel concepts and approaches.

Defining a societally accepted level of risk remains a challenge for manufacturers and operators of automated driving systems as communication about that level and associated kinds of risks is a delicate task. How these stakeholders can translate such risk acceptance criteria into high-level engineering requirements requires further investigation. A rigorous and partially quantitative risk model could leverage the management of such risks and assumptions that inform the risk assessment during verification and validation.

The perception chain is an important sub-system of an ADS in that regard. In order to verify and validate perception systems, appropriate test environments need to be designed and supplied with adequate test data, i.e. scenarios in the context of automated driving. These test environments still pose several challenges, such as assessing and managing the level of validity that these test benches exhibit. On one hand, selecting and specifying scenarios and associated data for test cases based on the ODD and use cases is an open question. On the other hand, specifying suitable metrics highly depends on the system under test. These metrics as well as that possibly recursive relationship between the two requires further research. Furthermore, the discussion should be widened about what baseline in testing methods and test data society expects from actors involved in V&V in order to ascertain a minimum level of quality, e.g. expressed through guarantees.

Although the validity of simulation models was addressed before in chapter 6, the need for investigating how to conceive credible and sufficiently valid test environments and test benches was identified. Leveraging a flexible combination of different environments and even novel techniques like simulation-as-a-service is key in order to improve the efficiency of a rigorous V&V. The research direction of how to use model-based systems engineering approaches to manage the trade-off between validity and cost and to assign test cases to test environments seems promising and worth pushing forward. This becomes especially evident as decompositional approaches to managing the systems' complexity currently suffer from a lack of established methods that adequately capture the complex interdependent and often probabilistic relationships of architectural components.

8. Architectural Requirements

The design of system architectures for safety-critical automated systems in highly uncertain contexts is still a wide field of research. Especially when the systems are subject to epistemic uncertainty and highly context-dependent requirements, the design of such architectures becomes challenging.

A broadly applied approach in the literature to handle uncertainty is the realization of runtime monitoring components for properties that cannot be guaranteed at design time. To make system behavior and system configurations adaptable to situations that are unknown during design and development, approaches in the domain of self-adaptive systems often target the design and implementation of models for an explicit representation of design-time knowledge as a basis for runtime monitoring and runtime adaptation (also referred to as *models@runtime*) [KG2016, LPR+2016, TSW2018]. The represented knowledge shall allow a system to infer suitable actions in unknown contexts by adapting the system behavior and/or the system configuration, where fixed rules may be ill-defined [MAU2000, LPR+2016, HMS2019]. The models that are used for knowledge representation can include models of the system architecture, (formalized) requirements, dynamic models for predicting the system's behavior, models of the expectable context (i.e., the Operational Design Domain), and more [MAU2000, ST2013, AGJ+2014, LPR+2016, KG2016, RSS+2020, RGL+2022, SES2023].

Systems that apply such knowledge to perceive and assess the current situation as well as their current and future capabilities with respect to their mission objectives and are thus able to make decisions under uncertainty are referred to as *self-aware* systems in the literature [GLS16]. [ARH2023] structures scientific challenges concerning dynamic risk management of autonomous systems and provides a conceptual frame for monitoring problem classification.

Regarding the design of system architectures in the field of automated driving, a great part of recent research has been focused on concrete (mostly functional and/or logical) architecture views that include dedicated modules for runtime monitoring [KBR12, BT2016, THS+2017, URR+2017, BNE+2018, TZM+2018, SAE2022], while the explicit discussion of the integration between runtime and design time models – especially across different architecture viewpoints – are less common in the field [KBR12, SME+2017a, SME+2017b, TZM+2018, AAF+2019, RSS+2020].

To realize a system architecture that facilitates 1) the dynamic adaptation of the perception chain, 2) runtime monitoring, and 3) safe degradation in case of perception insufficiencies (cf. [Chapter 4](#) and [5](#)), a combination of established architecture design concepts for automated systems and rigorous model-centric knowledge representation that bridges design- and runtime will be required.

8.1 Related Work

Reconfigurable architectures for automated vehicles that support safe degradation by activating pre-defined hot and cold stand by nodes have e.g. been introduced by [KBR12]. By leveraging models for multi-viewpoint component and interface descriptions and contract-like mechanisms for describing assumptions and guarantees for components and interfaces, [SME+2017a, KAK+2019, AAF+2019] present concepts for reconfigurable architectures that allow to generate (optimal) system configurations under non-functional constraints and objectives, such as timing behavior or required data quality. In contrast to [KAK+2019, SME+2017a], [AAF+2019] explicitly apply their analyses to identify operating modes that restrict the system's operational design domain at runtime. [TZM+2018] present a doer-checker architecture in which a supervisor channel monitors a nominal channel based on knowledge that is derived from the results of hazard analyses and risk assessments.

Regarding concrete proposals for functional and/or logical architectures for automated vehicles, respectively, [BT2016, THS+2017, URR+2017] present architectures that include components for performance monitoring. [THS+2017] suggest an architecture that includes performance monitors for each individual functional element and adds "fusion components" that are responsible for creating a

coherent image of the system's performance. Similarly, [URR+2017] describe functional elements for self-perception and self-representation that are responsible for processing data to assess performance (self-perception) and applying models to create knowledge about the internal system states and the current system performance. The architecture explicitly provides interfaces to adapt decision making depending on the current system performance. [RSS2+020] present a concept to derive runtime monitors for a platooning application from design-time models and explicitly bridge the gap between model-based safety analyses and the instantiation of runtime monitors in a limited operational design domain. The issue of optimal, in the sense of as informed as possible under the prevailing uncertainties in system observation, monitoring of complex spatio-temporal safety properties has been addressed in [FFK+2022]. Optimizing resilience of such condition monitoring against misperceptions has also been subject of [FHD+2023]; both publications demonstrate that condition monitoring can be rendered considerably more reliable than atomic percepts.

With respect to the dynamic adaptation of the perception chain, early publications introduce system architectures that support "gaze control", i.e. the controlled actuation of sensor (or camera) arrays toward areas of interest in a system's environment depending on the systems mission [MBF+1996]. Conceptually related approaches have been published for grid-map-based environment perception as "attention maps" [HMG+2023]. Regarding the system architecture, this approach allows to dynamically (re)configure the processing chain to focus system resources depending on those areas of the environment that are most relevant for the system's mission and depending on the perception performance requirements within those areas.

8.2 Research Questions and Possible Ways Forward

While the related work addresses aspects of architectures that facilitate 1) the dynamic adaptation of the processing chain, 2) runtime monitoring, and 3) safe degradation, open questions remain. From the discussed approaches, only [HMG+2023] specifically focus on the perception sub-system. At the same time, while the authors state that performance requirements (and in turn the available performance of a perception module) can be used to guide reconfiguration, they also conclude that performance assessment for the perception system is still a challenging element in their concept. In addition, none of the discussed monitoring approaches specifically focus on monitoring, and/or modeling, the quality of a perception sub-system. The high-level concepts of architectures for self-adaptable systems and the domain-specific functional and logical architecture views can be a starting point for the CONTROL architecture – fundamental architectural concepts and interfaces of performance monitors to decision making and behavior generation modules must hence not be reinvented.

A key question in the context of the architecture design, however, will be how actual monitoring components for the perception chain can be defined. For this purpose, performance models must be defined, as well. For the purpose of designing reconfigurable architectures that target the perception chain, performance models for the individual sensors, as well as performance models for the respective processing modules in the perception chain are required. Regarding sensor performance models, the research questions discussed in Chapter 4 with respect to the diligent identification of sensor strengths and weaknesses as well as questions regarding sensor model verification and validation will be crucial to answer to apply suitable performance metrics from sensor properties (cf. Chapter 4) to the performance metrics for the processing modules (cf. Chapter 3) – that provide the required expressiveness to enable dynamic adaptation and degradation. This step is also required for the adverse conditions of AI-based classification components, see Chapter 5. A related question is how resilience of such a function chain against uncertainties can be optimized while retaining its intended functionality, and whether such resilience optimization can be achieved architecturally or algorithmically. Additionally, even if performance models are available, the definition of monitoring *thresholds* becomes an additional challenge for the definition of monitoring architectures. In combination with the need for a coherent risk-based safety assurance cases (cf. Chapter 9), it remains an open but crucial issue to answer the question how such monitoring thresholds

can be derived from the results of (possibly context-specific) risk analyses.

While much of the related work focuses on functional, logical or software architecture views, dealing with uncertainty requires more than that: As mentioned in the chapter introduction, a key challenge for systems that operate under uncertainty in an open context are context-dependent requirements. For the systems in the scope of the CONTROL approach, this relevant system context is the Operational Design Domain. Modern architecture frameworks explicitly address the definition of architecture viewpoints for the representation of the system (or operational) context, use-cases, scenarios, and requirements. Especially in a safety-critical context, regarding the definition of assurance arguments (cf. [Chapter 9](#)), these representations are key for the traceability of (performance) requirements and/or design decisions as well as -assumptions to system architecture elements. Even if the focus of the CONTROL approach is on functional architecture views, the required design- and runtime models of the ODD and suitable runtime representations for performance requirements and -metrics must be defined and represented in the system architecture.

In summary, the following research items are still unsolved from an architectural point of view:

- **Development of Performance Models:** Devising detailed performance models for both individual sensors and processing modules within the perception subsystem to guide the dynamic adaptation of the system.
- **Monitoring Quality of Perception Sub-System:** Establishing methodologies for continuous monitoring and/or modeling of the quality of the perception sub-system, which has been less emphasized in existing literature.
- **Defining Monitoring Thresholds:** Determining appropriate monitoring thresholds in conjunction with risk analysis outcomes, a crucial step for the suitable definition of monitoring architectures.
- **Resilience Optimization:** Investigating how the resilience of function chains against uncertainties can be optimized, exploring whether these solutions are architectural, algorithmic, or a combination of both.
- **Performance Metrics and Dynamic Adaptation:** Linking sensor properties with performance metrics for processing modules to empower dynamic adaptation and degradation under adverse conditions.
- **Architecture Frameworks and ODD Representations:** Integrating design and runtime models within modern architecture frameworks to ensure the traceability of performance requirements and design decisions within the ODD.
- **Knowledge Representation Across Architectural Viewpoints:** Enhancing the representation of knowledge to span across different architectural viewpoints, ensuring that decisions made at design time remain valid and operational at runtime.

9. Putting it all together: Deriving Safety Assurance Cases for highly automated systems exploiting digital twins

9.1 Motivation / Need for assurance cases

The previous sections have laid out the various open technological challenges, which need to be solved to construct and validate a highly automated system for operation with bounded risk in complex open environments. The overall assurance process involves a large number of assumptions and decisions, which are interrelated with each other. A statement about the acceptance of risks associated with a deployed system can only be made if the meaning and sufficiency of created evidence, e.g. concrete validation test results, for the claims to be demonstrated is made *explicit*.

Structured assurance cases are an established means to make the safety reasoning assessable for certification bodies and aid companies during constructive assurance. Technically, an assurance case is a “reasoned, auditable artefact that supports the contention that its top-level claim is satisfied, including systematic argumentation and its underlying evidence and explicit assumptions that support the claim. [It contains:]

- one or more claims about properties;
- arguments that logically link the evidence and any assumptions to the claim(s);
- a body of evidence and possibly assumptions supporting these arguments for the claim(s); and
- justification of the choice of top-level claim and the method of reasoning.” [ISO 25026-1]

Thus, assurance cases enable to tell a system’s “safety story” in a step-wise manner by explicitly laying out 1) the decomposition strategies of the top safety claim in smaller sub claims (e.g. related to problem space analysis or validation strategies) and 2) the strength of created evidence to support these claims. This abstract view on the safety story enables parties to identify systematic flaws in their assurance processes early. More importantly the abstract view provides a means to communicate the story to external stakeholders who are not involved deeply in the development and assurance process. While the idea of making the safety reasoning explicit for different stakeholders is the central driver for the use of assurance cases in the first place, there exist various concrete methods and notations to operationalize this idea. The most prominent representants are the *Claims-Argument-Evidence (CAE)*⁵⁷ and the *Goal Structuring Notation (GSN)*⁵⁸ tightly related to the OMG standardized *Structured Assurance Metamodel (SACM)*⁵⁹.

For the realization of automated vehicles, engineers have to cope with significantly higher complexity of system, environment and machine-learning based perception, compared to pre-existing software systems in the transportation domain. With this increased complexity and new software development paradigms, methods to develop comprehensive and convincing safety assurance cases are more important than ever. While past German and European flagship projects have focused on holistic (V&V methods) and specific (scenario-based validation: PEGASUS, simulation technology: SetLevel & Vivid, ML: KI-absicherung) assurance case frameworks, no proven assurance case patterns exist for addressing and confining risks stemming from uncertainties in perception components with digital twin-based continuous validation.

9.2 State-of-the-art assurance case frameworks

The projects PEGASUS, V&V Methods and SET Level, all part of the PEGASUS Project family, contributed methods and technologies to build a holistic safety assurance case that is based on a scenario- and simulation-based verification and validation approach.

⁵⁷ <https://claimsargumentevidence.org/>

⁵⁸ <https://goalstructuringnotation.info/>

⁵⁹ <https://www.omg.org/spec/SACM/>

PEGASUS⁶⁰ presents four layers to structure a safety argument for ADS: 1) an ADS acceptance model, based on established technology acceptance models, 2) a GSN-based safety assurance case including extensions of the GSN standard to rate the credibility of safety case elements, 3) a description of methods and tools used to generate evidence, and 4) the description of concrete evidence generated by methods and tools. The PEGASUS approach thus covers highly relevant aspects by addressing the credibility of claims and arguments in the safety case as well as the credibility of methods, tools, and evidence. However, the PEGASUS approach does not consider a dedicated confidence argument, why Methods, Tools are suitable to generate credible evidence. (Mazzege, J. & Maus, A., 2019).

V&V Methods⁶¹ covers similar aspects but builds the assurance case on an exhaustive product and confidence argument. The argument is structured by patterns that shall enable a sufficiently complete decomposition of assumptions with respect to the “open context” (i.e. the essential unknowns in the design of the Operational Design Domain). This decomposition is supported by arguments that are based on coverage of normative sources, objective evidence and the refutation of subjective doubts in the elements of the assurance case. The VVM assurance case argues from a risk-based perspective what allows to tailor the composition of the argument for different stakeholder perspectives on safety. While the assurance case contains a general “Top-Level” argument. Details are only provided for the concrete methods that have been developed in the project. The VVMMethods assurance case does not systematically consider aspects related to uncertainty in the perception chain.

While PEGASUS and V&V Methods focused on methodological frameworks to realize scenario-based safety assurance, the **SET Level**⁶² project had a technical focus on simulation technology. A major result is the *Credible Simulation Process Framework*⁶³, which enhances the development and validation of automated driving functions through the use of simulation. It provides a structured set of processes and procedures that assure traceability, adaptability to specific organizational needs, and promotes collaboration, offering its users a modular and reusable approach to integrating credible simulation tasks into their engineering workflows. The framework provides an extensive basis for structuring assurance arguments specifically focusing on simulation credibility.

While an explicit focus was placed on *pre-deployment assurance* in the PEGASUS family, continuous *post-deployment* safety monitoring and update processes were the subject of the **Step-Up!CPS**⁶⁴ project and the ongoing **AutoDevSafetyOps**⁶⁵ project.

Specifically for the perception chain including machine-learning-based components, the following initiatives contributed with assurance approaches and argumentation fragments.

Within the **KI-Absicherung**⁶⁶ project, an approach⁶⁷ for safety assurance of machine-learning-based perception functions was developed. The need for an understanding of the potential causes of insufficiencies in the ML models is emphasized such that this can be used to identify appropriate design-time measures to minimize the risk of insufficiencies as well as operation-time measures to mitigate against inevitable residual errors in the ML model. Design-time measures could include improving the quality of training data, appropriate selection of ML technologies architectures and development approaches as well as suitable testing procedures. Operation-time measures to identify and compensate for residual insufficiencies in the model could include out-of-distribution detection, use of ensembles and online monitoring of environment assumptions. Furthermore, the following categories of evidence were proposed to support a safety assurance argument:

⁶⁰ <https://www.pegasusprojekt.de/en>

⁶¹ <https://www.vvm-projekt.de/en/>

⁶² <https://setlevel.de/en/>

⁶³ https://gitlab.setlevel.de/open/processes_and_traceability/credible_simulation_process_framework

⁶⁴ <https://www.stepup-cps.de/>

⁶⁵ <https://www.linkedin.com/company/autodevsafeops>

⁶⁶ <https://www.ki-absicherung-projekt.de/en/>

⁶⁷ [BHH+2022]

1. direct confirmation of residual error rates of the ML-based function based on a detailed definition of acceptance criteria,
2. Evaluation of the relevance of ML-specific insufficiencies such as robustness, generalization, brittleness, fairness and explainability and their potential impact on safety requirements,
3. Evaluation of the effectiveness of design-time controls to reduce the presence of insufficiencies and
4. Evaluation of the effectiveness of operation-time controls to mitigate the effects of residual insufficiencies in the model.

However, it should be noted that due to the lack of explainability of many ML models as well as the gaps between the semantic specification space (in terms of ODD ontologies), the syntactic space (in terms of inputs to the model such as individual pixel values) and the latent space representing the learnt concepts, it will typically not be possible to perform a causal safety analysis per failure. Therefore, such causes need to be hypothesized based on various observations and the effectiveness of counter-measures equally observed via a number of indirect measurements.

The ongoing **SAFE.Train**⁶⁸ project is developing approaches, which build upon the results of KI-Absicherung, applied to perception tasks for driverless trains. Current work on assurance arguments within the project is focused on the creation of an appropriate set of safety requirements including a mapping to measurable safety-related properties of the ML-models including the identification of suitable metrics and threshold values. These in turn will be used to construct a reference assurance argument linking high-level safety requirements to ML-specific properties and related evidence.

The Guidance on the Assurance of Machine Learning in Autonomous Systems (**AMLAS**)⁶⁹ provides an overview of different ML-lifecycle stages and guides the development of assurance cases for ML components by examining each stage in turn. The guideline emphasizes that the development of an effective safety argument requires an iterative process involving a large number of stakeholders. Furthermore, it stresses the importance that the safety considerations are meaningful only when scoped within the wider system and operational context. The complementary Guidance on the Safety Assurance of Autonomous Systems in Complex Environments (**SACE**)⁷⁰ provides assurance argument patterns for the identification, decomposition, verification and validation of system safety requirements at the system level.

A recent publication⁷¹ has explicitly addressed how properties of open context systems and the use of machine learning introduces uncertainty into the safety assurance process. The paper explored how the resulting uncertainty associated with our understanding of the environment and task, our observations used to develop (train) and evaluate the system and the technical system itself can be used to inform the safety assurance task. Furthermore, it was proposed that definitions of types and severity of uncertainty could be used to evaluate the confidence with which arguments and supporting evidence can be evaluated for each of these dimensions. The paper also motivates an inherently iterative development assurance process as a necessary means for the continual reduction of uncertainty both within the system as well as the assurance process itself.

Many of the above-mentioned contributions to the state-of-the-art are informing the upcoming publication of **ISO PAS 8800 Road Vehicles – Safety and Artificial Intelligence**, the publicly available specification on safety and AI for use within road vehicle applications. The standard proposes a structured assurance argument and evidence to support demonstrate that a set AI-specific safety properties are met and can be traced to the system-level safety requirements. An iterative approach to the development of the argument, including continual re-evaluation during operation is further proposed.

⁶⁸ <https://safetrain-projekt.de/en/>

⁶⁹ [HPP+2021]

⁷⁰ [HOP+2022]

⁷¹ [BH2023]

9.3 Research needs

The surveyed landscape has revealed substantial advancements in assurance arguments for automated driving systems. Specifically, the focus rested on the most pressing challenges scenario-based system validation, credible simulation processes and the assurance of ML-based components. Despite this progress, the articulation of assurance case patterns that deliver comprehensive coverage of uncertainties within the perception chain using digital twins for continuous validation is an open research problem. The following research gaps highlight aspects, which need to be addressed.

Integration of Holistic Assurance Case Frameworks

A seamless integration of assurance case frameworks from the aforementioned projects is paramount. The aim is to encompass the entire system lifecycle, including pre-deployment digital twin-based virtual testing, real-world field testing and operational safety assurance.

The challenge here is twofold:

- Establishing a robust qualitative and quantitative link between the system-level risk acceptance criteria and the specific uncertainties within the perception chain.
- Developing a cohesive argument structure that controls uncertainty from the perception chain, collating evidence from perception-specific sub-processes including model validation, virtual and field testing, and post-deployment safety monitoring. This methodology should aim to dynamically trace and statically confine the impact of uncertainty on critical decisions, supported by safety contracts across the prediction, decision, and maneuver execution layers.

Harmonizing Simulation Credibility Arguments with System Risk

Verification and validation methods (Chapter 7) provide a verification and validation concept putting all the measures into context to achieve an acceptable residual system risk. The simulation environment plays a major role in it, as it provides the evidence from digital twin-based virtual tests. Based on the work done in the SetLevel project, concretized arguments need to be created harmonizing the credible simulation processes with varying requirements and their associated system risks.

Concretizing Product and Process Arguments for the Perception Chain

Within this context, the intricacy lies in garnering solidified product and process arguments for the perception chain. This involves an in-depth application of digital twins to generate evidence, fortifying the safety argument. To that end, the concretization of the methods described in the previous sections focusing on model validation, virtual and field testing, online monitoring, architectural uncertainty control as well as safety-contract-based prediction, decision and maneuver execution will be the basis to concretize the decomposition of and required evidence in product and process arguments.

Addressing Uncertainty in Evidence and Arguments

Systematically and continuously identifying, assessing and handling possible concerns regarding the robustness of arguments, and strength of the underpinning evidence must form one focus of the next research frontier. Generalizable approaches on uncertainty handling in assurance arguments offer a foundation to build upon⁷². These accommodate two perspectives:

1. Assurance arguments need to effectively address environmental/task uncertainties, observation (sensor/data) reliability, and system/model certainty to ensure risks are constrained to an acceptable level.
2. A need exists to analyze uncertainty within the assurance argument's own framework, appraising the integrity of evidence, validity of assumptions, and the robustness of the argument structure itself.

⁷² [BH2023]

These approaches are not only pertinent to the perception chain of the system in question but also underscore the white paper's objectives.

Stakeholder-Centric Communication of Assurance Cases

The effective communication of safety evidence and engineering artifacts to a diverse array of stakeholders is a significant research area deserving exploration. Current practices largely focus on substantiating safety assurance for regulatory purposes rather than addressing information consumption needs across various stakeholders—type approval bodies, insurers, law enforcement, and the general public. Preliminary concepts such as 'assurance case views,' envisioned by the PEGASUS V&V Methods project, hint at the customization of safety argumentation to aid different stakeholder groups in fulfilling their roles.

Reusable Assets and Argumentation Contents

Another gap exists in the formulation of re-usable assets for broader application in the industry. The goal is to create a methodology that not only aids in the construction of an assurance case but also promotes its adoption across diverse use cases. This encompasses the curation of assurance case templates or patterns, annotated notations, and the development of supportive tools that align with industry patterns.

To summarize, the outlined areas of research are critical to close existing gaps and propel the development of robust safety assurance methods in automated driving systems using a digital twin paradigm:

1. **Holistic Frameworks Integration:** Bridging digital twin continuous validation with lifecycle coverage.
2. Simulation Credibility Argumentation:
3. **Perception Chain Arguments:** Product and process argument concretization, incorporating digital-twin generated evidence.
4. **Uncertainty in Assurance:** How to robustly address and encompass uncertainty in environmental understanding and within the argumentation process itself through systematic confidence argumentation.
5. **Stakeholder Communication:** Enhancing the delivery of assurance case content in a stakeholder-specific manner. This includes a model-based integration of stakeholder views based on a formalization of different assurance and argumentation terminology.
6. **Reusable Assurance Case Assets:** Developing and disseminating reusable assurance case components for industry application.

These research needs, once met, will advance the creation of compelling, comprehensive, and convincing safety assurance cases, critical for the widespread acceptance and deployment of highly automated systems.

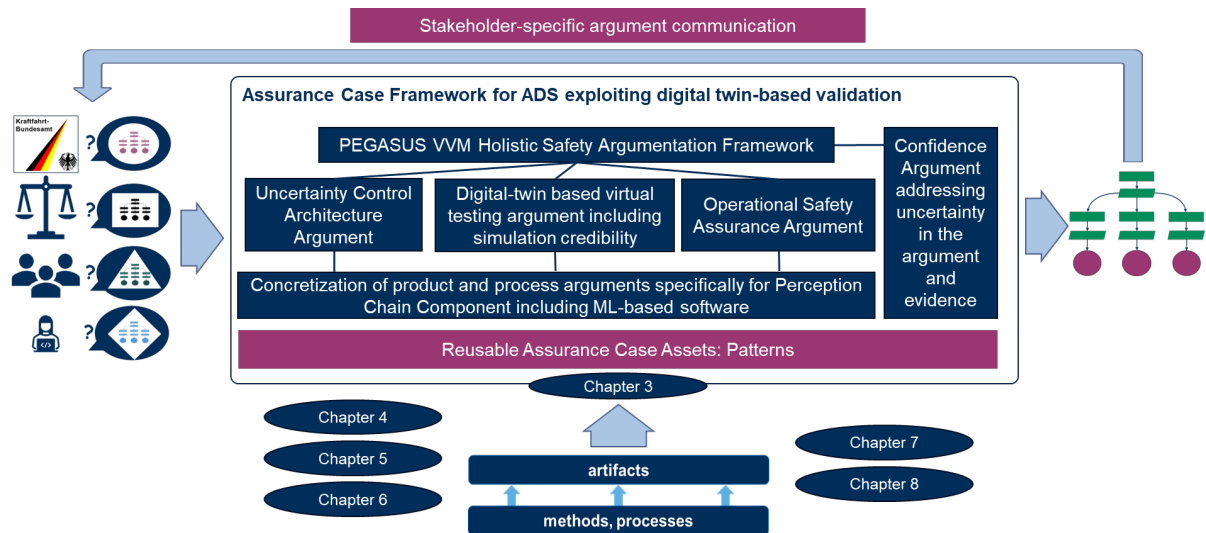


Figure 40 41 Assurance Case Framework for ADS exploiting digital twin-based validation

Figure 40 visually puts the research needs into the context of an assurance process. The chapters 4 (sensor characterization and modeling), 5 (sensor fusion and characterization), 6 (digital twins and simulation environments), 7 (verification and validation methods) and 8 (Uncertainty control architectures) of this whitepaper provide methods, which generate structure and evidence to be used in the argumentation. Chapter 3 (quality metrics and quality guarantees) is closely related to the argumentation, as quality metrics and required thresholds to achieve acceptable risk are influenced by the way in which the argumentation is built up. With this basis to build up the argumentation, the assurance case framework lays out the elements the identified research needs and how their hierarchical relationship looks like: The PEGASUS VVM safety argumentation framework as a top-level structure provides placeholders to integrate concretized arguments regarding uncertainty control architectures, digital-twin based virtual testing, simulation framework credibility, and operational safety assurance. While the dark blue elements represent concrete argumentation patterns based on the developed methods, the red elements are capabilities of the argumentation framework itself: Reusable assurance case assets like patterns realized in tools are important means for adoption of methodical approaches. The stakeholder-specific argument communication requires concepts and realization to present the same assurance evidence along different argumentation lines, dependent on the needs of different stakeholders.

References

- [AAF+2019] R. Adler ; M.N. Akram ; P. Feth ; T. Fukuda ; T. Ishigooka ; S. Otsuka ; D. Schneider ; K. Yoshimura: Engineering and Hardening of Functional Fail-Operational Architectures for Highly Automated Driving. In: *2019 IEEE International Symposium on Software Reliability Engineering Workshops (ISSREW)*. Berlin, Germany : IEEE, 2019 — ISBN 978-1-72815-138-0, pp.30–5
- [AGJ+2014] U. Aßmann ; S. Götz ; J.-M. Jézéquel ; B. Morin ; M. Trapp: A Reference Architecture and Roadmap for Models@run.time Systems. In: BENCOMO, N. ; FRANCE, R. ; CHENG, B. H. C. ; AßMANN, U. (eds.): *Models@run.time, Lecture Notes in Computer Science*. vol. 8378. Cham : Springer International Publishing, 2014 — ISBN 978-3-319-08914-0, pp.1–18
- [AGL2023] Adee, Ahmad; Gansch, Roman; Liggesmeyer, Peter (2023): Systematic Modeling Approach for Environmental Perception Limitations in Automated Driving. DOI: 10.48550/ARXIV.2303.04029.
- [AHD+2023a] Aust, Philip; Hau, Florian; Dickmann, Jürgen; Hein, Matthias A. (2023): Numerical Synthesis of Radar Target Detections Based on Measured Reference Data. In : 2023 20th European Radar Conference (EuRAD), pp. 26–29.
- [AHD+2023b] P. Aust, F. Hau, J. Dickmann and M. A. Hein, "A Data-driven Approach for Stochastic Modeling of Automotive Radar Detections for Extended Objects," 2022 14th German Microwave Conference (GeMiC), Ulm, Germany, 2022, pp. 80-83.
- [AHH+2020] Ahn, N.; Höfer, A.; Herrmann, M.; Donn, C. Real-time Simulation of Physical Multi-sensor Setups. *ATZelectron. Worldw.* 2020, 15, 8–11
- [ALF2022] M. Ahmann, V. T. Le, F. Eichenseer, F. Steimann, and M. Benedikt, "Towards Continuous Simulation Credibility Assessment," Proceedings of Asian Modelica Conference 2022, Tokyo, Japan, Nov. 2022.
- [ARH2023] Adler, Rasmus, Jan Reich, and Richard Hawkins. "Structuring Research Related to Dynamic Risk Management for Autonomous Systems." In *Computer Safety, Reliability, and Security. SAFECOMP 2023 Workshops*, edited by Jérémie Guiochet, Stefano Tonetta, Erwin Schoitsch, Matthieu Roy, and Friedemann Bitsch, 14182:362–68. Lecture Notes in Computer Science. Cham: Springer Nature Switzerland, 2023. https://doi.org/10.1007/978-3-031-40953-0_30.
- [ASA2022] ASAM e.V., "Evolving Landscapes of Collaborative Testing for ADAS & AD: ASAM Test Specification Group Report," 2022. Accessed: Dec. 11, 2023. [Online]. Available: <https://www.asam.net/index.php?eID=dumpFile&t=f&f=4763&token=b18f213f18bf45722faedacd9818e1153ce6ebe1>.

- [BEN2018] A. Benveniste et al., "Contracts for System Design," *Found. Trends Electron. Des. Autom.*, vol. 12, no. 2–3, pp. 124–400, 2018, doi: 10.1561/1000000053.
- [BER2004] G. Berry, "Synchronous methodology for designing hardware, software and mixed embedded systems," 17th International Conference on VLSI Design. Proceedings., Mumbai, India, 2004, pp. 24–25, doi: 10.1109/ICVD.2004.1260897..
- [BH2023] Simon Burton, Benjamin Herd, Addressing uncertainty in the safety assurance of machine-learning, *Front. Comput. Sci.* 5;1132580, doi:10.3389/fcomp.2023.1132580, April 2023
- [BHH+2022] Burton, Simon, Christian Hellert, Fabian Hüger, Michael Mock, and Andreas Rohatschek. "Safety assurance of machine learning for perception functions." In *Deep Neural Networks and Data for Automated Driving: Robustness, Uncertainty Quantification, and Insights Towards Safety*, pp. 335–358. Cham: Springer International Publishing, 2022.
- [BIL2022] Bilik, I. Comparative Analysis of Radar and Lidar Technologies for Automotive Applications. *IEEE Intell. Transp. Syst. Mag.* 2022, 15, 244–269.
- [BLI2022] Blickfeld "Cube 1 Outdoor v1.1", datasheet, 2022. Available online: https://www.blickfeld.com/wp-content/uploads/2022/10/549_blickfeld_Datasheet_Cube1-Outdoor_v1.1.pdf (accessed on 10 Jan. 2023).
- [BLI2023] Blickfeld Gmbh. Crowd Analytics Privacy-Sensitive People Counting and Crowd Analytics. Available online: <https://www.blickfeld.com/applications/crowd-analytics/> (accessed on 5 March 2023).
- [BNE+2018] G. Bagschik ; M. Nolte ; S. Ernst ; M. Maurer: A System's Perspective Towards an Architecture Framework for Safe Automated Vehicles. In: *2018 IEEE International Conference on Intelligent Transportation Systems (ITSC)*. Maui, HI, USA, 2018. — citation key: bagschik2018, pp.2438–45
- [BNZ2022] Bogdoll, Daniel; Nitsche, Maximilian; Zöllner, J. Marius (2022): Anomaly Detection in Autonomous Driving: A Survey. In : 2022 IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (CVPRW), pp. 4487–4498.
- [BT2016] S. Behere ; M. Törngren: A functional reference architecture for autonomous driving. In: *Information and Software Technology* vol. 73 (2016), pp.136–50
- [DC2022] R. Dona and B. Ciuffo, "Virtual Testing of Automated Driving Systems. A Survey on Validation Methods," *IEEE Access*, vol. 10, pp. 24349–24367, 2022, doi: 10.1109/ACCESS.2022.3153722.

- [DCL2011] B. Delahaye, B. Caillaud, and A. Legay, "Probabilistic contracts: a compositional reasoning methodology for the design of systems with stochastic and/or non-deterministic aspects," *Form Methods Syst Des*, vol. 38, no. 1, pp. 1–32, Feb. 2011, doi: 10.1007/s10703-010-0107-8.
- [DFF2023] T. Düser, D. Fischer, and J. Freyer, "A Study on the Paradigm Shift in the Validation of Automated Vehicles," presented at the IEEE International Automated Vehicle Validation Conference 2023, Austin, USA, Oct. 2023.
- [DFH+2019a] Damm, Werner; Fränzle, Martin; Hagemann, Willem; Kröger, Paul; Rakow, Astrid (2019): Dynamic Conflict Resolution Using Justification Based Reasoning. In *Electron. Proc. Theor. Comput. Sci.* 308, pp. 47–65. DOI: 10.4204/EPTCS.308.4.
- [DFH+2019b] Damm, Werner; Fränzle, Martin; Hagemann, Willem; Rakow, Astrid, and Swaminathan, Mani (2019b): Assuring Confidence in the Perception Chain of Highly Automated Vehicles, technical report, University of Oldenburg, 2019
- [DHS+2024] Werner Damm, David Hess, Mark Schweda, Janos Sztipanovits, Klaus Bengler, Bianca Biebl, Martin Fränzle, Willem Hagemann, Moritz Held, Klas Ihme, Severin Kacianka, Alyssa J. Kerscher, Sebastian Lehnhoff, Andreas Luedtke, Alexander Pretschner, Astrid Rakow, Jochem Rieger, Daniel Sonntag, Maike Schwammberger, Benedikt Austel, Anirudh Unni, Eric Veith, *A Reference Architecture of Human Cyber-Physical Systems – Part I: Fundamental Concepts*, [ACM Transactions on Cyber-Physical Systems Volume 8 Issue 1](https://doi.org/10.1145/3622879) Article No.: 2pp 1–32, 2024, <https://doi.org/10.1145/3622879>
- [DS2019] C. van Driesten and T. Schaller, "Overall approach to standardized sensor interfaces: Simulation and real vehicle," in *Fahrerassistenzsysteme 2018*, T. Bertram, Ed., Wiesbaden: Springer Fachmedien Wiesbaden, 2019, pp. 47–55, isbn: 978-3-658-23751-6.
- [ED2022] A. Elmquist and D. Negrut, "Methods and Models for Simulating Autonomous Vehicle Sensors," in *IEEE Transactions on Intelligent Vehicles*, vol. 5, no. 4, pp. 684-692, Dec. 2020, doi: 10.1109/TIV.2020.3003524
- [EHR2023] L. Elster, M. F. Holder and M. Rapp, "A Dataset for Radar Scattering Characteristics of Vehicles Under Real-World Driving Conditions: Major Findings for Sensor Simulation," in *IEEE Sensors Journal*, vol. 23, no. 5, pp. 4873-4882, 1 March 2023, doi: 10.1109/JSEN.2023.3238015
- [ENS2019] "European Initiative to Enable Validation for Highly Automated Safe and Secure Systems", ENABLE-S3, Accessed Jan. 14, 2024. [Online]. Available: <https://cordis.europa.eu/project/id/692455>

- [ESN2022a] A. Elmquist, R. Serban, and D. Negrut, "Camera simulation for robot simulation: how important are various camera model components?," <https://arxiv.org/abs/2211.08599>, 2022
- [ESN2022b] A. Elmquist, R. Serban, and D. Negrut, "A performance contextualization approach to validating camera models for robot simulation," arXivpreprint arXiv:2208.01022, 2022
- [ESP2023] Elster, Lukas; Staab, Jan Philipp; Peters, Steven (2023): Making Automotive Radar Sensor Validation Measurements Comparable. DOI: 10.20944/preprints202308.2045.v1.
- [ERH+2023] L. Elster, P. Rosenberger, M. F. Holder, K. Mori, J. Staab, S. Peters, "Introducing the Double Validation Metric for Radar Sensor Models," Preprint, 2023
- [FBK+2008] Fiorino, Steven & Bartell, Richard & Krizo, Mathew & Caylor, Gregory & Moore, Kenneth & Harris, Thomas & Cusumano, Salvatore. (2008). A First Principles Atmospheric Propagation & Characterization Tool—the Laser Environmental Effects Definition and Reference (LEEDR). Proceedings of SPIE - The International Society for Optical Engineering. 6878. 10.1117/12.763812.
- [FD2023] J. Freyer and T. Düser, "A Study on the transformation of virtual validation methods in the development of new mobility solutions," presented at the 9th International Symposium on Transportation Data & Modelling, Ispra, Italy, Jun. 2023.
- [FFK+2022] Bernd Finkbeiner, Martin Fränzle, Florian Kohn, Paul Kröger: A Truly Robust Signal Temporal Logic: Monitoring Safety Properties of Interacting Cyber-Physical Systems under Uncertain Observation. Algorithms 15(4): 126 (2022)
- [FHD+2023] Fränzle, Martin; Hagemann, Willem; Damm, Werner; Rakow, Astrid; Swaminathan, Mani (2023): Safer Than Perception: Assuring Confidence in Safety-Critical Decisions of Automated Vehicles. In Anne E. Haxthausen, Wen-ling Huang, Markus Roggenbach (Eds.): Applicable Formal Methods for Safe Industrial Products: Essays Dedicated to Jan Peleska on the Occasion of His 65th Birthday. Cham: Springer Nature Switzerland, pp. 180–201.
- [FHW+2022] Feng, D., Harakeh, A. Waslander, S., Dietmayer, K. A Review and Comparative Study on Probabilistic Object Detection in Autonomous Driving, IEEE Intelligent Vehicle Symposium 2022
- [FJG+2020] Jamil Fayyad, Mohammad A. Jaradat, Dominique Gruyer, and Homayoun Najjaran, Deep Learning Sensor Fusion for Autonomous Vehicle Perception and Localization: A Review, Sensors 2020, 20, 4220; doi:10.3390/s20154220
- [FSB+2022] Fink, M.; Schardt, M.; Baier, V.; Wang, K.; Jakobi, M.; Koch, A.W. Full-Waveform Modeling for Time-of-Flight Measurements based on Arrival Time of Photons. arXiv 2022, arXiv:2208.03426.

- [GAL2015] Gálvez del Postigo Fernández, C. Grid-Based Multi-Sensor Fusion for On-Road Obstacle Detection: Application to Autonomous Driving. Thesis, KTH Royal Institute of Technology School of Computer Science and Communication, Stockholm, Sweden, 2015.
- [GLU2012] Geiger, A.; Lenz, P.; Urtasun, R. Are we ready for autonomous driving? The KITTI vision benchmark suite. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Providence, RI, USA, 16–21 June 2012; pp. 3354–3361.
- [GMS+2021] Genser, S.; Muckenhuber, S.; Solmaz, S.; Reckenzaun, J. Development and Experimental Validation of an Intelligent Camera Model for Automated Driving. *Sensors* 2021, 21, 7583. <https://doi.org/10.3390/s21227583>
- [GSB+2020] R. Graubohm, T. Stolte, G. Bagschik, and M. Maurer, “Towards Efficient Hazard Identification in the Concept Phase of Driverless Vehicle Development,” in 2020 IEEE Intelligent Vehicles Symposium (IV), Las Vegas, NV, USA: IEEE, Oct. 2020, pp. 1297–1304. doi: 10.1109/IV47402.2020.9304780.
- [HB2014] Holmstrom, S.T.S.; Baran, U.; Urey, H. MEMS Laser Scanners: A Review. *J. Microelectromech. Syst.* 2014, 23, 259–275.
- [HCP+2023] Haider, A.; Cho, Y.; Pigniczki, M.; Köhler, M.H.; Haas, L.; Kastner, L.; Fink, M.; Schardt, M.; Cichy, Y.; Koyama, S.; et al. Performance Evaluation of MEMS-Based Automotive LiDAR Sensor and Its Simulation Model as per ASTM E3125-17 Standard. *Sensors* **2023**, 23, 3113. <https://doi.org/10.3390/s23063113>
- [HFS+2019] Höfer, M.; Fuhr, F.; Schick, B.; Pfeffer, P.E. Attribute-based development of driver assistance systems. In: Pfeffer, P. (eds) 10th International Munich Chassis Symposium 2019
- [HHR+2015] Hirsenkorn, N.; Hanke, T.; Rauch, A.; Dehlink, B.; Rasshofer, R.; Biebl, E. A non-parametric approach for modeling sensor behavior. In Proceedings of the 16th International Radar Symposium (IRS), Dresden, Germany, 24–26 June 2015; pp. 131–136.
- [HMG+2023] M. Henning ; J. Muller ; F. Gies ; M. Buchholz ; K. Dietmayer: Situation-Aware Environment Perception Using a Multi-Layer Attention Map. In: *IEEE Transactions on Intelligent Vehicles* vol. 8 (2023), Nr. 1, pp.481–91
- [HMS2019] D. Harel ; A. Marron ; J. Sifakis: Autonomics: In Search of a Foundation for Next Generation Autonomous Systems. In: *arXiv:1911.07133 [cs]* (2019). — arXiv: 1911.07133
- [HOL2023] M. F. Holder, “Synthetic Generation of Radar Sensor Data for Virtual Validation of Autonomous Driving”, 2023, Darmstadt, Technische Universität Darmstadt, DOI: 10.26083/tuprints-00017545

- [HOP+2022] Hawkins, Richard David, Matthew Osborne, Michael Stephen Parsons, Mark Nicholson, John Alexander McDermid, and Ibrahim Habli. "Guidance on the Safety Assurance of Autonomous Systems in Complex Environments (SACE)." (2022)
- [HPK+2023] Haider, A.; Pigniczki, M.; Koyama, S.; Köhler, M.H.; Haas, L.; Fink, M.; Schardt, M.; Nagase, K.; Zeh, T.; Eryildirim, A.; et al. A Methodology to Model the Rain and Fog Effect on the Performance of Automotive LiDAR Sensors. *Sensors* **2023**, *23*, 6891. <https://doi.org/10.3390/s23156891>
- [HPP+2021] Hawkins, Richard, Colin Paterson, Chiara Picardi, Yan Jia, Radu Calinescu, and Ibrahim Habli. "Guidance on the Assurance of Machine Learning in Autonomous Systems (AMLAS)."
- [HS2022] H.-M. Heinkel and K. Steinkirchner, "Credible Simulation Process -- With Example," SET Level, Aug. 2022. Accessed: Dec. 11, 2023. [Online]. Available: https://gitlab.setlevel.de/open/processes_and_traceability/credible_simulation_process_framework/-/raw/main/Credible-Simulation-Process-v1-3.pdf.
- [IAM2021] International Alliance for Mobility Testing Standardization (IAMTS): Best Practice for A Comprehensive Approach for the Validation of Virtual Testing Toolchains IAMTS0001202104, 2021
- [IPS2021] IPG CarMaker. Reference Manual Version 9.0.1; IPG Automotive GmbH: Karlsruhe, Germany, 2021.
- [ISO26262] "Road vehicles -- Functional safety." International Organization for Standardization, Geneva, Switzerland, 2018.
- [ISOTC22] ISO TC 22 SC33: ISO NWIP 11010-2 Passenger Cars - Simulation model classification - Part 2: Perception sensor models for ADAS /AD
- [KAK+2019] A. Kampmann ; B. Alrifaae ; M. Kohout ; A. Wüstenberg ; T. Woopen ; M. Nolte ; L. Eckstein ; S. Kowalewski: A Dynamic Service-Oriented Software Architecture for Highly Automated Vehicles. In: , 2019, pp.2101–8
- [KD2009] Kiureghian, Armen Der; Ditlevsen, Ove (2009): Aleatory or epistemic? Does it matter? In *Structural Safety* 31 (2), pp. 105–112. DOI: 10.1016/j.strusafe.2008.06.020.
- [KG2016] D. Kirchner ; K. Geihs: Adaptive Model-Based Monitoring for Robots. In: *Intelligent Autonomous Systems* : Springer, Cham, 2016, pp.43–56
- [KHM+2019] Keidler, S.; Haselberger, J.; Mayannavar, K.; Schick, B.; Schneider, D. Development of lane-precise "Ground Truth" maps for the objective Quality Assessment of automated driving functions, In: *VDI Reifen – Fahrwerk – Fahrbahn*, 2019
- [LPR+2016] LEWIS, P. R. ; PLATZNER, M. ; RINNER, B. ; TØRRESEN, J. ; YAO, X. (eds.): *Self-Aware Computing Systems: An Engineering Approach, Natural*

Computing Series: Springer International Publishing, 2016
— ISBN 978-3-319-39674-3

- [LRS+2021] Linnhoff, Clemens; Rosenberger, Philipp; Schmidt, Simon; Elster, Lukas; Stark, Rainer; Winner, Hermann (2021): Towards Serious Perception Sensor Simulation for Safety Validation of Automated Driving - A Collaborative Method to Specify Sensor Models. In : 2021 IEEE International Intelligent Transportation Systems Conference (ITSC). 2021 IEEE International Intelligent Transportation Systems Conference (ITSC). Indianapolis, IN, USA, 19.09.2021 - 22.09.2021: IEEE, pp. 2688–2695.
- [LVC+2017] Lang, A.H.; Vora, S.; Caesar, H.; Zhou, L.; Yang, J.; Beijbom, O. Pointpillars: Fast encoders for object detection from point clouds. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, Honolulu, HI, USA, 21–26 July 2017; pp. 12697–12705.
- [LYW2005] F. Liu, M. Yang, and Z. Wang, "Study on Simulation Credibility Metrics," in Proceedings of the Winter Simulation Conference, 2005, pp. 2554–2560
- [MAU2000] M. Maurer: EMS-vision: knowledge representation for flexible automation of land vehicles. In: *2000 IEEE Intelligent Vehicles Symposium*. Dearborn, MI, USA, 2000, pp.575–80
- [MBF+1996] M. Maurer ; R. Behringer ; S. Furst ; F. Thomanek ; E.D. Dickmanns: A compact vision system for road vehicle guidance. In: *Proceedings of 13th International Conference on Pattern Recognition*. vol. 3, 1996, pp.313–7 vol.3
- [MDM2010] Munz, M. Dietmayer, K., Mählich, M. Generalized Fusion of Heterogeneous Sensor Measurements for Multi Target Tracking, DOI: 10.1109/ICIF.2010.5711926 · Source: IEEE Xplore, August 2010
- [MLR+2022] Z.F. Magosi, H. Li, P. Rosenberger, L. Wan and A. Eichberger, "A Survey on Modelling of Automotive Radar Sensors for Virtual Test and Validation of Automated Driving", *Sensors* 2022, 22, 5693. <https://doi.org/10.3390/s22155693>
- [MLW2000] Muessig, Paul R., Dennis R. Laack, and J. L. Wroblewski. "An integrated approach to evaluating simulation credibility." *SUMMER COMPUTER SIMULATION CONFERENCE*. Society for Computer Simulation International; 1998, 2000.
- [MSP2023] Mori, Ken; Storms, Kai; Peters, Steven (2023): Conservative Estimation of Perception Relevance of Dynamic Objects for Safe Trajectories in Automotive Scenarios. In : 2023 IEEE International Conference on Mobility, Operations, Services and Technologies (MOST), pp. 83–95.
- [NAS2013] National Aeronautics and Space Administration (NASA), "NASA-STD-7009A," 2013.

- [NAS2016] "STANDARD FOR MODELS AND SIMULATIONS.", National Aeronautics and Space Administration, 2016.
- [NGO2023] A. Ngo, "A methodology for validation of a radar simulation for virtual testing of autonomous driving", 2023, Stuttgart, Germany
- [MWT+2022] Z.F. Magosi, C. Wellershaus, V.R. Tihanyi, P. Luley and A. Eichberger, "Evaluation Methodology for Physical Radar Perception Sensor Models Based on On-Road Measurements for the Testing and Validation of Automated Driving", *Energies* 2022, 15, 2545. <https://doi.org/10.3390/en15072545>
- [NKL+2021] Neuwirthová, E.; Kuusk, A.; Lhotáková, Z.; Kuusk, J.; Albrechtová, J.; Hallik, L. Leaf Age Matters in Remote Sensing: Taking Ground Truth for Spectroscopic Studies in Hemiboreal Deciduous Trees with Continuous Leaf Formation. *Remote Sens.* 2021, 13, 1353.
- [NKM+2023] Neurohr, Birte; Koopmann, Tjark; Möhlmann, Eike; Fränzle, Martin (2023): Determining the Validity of Simulation Models for the Verification of Automated Driving Systems. In *IEEE Access* 11, pp. 102949–102960. DOI: 10.1109/ACCESS.2023.3316354.
- [NSB2021] C. Neimröck, W. Schlecht, M. Berlin, H. Ehrich, and J.-E. Stavesand, "C3 Whitepaper: Digital Loop -- Data-Driven Development of Driving Functions," Jul. 2021. Accessed: Jan. 14, 2024. [Online]. Available: https://www.dspace.com/shared/data/pdf/2022/dSPACE_Digital_Loop_Whitepaper_EN_07-2021.pdf
- [PEG2024] "PEGASUS Method." Accessed Jan. 04, 2024. [Online]. Available: <https://www.pegasusprojekt.de/en/home>
- [PET2022a] Petit, F. Myths about LiDAR Sensor Debunked. Available online: <https://www.blickfeld.com/de/blog/mit-den-lidar-mythenaufgeraemt-teil-1/> (accessed on 5 July 2022).
- [PET2022b] P. Petersen et al., "Towards a Data Engineering Process in Data-Driven Systems Engineering," in 2022 IEEE International Symposium on Systems Engineering (ISSE), Vienna, Austria: IEEE, Oct. 2022, pp. 1–8. doi: 10.1109/ISSE54508.2022.10005441.
- [PPF2023] Peters, S., Peters, J. & Findeisen, R. Quantifying Uncertainties along the Automated Driving Stack. *ATZ Worldw* 125, 62–65 (2023). <https://doi.org/10.1007/s38311-023-1489-8>
- [PT2020] Ana Pereira and Carsten Thomas, Challenges of Machine Learning Applied to Safety-Critical Cyber-Physical Systems, *Mach. Learn. Knowl. Extr.* 2020, 2, 579–602; doi:10.3390/make2040031
- [RAK2023] Rakow, Astrid (2023): Framing Relevance for Safety-Critical Autonomous Systems. Available online at <https://arxiv.org/pdf/2307.14355.pdf>.
- [RB2012] Roy, C. J.; Balch, M. (2012): A HOLISTIC APPROACH TO UNCERTAINTY QUANTIFICATION WITH APPLICATION TO SUPERSONIC NOZZLE

- THRUST, in: International Journal for Uncertainty Quantification, Vol. 2, pp. 363–381, 2012
- [RB2019] Royo, S.; Ballesta-Garcia, M. An Overview of Lidar Imaging Systems for Autonomous Vehicles. *Appl. Sci.* **2019**, *9*, 4093. <https://doi.org/10.3390/app9194093>
- [RCG2021] Roriz, R.; Cabral, J.; Gomes, T. Automotive LiDAR technology: A survey. *IEEE Trans. Intell. Transp. Syst.* **2021**, *23*, 6282–6297.
- [RGL+2022] Reich, Jan, Pascal Gerber, Nishanth Laxman, Daniel Schneider, Takehito Ogata, Satoshi Otsuka, and Tasuku Ishigooka. “Engineering Dynamic Risk and Capability Models to Improve Cooperation Efficiency Between Human Workers and Autonomous Mobile Robots in Shared Spaces.” In *Model-Based Safety and Assessment*, edited by Christel Seguin, Marc Zeller, and Tatiana Prosvirnova, 13525:237–51. Lecture Notes in Computer Science. Cham: Springer International Publishing, 2022. https://doi.org/10.1007/978-3-031-15842-1_17.
- [RIG2022] P. Rigoll, “Augmentation von Kameradaten mit Generative Adversarial Networks (GANs) zur Absicherung automatisierter Fahrfunktionen,” 35. VDI-Tagung: Fahrerassistenzsysteme und automatisiertes Fahren, May 2022.
- [RLS2023] P. Rigoll, J. Langner, and E. Sax, “Unveiling Objects with SOLA: An Annotation-Free Image Search on the Object Level for Automotive Data Sets,” 2023, doi: 10.48550/ARXIV.2312.01860.
- [ROS2022] Rosenberger, Philipp (2022): Metrics for Specification, Validation, and Uncertainty Prediction for Credibility in Simulation of Active Perception Sensor Systems, PhD Thesis, Technische Universität Darmstadt, 2022
- [RPS+2023] P. Rigoll, P. Petersen, H. Stage, L. Ries, and E. Sax, “Focus on the Challenges: Analysis of a User-friendly Data Search Approach with CLIP in the Automotive Domain.” arXiv, Apr. 21, 2023. Accessed: Jan. 07, 2024. [Online]. Available: <http://arxiv.org/abs/2304.10247>
- [RSS+2020] J. Reich ; D. Schneider ; I. Sorokos ; Y. Papadopoulos ; T. Kelly ; R. Wei ; E. Armengaud ; C. Kaypmaz: Engineering of Runtime Safety Monitors for Cyber-Physical Systems with Digital Dependability Identities. In: CASIMIRO, A. ; ORTMEIER, F. ; BITSCH, F. ; FERREIRA, P. (eds.): *Computer Safety, Reliability, and Security, Lecture Notes in Computer Science*. vol. 12234. Cham : Springer International Publishing, 2020 — ISBN 978-3-030-54548-2, pp.3–17
- [RSS2022] P. Rigoll, L. Ries, and E. Sax, “Scalable Data Set Distillation for the Development of Automated Driving Functions,” in 2022 IEEE 25th International Conference on Intelligent Transportation Systems (ITSC), Macau, China: IEEE, Oct. 2022, pp. 3139–3145. doi: 10.1109/ITSC55140.2022.9921868.
- [RWR+2024] Reckenzaun Jakob; Wiegand, Christopher; Rott, Relindis; Solmaz, Selim; Simkin, Barnaby; Düser, Tobias; Gutenkunst, Christian (2024):

A classification scheme for sensors models with related validation measures and application examples for automated driving systems. IAMTS (to appear)

- [SAE2018] J3016 201806: Taxonomy and Definitions for Terms Related to Driving Automation Systems for On-Road Motor Vehicles - SAE International, https://www.sae.org/standards/content/j3016_201806/.
- [SAE2021] SAE J3016? "Taxonomy and Definitions for Terms Related to Driving Automation Systems for On-Road Motor Vehicles." SAE International, Warrendale, PA, USA, Apr. 2021.
- [SAE2022] SAE: J3131:Definitions for Terms Related to Automated Driving Systems Reference Architecture (Standard), 2022
- [SCH+2021] M. Scholtes et al., "6-Layer Model for a Structured Description and Categorization of Urban Traffic and Environment," IEEE Access, vol. 9, pp. 59131–59147, 2021, doi: 10.1109/ACCESS.2021.3072739.
- [SCK+2021] Salay, Rick; Czarnecki, Krzysztof; Kuwajima, Hiroshi; Yasuoka, Hirotohi; Nakae, Toshihiro; Abdelzad, Vahdat et al. (2021): The missing link: Developing a safety case for perception components in automated driving. Available online at <http://arxiv.org/pdf/2108.13294v4>.
- [SES2023] SESAME Project Consortium: Deliverable D7.1 Runtime Safety and Security Concept – EDDI Runtime Model Specification. H2020 Secure and Safe Multi-Robot Systems (SESAME) Project (Grant 101017258). URL: <https://www.sesame-project.org/>
- [SHL+2018] Schneider, D.; Huber, B.; Lategahn, H.; Schick, B. Measuring method for function and quality of automated lateral control based on high-precision digital "Ground Truth" maps, In: VDI/VW-Gemeinschaftstagung Fahrerassistenzsysteme und Automatisiertes Fahren, 2018
- [SLB2017] Schaefer, A.; Luft L.; Burgard, W. An Analytical Lidar Sensor Model Based on Ray Path Information. IEEE Robot. Autom. Lett. 2017, 2, 1405–1412.
- [SME+2017a] J. Schlatow ; M. Mostl ; R. Ernst ; M. Nolte ; I. Jatzkowski ; M. Maurer: Towards model-based integration of component-based automotive software systems. In: *IECON 2017 – 43rd Annual Conference of the IEEE Industrial Electronics Society*. Beijing, China, 2017 — ISBN 978-1-5386-1127-2, pp.8425–32
- [SME+2017b] J. Schlatow ; M. Möstl ; R. Ernst ; M. Nolte ; I. Jatzkowski ; M. Maurer ; C. Herber ; A. Herkersdorf: Self-Awareness in Autonomous Automotive Systems. In: *2017 Design, Automation Test in Europe Conference Exhibition (DATE)*, 2017, pp.1050–5
- [SMH+2022] Strohbeck, Jan; Muller, Johannes; Herrmann, Martin; Buchholz, Michael (2022): Deep Kernel Learning for Uncertainty Estimation in Multiple Trajectory Prediction Networks. In : 2022 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS).

- 2022 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS). Kyoto, Japan, 23.10.2022 - 27.10.2022: IEEE, pp. 11396–11402.
- [SMP2023] Storms, Kai; Mori, Ken; Peters, Steven (2023): SURE-Val: Safe Urban Relevance Extension and Validation. Available online at <https://arxiv.org/pdf/2308.02266.pdf>.
- [SN2018] Stolz, M.; Nestlinger, G. Fast generic sensor models for testing highly automated vehicles in simulation. *Elektrotech. Inf.* 2018, 135, 365–369.
- [SPR2000] Sproston, J.: Decidable model checking of probabilistic hybrid automata. In: *Formal Techniques in Real-Time and Fault-Tolerant Systems*. pp. 31-45. Springer (2000)
- [SRL+2022] H. Stage, L. Ries, J. Langner, S. Otten, and E. Sax, "Analysis and Comparison of Datasets by Leveraging Data Distributions in Latent Spaces," in *Deep Neural Networks and Data for Automated Driving*, T. Fingscheidt, H. Gottschalk, and S. Houben, Eds., Cham: Springer International Publishing, 2022, pp. 107–126. doi: 10.1007/978-3-031-01233-4_3.
- [ST2013] Schneider, Daniel, and Mario Trapp. "Conditional Safety Certification of Open Adaptive Systems." *ACM Transactions on Autonomous and Adaptive Systems* 8, no. 2 (July 2013): 1–20. <https://doi.org/10.1145/2491465.2491467>.
- [STE2016] Jan Erik Stellet, Statistical modelling of algorithms for signal processing in systems based on environment perception, PhD Thesis, Institute of Industrial Information Technology (IIIT), Karlsruhe Institute of Technology (KIT), 2016, <https://publikationen.bibliothek.kit.edu/1000056749>
- [SWH2023] Stierand, Ingo; Westhofen, Lukas; Hagemann, Willem (2023): On Using Ontologies in the Engineering of Intelligent Cyber-Physical Systems.
- [THA2016] Thakur, R. Scanning LIDAR in Advanced Driver Assistance Systems and Beyond: Building a road map for next-generation LIDAR technology. *IEEE Consum. Electron. Mag.* 2016, 5, 48–54.
- [THS+2017] Ö.S. Tas ; S. Hörmann ; B. Schäufole ; F. Kuhnt: Automated Vehicle System Architecture with Performance Assessment. In: *2017 IEEE 20th International Conference on Intelligent Transportation Systems (ITSC)*, 2017. — Citation Key: tas_functionalperformance_2017
- [TPB+2008] S. Tripakis, C. Pinello, A. Benveniste, A. Sangiovanni-Vincent, P. Caspi and M. Di Natale, "Implementing Synchronous Models on Loosely Time Triggered Architectures," in *IEEE Transactions on Computers*, vol. 57, no. 10, pp. 1300-1314, Oct. 2008, doi: 10.1109/TC.2008.81
- [TSW2018] M. Trapp ; D. Schneider ; G. Weiss: Towards Safety-Awareness and Dynamic Safety Management. In: *2018 14th European Dependable*

- Computing Conference (EDCC)*. Iasi, Romania : IEEE, 2018 — ISBN 978-1-5386-8060-5, pp.107–11
- [TZM+2018] M. Törngren ; X. Zhang ; N. Mohan ; M. Becker ; L. Svensson ; X. Tao ; D.-J. Chen ; J. Westman: Architecting Safety Supervisors for High Levels of Automated Driving. In: *2018 21st International Conference on Intelligent Transportation Systems (ITSC)*. Maui, HI : IEEE, 2018 — ISBN 978-1-72810-321-1, pp.1721–8
- [UN2021] United Nations, Uniform provisions concerning the approval of vehicles with regard to Automated Lane Keeping Systems, addendum 156, UN Regulation No. 157 (2021)
- [UNA2024] US National Academy of Sciences: Machine Learning for Safety-Critical Applications: opportunities, challenges, and a research agenda, National Academic Press, Washington DC, 2024,
- [UNE2022] UNECE, Ed., “New Assessment/Test Method for Automated Driving (NATM) Guidelines for Validating Automated Driving System (ADS) – amendments to ECE/TRANS/WP.29/2022/58.” Jun. 2022. Accessed: Dec. 11, 2023. [Online]. Available: <https://unece.org/sites/default/files/2022-05/WP.29-187-08e.pdf>
- [URR+2017] S. Ulbrich ; A. Reschka ; J. Rieken ; S. Ernst ; G. Bagschik ; F. Dierkes ; M. Nolte ; M. Maurer: Towards a Functional System Architecture for Automated Vehicles. In: *arXiv:1703.08557 [cs]* (2017). — arXiv: 1703.08557citation-key: ulbrich2017a
- [VVM2023] “Verification Validation Methods.” Accessed: Dec. 11, 2023. [Online]. Available: <https://www.vvm-projekt.de/>
- [WMG+2024] C. Wiecher et al., “Model-based Analysis and Specification of Functional Requirements and Tests for Complex Automotive Systems.” arXiv, Nov. 15, 2023. Accessed: Jan. 14, 2024. [Online]. Available: <http://arxiv.org/abs/2209.01473>
- [WNB2022] Westhofen, Lukas; Neurohr, Christian; Butz, Martin; Scholtes, Maïke; Schuldes, Michael (2022): Using Ontologies for the Formalization and Recognition of Criticality for Automated Driving. In *IEEE Open J. Intell. Transp. Syst.* 3, pp. 519–538. DOI: 10.1109/OJITS.2022.3187247.
- [WNK2023] Westhofen, Lukas; Neurohr, Christian; Koopmann, Tjark; Butz, Martin; Schütt, Barbara; Utesch, Fabian et al. (2023): Criticality Metrics for Automated Driving: A Review and Suitability Analysis of the State of the Art. In *Arch Computat Methods Eng* 30 (1), pp. 1–35. DOI: 10.1007/s11831-022-09788-7.
- [WW2016] W. Wachenfeld and H. Winner, “The Release of Autonomous Vehicles,” in *Autonomous Driving*, Springer Berlin Heidelberg, 2016, pp. 425–449. doi: 10.1007/978-3-662-48847-8_21.

Authors

Name	Organisation
Dr. Martin Benedikt	Virtuelles Fahrzeug ViF
Eckard Böde	DLR e.V.
Andreas Bossert	ITK Engineering
Prof. Dr. Jens Braband	Siemens Mobility GmbH (Steuerkreis)
Tino Brade	Robert Bosch GmbH
Niklas Braun	TU Braunschweig
Tobias Braun	Fraunhofer IESE
Dr. Simon Burton	Fraunhofer IKS
Prof. Dr. Thomas Dallmann	TU Illmenau
Prof. Dr. Werner Damm	SafeTRANS (Steuerkreis)
Prof. Dr. Tobias Düser	KIT, IPEK
Lukas Elster	TU Darmstadt
Prof. Dr.-Ing. Tim Fingscheidt	TU Braunschweig
Marco Fistler	IAV
Marzena Franek	Robert Bosch GmbH
Prof. Dr. Martin Fränzle	Universität Oldenburg
Jonas Freyer	KIT, IPEK
Roland Galbas	Robert Bosch GmbH (Steuerkreis)
Roman Gansch	Robert Bosch GmbH
Dirk Geyer	AVL Software and Functions GmbH
Lukas Haas	Hochschule Kempten
Arsalan Haider	Hochschule Kempten
Peter Heidl	ehemals Robert Bosch GmbH (Steuerkreis)
Prof. Dr. Matthias Hein	TU Illmenau (Steuerkreis)
Dr. Andreas Heyl	Robert Bosch GmbH
Johannes Hiller	RWTH Aachen
Dr. Hardi Hungar	DLR e.V.
Prof. Dr. Dieter Hutter	DFKI
Prof. Dr. Rolf Jung	Hochschule Kempten
Cornel Klein	Siemens AG (Steuerkreis)
Jörg Krüger	IngenX Technologies GmbH

Dr. Thomas Kuhn	Fraunhofer IESE
Jacob Langner	FZI
Prof. Dr. Markus Maurer	TU Braunschweig
Kerstin Mayr	AVL (Steuerkreis)
Dr. André Meyer-Vitali	DFKI
Dr. Eike Möhlmann	DLR e.V.
Dr. Adam Molin	Denso
Björn Möller	TU Braunschweig
Jürgen Niehaus	SafeTRANS
Bastian Nolte	TU Braunschweig
Marcus Nolte	TU Braunschweig
Dr. Stefan Otten	FZI
Prof. Dr. Jan Peleska	Universität Bremen
Prof. Dr. Steven Peters	TU Darmstadt
Prof. Dr. Tim Poguntke	Hochschule Kempten
Florian Poprawa	Siemens Mobility GmbH
Jan Reich	Fraunhofer IESE
Dr. Philipp Rosenberger	Persival GmbH
Nayel Fabian Salem	TU Braunschweig
Prof. Bernhard Schick	Hochschule Kempten
Dr. Daniel Schneider	Fraunhofer IESE
Prof. Dr. Stefan-Alexander Schneider	Hochschule Kempten
Dr. Christian Schyr	AVL Deutschland GmbH
Prof. Dr. Carsten Thomas	HTW Berlin
Prof. Dr. Mario Trapp	Fraunhofer IKS
Florence Wagner	Hochschule Kempten
Nicolas Wagener	RWTH Aachen
Timo Woopen	RWTH Aachen
Prof. Dr. Thomas Zeh	Hochschule Kempten

Annex 1: References in Tables 1, 2, and 3 of Section 5.2 from [FJG+2020]

- [32] Gruyer, D.; Belaroussi, R.; Revilloud, M. Accurate lateral positioning from map data and road marking detection. *Expert Syst. Appl.* 2016, 43, 1–8.
- [34] Schlosser, J.; Chow, C.K.; Kira, Z. Fusing LIDAR and images for pedestrian detection using convolutional neural networks. In *Proceedings of the 2016 IEEE International Conference on Robotics and Automation (ICRA)*, Stockholm, Sweden, 16–21 May 2016; IEEE: Stockholm, Sweden, 2016; pp. 2198–2205.
- [35] Melotti, G.; Premebida, C.; Gonçalves, N.M.D.S.; Nunes, U.J.; Faria, D.R. Multimodal CNN Pedestrian Classification: A Study on Combining LIDAR and Camera Data. In *Proceedings of the 2018 21st International Conference on Intelligent Transportation Systems (ITSC)*, Maui, HI, USA, 4–7 November 2018; IEEE: Maui, HI, USA, 2018; pp. 3138–3143.
- [36] Labayrade, R.; Gruyer, D.; Royere, C.; Perrollaz, M.; Aubert, D. Obstacle Detection Based on Fusion between Stereovision and 2D Laser Scanner. In *Mobile Robots: Perception & Navigation*; Kolski, S., Ed.; Pro Literatur Verlag: Augsburg, Germany, 2007.
- [37] Liu, J.; Zhang, S.; Wang, S.; Metaxas, D. Multispectral Deep Neural Networks for Pedestrian Detection. *arXiv2016*, arXiv:1611.02644
- [38] Hou, Y.-L.; Song, Y.; Hao, X.; Shen, Y.; Qian, M.; Chen, H. Multispectral pedestrian detection based on deep convolutional neural networks. *Infrared Phys. Technol.* 2018, 94, 69–77.
- [39] Wagner, J.; Fischer, V.; Herman, M.; Behnke, S. Multispectral Pedestrian Detection using Deep Fusion Convolutional Neural Networks. In *Proceedings of the ESANN*, Bruges, Belgium, 27–29 April 2016.
- [40] Lee, Y.; Bui, T.D.; Shin, J. Pedestrian Detection based on Deep Fusion Network using Feature Correlation. In *Proceedings of the 2018 Asia-Pacific Signal and Information Processing Association Annual Summit and Conference (APSIPA ASC)*, Honolulu, HI, USA, 12–15 November 2018; IEEE: Honolulu, HI, USA, 2018; pp. 694–699
- [41] Zheng, Y.; Izzat, I.H.; Ziaee, S. GFD-SSD: Gated Fusion Double SSD for Multispectral Pedestrian Detection. *arXiv 2019*, arXiv:1903.06999
- [42] Shopovska, I.; Jovanov, L.; Philips, W. Deep Visible and Thermal Image Fusion for Enhanced Pedestrian Visibility. *Sensors* 2019, 19, 3727.
- [43] Gu, S.; Lu, T.; Zhang, Y.; Alvarez, J.M.; Yang, J.; Kong, H. 3-D LiDAR + Monocular Camera: An Inverse-Depth-Induced Fusion Framework for Urban Road Detection. *IEEE Trans. Intell. Veh.* 2018, 3, 351–360
- [44] Yang, F.; Yang, J.; Jin, Z.; Wang, H. A Fusion Model for Road Detection based on Deep Learning and Fully Connected CRF. In *Proceedings of the 2018 13th Annual Conference on System of Systems Engineering (SoSE)*, Paris, France, 19–22 June 2018; IEEE: Paris, France, 2018; pp. 29–36.
- [45] Lv, X.; Liu, Z.; Xin, J.; Zheng, N. A Novel Approach for Detecting Road Based on Two-Stream Fusion Fully Convolutional Network. In *Proceedings of the 2018 IEEE Intelligent Vehicles Symposium (IV)*, Changshu, China, 26–30 June 2018; IEEE: Changshu, China, 2018; pp. 1464–1469.

- [46] Caltagirone, L.; Bellone, M.; Svensson, L.; Wahde, M. LIDAR-Camera Fusion for Road Detection Using Fully Convolutional Neural Networks. *Robot. Auton. Syst.* 2019, 111, 125–131.
- [47] Zhang, Y.; Morel, O.; Blanchon, M.; Seulin, R.; Rastgoo, M.; Sidibé, D. Exploration of Deep Learning-based Multimodal Fusion for Semantic Road Scene Segmentation. In *Proceedings of the 14th International Joint Conference on Computer Vision, Imaging and Computer Graphics Theory and Applications; SCITEPRESS—Science and Technology Publications: Prague, Czech Republic, 2019; pp. 336–343.*
- [48] Kato, T.; Ninomiya, Y.; Masaki, I. An obstacle detection method by fusion of radar and motion stereo. *IEEE Trans. Intell. Transp. Syst.* 2002, 3, 182–188.
- [49] Bertozzi, M.; Bombini, L.; Cerri, P.; Medici, P.; Antonello, P.C.; Miglietta, M. Obstacle detection and classification fusing radar and vision. In *Proceedings of the 2008 IEEE Intelligent Vehicles Symposium, Eindhoven, The Netherlands, 4–6 June 2008; pp. 608–613.*
- [50] Du, X.; Ang, M.H.; Rus, D. Car detection for autonomous vehicle: LIDAR and vision fusion approach through deep learning framework. In *Proceedings of the 2017 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS), Vancouver, BC, Canada, 24–28 September 2017; IEEE: Vancouver, BC, Canada, 2017; pp. 749–754.*
- [51] Valente, M.; Joly, C.; de La Fortelle, A. Deep Sensor Fusion for Real-Time Odometry Estimation. *arXiv* 2019, arXiv:1908.00524
- [52] Alatise, M.B.; Hancke, G.P. Pose Estimation of a Mobile Robot Based on Fusion of IMU Data and Vision Data Using an Extended Kalman Filter. *Sensors* 2017, 17, 2164.
- [53] Bresson, G.; Alsayed, Z.; Yu, L.; Glaser, S. Simultaneous Localization and Mapping: A Survey of Current Trends in Autonomous Driving. *IEEE Trans. Intell. Veh.* 2017, 2, 194–220
- [54] Jaradat, M.A.K.; Abdel-Hafez, M.F. Non-Linear Autoregressive Delay-Dependent INS/GPS Navigation System Using Neural Networks. *IEEE Sens. J.* 2017, 17, 1105–1115.
- [55] Rohani, M.; Gingras, D.; Gruyer, D. A Novel Approach for Improved Vehicular Positioning Using Cooperative Map Matching and Dynamic Base Station DGPS Concept. *IEEE Trans. Intell. Transp. Syst.* 2016, 17, 230–239.
- [64] Castanedo, F. A Review of Data Fusion Techniques. *Sci. World J.* 2013, 2013, 1–19.
- [65] Pires, I.; Garcia, N.; Pombo, N.; Flórez-Revuelta, F. From Data Acquisition to Data Fusion: A Comprehensive Review and a Roadmap for the Identification of Activities of Daily Living Using Mobile Devices. *Sensors* 2016, 16, 184.
- [67] Santoso, F.; Garratt, M.A.; Anavatti, S.G. Visual–Inertial Navigation Systems for Aerial Robotics: Sensor Fusion and Technology. *IEEE Trans. Autom. Sci. Eng.* 2017, 14, 260–275
- [68] Jaradat, M.A.K.; Abdel-Hafez, M.F. Enhanced, Delay Dependent, Intelligent Fusion for INS/GPS Navigation System. *IEEE Sens. J.* 2014, 14, 1545–1554
- [69] Alkhawaja, F.; Jaradat, M.; Romdhane, L. Techniques of Indoor Positioning Systems (IPS): A Survey. In *Proceedings of the 2019 Advances in Science and Engineering Technology International Conferences (ASET), Dubai, UAE, 26 March–10 April 2019; pp. 1–8.*

- [70] Luo, R.C.; Chang, C.-C. Multisensor Fusion and Integration: A Review on Approaches and Its Applications in Mechatronics. *IEEE Trans. Ind. Inform.* 2012, 8, 49–60
- [71] Khaleghi, B.; Khamis, A.; Karray, F.O.; Razavi, S.N. Multisensor data fusion: A review of the state-of-the-art. *Inf. Fusion* 2013, 14, 28–44.
- [72] Nagla, K.S.; Uddin, M.; Singh, D. Multisensor Data Fusion and Integration for Mobile Robots: A Review. *IAES Int. J. Robot. Autom. IJRA* 2014, 3, 131–138.
- [73] Vincke, B.; Lambert, A.; Gruyera, D.; Elouardi, A.; Seignez, E. Static and dynamic fusion for outdoor vehicle localization. In *Proceedings of the 2010 11th International Conference on Control Automation Robotics Vision, Singapore, 7–10 December 2010*; pp. 437–442.
- [74] Kueviakoe, K.; Wang, Z.; Lambert, A.; Frenoux, E.; Tarroux, P. Localization of a Vehicle: A Dynamic Interval Constraint Satisfaction Problem-Based Approach. Available online: <https://www.hindawi.com/journals/js/2018/3769058/>
- [75] Wang, Z.; Lambert, A. A Reliable and Low Cost Vehicle Localization Approach Using Interval Analysis. In *Proceedings of the 2018 IEEE 16th Intl Conf on Dependable, Autonomic and Secure Computing, 16th Intl Conf on Pervasive Intelligence and Computing, 4th Intl Conf on Big Data Intelligence and Computing and Cyber Science and Technology Congress(DASC/PiCom/DataCom/CyberSciTech), Athens, Greece, 12–15 August 2018*; pp. 480–487.
- [82] Matsugu, M.; Mori, K.; Mitari, Y.; Kaneda, Y. Subject independent facial expression recognition with robust face detection using a convolutional neural network. *Neural Netw.* 2003, 16, 555–559
- [83] Gao, H.; Cheng, B.; Wang, J.; Li, K.; Zhao, J.; Li, D. Object Classification Using CNN-Based Fusion of Vision and LIDAR in Autonomous Vehicle Environment. *IEEE Trans. Ind. Inform.* 2018, 14, 4224–4231.
- [84] Melotti, G.; Asvadi, A.; Premebida, C. CNN-LIDAR pedestrian classification: Combining range and reflectance data. In *Proceedings of the 2018 IEEE International Conference on Vehicular Electronics and Safety (ICVES), Madrid, Spain, 12–14 September 2018*; IEEE: Madrid, Spain, 2018; pp. 1–6.
- [85] Xiong, W.; Wu, L.; Alleva, F.; Droppo, J.; Huang, X.; Stolcke, A. The Microsoft 2017 Conversational Speech Recognition System. In *Proceedings of the 2018 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), Calgary, AB, Canada, 15–20 April 2018*; pp. 5934–5938.
- [86] Mao, J.; Xu, W.; Yang, Y.; Wang, J.; Huang, Z.; Yuille, A. Deep Captioning with Multimodal Recurrent Neural Networks (m-RNN). *arXiv* 2014, arXiv:1412.6632.
- [87] Shi, H.; Xu, M.; Li, R. Deep Learning for Household Load Forecasting—A Novel Pooling Deep RNN. *IEEE Trans. Smart Grid* 2018, 9, 5271–5280.
- [88] Conneau, A.; Schwenk, H.; Barrault, L.; Lecun, Y. Very Deep Convolutional Networks for Text Classification. *arXiv* 2016, arXiv:1606.01781.
- [89] Hongliang, C.; Xiaona, Q. The Video Recommendation System Based on DBN. In *Proceedings of the 2015 IEEE International Conference on Computer and Information Technology; Ubiquitous Computing and Communications; Dependable, Autonomic and Secure Computing; Pervasive Intelligence and Computing, Liverpool, UK, 26–28 October 2015*; pp. 1016–1021.

- [90] Sazal, M.M.R.; Biswas, S.K.; Amin, M.F.; Murase, K. Bangla handwritten character recognition using deep belief network. In Proceedings of the 2013 International Conference on Electrical Information and Communication Technology (EICT), Khulna, Bangladesh, 13–15 February 2014; pp. 1–5.
- [91] Mohamed, A.; Dahl, G.; Hinton, G. Deep belief networks for phone recognition. In Proceedings of the NIPS Workshop on Deep Learning for Speech Recognition and Related Applications; MIT Press: Whister, BC Canada, 2009; Volume 1, p. 39.
- [92] Hinton, G.E. Reducing the Dimensionality of Data with Neural Networks. *Science* 2006, 313, 504–507
- [93] Krizhevsky, A.; Hinton, G.E. Using very deep autoencoders for content-based image retrieval. In Proceedings of the ESANN, Bruges, Belgium, 27–29 April 2011.
- [94] Lu, X.; Tsao, Y.; Matsuda, S.; Hori, C. Speech enhancement based on deep denoising autoencoder. In Proceedings of the Annual Conference of International Speech Communication Association; INTERSPEECH, Lyon, France, 25–29 August 2013; pp. 436–440.

Annex 2: Glossary

This glossary defines resp. explains some of the terms used within this roadmap with respect to validation of highly automated cars. It is meant to be a supplement to the already very extensive PEGASUS family glossar [PEG2023], which we recommend for further explanations.

Automated Driving System (ADS)

Is the hardware and software that are collectively capable of performing the entire Dynamic Driving Task (DDT) on a sustained basis, regardless of whether it is limited to a specific operational design domain (ODD); this term is used specifically to describe a Level 3, 4, or 5 driving automation system.

Reference: SAE J3016:2018, 3.2

Architecture

Fundamental concepts or properties of an entity in its environment (3.13) and governing principles for the realization and evolution of this entity and its related life cycle processes

Reference: ISO 42010:2022

Artifact

An artifact in a sensor measurement is a noticeable deviation from ground truth in the sensor readings that is inherent in the sensor measurement principle and its system design.

Conceptual validation

The process of determining the degree to which a conceptual model (as defined in this NASA Technical Standard) or model design adequately represents the real world from the perspective of the intended uses of the model or the simulation.

Reference: NASA STD 7009A

Coverage

Degree of covering the possible concrete instances of an abstract item, like an ODD or an operational scenario; esp. w.r.t. V&V activities: test coverage, simulation coverage

Credibility

The “quality to elicit belief or trust in [modeling and simulation] results”

Reference: NASA STD 7009A

Decomposition (artifact)

Set of components and interfaces obtained by applying the process of decomposition to a system

Decomposition (process)

Process of splitting a system into its components with defined interfaces, usually decreasing the level of abstraction

Dependability

The persistence of the avoidance of failures that are unacceptably frequent or severe, when facing changes.

Reference: Laprie, J.-C.: "From Dependability to Resilience". In: Dependable Systems and Networks, 2008

Electronic horizon

The electronic horizon refers to an advanced system that extends the perception of the vehicle beyond the range of its onboard sensors. It integrates detailed digital map data and GPS positioning to provide a predictive view of the road network ahead.

Environment simulation

Environment simulations represent the vehicle's physical environment in a virtual world. They include roads, buildings, other vehicles, pedestrians and all potential obstacles.

Failure

Termination of an intended behaviour of an element (3.41) or an item (3.84) due to a fault (3.54) manifestation

Reference: ISO 26262-1:2018, Def. 3.50

Failure Mode and Effects Analysis (FMEA)

FMEA is a systematic method for evaluating a process to identify where and how it might fail and to assess the relative impact of different failures. It aims to identify potential failure modes, determine their effect on the operation of the product or process, and identify actions to mitigate the failures.

Field of view (FoV)

The angle, from which a sensor is receiving information, for which a specified detection performance is reached. Typically, the angle is given as azimuth and elevation angle.

Functional Mockup Unit (FMU)

Functional Mockup Unit (FMU) is a component modeling the behavior of a dynamic entity within an concrete scenario while being executed. It contains an abstract model and can contain a solver. It connects via the Functional Mockup Interface (FMI) to the tool.

Ground truth

Ground truth refers to a set of measures known to be more accurate than the measurements of the SUT. The ground truth represents a reference which is used as a standard for comparison. It is possible that the ground truth was not or cannot be checked.[1, pp.28-30] In context of driving datasets, ground truth typically refers to human annotation.[2, p.3357][3, p.4]

Hardware-in-the-Loop (HiL)

The third in-the-loop method is used to transfer the developed models from the SiL environment to the real components or be replaced by them respectively. The method is referred to as Hardware-in-the-Loop (HiL). In distributed systems, this stage is typically performed in several steps. First, the individual components are tested independently against their respective specifications. Here, a simulation environment is used that provides the interfaces of the components that are to be tested. Once all components are verified with this method, they are partially integrated using the same method to also verify their interaction. At the end of this stage, the entire system exists in real components and is tested against its specification up to the level of the logical architecture.[PEG2023]

Operational context

Context: [...] relationships [...], resolved around a selected [entity]-of-interest - Vorschlag
Operational Context: "relationships resolved around a selected entity of interest relating to the operation of the entity"

Reference: Flood, R.L. and E.R. Carson. 1993. Dealing with complexity: An introduction to the theory and application of systems science, 2nd ed. New York, NY, USA: Plenum Press. (Nach SEBoK)

Operational design domain (ODD)

ODD is defined as the set of all "operating conditions for which a given SUT (driving automation system) is designed, including all restrictions regarding environmental, geography and time of day and/or the required presence or absence of certain traffic or road features". The ODD is the design area of a SUT with regard to its operation.

Operational validation

The process of determining the degree to which a simulation model adequately represents the real world in a specific application

P-Box

A p-box is a probabilistic box or a range of uncertainty associated with a probability distribution. Specifically, it represents uncertainty in the form of a range of possible probability distributions, rather than a single, fixed distribution. Instead of assuming a precise probability distribution for uncertain parameters or variables, which may be challenging or even impossible to determine accurately, a p-box defines a range of possible distributions that encapsulate the uncertainty.

Risk

Combination of the probability of occurrence of harm and the severity of that harm

Reference: ISO 26262, ISO 21448, ISO/IEC Guide 51

Safety

Absence of unreasonable risk

Reference: ISO 26262, ISO/TR 4804, ISO 21448, ISO/IEC Guide 51

Safety of the intended functionality (SOTIF)

The absence of unreasonable risk due to a hazard caused by: a. the insufficiencies of specification of the intended functionality at the vehicle level, or b. the insufficiencies of specification or performance limitations in the implementation of E/E elements in the system

Reference: ISO 21448

Scenario

Description of the temporal development between several scenes in a sequence of scenes.

Every scenario starts with an initial scene. Actions/events, as well as goals/values, can be specified to characterise this temporal development within a scenario. SOTIF

Reference: [PEG2023]

Sensor Model

Virtual entity of a sensor based on effects and uncertainties containing signal propagation and processing. Therefore, the sensor interfaces are identical.

Simulation model

The operational or usable implementation of the conceptual model, including all mathematical, numerical, logical, and qualitative representations.

Reference: NASA STD 7009A

Software-in-the-Loop

The Software-in-the-Loop method (SiL) allows for an assurance up to the level of the individual components. This is achieved by transferring the previously created models into a simulation environment that is very similar to the technical characteristics of the target system in terms of computing power, real-time behavior, or resolution accuracy but is still hardware independent (Martinus et al. 2013). Therefore, the software in the loop (SiL) method offers the possibility to check the specifications of the individual components of a system prior to its implementation and adjust them if necessary.[PEG2023]

System level

Level of abstraction where the ADS-equipped vehicle is considered as a whole and the operational situation is the system's operational context.

System under test

The system under test (SUT) is, like components, not necessarily part of the simulation framework. While the simulation must be able to execute without the SUT, the SUT acts as an

independent agent. [SL] The system to be tested is called the system under test. The complexity of the SUT used depends on the tests to be carried out. [PEG2023]

Test bench

Technical device' (consisting of hardware and software) that provides test objects and elements intended to execute test cases; common test bench types are hardware-in-the-loop, model-in-the-loop, software-in-the-loop, vehicle-in-the-loop

Reference: M. Steimle, T. Menzel, and M. Maurer, "Toward a Consistent Taxonomy for Scenario-Based Development and Test Approaches for Automated Vehicles: A Proposal for a Structuring Framework, a Basic Vocabulary, and Its Application," IEEE Access, vol. 9, pp. 147828–147854, 2021, doi: 10.1109/access.2021.3123504

Testing

Process of creating objective evidence concerning the actual properties of a system or a model hereof

Time-To-Collision

Time-To-Collision (TTC) is a measure for evaluating traffic scenarios (e.g. in simulation). It predicts the time until a collision occurs between objects based on dynamic models.

Uncertainty (development)

"broad and general term used to describe an imperfect state of knowledge or a variability [...]"

Reference: based on NASA STD 7009A

Uncertainty (technical)

1. "estimated amount [...] by which an observed or calculated value may differ from the true value"

2. "non-negative parameter characterizing the dispersion of values attributed to a measured quantity"

Reference: based on NASA STD 7009A

Unreasonable (level of) risk

Risk judged to be unacceptable in a certain context according to valid societal moral concepts"

Reference: ISO 26262, ISO/TR 4804, ISO 21448, ISO/IEC Guide 51

Validation

Confirmation, through the provision of objective evidence, that the requirements for a specific intended use or application have been fulfilled"

Reference: ISO/IEC 9000

Validity (of a simulation model)

Validity of a simulation method: The uncertainty (taking into account the precision of the simulation framework) which affects the result

Vehicle-in-the-Loop

Vehicle-in-the-Loop (ViL) is a newer method for usefully complementing and enhancing the development of advanced driver assistance systems with the V-model. It addresses the need of many driver assistance functions for a complex test drive and a high standard of functional safety. This group of driver assistance functions will progress in importance and size. A major reason for this is the growing number of vehicle variants that offer driver assistance functions and which must remain safe even with the ever-increasing degree of automation and network integration. The ViL method allows the operation of the real test vehicle in a virtual environment. The coupling between the vehicle and the virtual environment can be done in two ways. One way is by creating an interface to the available environment sensors and, thus, replacing the real sensors. At this interface, the simulation environment is feeding simulated sensor signals, which correspond to the sensor response from a real environment. [HK15a, pp.166,167] Otherwise it is possible to maintain the real sensors and stimulate them

artificially, as it is possible for Radar sensors [BAB+21], Lidar sensors, camera and ultrasonic sensors [RGN17]. In both variants, the real test vehicle responds to attributes and events of the virtual environment. This way, critical driving maneuvers with obstacles or objects on a collision course can be tested reliably and reproducibly. The created interface can also be used to generate the sensor signals as they would occur due to a changed position in a vehicle variant or due to different tolerances. This method therefore offers the possibility to test these variants or tolerances with a single test vehicle. In addition to the considerably more safe test operation, this allows efficient testing and application of advanced driver assistance systems. This results in a substantial economic gain with respect to the test drive when it comes to driver assistance systems. [HK15a, pp.166,167]

Reference: • Stephan Hakuli and Markus Krug. *Virtuelle integration*. In Hermann Winner, Stephan Hakuli, Felix Lotz, and Christina Singer, editors, *Handbuch Fahrerassistenzsysteme*, pages 125–138. Springer Fachmedien Wiesbaden, Wiesbaden, 2015
- Sreehari Buddappagari, M.E. Asghar, F. Baumgartner, S. Graf, F. Kreuz, A. Löffler, J. Nagel, T. Reichmann, R. Stephan, and Matthias A. Hein. *Over-the-air vehicle-in-the-loop test system for installed-performance evaluation of automotive radar systems in a virtual environment*. In *2020 17th European Radar Conference (EuRAD)*, pages 278–281, 2021
• Romain Rossi, Clement Galko, and Hariharan Narasimman. *11 vehicle hardware-in-the-loop system for adas virtual testing*. 2017

Verification

confirmation, through the provision of objective evidence, that specified requirements have been fulfilled

Reference: ISO/IEC 9000

Verification (of simulation model)

The process of determining the extent to which an M&S is compliant with its requirements and specifications as detailed in its conceptual models, mathematical models, or other constructs.

Reference: NASA STD 7009A

Virtual environment

Being on or simulated on a computer or a computer network. [Mer21]

Vulnerable Road User

Traffic participants such as pedestrians or cyclists that are not protected by a metal vehicle body and are thus more susceptible to injury or death in the event of a collision.