Measurable Safety of Automated Driving Functions in Commercial Motor Vehicles

Technological and Methodical Approaches

Zur Erlangung des akademischen Grades eines
DOKTORS DER INGENIEURWISSENSCHAFTEN (Dr.-Ing.)

von der KIT-Fakultät für Maschinenbau des
Karlsruher Instituts für Technologie (KIT)

genehmigte

DISSERTATION

von

M.Eng. Mohamed Elgharbawy

Tag der mündlichen Prüfung: 14.09.2022

Hauptreferent: Prof. Dr. rer. nat. Frank Gauterin
Korreferent: Prof. Dr.-Ing. Eric Sax

Vorwort des Herausgebers

Die Fahrzeugtechnik ist kontinuierlich Veränderungen unterworfen. Klimawandel, die Verknappung einiger für Fahrzeugbau und -betrieb benötigter Rohstoffe, globaler Wettbewerb, gesellschaftlicher Wandel und das rapide Wachstum großer Städte erfordern neue Mobilitätslösungen, die vielfach eine Neudefinition des Fahrzeugs erforderlich machen. Die Forderungen nach Steigerung der Energieeffizienz, Emissionsreduktion, erhöhter Fahr- und Arbeitssicherheit, Benutzerfreundlichkeit und angemessenen Kosten sowie die Möglichkeiten der Digitalisierung und Vernetzung finden ihre Antworten nicht aus der singulären Verbesserung einzelner technischer Elemente, sondern benötigen Systemverständnis und eine domänenübergreifende Optimierung der Lösungen.

Hierzu will die Karlsruher Schriftenreihe für Fahrzeugsystemtechnik einen Beitrag leisten. Für die Fahrzeuggattungen Pkw, Nfz, Mobile Arbeitsmaschinen und Bahnfahrzeuge werden Forschungsarbeiten vorgestellt, die Fahrzeugsystemtechnik auf vier Ebenen beleuchten: das Fahrzeug als komplexes, digitalisiertes mechatronisches System, die Mensch-Fahrzeug-Interaktion, das Fahrzeug in Verkehr und Infrastruktur sowie das Fahrzeug in Gesellschaft und Umwelt.

Die Absicherung hoch automatisierter Fahrfunktionen ist allein im Fahrversuch nicht realisierbar. Die schier unendlich hohe Anzahl verschiedener Verkehrsszenarien würde bei einer reinen Teilnahme am Straßenverkehr sehr hohe Fahrstrecken erfordern, um jeder möglichen Konstellation zu begegnen. Dabei würde der Großteil der gefahrenen Strecke keine besonderen Herausforderungen für Sensorik, Algorithmik und Aktorik darstellen. Nur ein kleiner Teil der Fahrsituationen wäre zum Überprüfen eines angemessenen Fahrzeugverhaltens geeignet. Zur Steigerung der Effektivität und Effizienz von Absicherungsverfahren ist die Identifikation sicherheitskritischer Verkehrssituationen und deren gezielten Verwendung zum Funktionstest erforderlich.

Hier setzt die Arbeit von Herrn Elgharbawy an, in der er im Fahrversuch gewonnene wissensbasierte Sammlungen risikobehafteter Szenarien durch automatisiert generierte Testfälle ergänzt, um so den vorgesehenen Betriebsbereich des Fahrzeugs möglichst vollständig abzudecken. Das Fahrzeugverhalten in diesen Testfällen bestimmt er in der Simulation oder auch ergänzt durch Hardware-in-the-Loop-Verfahren und leitet daraus die Wahrscheinlichkeit definierter Fehlerfälle sowie über eine Akzeptanzschwelle den Abbruch des Verifikationslaufs nach Erreichen der gewünschten Vorhersagegüte ab.

Frank Gauterin

Karlsruhe, im September 2022

Kurzfassung

Fahrerassistenzsysteme sowie automatisiertes Fahren leisten einen wesentlichen Beitrag zur Verbesserung der Verkehrssicherheit von Kraftfahrzeugen, insbesondere von Nutzfahrzeugen. Mit der Weiterentwicklung des automatisierten Fahrens steigt hierbei die funktionale Leistungsfähigkeit, woraus Anforderungen an neue, gesamtheitliche Erprobungskonzepte entstehen. Um die Absicherung höherer Stufen von automatisierten Fahrfunktionen zu garantieren, sind neuartige Verifikations- und Validierungsmethoden erforderlich.

Ziel dieser Arbeit ist es, durch die Aggregation von Testergebnissen aus wissensbasierten und datengetriebenen Testplattformen den Übergang von einer quantitativen Kilometerzahl zu einer qualitativen Testabdeckung zu ermöglichen. Die adaptive Testabdeckung zielt somit auf einen Kompromiss zwischen Effizienz- und Effektivitätskriterien für die Absicherung von automatisierten Fahrfunktionen in der Produktentstehung von Nutzfahrzeugen ab. Diese Arbeit umfasst die Konzeption und Implementierung eines modularen Frameworks zur kundenorientierten Absicherung automatisierter Fahrfunktionen mit vertretbarem Aufwand. Ausgehend vom Konfliktmanagement für die Anforderungen der Teststrategie werden hochautomatisierte Testansätze entwickelt. Dementsprechend wird jeder Testansatz mit seinen jeweiligen Testzielen integriert, um die Basis eines kontextgesteuerten Testkonzepts zu realisieren. Die wesentlichen Beiträge dieser Arbeit befassen sich mit vier Schwerpunkten:

- Zunächst wird ein Co-Simulationsansatz präsentiert, mit dem sich die Sensoreingänge in einem Hardware-in-the-Loop-Prüfstand mithilfe synthetischer Fahrszenarien simulieren und/ oder stimulieren lassen. Der vorgestellte Aufbau bietet einen phänomenologischen Modellierungsansatz, um einen Kompromiss zwischen der Modellgranularität und dem Rechenaufwand der Echtzeitsimulation zu erreichen. Diese Methode wird für eine modulare Integration von Simulationskomponenten, wie Verkehrssimulation und Fahrdynamik, verwendet, um relevante Phänomene in kritischen Fahrszenarien zu modellieren.

- Danach wird ein Messtechnik- und Datenanalysekonzept für die weltweite Absicherung von automatisierten Fahrfunktionen vorgestellt, welches eine Skalierbarkeit zur Aufzeichnung von Fahrzeugsensor- und/ oder Umfeldsensordaten von spezifischen Fahrereignissen einerseits und permanenten Daten zur statistischen Absicherung und Softwareentwicklung andererseits erlaubt. Messdaten aus länderspezifischen Feldversuchen werden aufgezeichnet und zentral in einer Cloud-Datenbank gespeichert.

- Anschließend wird ein ontologiebasierter Ansatz zur Integration einer komplementären Wissensquelle aus Feldbeobachtungen in ein Wissensmanagementsystem beschrieben. Die Gruppierung von Aufzeichnungen wird mittels einer ereignisbasierten Zeitreihenanalyse mit hierarchischer Clusterbildung und normalisierter Kreuzkorrelation realisiert. Aus dem extrahierten Cluster und seinem Parameterraum lassen sich die Eintrittswahrscheinlichkeit jedes logischen Szenarios und die Wahrscheinlichkeitsverteilungen der zugehörigen Parameter ableiten. Durch die Korrelationsanalyse von synthetischen und naturalistischen Fahrszenarien wird die anforderungsbasierte Testabdeckung adaptiv und systematisch durch ausführbare Szenario-Spezifikationen erweitert.

- Schließlich wird eine prospektive Risikobewertung als invertiertes Konfidenzniveau der messbaren Sicherheit mithilfe von Sensitivitäts- und Zuverlässigkeitsanalysen durchgeführt. Der Versagensbereich kann im Parameterraum identifiziert werden, um die Versagenswahrscheinlichkeit für jedes extrahierte logische Szenario durch verschiedene Stichprobenverfahren, wie beispielsweise die Monte-Carlo-Simulation und Adaptive-Importance-Sampling, vorherzusagen. Dabei führt die geschätzte Wahrscheinlichkeit einer Sicherheitsverletzung für jedes gruppierte logische Szenario zu einer messbaren Sicherheitsvorhersage.

Das vorgestellte Framework erlaubt es, die Lücke zwischen wissensbasierten und datengetriebenen Testplattformen zu schließen, um die Wissensbasis für die Abdeckung der Operational Design Domains konsequent zu erweitern. Zusammenfassend zeigen die Ergebnisse den Nutzen und die Herausforderungen des entwickelten Frameworks für messbare Sicherheit durch ein Vertrauensmaß der Risikobewertung. Dies ermöglicht eine kosteneffiziente Erweiterung der Validität der Testdomäne im gesamten Softwareentwicklungsprozess, um die erforderlichen Testabbruchkriterien zu erreichen.

Abstract

Driver assistance systems and automated driving make a significant contribution in improving the road safety for vehicles, particularly the commercial motor vehicles. With the further development of automated driving, the functional performance increases resulting in the need for new and comprehensive testing concepts. New verification and validation methods are therefore required to cope up with the testing of higher levels of the automated driving functions.

This doctoral thesis aims to enable the transition from quantitative mileage to qualitative test coverage by aggregating the results of both knowledge-based and data-driven test platforms. The adaptive test coverage thus seeks to achieve a compromise between efficiency and effectiveness criteria of the assessment of automated driving functions in the product development of commercial motor vehicles. The systematic approach of this work includes the conception and implementation of a modular framework for the customer-oriented testing of the automated driving functions with reasonable efforts. Based on conflict management for the requirements of the test strategy, highly automated test approaches are developed. Accordingly, each test approach is integrated with its respective test objectives to realize the basis of a context-driven test concept. The main contributions of this thesis are fourfold:

- First, a co-simulation approach is presented which allows the perception sensor inputs to be simulated and/or stimulated in a Hardware-in-the-Loop test bench using synthetic driving scenarios. The presented setup offers a phenomenological modeling approach to achieve a compromise between the model granularity and the computational effort of the real-time simulation. This approach is used in a modular integration of simulation components such as traffic simulation and vehicle dynamics to model the relevant phenomena in critical driving scenarios.

- Next, a data-logging and analysis approach for the worldwide validation of the automated driving functions is presented. This approach provides

scalability for the acquisition of vehicle sensor and/or environment-perception-sensor data of specific driving events on one hand and permanent data for statistical coverage and software development on the other hand. A cloud database centrally stores the acquired measurement data from the country-specific field tests.

- Subsequently, an ontology-based approach is described for integrating a complementary knowledge source from field observations into a knowledge management system. The clustering of recordings is realized by the means of an event-based time-series analysis with hierarchical clustering and normalized cross-correlation. The extracted clusters and their parameter space define the probability of occurrence of each logical scenario and the probability distributions of the associated parameters. Thereby, the correlation analysis of synthetic and naturalistic driving scenarios enlarges the requirements-based test coverage adaptively and systematically by executable scenario specifications.

- Eventually, a prospective risk assessment is carried out as an inverted confidence level of measurable safety using sensitivity and reliability analyses. The failure region is identified in the parameter space to predict the failure probability for each extracted logical scenario using sampling methods such as Monte-Carlo simulation and Adaptive Importance Sampling. The estimated probability of a safety violation for each clustered logical scenario results in a measurable safety prediction.

The presented framework allows a patching of the gap between knowledge-based and data-driven test platforms, thus consistently expanding the knowledge database of the Operational Design Domain coverage. In summary, the results show the benefits and challenges of the developed framework for measurable safety through a risk assessment confidence level. As a result, the validity of the test domain can be extended cost-effectively throughout the software development process to achieve meaningful test termination criteria.

Acknowledgement

The proposed work is the result of a fruitful cooperation between the *Truck Production Engineering* at the Daimler Truck AG and the *Institute of Vehicle System Technology* (FAST) at the Karlsruhe Institute of Technology (KIT).

First and foremost, I would like to express my sincere gratitude to my academic supervisor Professor Dr. rer. nat. Frank Gauterin for his guidance, encouragement and support throughout my scientific work. I'm deeply indebted to Dr.-Ing. Michael Frey for his support and inspiring ideas for this dissertation. I am extremely grateful to Professor Dr.-Ing Eric Sax from the *Institute for Information Processing Technologies* (ITIV) at the KIT for his helpful suggestions and constructive criticism. I am also grateful to Professor Dr.-Ing Xu Cheng from the *Institute for Applied Thermofluidics* (IATF) at the KIT for chairing the examination committee.

I would like to extend my sincere thanks to Dr.-Ing. Christof M. Weber, head of global testing at Daimler Trucks, for his interest in my dissertation. I would like to express my deepest appreciation to my industrial supervisors Dr.-Ing. Andreas Schwarzhaupt and Ingo Scherhaufer for their great enthusiasm and inspiring feedback during the four years of my work on this project.

Several people have played a key role in the completion of this work. I am grateful to my colleagues from Daimler Truck AG for the great atmosphere and on a personal level, Matthias Gut, Hans-Jürgen Gutmayer, Christina Werner and Dr.-Ing. Urs Wiesel and my colleagues from Mercedes-Benz Group AG, Dr.-Ing. Jürgen Dickmann and Christoph Wohlfahrt. Furthermore, I would like to express my warmest thanks to my colleagues from the FAST institute for their support; i.e.: Adam Birlet, Alexander Brunker and Dr.-Ing. Rayad Kubaisi. I would also like to highlight the company VIRES Simulationstechnologie GmbH, which with its managing director, Mr. Marius Dupuis, carried out the technical implementation of the Virtual Test Drive (VTD) integration into the Hardware-in-the-Loop test bench with professional passion.

I gratefully acknowledge the effort of my students who contributed with their Bachelor's or Master's Theses to this work those being: Max Burkhard Ammar, Martin Kanberger, Rainer Arenskrieger, Gerrit Scheike and Hesham Elsayed.

Last but not least, a great deal of thanks goes to my family and closest friends, especially my beloved mother and my adorable wife, for their unlimited care, support, and encouragement. To both my sisters, who were of great help to me in discussing some of the equations used in the dissertation and applied in pharmaceutical applications, I express my gratitude for their genuine and heartfelt support. Many thanks to my brother, as a radiologist, for inspiring me to explore the applications of health risk assessments and their use in healthcare. This thesis is dedicated to the memory of my late father, who was and remains my constant source of inspiration.

Mohamed Elgharbawy

Karlsruhe, 11. Juli 2022

Contents

1 Introduction

The so-called Commercial Motor Vehicles (CMVs) are motor vehicles specially designed to acquire economic value by transporting goods or individuals [FL+20]. Therefore, CMVs are highly specialized in fulfilling specific tasks and are primarily controlled by economic efficiency [Abe08]. In addition, the CMVs are characterized by a large number of series and models with tractors, semi-trailers or trailer combinations [T+17]. In most nations, the legislation regulates the concepts and functions of CMVs up to a specific vehicle system [ZMMFm13]. Moreover, the current challenges consist of improving road safety, addressing the scarcity of CMV drivers and making the profession of a CMV driver more attractive [Kir15]. The vision of autonomous and accident-free driving thus, forces the further development of the automated driving technologies to be a current topic of high relevance in the transport sector [WRM+19]. Consequently, a broad range of political and economic stakeholders support this innovation in the expectation of reducing congestion, fuel consumption and accidents [MV19].

1.1 Background

In accordance with the World Health Organization (WHO), road accidents cause almost 1.35 million deaths and 20 to 50 million injuries every year [WHO20]. According to truck accident statistics, most accidents are caused by driver fatigue, lack of route information, as well as job pressure and aggressive driving [Kop22]. As reported by the United States Department of Transportation (USDOT), driver fatigue is a major cause of nearly 4,000 fatalities in truck accidents on United States (U.S.) roads annually [Ass20]. Therefore, the United Nations General Assembly (UNGA) has launched a decade of road safety measures between 2011 and 2020 in order to reduce the risk of road accidents and injuries [fSIRTA22].

In 2015, the Federal STATIStical Office of Germany (DESTATIS) recorded a total of 29,480 accidents involving personal injury with the participation of at least one heavy-duty truck in Germany. In spite of the accident variety with heavy-duty trucks, the statistics shows that rear collisions and unintended lane departures are the common types of CMV accidents with 68% of 32,500 truck drivers [DES16]. Figure 1.1 depicts that the number of killed road users represent 2% of the total number of injuries and deaths with 787 fatalities. The three pie charts illustrate the percentage of types and causes of accidents involving the commercial trucks and the individuals involved on German roads in 2015.

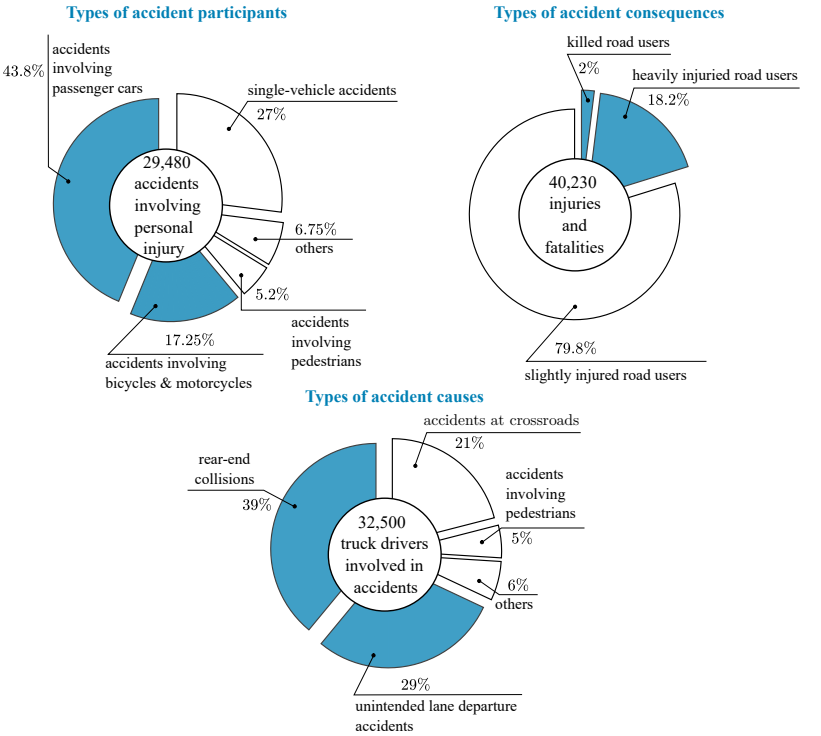

Figure 1.1: Retrospective accident analysis of heavy-duty trucks on German roads according to the DESTATIS report for 2015 [DES16].

On one hand, the truck accidents often have serious consequences such as injury and death, along with considerable financial impacts and environmental risks. Beyond the health and economic consequences, the truck accidents potentially cause significant losses on several levels such as reducing the efficiency of traffic flow and causing congestion [SWZ12]. On the other hand, freight traffic continues to increase globally and is the dominant means of transport. According to the traffic forecast for 2030, the volume of road freight transport in Germany will increase by 38% compared to the level of 2010 [M+20a]. Between 1992 and 2021, the number of truck accidents involving seriously injured road users has fallen by more than 59.0%. While the volume of truck traffic increased by 92.7% over the same period, the number of people who died in these accidents fell by more than 67.3%, as shown in figure 1.2.

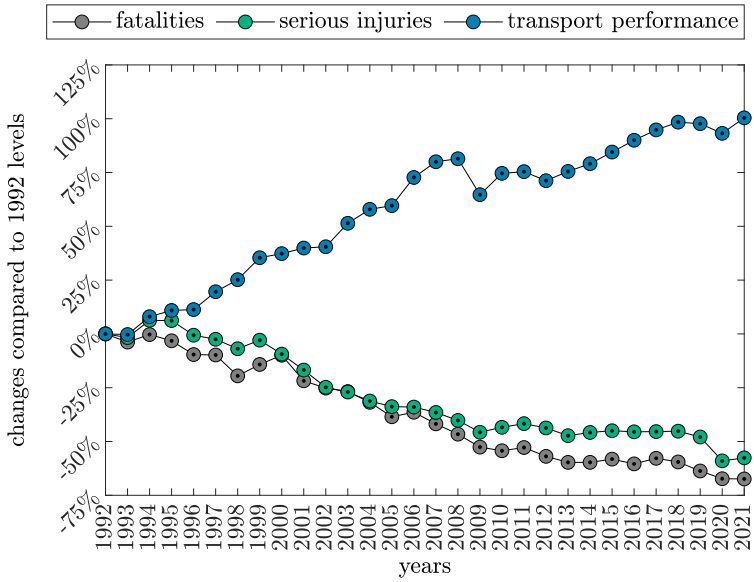

Figure 1.2: Fatalities and seriously injured persons in truck accidents on German roads compared to truck transport performance between 1992 and 2021 [uEB22].

From a legal perspective, the road safety of commercial vehicles is an essential aspect of civil society. The European Commission (EC) has adopted the United Nations Regulation (UNR) nos. 130 and 131 to improve the safety of heavy-duty vehicles in the framework of the general safety regulation no. 661/2009. As a result, the installation of Autonomous Emergency Braking (AEB) and Lane Departure Warning (LDW) systems has become mandatory in Europe for all the new heavy-duty vehicles with a permissible gross weight of more than 3.5 tonnes. Only the off-road vehicles and trucks with more than three axles are exempt from these regulations [Elg12]. Moreover, the United Nations Economic Commission for Europe (UN/ECE) has introduced amendments to the Vienna Convention on Road Traffic (VCRT) in 2016 to explicitly allow the transfer of Dynamic Driving Tasks (DDTs) to Automated Driving Functions (ADFs) under the condition that these functions can be bypassed or switched off by the driver [MO17].

Definition 1.1 (AEB)**:** An Advanced Driver Assistance System (ADAS) that can automatically determine the time required to perform the warning cascade and emergency braking in order to prevent the collision. Therefore, the AEB system represents an active safety function in the forward control of the CMV, in particular the control of the truck's braking system. The AEB function includes a detection unit for measuring a distance between the Ego-vehicle and the relevant object ahead [ESS$^+$19c].

Definition 1.2 (LDW)**:** An ADAS that can warn the driver to prevent unintentional lane departure due to driver inattention or distraction. For this purpose, the LDW function utilizes the inputs from a front camera Electronic Control Unit (ECU) installed in the middle of the CMV behind the windshield. The camera ECU detects the lane markings and operates at speeds above 60 [km/h] to minimize false alarms related to construction sites and urban traffic [ESF19].

The VCRT of 1968 provides governments with a regulatory automobile framework for their national highways to ensure a high level of road safety for the contracting parties. For this reason, the Automatically Commanded Steering Function (ACSF) Category E within the UN/ECE Regulation no. 79 corresponds to a function which is activated by the driver and which can continuously identify the possibility of a maneuver (e.g. lane change) and perform these maneuvers over extended periods of time without driver confirmation [BHS17].

In addition, the USDOT published a framework in 2016 to support the safe development, testing and integration of automated vehicles [NHT17a]. Moreover, in 2018, the Ministry of Industry and Information Technology (MIIT) of the Chinese government launched guidelines for the building of intelligent connected vehicles to expedite the development and review of standards for autonomous driving safety [WRM+19].

From a commercial perspective, heavy-duty and passenger vehicles differ in both their economic significance and the vehicle technology. The trucking business focuses on economic factors such as fuel consumption, truck utilization, driver demand and hours of service [LBS19]. A potential cut off in the cost therefore is a motivation for high driving automation in the long distance trucking. For fleet owners, financial benefits are of paramount importance in this context concerning the proposal of a new business case or insurance incentives [P+19a]. Thus, one of the key business cases is to achieve more service times for trucks and buses by supporting drivers or even substituting them with automated driving technologies [LSW18]. In this context, the niche strategy requires starting an application domain with a controllable environment in order to gradually expand the intended Operational Design Domain (ODD). According to the American Trucking Associations (ATA), the annual turnover rate of CMV fleet drivers is increasing, which indicates a strong demand for truck drivers [McN18]. In addition, it is expected that by 2026 the scarcity of truck drivers in the U.S. will reach over 174,000 drivers [CS15]. Accordingly, fleets respond to the scarcity of drivers by increasing the salaries to make the truck driver's profession more attractive [McN19, BLS22].

From a technical perspective, there are several challenges to be overcome that affect the functionality of driving automation, such as significant variations in the state of vehicle loading, dimensions, weight, center of gravity position and braking performance of the heavy-duty truck. Despite the rapid advances in driving dynamics for enhanced safety, the requirements for heavy-duty vehicles differ considerably from those for passenger cars. In addition, CMV manufacturers face special challenges of the relatively large numbers of variants with significantly lower production volumes and longer product life cycles. Table 1.1 summarizes the different technical requirements for automation of passenger cars and CMVs. Therefore, long-haul CMVs are heavier, larger and less maneuverable than passenger cars [Har03].

CMV characteristics (e.g. dimension, low-speed transient off-tracking, braking distance, number of variants, etc.) therefore pose new challenges for ADFs. The off-tracking refers to a phenomenon in which the rear wheels track inside the path, traced by the front wheels, when a vehicle turns [ESS+19b]. However, automated driving systems in the CMV sector are the most important business cases due to long-distance journeys, which sometimes reach more than 100,000 kilometers per year on long and monotonous routes [Kir15].

Property	**Automated passenger car**	**Automated CMV**
Niche market	urban automation	highway automation
Mainstream market	vehicle on demand services	road freight transport services
Potential user groups	people with age-related or medical constraints, teenagers and long-distance commuters	logistics and haulage companies
360° surround detection	feasible	full rear view not feasible
Detection range	well standardized	extended range due to longer braking distance
Vehicle setup	single body with well-known structure	multi-body with decoupled cabin and unknown trailer structures
Number of variants	limited	numerous
Off-tracking	negligible effect	non-negligible effect within curve driving scenarios
Tire type	single	mixture of single and twin
Trailer	hardly probable	articulated trailers without environment sensing
Supposed driver skills	basic	professional
Driving scenarios	complex	selected routes on highway

Table 1.1: Overview of the differences in automation requirements between passenger cars and CMVs.

1.2 Automated Truck Driving

Driver assisted trucks seek to improve the road safety either by warning the driver to avoid truck accidents or by directly taking control of the vehicle. With no claim to completeness compared with the relevant market, in 2006, Mercedes-Benz Trucks launched Active Brake Assist (ABA) 1 [TZ15] as the first emergency braking assistant in trucks that can prevent rear-end collisions.

Consequently, since 2011 the Adaptive Cruise Control (ACC) has been supplemented with the Stop and Go function as an automatic distance control system. The ACC can identify the relevant preceding vehicles and calculate both the deceleration as well as the possible acceleration required to maintain a safe distance. Due to the protection of Vulnerable Road Users (VRUs), such as pedestrians and cyclists, being a central aspect of safety development, ABA 4 and Side Guard Assist (SGA), both featuring pedestrian detection, were launched in 2016. The SGA can assist the truck driver in turning at low speeds when an object is laterally next to the heavy-duty truck or when the visibility is restricted by the vehicle length or due to adverse weather conditions. Furthermore, the SGA can detect moving and stationary objects in the warning zone on the co-driver's side or in the turning curve and warn the driver visually and acoustically in critical driving situations [HSP19].

The ABA 4 can detect obstacles and VRUs by using multi-mode RAdio Detection And Ranging (RADAR) systems for short and long ranges and warn the driver of imminent collisions with pedestrians and simultaneously automatically initiate partial braking. The ABA 4 therefore allows the driver to avoid a collision by means of an emergency braking or a steering maneuver. In addition, the driver can warn pedestrians at risk by sounding the horn [Gol18]. In 2019, the fifth generation of the ABA has been launched into series production with improved pedestrian detection using a camera and RADAR system to monitor and detect the relevant preceding objects. The ABA 5 can warn the driver by a combination of optical, acoustic or haptic signals. Furthermore, the system can automatically determine the time required to perform the warning cascade and emergency braking in order to prevent the collision [Tru18].

In parallel, a partially automated driving system called Active Drive Assist (ADA) was launched in the market which can assist the driver in steering, accelerating and decelerating. In the case of lateral control, ADA can detect lane markings to keep the vehicle actively in lane via an electronically controlled steering system. As a result, the steering torque is designed to allow the driver to retain control of the vehicle at all times. The ADA also includes another function, called Lane Departure Protection (LDP), which can smoothly guide the vehicle back to the center of the lane in the event of an unintended crossing of lane markings following an acoustic warning. In the case of longitudinal control, ADA allows the truck to slow down on the approach to other vehicles in traffic and to accelerate again as the preceding traffic moves away.

Accordingly, Daimler Trucks and Buses already met many important legal requirements in the area of road safety for CMVs many years prior to their entry into force [AG19, WRM+19, SB08]. For the self-driving trucks, the National Highway Traffic Safety Administration (NHTSA) has defined a set of safety design elements for the development and safeguarding of ADFs [NHT17b]. NHTSA encourages the automotive industry to provide these elements in the form of Voluntary Safety Self Assessment (VSSA) reports as guidelines for a safe deployment of ADFs on U.S. roadways. Potential users of truck automation technology are logistics and freight forwarding companies as buyers of trucks [Flä16]. One of the motivations for a higher degree of automation in long-distance haulage is therefore a possible reduction in operating costs. The greatest potential benefit for users therefore arises when the driver is replaced by the technology. In fact, the prevailing shortage of truck drivers, rising cost pressure and low margins, as well as the growing need for efficient logistics processes, are driving the demand for automated trucks [MV19].

By law, the companies that actively test their self-driving trucks on California's public highways are required to disclose the number of kilometers driven in autonomous mode and the number of disengagements in which the test driver disengages the autonomous mode and takes immediate manual control of the self-driving truck. Accordingly, the vision of automated road transport in the logistics industry can serve to address the driver shortage, reduce costs and thus increase profit margins, and achieve greater process reliability by leveraging business cases, e.g. truck platooning and hub-to-hub autonomous trucking [MV19].

Definition 1.3 (**Truck Platooning**)**:** A linkage of two or more vehicles in convoy, where the leading vehicle would be manned by a driver and the following vehicles would be electrically coupled trailers for certain parts of a journey. Truck platooning enables a business case for truck automation to increase road capability and reduce fuel consumption [ESFG18].

Definition 1.4 (**Hub-to-Hub Autonomous Trucking**)**:** An application of the Transport-as-a-Service (TaaS) model that deploys trucks on predefined routes between depots or scheduled waypoints for shippers, carriers, logistics service providers and freight brokers [Sjo22].

With no claim to completeness compared with the relevant market, various autonomous truck companies have recently introduced concepts and prototypes for self-driving trucks and released their VSSA disclosures or disengagement reports according to the safety design elements outlined by the NHTSA, such as Daimler Trucks[1], Waymo[2], Volvo[3], Tesla[4], Einride[5], Embark[6] [Rod19], TuSimple[7] [TuS19], Kodiak Robotics[8] [Kod20] and Aurora[9] [Aur21].

1.3 Problem Statement

The digital transformation has led to rapid and profound changes in the CMV industry. As a result, the future of road freight transport is changing in the era of connected and automated driving. From today's perspective, global research and development activities are encouraging the Technology Readiness Level (TRL) of automated driving in the automotive industry. The digital transformation journey is thus fueled by the new driving force that is the data. The use of Field Operational Tests (FOTs) is indispensable to demonstrate the safety and reliability of these innovative technologies. Furthermore, the test CMVs equipped with data logging devices produce massive quantity of data on regular from public roads with the corresponding geographical distribution patterns.

On one hand, their validation methods necessitate a higher level of bandwidth and storage of measurement recordings (e.g. automotive data communication buses, raw sensor detection lists, reference video streams, etc.). On the other hand, the use of these recorded data-sets is essential within software reprocessing and repeatable regression testing for a data-driven development of driving automation.

[1] https://www.daimlertruck.com/innovation/autonomous-driving/our-path-to-autonomous-trucks.html
[2] https://waymo.com/waymo-via/
[3] https://www.volvogroup.com/en-en/innovation/automation.html
[4] https://www.tesla.com/semi
[5] https://www.einride.tech/
[6] https://embarktrucks.com/
[7] https://www.tusimple.com/
[8] https://kodiak.ai/
[9] https://aurora.tech/

Apart from random hardware failures, logical and statistical failures can be distinguished as two categories of systematic software failures. The identification of logical errors demands meticulous analysis, convenient test equipment and proper functional decomposition. In contrast, the estimation of statistical errors often requires a stochastic analysis of unintended reactions of the ADF in the recorded data-sets with the dynamic traffic situations of daily life. Consequently, the new highly complex technologies for automated driving need appropriate test strategies for a reliable and safe heavy-duty vehicle.

The Automotive Systems Engineering (ASE) has established data-driven and knowledge-based test methods to ensure the required reliability and safety of ADFs. Data-driven approaches provide empirical evidence for the validation based on the Key Performance Indicators (KPIs). Moreover, various X(something)-in-the-Loop (XiL) simulations, ranging from microscopic to macroscopic traffic simulations and proving grounds, are used to enable efficient verification within the process of software product engineering. These knowledge-based approaches are often applied in a closed-loop setting to ensure a requirement-based test coverage, defined by expert knowledge in the form of Natural Language (NL) statements. Consequently, implicit knowledge sources are extracted to transform them into explicit test specifications with the required granularity. Despite the systematic of the top-down approach, one of its major drawbacks is the assumption of knowledge completeness with restricted change of requirements during development. Besides, route-based assessment procedures offer a variety of situations with real traffic conditions [KRK+19]. Nevertheless, a significant reduction of the required driving mileages is unavoidable [JS+19a].

According to the Society of Automotive Engineers (SAE) J3016, the status quo of ADFs extends to partially automated driving (SAE L2) [Int18]. In these systems, the driver retains control of the vehicle and remains obliged to monitor the functional intervention on a regular basis and, if necessary, to take over vehicle control. The decisive factors for the series development of these systems are the controllability of system interventions and the effectiveness in real traffic with minimal unintended reactions [WW16]. But a heavy-duty truck equipped with automated drive controlling can still be identified as a cyber-physical vehicle system in which the driving functions are designed to cope with dynamic traffic situations in an extremely safety-critical context.

Starting from SAE L3, a new type of vehicle guidance is implemented, in which the established test methods are neither sufficient nor appropriate [Win16b]. The main difference is that driver assistance may have unintended interventions that can be corrected by the driver. In the presence of functional inadequacies, the driver supervises the automated driving controller and performs the Object and Event Detection and Response (OEDR) sub-task [BDF+14]. Accordingly, the human assisted driving functions are designed to be controllable at any time, but this can reduce their benefits [Wei13]. The controllability of system interventions and the effectiveness in the field with minimal undesired consequences are therefore decisive for the series development of these driving systems [W+18]. As a result, the ASE requires state-of-the-art evaluation procedures to verify and validate these systems. The FOT is carried out to define thresholds for intervening systems based on the collected data. On one hand, trigger algorithms can be optimized to minimize the frequency and impact of falsely triggered interventions and, on the other hand, to maximize the number of legitimate responses. Nevertheless, driving automation requires the system to exploit the limits of DDTs and to master most environmental conditions controlled by a human driver [Sch15].

The International Organization for Standardization (ISO) 26262:2018 standard extends the functional safety regulations of Electrical and/or Electronic (E/E) systems for motorcycles and CMVs. However, the safety standard is limited to avoiding potentially safety-critical situations caused by systematic software and random hardware failures [AW17]. Safety violations due to technological and technical deficiencies remain outside the scope of ISO 26262:2018 (e.g. insufficient robustness, uncertainty issues with perception sensors, etc.) [BGH17]. Specifically, automated driving without driver monitoring can also lead to potentially safety-critical situations resulting from deficiencies in the estimation, interpretation and perception processes. Therefore, critical driving situations due to systematic software and random hardware failures can be handled within the ISO 26262:2018 standard. The ISO/ Publicly Available Specification (PAS) 21448:2019 standard regulates the absence of unreasonable risk due to hazards resulting from functional insufficiencies of the intended functionality or by reasonably foreseeable misuse by persons [SH19b].

Definition 1.5 (ISO **26262):** It specifies a development process for the functional safety of E/E systems in the automotive industry. It outlines a risk classification system and aims to reduce possible hazards caused by the malfunctioning behavior of E/E systems [Hil12].

Definition 1.6 (ISO/PAS **21448):** It presents a guidance on the applicable design, verification and validation measures needed to demonstrate that there are no unreasonable risks arising from hazards due to performance limitations of the intended behavior [fS+19b].

While there are no generally accepted test procedures at present that enable ADFs to be validated in affordable efforts, ongoing and completed German Federal Ministry for Economic Affairs and Energy (BMWi) research projects (e.g. PEGASUS [WLFM19], AdaptIVe [Ete17], SafeMove [GA+18a, ADPZ19], KI-Absicherung [GGSB19], SET Level4to5 [HTR+20] and VVM [ZRBE20]) show the relevance of research for new test methods. Accordingly, the main research questions are formulated as follows:

RQ1: **How can the knowledge-based and data-driven test platforms be combined in a complementary and collaborative manner?**

RQ2: **How can the ADFs be effectively and efficiently tested?**

RQ3: **How can the prospective risk be measured in order to achieve reasonable termination criteria for the testing of the ADFs?**

1.4 Research Objectives

The statistical analysis of road accidents predicts the required mileage for levels of automation without driver involvement as a basis for the safety of new systems compared to their predecessors [JWKW18]. These technologies face an unsolved challenge when it comes to proving safety during the development phase by means of FOTs. While the ADF uncertainties remain before automated driving is released for widespread use, it is essential to develop performance assessments for the safety confidence. Furthermore, highly automated test approaches are integrated to safeguard the ADFs. Black-box, grey-box and white-box tests are associated with their respective test objectives and form the basis for a context-driven test concept, as depicted in figure 1.3.

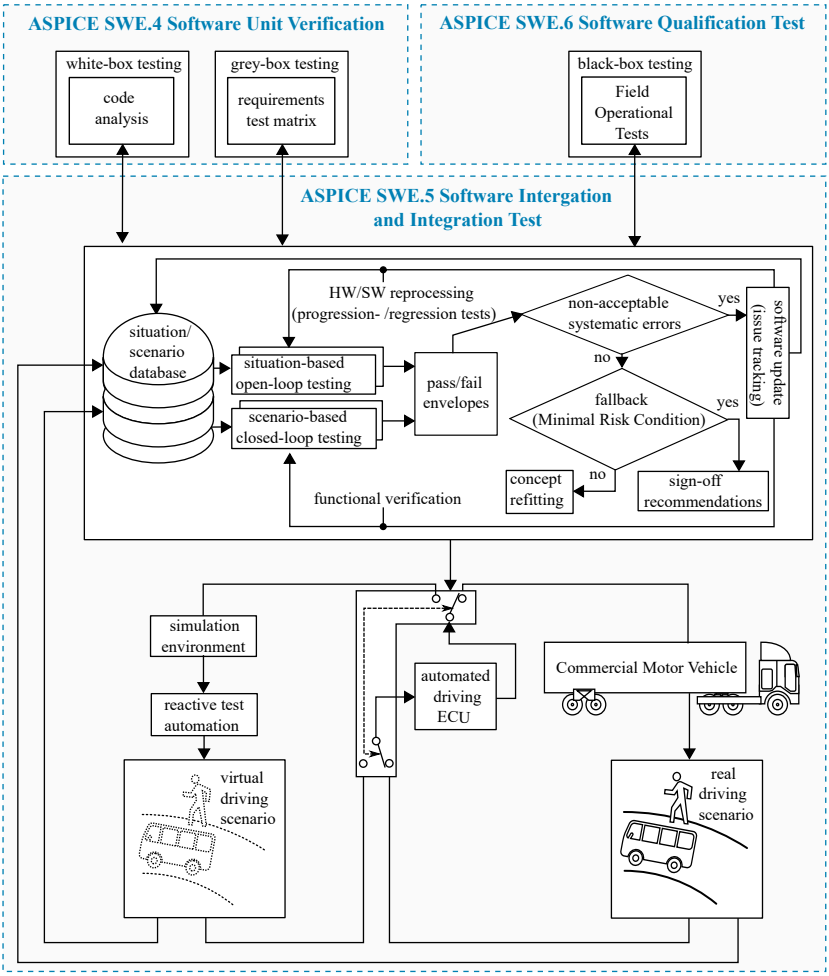

Figure 1.3: Development and testing of ADFs using a trade-off between the efficiency and effectiveness criteria of a context-driven test concept [ESS+19c].

The safeguarding process of the ADFs takes place in three stages in compliance with Automotive Software Process Improvement and Capability dEtermination (ASPICE) [VDA17, fS+19a]. In the first stage, there are two parts, namely white-box testing and grey-box testing according to the process of ASPICE Software Engineering Process Group (SWE).4 software unit verification. The white-box testing requires knowledge of internal software structures for functional and non-functional tests. The grey-box testing is the typical combination of the black-box testing and the white-box testing assessments that verify the functions of various components, systems or sub-systems to ensure maximized test coverage.

The second stage is the database enrichment for open-loop and closed-loop tests according to the process of ASPICE SWE.5 software integration and integration test. The data of situation/scenario database goes through the open-loop test as well as the closed-loop test depending on the type of data extracted [P+17a]. The ADF shall be adequately safe in the event of unintended reactions that could violate the safety goals. For this reason, the driver's controllability or the Minimal Risk Condition (MRC) fallback shall confirm the probability of overcoming the system limits and failures. Also, the regression and progression testing ensure that the previously tested software remains at the same performance level, even if it is modified or combined with other software. Regression testing ensures that the maturity level of the software is retained while adding new logic and fixing issues. Meanwhile, progression testing tracks progress against specific ODD features. The progression tests serve as a benchmark against which developers can measure their progress. Based on the Minimum Viable Product (MVP) levels, the result of this stage decides whether more simulation tests are required or more test kilometers are needed. The MVP defines a version of a product with features to be usable by early customers who can then provide feedback for future product development.

According to the process of ASPICE SWE.6 software qualification test, the development of ADFs is carried out using field based-observation. In the black-box testing, the data from various on-road tests, which are both functional as well as non-functional, is extracted. Depending on these test results, data-driven development provides the decision as to whether type approval is recommended or not [P+17b].

The presented framework uses the standard Quality Gates (QGs)[10] for development of ADFs within an agile development process. Accordingly, the proposed concept aims to bridge the gap between knowledge-based and data-driven test approaches to enable continuous extensibility of experience in an adaptive test coverage manner. The final sign-off is carried out at the end of the development process to ensure that the system meets the specified and intended requirements. This thesis describes a risk-based framework that provides a test process of sensing, perception, prediction and planning and motion control software modules.

Definition 1.7 (**Sensing**)**:** A software module indicates the ability of an ADF to receive adequate information from the vehicle's internal and external environment through connected sensors [Ste16].

Definition 1.8 (**Perception**)**:** A software module denotes the ability of an ADF to interpret information about its environment obtained through its sensors [Ber19].

Definition 1.9 (**Planning**)**:** A software module defines the ability of an ADF to establish and navigate the route it will take on the way to its destination [SHE+17].

Definition 1.10 (**Motion Control**)**:** A software module characterizes the system's ability to execute the driving functions necessary to carry out a continuously updated driving plan by delivering appropriate control inputs such as steering and braking [AKM17].

The modular framework represents a virtual testing for verifying ADFs on the ECU in the laboratory. The test bench offers an efficient compromise between the requirements of simulation realism and the real-time performance of the simulation environment. In this scheme, the real-world testing includes hierarchical clustering of recorded time-series signals to identify and assign the necessary test cases for different appropriate test environments. In addition, the structure employed utilizes a back-end database that is filled with catalogs of extracted driving scenarios from field-based observations.

[10] The QG indicates a special milestone that highlights the qualitative aspect of the deliverables at a defined point in a project.

Using an ontology-based method, a category of adequate and relevant logical scenarios for existing field tests is extracted. A semantic representation of concrete scenarios can be obtained using data mining techniques, and systematically processed into executable requirements for ODD coverage. These new test cases then complement the existing test cases, which were developed from expert knowledge with NL based statements, in an adaptive test coverage manner. Moreover, the extracted scenarios and their parameter space define the probability of occurrence of each extracted logical scenario and the probability distributions of the associated parameters. The stochastic analysis methods employ surrogate models and sampling methods to calculate the probability of failure for each clustered logical scenario [Tar12]. The surrogate model refers to an interpretable model that is trained using input-output data to approximate the predictions of a black-box model [M+17]. The proposed procedure therefore offers an optimized test strategy for extending the requirement-based test coverage in an adaptive manner.

1.5 Structure of the Thesis

The thesis is structured based on the flow required to support the above-mentioned research objectives. In this chapter, the thesis and its contributions are outlined. Chapter 2 discusses the safety assessment concepts of ADFs for long-distance CMVs. Furthermore, the ASE processes are reviewed. The fundamental concepts of software fault tolerance, which take deficiencies in environmental perception and their management by multi-sensor data fusion and fail-operational architectures into account, are presented [Tun19]. In addition, the requirements for a measurable safety framework for different levels of automated driving are described. Chapter 3 presents an overview of the current state of scientific and technical knowledge on the test methods used in the work, which leads to all subsequent chapters. In addition, the challenges of test procedures for ADFs are discussed with respect to current research projects and standardization activities.

Chapter 4 presents the integration of traffic simulation and truck dynamics into a co-simulation configuration based on trajectory-based control of road users. The sensor systems for the perception of the vehicle environment are simulated and evaluated. Thereafter, a round-trip latency in the distributed heterogeneous co-simulation environment is estimated.

Chapter 5 illustrates various measurement methods for scenario mining. The data handling process of the in-vehicle data logging system and its main elements are explained with regard to event-based time-series analysis. Subsequently, various methods of cluster analysis are discussed based on case studies of customer-oriented testing.

Chapter 6 proposes the systematic process to complement virtual testing by extracting insights from field testing database using sensitivity analysis. In this thesis, the prototypical safety margin is the minimum Time To Collision (TTC). Thereby, the sensitivity analysis employs meta-model techniques, where the parameters extracted from cluster analysis are considered as input and the safety margins as output. In addition, the integration of ontology-based test scenario synthesis enables a systematic scenario enlargement. Chapter 7 concludes the framework implementation with the use of reliability analysis to estimate the probability of exceeding the safety margin. Accordingly, the reliability analysis explores the unsafe region in the parameter space to predict the failure probability for each logical scenario using sampling methods. Furthermore, some conclusions on the overall approach, limitations and recommendations for further possible developments are described in chapter 8 of this work. The mentioned flow is portrayed in figure 1.4.

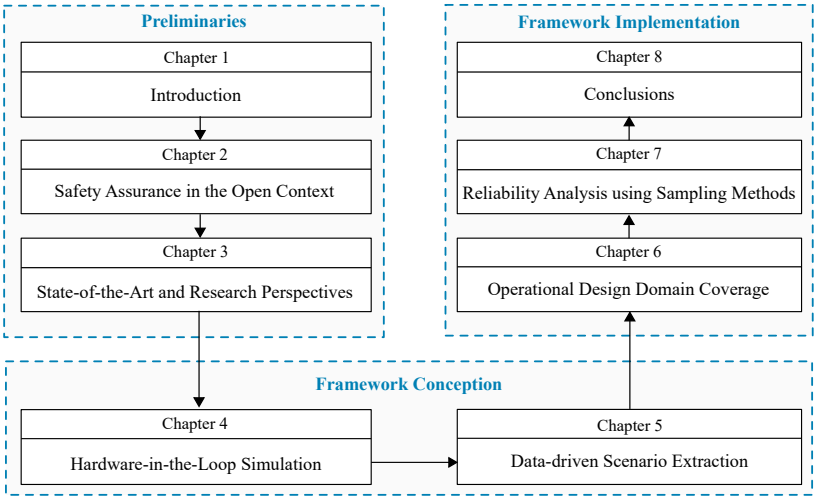

Figure 1.4: Comprehensive overview of the thesis research process with three key parts (preliminaries, framework conception and implementation).

2 Safety Assurance in the Open Context

The active and passive safety systems in the automotive industry differ in their assessment methods. A standard evaluation approach for the passive safety systems has been developed to assess the behavior in an appropriate number of crash test cases under certain critical conditions [Win16a]. Considering the active safety systems, there are many challenges to be overcome in safety Verification and Validation (V&V) under real traffic conditions, namely the diversity of scenarios and environmental conditions, system complexity, and functional deficiencies [HPT10]. Despite strong support from the industries and academia, questions are often raised about the business cases, ethical dilemmas, legal liability and safety regarding automated driving. The SAE J3016 automation levels are therefore expected to overlap and not be available on the market back to back. Due to a complex, uncertain and unpredictable traffic environment, the motion control and path planning algorithms have to deal with uncertainties in measurements and predictions [AB15]. Therefore, the automated driving relies on intelligent algorithms that receive input from a variety of environmental perception sensors to control real-time actions in a highly safety-critical context [ESW+16].

2.1 Definition of Terms and Categories

This section gives an overview of the basic terminologies and highlights the meaning of the basic terms used in this thesis. The essential phrases are categorized thematically to clarify the scope of the work.

2.1.1 Scene, Situation, Scenario and Test Case

The interaction of ADFs, in conjunction with the traffic environment, increases the complexity of functional testing [M+18]. Thereby, the use case specifies the application, its desired behavior and its functional system boundaries. The use case description typically does not include a detailed list of all relevant scenarios for this use case. Instead, a rather abstract description of the deployed scenarios is used.

Definition 2.1 (**Scene):** A snapshot of an environment with a certain state containing the scenery, dynamic elements, all actors and their relationships. The scenery includes all spatial stationary elements such as lane markings, traffic signs, obstacles and traffic lights [S+18a].

Definition 2.2 (**Situation):** A selection of behavioral patterns for a specific triggering event. The behavior reflects the interaction of the ADF with the environment. While the behavioral patterns deal with the allocation of responsibilities between the Ego-vehicle and its surroundings, the triggering event induces a traffic situation with certain conditions and subsequent system reactions [HSSB14].

Definition 2.3 (**Scenario):** A historical development between several scenes in a chronological sequence of situations leading to a potentially hazardous consequence [BLO+17].

Definition 2.4 (**Test Case):** A specification of the inputs, execution conditions, testing procedure and expected results that aims to prove a particular property of a test object. Therefore, the test case contains a logical scenario with a set of parameters to determine whether a function operates according to its intended functionality [L+18a].

2.1.2 Functional, Logical and Concrete Scenarios

The layered structure of the automated driving system can be described as a cyber-physical system with a sensor system, an automated driving ECU and an actuator system, whereby the responsibility between the driver and the ADF is classified according to the SAE J3016 [Int18].

Both the road and environment constraints consist of four layers of scenario description structured as follows:

- Layer 1: road geometry (e.g. road curvature, lane marking, etc.),
- Layer 2: static objects (e.g. speed limits, construction barriers, etc.),
- Layer 3: dynamic objects (e.g. cars, pedestrians, trucks, etc.) and
- Layer 4: weather conditions (e.g. fog, rain, snow, etc.).

The scenarios can be classified into three levels of abstraction. One is functional, the second is logical and the third one is concrete. The functional scenarios specify the application, the desired behavior and the functional system boundaries. The description of the functional scenarios doesn't typically contain a detailed list of all relevant scenarios. Therefore, the functional scenarios illustrate the most abstract level of the scenario representations as high-level requirements with a textual or graphical description. The entities and their relationships are represented within the functional scenarios in NL statements. The logical scenarios represent precise specifications based on the parameter spaces in the stated space and contain a formal scenario description [BOS17]. The concrete scenarios define test cases and describe the concrete representation of a logical scenario with predefined values in the stated space [Web19].

2.1.3 Structure-based, Open-loop and Closed-loop Testing

The ADF contains software components that interact with an unstructured, public real-world environment to support and automate DDTs. Figure 2.1 depicts the differences between structure-based, situation-based open-loop and scenario-based closed-loop testing in the time sequence. The particular function provides the in-/output behavior or the reaction to sensor input variables in the automated driving system [SBWE19].

Definition 2.5 (**Structure-based Testing):** It is performed to determine the code coverage of software structure components such as specific software functions and code parts [Wil16].

Definition 2.6 (**Situation-based Open-loop Testing**)**:** It generates driving situations from the required behavior and evaluates the function reaction without referring to future conditions [U+15].

Definition 2.7 (**Scenario-based Closed-loop Testing**)**:** It is used to test the behavior in a closed loop within a traffic sequence of scenes associated with actions of the Ego-vehicle, events from the environment and goals for the ADF [FHW16].

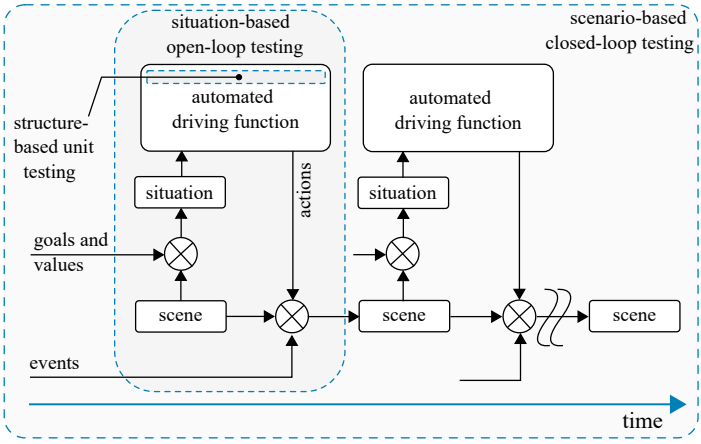

Figure 2.1: Illustration of the logical relationships between structure-based, situation-based open-loop and scenario-based closed-loop testing [ESO+19].

2.1.4 Operational Design Domain and Fallback

The fallback method provides a MRC to reduce the risk of a collision if a particular route cannot be completed within the ODD.

Definition 2.8 (ODD)**:** According to SAE J3016, the ODD refers to the operating conditions under which a given driving automation system or feature thereof is specifically designed to function, including environmental, geographical and time-of-day restrictions, and/or the requisite presence or absence of certain traffic or roadway characteristics [Cza18].

Definition 2.9 (MRC)**:** The MRC allows a driver or an ADF to switch to a low-risk operating condition using a DDT fallback. In case of a MRC, various fault tolerance strategies can be implemented to avoid a hazard to the system, so that the system continues with fail-safe, fail-degraded or fail-operational characteristic [CFHL07].

The fail-safe behavior characterizes the transition of the system to a safe state despite the presence of hardware or software failures. The fail-degraded behavior provides the safe-degraded property to run an intended degraded safe operation. In this context, degradation is defined as the reduced performance of the function which can still provide safe operation in the presence of hazardous events [Xi08]. The fail-operational behavior describes the ability to continue normal operation through redundancy of the components so that the loss of safety-related functions does not lead to a hazard [WRM+19].

2.1.5 Verification, Validation and Accreditation

In accordance with the Institute of Electrical and Electronics Engineers (IEEE) Std 610-1990, software verification denotes the process of evaluating software to determine whether the products of a given development phase satisfy the conditions imposed at the start of that phase. Meanwhilst, software validation describes the process of evaluating software during or at the end of the development process to determine whether it meets specified requirements. The ISO/ International Electrotechnical Commission (IEC)/IEEE 15288 describes a framework for characterizing the life cycle of systems and software engineering. The testing process requires three interrelated but distinct procedures, namely Verification, Validation and Accreditation (VVA), in order to develop a realization for a purpose to be fulfilled in a specific context [JS+19b].

Definition 2.10 (**Functional Verification**)**:** According to the ISO/IEC/IEEE 15288:2015, the functional verification refers to a procedure of confirmation by objective evidence that the specified requirements have been met [fS+15].

Definition 2.11 (**Functional Validation**)**:** It concerns a confirmation procedure by objectively demonstrating that the requirements of a specific intended use or application have been fulfilled [P+17b].

Definition 2.12 (Functional Accreditation): It relates to a confirmatory procedure by objective evidence that the function is acceptable for use in a specific purpose [fS+15].

These procedures collect and evaluate evidence to credibly determine whether an ADF can be safely used in a real-world ODD. Thereby, the deductive gap between required, specified and implemented behaviors refers to the presence of invalid hypotheses on different levels of abstraction that cause unintended functionality. Figure 2.2 represents three sets of $\mathcal{S}$, $\mathcal{M}$, and $\mathcal{N}$ in the set diagram, which create several overlapping areas when they intersect. The verification refers to a procedure that proves the correct implementation of each individual requirement to minimize the intersection area, $\mathcal{K}$, so it is the set $(\mathcal{M} \cap \mathcal{S} \cap \bar{\mathcal{N}})$. The validation is a procedure that compares the results with observed empirical data to confirm the correctness of the requirements to minimize the intersection area, $\mathcal{J}$, so it is the set $(\mathcal{M} \cap \overline{(\mathcal{S} \cup \mathcal{N})})$. The accreditation is aimed at minimizing the areas $\mathcal{A}$ and $\mathcal{D}$, so they are the sets $(\mathcal{N} \cap \overline{(\mathcal{M} \cup \mathcal{S})})$ and $(\mathcal{N} \cap \mathcal{S} \cap \bar{\mathcal{M}})$, respectively. The adaptive functional testing aims to maximize the optimized behavior with the intersection area, $\mathcal{C}$, so it is the set $(\mathcal{S} \cap \mathcal{M} \cap \mathcal{N})$ and thus minimizes the deductive gap. If the deductive gaps are less than the reasonable risk, the continuous test process can be terminated according to the optimization goals by sign-off recommendations of ADFs. The areas $\mathcal{B}$ and $\mathcal{L}$ are not critical where they have no safety-related impacts, so they are the sets $(\mathcal{N} \cap \mathcal{M} \cap \bar{\mathcal{S}})$ and $(\mathcal{S} \cap \overline{(\mathcal{M} \cup \mathcal{N})})$, respectively [ESFG19a].

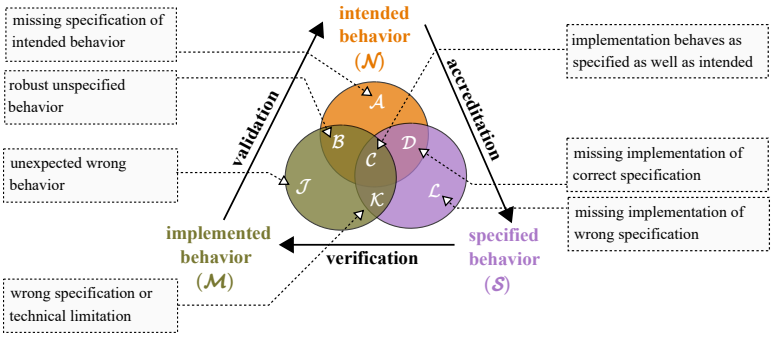

Figure 2.2: Three-circles model of the VVA challenges due to the deductive gap between required, specified, and implemented behaviors [SBP+19].

2.1.6 Top-Down, Bottom-Up and Middle-Out Approaches

Software engineering is the systematic application of methods, principles and techniques to develop software-based systems in a consistent manner. Consequently, the top-down and bottom-up approaches refer to the design philosophies that are executed either from the top or the bottom for the requirements engineering [SS06]. In the software development process, the top-down approach requires a detailed design perspective of the system before the actual implementation can begin. In contrast, the bottom-up approach emphasizes developing the system by integrating the designed components. The middle-out approach is a mixture of the top-down and bottom-up approaches and offers advantages over these when evaluating automated driving applications. The hardware and software development process is divided into several layers according to a V-shaped model (component, subsystem, system and vehicle), as shown in figure 2.3.

Definition 2.13 (V**-shaped model):** It is a software development process model that combines the requirements and design on the left side with V&V on the right side. While the V-shaped model has established itself for series development in the automotive industry, it provides a procedure of abstraction and does not necessarily describe the test methods directly [Bac18].

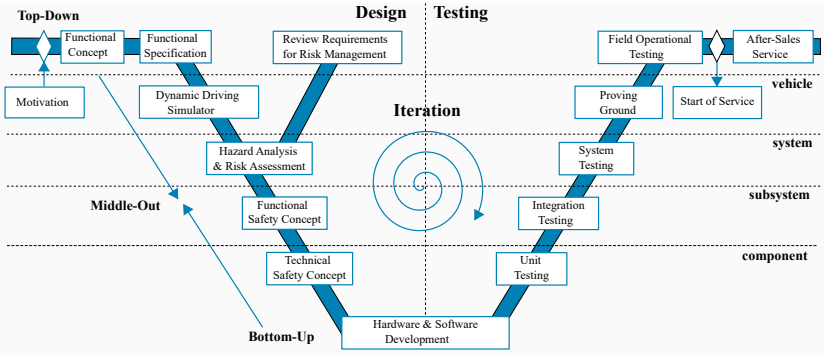

Figure 2.3: Adaption of the safety life cycle according to the ASPICE process model as a development process for ADFs [Sax08].

The components consist of hardware as well as of software elements on the bottom layer. The subsystem therefore includes more than one component. Thus, the system can be divided into sub-systems with a hierarchical structure. The vehicle is comprised of one or more systems, whereby one system consists of at least one sensor, one processing unit and one actuator. A system implements one or more functions, but a function can also be applied in several systems.

2.1.7 Black-Box, White-Box and Grey-Box Testing

The closed-loop testing uses XiL techniques for functional verification of software in synthetic simulation environments. For example, Hardware-in-the-Loop (HiL) platforms describe the functional verification of the embedded software integrated on the target ECU via technical interfaces. At the same time, open-loop recomputing performs logged data simulations and repeatable regression tests with many iterations to achieve functional improvements and parameter calibrations using recorded real-world data. The logged data simulations are based on sensor data-sets collected from physical test drives to select the most appropriate release version. The extensive data-sets are collected worldwide under realistic driving conditions with CMV fleets. The evaluation of algorithms requires an efficient search and interpretation of relevant traffic situations with the help of Root Cause Analysis (RCA) to modify the algorithm accordingly.

The automated driving ECU shall be sufficiently safe in the event of unintended reactions that could violate the safety goals. For this reason, the driver's controllability or the fail-operational modes must confirm the probability of overcoming the system's limits and failures. Moreover, the regression testing ensures that the previously tested software remains at the same performance level even if it is modified or combined with other software. A correlation between the various test coverage criteria intends to support the controlling whether to collect more kilometers or to carry out more simulations. To ensure that the system meets the specified requirements, a sign-off is carried out at the end of the development process [Lat08].

Definition 2.14 (Black-box Testing): It involves functional and non-functional tests with system description regardless of the component or system internal structures, such as on-road vehicle testing [Sax08].

Definition 2.15 (White-box Testing): It requires knowledge of internal software structures for functional and non-functional tests. The code coverage then defines the parts of the software that are executed and those that are not, such as the Modified Condition/Decision Coverage (MC/DC) [Wil15].

Definition 2.16 (Grey-box Testing): It is a mixture of white-box and black-box assessments that verify the functional specifications of a component, subsystem, or system to ensure maximized test coverage [ESO+19].

2.1.8 Software Development Process Models

The software process is a set of activities for specifying, designing, implementing and testing software systems. The software process model is an abstract representation of a process that presents a description of a process from particular perspective [P+19b]. The Software Development Life Cycle (SDLC) models specify the various stages of the process and the order in which they are carried out. Thereby, the SDLC models can essentially be divided into the following model types (Waterfall, V-shaped, Incremental, Spiral and Agile). The Waterfall model is a breakdown of project activities into sequential linear phases, where each phase depends on the deliverables of the previous one and corresponds to a specialization of tasks. In the V-shaped model, the relationships between each phase of the SDLC and its associated phase of testing are represented. The Incremental model is a method of software development in which the model is designed, implemented and tested incrementally until the product is finalized. The Spiral model is a risk-driven SDLC model in which the project is executed in loops. Each loop of the Spiral is referred to as a phase of the SDLC. In the Agile model, iterative development is used where iterations and continuous feedback are deployed to refine and deliver a software system.

2.1.9 Requirements Specification Notations

Requirements engineering refers to the process of eliciting, specifying and evaluating the desired behavior of a software-intensive system. The functional requirements form the backbone of a comprehensive technical understanding of the developed system.

Requirements as such, therefore, need to be unambiguous and understandable to allow an external testing organization to perform independent tests of the system. The NL-based requirements can be engineered in a straight-forward way without explicit knowledge of the syntax. Meanwhile, the model-based requirements facilitate the clarity of a complex software product and enable a simplified representation of the system with diagrams and axioms. The approaches of requirements management can essentially be divided into five notation types by using natural language and model-based notations as follows:

- Ad-hoc notation in NL is a free writing style that offers a high degree of flexibility in specifying requirements, but at the same time leads to ambiguity and a lack of expressiveness, completeness and consistency.
- Structured NL notation has less potential for ambiguity than an ad-hoc informal approach by requiring the use of NL in a structured format [CI14].
- Ontology-based NL notation uses a formal description language to describe an ontological knowledge base with a glossary of terms and relationships specified by a set of rules [STZ+11]. The ontology-based NL notation is a semantic human- and machine-understandable representation of knowledge terms and their inference rules. The ontologies support the verification of the consistency and completeness of the requirements.
- Graphical modeling notation, based on finite state machines and sets using Unified Modeling Language (UML), represents use cases and/or sequence diagrams. The graphical model-based notation facilitates the clarity of a complex software product and enables a simplified representation of the system with diagrams or axioms. In spite of improved human readability, it may not be suitable for large systems where the graphical modeling notations are not easily maintainable.
- Mathematical notation provides a formal machine-readable format based on a set theory with Z notation specification [Bow01]. The Z notation is a formal specification language used for describing and modeling functional specifications. Therefore, the Z notation is a mixture of formal mathematical statements and informal text. The formal mathematical statements give a precise description of the system.

The informal text describes the meaning of the mathematical statements in NL to make the specification more readable. Large systems with a complex domain may not be easily specified in the Z notation, where formal specification is probably not readable and understandable to the client.

Table 2.1 refers to a comparable evaluation where the scale ranges in requirement notations from poor to optimal. The development of functional requirements presents a joint process between the client and the contractor, in which the technical knowledge of the client and the software development competence of the contractor become accessible. Siegemund et al. [STZ+11] introduced a meta model for ontology-driven goal-oriented requirements engineering. In his dissertation, Siegemund [Sie14] derives the following quality criteria, which serve to analyze the requirements specification notations.

Quality criteria	**Requirements specification notations**				
	Ad-hoc NL notation	Structured NL notation	Ontology-based NL notation	Graphical modeling notation	Mathematical notation
Human readability	●	●	◒	◒	○
Expressiveness	○	◒	●	●	●
Completeness	○	◒	●	◒	●
Traceability	○	◒	●	◒	●
Consistency	○	●	●	◒	○
Verifiability	○	●	●	◒	●
Maintainability	○	◒	●	◒	○
Usability	○	◒	◒	◒	○

● : optimal ◒ : fairly optimal
◒ : natural ◒ : fairly poor
○ : poor

Table 2.1: Comparable evaluation of notation types in functional requirements engineering [ESO+19].

2.1.10 Software Process Assessments

In safety-critical automotive software engineering, some quality process standards apply to organizational processes to achieve a high degree of software quality. ASPICE is a process reference model developed by organizations within the automotive industry to create a more automotive-focused reference model compared to ASPICE or Capability Maturity Model Integration (CMMI) [Bär08]. There are two dimensions in ASPICE: a process dimension and a maturity level dimension according to ISO/IEC 33001:2015.

The ISO/IEC 33001 contains a set of technical standards for process evaluation to assess the achievement of process quality characteristics. The process dimension includes various software development processes. The second dimension allows an averaged assessment of the maturity of a single process on a scale of 0 to 5, as illustrated in figure 2.4. The maturity level is determined by certified assessors who perform a comparison based on the processes defined in ASPICE [Win13]. The ASPICE enables the assessment and classification of their own process in relation to the state-of-the-art. Furthermore, the suppliers can be selected in a qualified manner and the improvement potential can be identified [BOS19].

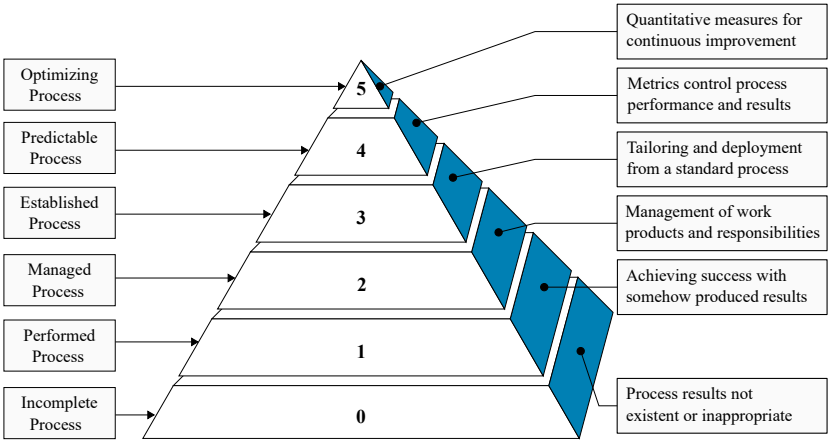

Figure 2.4: ASPICE capability levels according to ISO/IEC 33001:2015 [Sch16a].

2.1.11 Safety, Reliability and Availability

The safety requirements are intended to ensure the absence of unreasonable risks at each stage of the development process of safety-critical automotive systems [IEC+10]. Consequently, operational safety includes passive and active safety technologies to minimize the occurrence and consequences of traffic conditions. For this purpose, operational safety is divided into three main aspects: Functional safety, Safety Of the Intended Functionality (SOTIF) and behavioral safety.

The functional safety describes the probability that a function does not go into an unsafe condition if an independent event may cause an accident. The functional safety focuses on system design to identify hazards using Hazard Analysis and Risk Assessment (HARA) and to alleviate the consequences of E/E malfunctions that may occur in the components of an automated driving system. However, the goal of SOTIF is to validate the ADFs, including perception and decision making, in all the relevant environmental scenarios. Behavioral safety, meanwhile, focuses on the system design to behave safely in its environment to avoid hazards and reduce the risk of failures.

Definition 2.17 (**Reliability**)**:** It can be defined as the probability that a function fulfills its intended functionality in an ODD and over a certain period of time. The reliability requirements shall ensure that the system does not reach an unreasonable number of failures [Mul85]. Therefore, the reliability can be defined by the Mean Time Between Failures (MTBF) and the failure rate. While the MTBF describes the expected time between two failures of system components, the failure rate, in contrast, indicates the frequency with which a system component fails, in terms of failures per unit time [BA92].

Definition 2.18 (**Availability**)**:** It is the probability that a system is available for operation at a certain point in time, i.e. the time period during which a system is actually in operation as a percentage of the total time that the system should be in operation. In parallel, robustness defines the degree to which a system or subsystem or component can operate correctly under invalid inputs or stressful environmental conditions [IEC+90].

2.1.12 Fault, Error and Failure

The error propagation model follows the subsequent steps: fault, error, failure, hazard and accident. The active errors are error classes that are caused by faults and cause failures, but the latent errors are error classes that are caused by faults and don't cause failures [KM05].

Definition 2.19 (Fault): An abnormal condition or defect that can cause an element or an item to fail. The faults occur as logical, Type I or Type II errors, which either lead to active errors or remain in the test object as latent errors [P+01].

The Type I error, also called False Positive (FP) event, is an error type that occurs as a false alarm if the null hypothesis (H_0) is true, although the given condition doesn't exist. While the Type II error, also called False Negative (FN) event, is an error type that occurs as missed detection if the null hypothesis (H_0) is false but the given condition erroneously fails to be recognized.

Definition 2.20 (Error): A discrepancy from the intended design, which can lead to an unintended condition at the test object boundary. After that, a failure depicts the deviation from specified behavior due to undetected errors within the component, subsystem or system [SWB+20].

The occurrence of Type I and Type II errors summarizes the uncertainties resulting from the restriction of environmental perception and the variability of predictive situations. The Type I errors can provoke a new critical situation for endangering road safety. The Type II errors lead to a loss of the safety benefit concerning the system specification. However, each Type I or Type II detection error does not change the assessment of a situation as safe or hazardous. Therefore, the error criterion defines the formulation of the safety envelope that determines whether a situation is safe or hazardous. In addition, a safety-critical ghost or miss can be caused by measurement errors, e.g. a noisy distance measurement, or by detection errors, e.g. missing detection. For each detection or measurement error, there are two types of error for each signal value. The first describes the systematic error and the second the statistical error. The systematic error represents an error in the algorithm according to knowledge-based design. In contrast, the statistical error represents an error that occurs during operation due to environmental influences.

Safety-critical ghosts represent the rate of situations that are erroneously considered hazardous, where ghosts can occur at any time. Safety-critical misses represent the situations that are erroneously considered safe, where safety-critical misses can only occur when a hazardous situation exists [SWB+20].

2.1.13 Microscopic and Macroscopic Risk Metrics

The risk of the system can be estimated using a statistical approach to the frequency of accidents. Hazards at the system level are the physical situations that may cause an accident. The accident is an unplanned or undesirable event that leads to an unrecoverable loss of service that typically brings loss and/or injury. Due to the long distance between two road traffic accidents, the macroscopic risk cannot be estimated without large field data. Therefore, an enormous mileage is required to collect enough accident data for a meaningful statistical analysis [JSW19]. Microscopic risk refers to the risk for single vehicle type or fleet of identical vehicles, e.g. a prospective risk of an ADF using Time To React (TTR) metrics. The macroscopic risk represents the average risk in road traffic within a fleet of different vehicles, e.g. the occurrence rate of fatal accidents. If the investigated event is a critical scenario rather than an accident, the accident frequency rate increases and therefore less mileage is required for the equivalent statistical significance. Junietz [Jun19] applies microscopic risk metrics to evaluate critical scenarios and uses extreme value theory to extrapolate frequency of accidents using macroscopic risk metrics. The extrapolation is assumed with a hypothesis that is applied to ADFs to estimate the probability of accidents using macroscopic risk assessments.

2.1.14 Sensitivity, Specificity and Precision

The intended events contribute towards improving the safety of the ADFs. The True Positive (TP) event represents an intended reaction in which the system responds correctly to a critical situation. Similarly, the True Negative (TN) event refers to an intended reaction in which the system reacts correctly to a non-critical situation. The unintended events, in contrast, have different consequences: The FP event, also called Type I error, refers to an unintended reaction in which the system reacts incorrectly to a non-critical situation [Ste16]. As a result, the safety benefit is lost in this critical situation with regard to the system specifications.

The FN event, also called Type II error, means an unintended reaction that is not associated with a critical situation but can provoke a new critical situation [Hel14].

Definition 2.21 (Sensitivity): It evaluates how good the test is at detecting a critical situation and indicates the conditional probability that the system performs intended reactions to critical situations according to the system specifications.

The sensitivity, also called TP-rate, is calculated by dividing the number of TP events by the total number of TP and FN events, as defined in Equation 2.1. High sensitivity then refers to a low number of FN events [Ber16].

$$\text{sensitivity} = P(\text{positive reactions}|\text{critical situations}) = \frac{\text{TP}}{\text{TP} + \text{FN}} \tag{2.1}$$

Definition 2.22 (Specificity): It estimates how likely non-critical situations can be correctly ruled out and describes the conditional probability for the proportion of non-critical situations treated by the system.

The specificity, also called TN-rate, is calculated by dividing the number of TN events by the total number of TN and FP events, as defined in Equation 2.2. High specificity then refers to a low number of FP events.

$$\text{specificity} = P(\text{negative reactions}|\text{non-critical situations}) = \frac{\text{TN}}{\text{TN} + \text{FP}} \tag{2.2}$$

The Receiver Operating Characteristic (ROC) curve is a graphical representation of the relationship between sensitivity and specificity as performance measures for model selection. According to the German Insitute for Standardization (DIN) ISO 5725-1:1997-11, precision is described by the standard deviation of the measured signal. In contrast, the accuracy is used to describe the signal bias to the ground truth value. The precision, also called Positive Predictive Value (PPV), is calculated by dividing the number of TP events by the total number of TP and FP events, as described in Equation 2.3.

$$\text{precision} = \frac{\text{TP}}{\text{TP} + \text{FP}} = 1 - \frac{\text{FP}}{\text{TP} + \text{FP}} \tag{2.3}$$

2.1.15 Supervised, Unsupervised and Reinforcement Learning

Machine learning approaches rely on computational statistics to make accurate predictions. In general, there are three categories of machine learning approaches: supervised learning, unsupervised learning and reinforcement learning.

Definition 2.23 (**Supervised Learning):** It focuses on approximating functions from a labeled data-set that can be used for classification and regression, i.e. Support Vector Machine (SVM), Naive Bayes, Logistic Regression, Decision Trees, Random Forest and K nearest neighbors. The supervised learning approaches utilize inductive learning, in which a run-time algorithm uses the results of a learning process to perform algorithmic operations [K+19b].

Therefore, machine learning algorithms are evaluated to avoid systematic failures during the training and validation process. Their specifications are therefore not in the V-shaped model of a set of functional requirements for the system itself, but rather in a set of training data or a plan for capturing the set of training data [KW16]. If the machine learning process is not trained with a certain driving situation, the algorithm cannot recognize this situation and therefore the respective situation is considered as a Black Swan[1] situation. Therefore, the verification philosophy should consider Black Swan situations by robustness testing and run-time fault injection techniques. Thereby, the data collection process needs to reduce hazards such as unintended bias or distortion in the gathered data. The offline machine learning process starts with the acquisition of a training database using a collection of data from the on-road and proving ground tests. The data is then annotated to learn specific features (e.g. cars, pedestrians, trucks, etc.) and the scene labeling is used for the determination of parameters through training. The process of scene labeling is implemented by introducing artificial markers into a scene with the purpose of marking positions in a three-dimensional space.

[1] Black Swan refers to a surprising event that is not observed by the system. Behind this term lies the story that in medieval Europe, people believed that there were no swans except white ones. After the discovery of the Black Swans in Australia, this surprising event broke the previous thinking [Tal10]. Therefore, the Black Swan is a metaphor that describes an event that comes as a surprise.

Concerning self-verification, the training result is then verified with the training data on the basis of acceptance criteria, such as acceptable Type I and Type II error rates. The acceptance criteria refer to conditions that are defined to determine whether the test object can be delivered to the respective next test level. If the self-verification fails, the process can be restarted after collecting the additional data for the test abort criteria. The abort criteria relate to conditions under which the tests are terminated prematurely as continuation is not advisable. The self-assessment could be insufficient because it is difficult to ensure that the learning system has been trained for the essential features of training data rather than coincidental correlations.

The coincidental correlation relates to a non-caused correlation of self-verification assessment criteria with training and validation data, which are consistent but are not caused by each other. Therefore, cross-verification is used to verify the learning process using a separately collected and annotated database with acceptable pass and fail criteria. The adaptive machine learning systems that change weights at run-time are relevant for the high and full automated driving systems and go beyond the scope of the annotation and labeling process.

Definition 2.24 (**Unsupervised Learning):** It focuses on discovering pattern in unlabeled data-sets with a minimum human supervision, i.e. hierarchical clustering, k-means, Density-Based Spatial Clustering of Applications with Noise (DBSCAN) and Principal Component Analysis (PCA) [KWB18].

Definition 2.25 (**Reinforcement Learning):** It focuses on learning interaction from trial and error to take a decision in an environment based on maximizing the cumulative rewards, i.e. Markov Decision Process.

The uncertainty quantification can provide information that is employed in object plausibility within sensor fusion algorithms [Min17]. There are two types of uncertainty that can be distinguished (Aleatoric and Epistemic uncertainties) [GB12]. Moreover, Type I errors (e.g. ghost objects, etc.), Type II errors (e.g. not detected objects, etc.) and misclassification issues can be tuned by the coverage of the training data-set. The performance of machine learning algorithms relies on the amount of training data-sets. The statistically relevant spread of operational situations can ensure adequate coverage during training for environmental perception tasks.

Definition 2.26 (**Aleatoric Uncertainty**)**:** It comprises noises that are inherent to the observation (e.g. sensor or motion noises) [ODR+02]. Since the Aleatoric uncertainty refers to the self-noise of the observation, this uncertainty cannot be reduced by increasing the training data [R+20]. Therefore, Aleatoric uncertainty captures the observation noises that are inherent in the sensor systems, while the detection of a distant object may result in high Aleatoric uncertainty [VR15]. The Aleatoric uncertainty may lead to Type II errors and can be resolved by the complementarity of the environmental perception sensor systems [FRD18].

Definition 2.27 (**Epistemic Uncertainty**)**:** It has the effect that the system operates inconsistently for a given input class within a certain error range [VR15]. Therefore, the Epistemic uncertainty indicates that the model hypotheses may not adequately reflect reality for the intended functionality [HALZ19]. The Epistemic uncertainty or model uncertainty indicates how uncertain an object detector is to explain the object from the observed training data-set. The Epistemic uncertainty may lead to Type I errors and can be addressed by more training data. For example, the detection of an abnormal object, different from the training data-set, may result in high Epistemic uncertainty [SKOK18].

2.1.16 Known, Unforeseeable and Unknowable Black Swans

Safety reflects a result of the absence of unreasonable risks. Beyond the challenges of complying with accepted safety engineering procedures, a key challenge in the safety validation of automated driving is to behave appropriately in the presence of surprises. Thereby, it can be difficult or impossible for human reviewers to experience these events in advance [KW16]. The Black Swan events comprise three major categories of surprising critical scenarios in relation to present knowledge to verify ADFs, as follows:

- The first type of Black Swan events comprises known critical scenarios that occur despite the fact that the probability of occurrence is judged to be negligible. Thus, acceptable risks shall not be determined exclusively by the probability assessment. The risk is a combination of the probability of occurrence of harm and the severity of that harm.

In addition, the prudence and precautionary principles are fundamental elements of the risk management linked to such Black Swans.

- The second category of Black Swan events is unforeseen scenarios that are not covered by the corresponding risk assessment and management. Dealing with unforeseeable Black Swans demands improved risk assessment as well as knowledge discovery in order to identify unforeseeable surprises and include them in the relevant verification process to complement the present knowledge base.
- Unknowable Black Swan events represent the third category of uncertainty scenarios, which are completely unknown and only become known after the deployment of ADFs. The handling of unknowable critical scenarios demands that the ADF can automatically detect surprises in order to behave robustly and resiliently to uncertainties [WK15].

2.1.17 Safety Integrity Requirements for Automated Driving

The SAE J3016 represents the levels of automation from 0 to 5, whereby SAE L0 means no automation, SAE L1 applies to driver assistance, SAE L2 stands for partially automated driving, SAE L3 for conditional automated driving, SAE L4 for highly automated driving and SAE L5 for fully automated driving [Int18]. The ADF is a combination of both hardware and software that can be equipped with one or more systems. ADFs with an automation level below SAE L3 enhance safety or assist the driver, but are not capable of controlling or operating the truck without active physical control or monitoring of a human driver. ADFs with an automation level higher than SAE L2 are capable of performing the DDTs without active physical control or supervision by a human driver physically present in the truck cab. The safety integrity describes the reliability of a confidence indicator in conjunction with the functional requirements for the various automated driving levels. The safety integrity requirements for the respective automation levels are described, as follows:

- The driver assistance and partially automated driving (SAE L0-L2) handle tasks of limited complexity autonomously in a precisely specified context. These functions perform limited tasks in a defined context and do not learn during operation.

Since the collaboration is a restricted task in a determined context, co-operation is therefore limited to the exchange of information on the system context. The safety integrity of ADAS and semi-automated driving is aimed at ensuring the ISO 26262 HARA [SH19a] and ISO 21448 SOTIF requirements [Hil12].

- The highly automated driving (SAE L3) accomplishes a sequence of tasks, in which every single task is controllable, but the sequence and transitions between them are situation-dependent. While the system is not learning during operation, it optimizes its trajectories during the control process according to the defined objectives such as time or other resources. The co-operation with other systems is therefore limited to the exchange of information about the system context and the system itself. The safety integrity of highly automated driving shall ensure the feedback model from ISO 26262 HARA, SOTIF triggering events, and post-deployment driver experience to deal with the known Black Swan events in fail-operational mode using the cautionary and precautionary risk management principles [DH17].

- The fully automated driving (SAE L4) is able to work together with other systems to perform their task. They negotiate their goals, plans and actions with other systems and adapt their behavior to the negotiated procedure. Since the system boundaries change dynamically due to the collaborative relationship, mechanisms for distributed planning and co-ordination of interpretations are required to ensure safe system functionality. Beyond the need to follow accepted safety engineering practices, the fully automated driving focuses on the safety requirements of ISO 26262 HARA, ISO 21448 SOTIF, ISO/DIS 34502 [fS$^+$22a] and ODD coverage [fS$^+$22b] to ensure reasonable behavior against unforeseeable Black Swans using improved risk assessment and knowledge discovery principles [DH17].

- Full automation (SAE L5) can expand the environmental perception, situational awareness and actions with the ability of unsupervised learning and has some sort of fail-operational autonomy capability. The ANSI/UL 4600 [UL22] safety case and continuous improvement feedback shall be required for the Autopoietic driving to safeguard a possible online expansion.

The Safety Performance Indicators (SPIs) and lifecycle feedback demonstrate the self-recognition and deal with unknowable Black Swans by defining the reasonable behavior. The SPIs are used to assess operational safety performance through monitoring and measure validity of a safety case claim, e.g. violation of a safe clearance limit [KFFW19]. Therefore, the system needs to be good enough to recognize surprises and ensure that the behavior remains relatively modest until the uncertainty is resolved. One of the safety skills is that the human driver has enough self-awareness to recognize an unclear driving situation and to minimize the risk until the uncertainty is eliminated. Therefore, continuous supervision and learning from field observations help in coping with rare and dangerous events. As a result, understanding the system's own limits is essential when dealing with unknowable Black Swans.

SAE J3016 level	**Description**	DDTs				ODD scope
		execution of steering and acceleration/deceleration	monitoring of driving environment	DDT failures	fallback performance of DDT	
0	no driving automation	○	○	○	○	not available
1	driver assistance	◒	○	○	○	limited
2	partially driving automation	●	○	○	○	limited
3	conditional driving automation	●	●	●	○	limited
4	high driving automation	●	●	●	●	limited
5	full driving automation	●	●	●	●	unlimited

○ : human driver
◒ : human driver and/or system
● : system

Table 2.2: ODD scope and role-sharing between the human driver and the system at each automation level, adapted from the SAE J3016 automation levels [Int18], [ESO+19], [DH17].

The ODD scope and DDT fallback requirements for each level of automation, as illustrated in table 2.2 [Smi17]. The DDTs comprise all real-time operational functions required for the operation of a vehicle in road traffic, including environment perception, longitudinal and lateral motion control and maneuver planning.

2.2 Uncertainties in the Environment Perception

Since automated driving depends on the perception of the vehicle's environment, safety violations can be caused by system restrictions due to physical or technical limitations of the intended functionality [Cho16]. Table 2.3 refers to a comparable evaluation with a scale from poor to optimal using the following twelve sensor capability criteria applied to the environment perception of CMVs. The perception and map-based prediction sensors have different measurement principles, which are generally divided into camera, RADAR, Light Detection And Ranging (LiDAR) and electronic Horizon (eHorizon) sensors.

Steinmeyer et al. [S+18b] introduced a method for improved data fusion in an environment detection. In his dissertation, Steinmeyer [Ste14] evaluates the environment perception sensors based on different sensor capability criteria, e.g. maximum longitudinal range, lateral Field of View (FOV), etc. The object recognition and classification tasks are performed by machine learning techniques to extract relevant characteristics in an unstructured operational context. While machine learning paradigms offer a promising perception performance, high levels of Type I and Type II error rates can decisively influence the functional safety of the overall system. Therefore, the performance evaluation of the environment perception should be defined to ensure a sufficiently safe level of residual risk associated with functional deficiencies in machine learning algorithms. For this, various sensors must be verified not only concerning their failure rates, but also about the possible causes of technical shortcomings in machine learning. Consequently, the quantitative evaluation of perception sensors and algorithms should consist of Type I and Type II error rates, in which some assumptions about the system context are implied. Thus, the robustness in real traffic can be achieved by the creative fusion of sensor data as well as appropriate system design.

Sensor capability criteria	Environment perception sensors			
	Camera-based perception	RADAR-based perception	LiDAR-based perception	eHorizon perception
Maximum longitudinal range	◒	●	●	●
Lateral field of view	◒	◒	●	●
Longitudinal range accuracy	◒	◒	●	○
Lateral range accuracy	●	○	◒	○
Relative object speed estimation	◒	●	◒	○
Moving object dimension	◒	○	◒	○
Moving object classification	◒	◒	◒	○
Adverse weather conditions	○	●	○	●
Behavior at darkness	◒	●	●	●
Sensor installation flexibility	◒	●	◒	●
Sensor cost requirements	◒	●	○	◒
Road classification	◒	◒	●	●

● : optimal ◒ : fairly optimal
◒ : natural ○ : fairly poor
○ : poor

Table 2.3: Comparable evaluation of environmental perception and situation prediction sensors [Ste14].

2.2.1 Camera-based Perception

The camera sensors measure the incident light using an optical system and capture visible cues similar to human perception. The exposure control regulates the exposure with constant contrast regardless of changes in brightness and direction. However, no depth or speed information can be measured directly, whereby the three-dimensional world is projected onto the two-dimensional image. The recognized object features are mapped to a vector representing an object hypothesis in the state space of the used classifier.

The stereo vision sensors consist of two monocular cameras with a certain distance between each other, typically known as base width, and measure an environment detail from different perspectives. The measurements are carried out synchronously, whereby a depth estimation is generated in the two images by comparing the displacement (disparity) of individual pixels or patterns. Besides, the distance accuracy is limited by the resolution, especially at long distances. Today's automotive cameras play a vital role in environment perception due to the high information density present in images. However, its drawbacks are, similar to automotive LiDAR, weather sensitivity and limited detection range. The following causes can lead to Type I errors when camera-based perception has low specificity:

- False object hypotheses due to ghost objects or bright lights.
- Ambiguities in the disparity calculation through repetitive patterns.
- Underexposed backgrounds due to color distortions.

Type II errors due to low sensitivity can have the following causes:

- Poorly illuminated objects.
- No pattern matching within the training data-set.
- Objects with low disparity due to homogeneous surfaces.
- Objects with low height due to no separation by layer.
- Overexposed backgrounds due to direct sunlight.
- Overexposed backgrounds due to adverse weather conditions (e.g. fog, snow, rain, etc.).

2.2.2 RADAR-based Perception

Range Sensors such as LiDAR and RADAR calculate the distance, angle and signal power to detect targets in a particular region of interest. Automotive RADAR sensors observe the position and velocity of moving objects as well as stationary roadside objects with precise range information and high resistance to adverse weather conditions.

However, RADAR detection is afflicted with a limited angular resolution in the case of stationary or longitudinally moving pedestrians. Automotive RADAR sensors transmit and receive radio waves to determine the velocity, range and angle of objects in the 76-81 GHz band. Its strengths derive from an extended longitudinal range, an optimal accuracy of the range rate with weather independence. RADAR sensors, however, can poorly resolve closely spaced objects over long distances. The Type I errors due to low specificity can have the following causes:

- Extended metallic objects that can be driven over (e.g. road sign gantries, road bridges and overpasses, tunnel fans, corrugated sheets, etc.).
- Extended metallic objects that can be driven under (e.g. guard rails, movable manhole covers, beverage cans, etc.).
- Ambiguity effects in object classification due to insufficient resolution in frequency tuning (e.g. through alley situations).
- Ambiguities with an extended activated field of view due to the higher deceleration time of commercial vehicles.
- Specular reflections and noise detections during Constant False Alarm Rate (CFAR) detection.

The following causes can lead to Type II errors when RADAR-based perception has low sensitivity:

- Objects with low Radar Cross Section (RCS) values.
- Aging affected radome behind the bumper with different damping characteristics. The radome is a plastic housing to shield the RADAR antenna from weather influences.

2.2.3 LiDAR-based Perception

Automotive LiDAR sensors are based on an optical measurement principle to locate and measure the distance of objects in space. The LiDAR sensors typically use the time of flight principle for distance measurement where a laser pulse is emitted and the elapsed time is measured until the reflected signal is received again.

The time delay between transmission and reception is directly proportional to the distance due to the proportionality between the time of flight and distance. The possible causes of Type I errors for LiDAR-based perception are poorly illuminated objects and high road inclination. At the same time, those for the Type II errors are light-absorbing objects, planer surface objects and adverse weather conditions.

2.2.4 eHorizon-based Perception

There are several standardized solutions related to High Definition (HD) maps to meet the challenges of further developing the driver assistance to a higher level of automated driving [Sas17]. The SENSOR Interface Specification (SENSORIS) protocol defines a standard to collect data from the sensors of environment perception to the map provider. The Ego-vehicle receives the necessary traffic information from the map provider via the Transport Protocol Experts Group (TPEG) protocol. The cloud-based exchange maintains incremental updates for digital maps in navigation and infotainment systems via the Navigation Data Standard (NDS) protocol, as demonstrated in figure 2.5.

The ADAS Interface Specification (ADASIS) protocol defines a standardized data model and interface that ensures regular interaction between the ADFs that generate or use the eHorizon. In the example of a Traffic Sign Recognition (TSR) function, the fusion of information sources, i.e. data from the eHorizon, image processing data and vehicle data, makes it possible to recognize the maximum admissible speed more reliably and conveniently [J+11].

The hybrid data representation of detailed digital maps and physical automotive sensors provide an extended view of the Ego-vehicle environment and thereby facilitate improved inferences and more competent decision-making [MPS+15]. Incidentally, digital maps are designed to be updated from time to time. Janssen et al. used the Dempster-Shafer fusion technique to merge digital map road signs with road signs detected by a video system [JN04]. Nienhuser et al. also employed the same fusion technique to give the corresponding priority to each data source [N+09]. The decision level fusion functions to validate inputs of higher levels and increase the robustness of the overall system in complex scenarios.

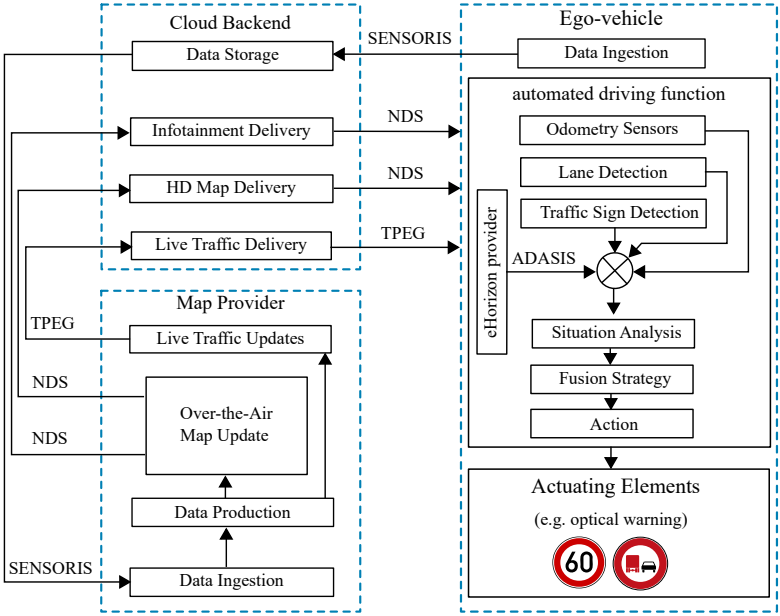

Figure 2.5: Proposed architecture by Open AutoDrive Forum in the context of HD 3D maps using the example of TSR function [Sas17].

Since today's standard vehicle sensor measurements contain relatively limited information content, the eHorizon data serves as a predictive sensor to anticipate the driving path. The eHorizon sensors employ the digital map data and Global Positioning System (GPS) sensors to predict the driving route. The GPS sensor determines the vehicle position in world co-ordinates [Bra13]. The map matching transforms this place into map coordinates and assigns it to a specific road on the map [ZBIG12]. Subsequently, the Most Probable Path (MPP) extracts and processes the relevant map characteristics along the most likely future route using prediction algorithms [Kv07]. The eHorizon data includes vehicle position data and road segment attributes such as road geometry, road class, number of lanes and speed limits [JS14]. On the transmitter side, eHorizon data is extracted by an eHorizon provider along the MPP and delivered to the vehicle bus system.

On the receiver side, an eHorizon re-constructor decodes the vehicle bus messages and transfers the data to the fusion module [DS11]. The discrepancies between GPS position data and matching maps are the probable causes of the Type I errors, while the usage of non-updated map data in memory due to the changes in traffic conditions are the possible causes of the Type II errors.

2.3 Data Fusion of Environment-Perception Sensors

Despite the recent rapid growth of human assisted driving systems to consolidate road safety, they still have various challenges when coping with dynamic traffic situations of daily life. Moreover, the issue of protecting VRUs is that their movement is mostly unpredictable [CAR15]. The VRUs are defined as non-motorized road users who use a road including sidewalk and other adjacent spaces, such as pedestrians, cyclists and persons with disabilities or limited mobility and orientation. Hence, there is an urgent need to locate the spatial and temporal coverage problems using an adequate environment model for situation evaluation, which is based on a synergistic approach between the existing sensors located on the body surface present in the truck's sensor network [OPLS11]. Thereby, the collaborative multi-sensor data fusion becomes increasingly critical to provide a better understanding of the monitored area. As a result, fusion algorithms analyze and interpret the traffic situation to provide a clear situation analysis and derive suitable control measures.

There are various ways to categorize the structures and methods of data fusion approaches. A distinction can be made between the architecture, the abstraction level of the input data and the sensor integration. Also, it is possible to distinguish between implicit and explicit fusion approaches as well as grid-based and parametric approaches. Basically, the fusion system can be divided into three primary fusion types: complementary to supplement incomplete sensor data, redundant to reduce erroneous measurements and cooperative to improve quality of environment model [Ott13]. The recognition of the vehicle environment for a comprehensive understanding of the scene is one of the most important elements for the realization of an ADF. Therefore, the use of different sensor technologies like camera, RADAR, LiDAR and eHorizon with different detection capabilities is necessary to provide both complementary and redundant information for the data fusion unit.

The fusion unit analyzes and evaluates the different sensor signals and generates a dynamic surround model with a good scene understanding. Although current ADFs (SAE L0-L2) can only use objects to generate a simple surround model, future stages of automated driving (SAE L3-L5) will need to combine not only the objects but also the sensor-specific features and characteristics of these objects. Hence, data fusion can often occur at three logical interface levels (detection, feature and object level). First, detection level fusion performs fused detections in the earlier steps of the processing chain to validate lower level inputs and increase the robustness of the overall system in complex scenarios. However, detection level fusion combines recognized detections without classification, model-based filtering and tracking. The origin of the detection co-ordinate system is the position of the sensor mounting in the vehicle co-ordinate system. Next, the feature level fusion provides recognized features that are merged based on classified sensor-specific attributes in the vehicle co-ordinate system without model-based filtering or tracking. Finally, the recognized object entities are tracked and filtered over time. Model-based algorithms classify the object entities (e.g. potentially moving objects, static objects and road markings) and evaluate them with existence probability values. Therefore, each recognized object entity has a unique object identification value that does not change over time. The object identification value can only be reused when the object is no longer visible.

The ISO/ Draft International Standard (DIS) 23150 is an automotive standard that addresses the logical interface of data communication between perception sensors and data fusion unit for ADFs [fS+21]. The data fusion unit intends to utilize the overlapping FOVs of the employed environment perception sensors and produces accurate information with a low rate of false-alarm detections [S+09]. Duraisamy and Schwarz have given a survey of the track-to-track data association methods using a decentralized sensor fusion architecture in order to achieve an optimal fused result [DS15]. Despite its significant contribution to situation awareness, one of its primary drawbacks is the increased software complexity caused by multiple interconnecting sensors [DFS+10]. The abstraction level at which the data fusion takes place is a trade-off between information context and complexity. While environmental perception sensors are still afflicted with adverse weather conditions, such as snowing or raining, HD maps can cope well with implicit traffic information such as speed limit changes through the highway-to-city transition [W+14].

Intelligent transportation systems use various sensor technologies to perceive the vehicle environment. Automotive RADAR sensors are commonly employed for object and pedestrian detection with precise range information and high resistance to adverse weather conditions. However, RADAR detection is afflicted with limited angular resolution in the case of stationary or longitudinally moving pedestrians. These days, the automotive cameras also play a decisive role in environment perception due to the high information density in images.

Figure 2.6 illustrates the functional components of an automated driving system relying on object-level data fusion. The tracked object properties with corresponding object list are shown in table 2.4. The layered structure of the multi-sensor fusion system can be described as a cyber-physical vehicle system with a sequence of sensor nodes N_s, where $k \in [1, N_s]$. Each sensor node $S^{(k)}$ maps a set of sensor detected objects $P^{(k)} = \left\{p_i^{(k)}\right\}_{i \in \left[1, N_p^{(k)}\right]}$ from the Ego-vehicle's environment to a set of tracked objects $Q^{(k)} = \left\{q_j^{(k)}\right\}_{j \in \left[1, N_q^{(k)}\right]}$ into the object-level data fusion algorithm, where the indices i and j enumerate the objects in the respective set. The detections of object and lane boundaries refer to unique IDs. The existence probability represents the detection quality and certainty of an object. A higher value indicates an object which exists with high probability and a lower value indicates a potential false alarm object.

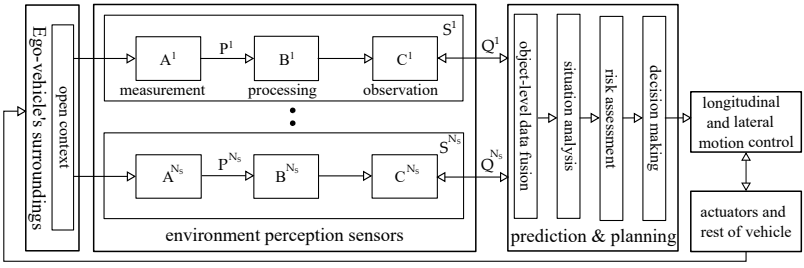

Figure 2.6: Functional components of an automated driving system including an object-level data fusion module [ESW+16].

Object property	**Object list** $Q^{(k)}$
Object ID	unique ID of the detected object or lane marking
Detection history	time stamp at the global synchronized time
Relative distance	longitudinal and lateral
Relative velocity	longitudinal and lateral
Relative acceleration	longitudinal and lateral
Object deviation	longitudinal and lateral
Relevance object selection	lane assignment (e.g. left, ego and right lane)
Bounding box dimension	width and height
Movement status	object state (e.g. stopped, moving, stationary, etc.)
Classification probability	object type (e.g. truck, pedestrian, etc.)
Existence probability	object detection quality and certainty
Lane boundary	type and color of the classified lane marking
Criticality level	confidence of the classified object or lane boundary type
Road specific data	route information (e.g. construction area, exit lane, etc.)
Additional data	sensor specific data (e.g. angular velocity, etc.)

Table 2.4: Data structure of object level fusion for environment perception sensors [ESFG19b].

In general, the existence probability depends on two factors: quality and confidence of the measurement used to generate the object and the object's observability over time. The measurement quality model is usually closely associated with a sensor's specific characteristics, for example RCS or a camera classifier. Equation 2.4 represents the signal flow chain of sensor $S^{(k)}$, where $A^{(k)}$, $B^{(k)}$ and $C^{(k)}$ represent characteristics of measuring principle, processing and observation, respectively.

$$S^{(k)} \left\{ B^{(k)}, C^{(k)} \right\} : P^{(k)} \rightarrow Q^{(k)} \tag{2.4}$$

Each of the elements $p_i^{(k)}$ in $P^{(k)}$ and $q_j^{(k)}$ in $Q^{(k)}$ is a vector containing the object properties. The Kalman filter is typically used for a multi-sensor data algorithm that employs Bayesian rules for the noisy environmental sensor measurements to produce reliable estimates of unknown quantities.

2.4 Retrospective and Prospective Safety Evaluation

The vehicle-based safety evaluation can be categorized as retrospective and prospective. The major difference is the timing of the assessment with regard to the life cycle of the development process. Real accident databases are used for the retrospective analysis. In contrast, prospective analysis estimates the number of critical situations that may occur in the future [FSLL19]. In a given ADF, human drivers encounter an average number of kilometers between events as a benchmark of human performance. The human-driver failure rate is assumed to have a Binomial distribution.

A safety case exists when the system under test has a failure rate lower than or equal to the benchmark reference with a particular level of confidence. Therefore, the number of test kilometers required for statistical evidence of an automated driving system can be calculated by a benchmark reference for the expected interval between accidents of equivalent severity. The total fatality rate in Germany caused by heavy-duty trucks in 2015 was 787 fatalities — as indicated in chapter 1 — , totaling 58, 93 billion kilometers [DES16]. According to the Binomial distribution, the equation 2.5 represents the confidence level $C[\%]$ for an ADF with m failures during a cumulative driving distance $d_c[km]$.

$$C_{(\zeta=m)} = 1 - \sum_{\zeta=0}^{m} \frac{d_c!}{\zeta!\,(d_c-\zeta)!} \lambda^{\zeta}\,(1-\lambda)^{d_c-\zeta} \tag{2.5}$$

If the failure rate of a CMV is $\lambda[1/km]$, then the reliability $\gamma[1/km]$ is $(1-\lambda)$ and can be interpreted as the probability that there is no failure in the route driven. A hypothesis about the scenario (failure-free driving) can be used to estimate a lower limit for the number of failure-free kilometers to determine the reliability of automated CMVs with a confidence level $C[\%]$. Consequently, the safety can be claimed for a certain number of failure-free kilometers at a particular confidence level, as shown in equation 2.6 [KP16].

$$C_{(\zeta=0)} = 1 - (1-\lambda)^{d_c} \tag{2.6}$$

Equation 2.7 gives the required driving distance d_c[km] without failures for given confidence C[%] and reliability γ[1/km].

$$d_c = \frac{\ln\left(1 - C_{(\zeta=0)}\right)}{\ln\left(1 - \lambda\right)} \tag{2.7}$$

The required driving distance d_c[km] is calculated by substituting the failure rate λ with $\frac{787}{58.934*10^9} = 1.34 * 10^{-8}$ [1/km] and the confidence level C with 95%, as indicated in equation 2.8.

$$d_c = \frac{\ln\left(1 - 0.95\right)}{\ln\left(1 - \left(1.34 * 10^{-8}\right)\right)} \approx 220 * 10^6 \text{km} \tag{2.8}$$

Figure 2.7 depicts the failure rate factor ($\lambda_A \div \lambda_H$), where λ_A[1/km] is the failure rate of an ADF and λ_H[1/km] is the benchmark failure rate of a human driver. For the CMVs these days, such long distance validations at which the controllability of the driver provides the necessary Proof of Safety (PoS) is unnecessary. However, in case of a fully automated driving system, the 2 million kilometers used to validate the current driver assistance systems are sufficient to prove the fatality factor Λ. The fatality factor $\Lambda_{(2*10^6\text{km})}$ is 25.5 times that of the humans, with about 50% confidence. Moreover, about 340 million kilometers are needed to prove that an automated driving system has a failure rate similar to that of human drivers in 2015 as the benchmark failure rate. This is done assuming that CMV has no failure ($m = 0$) during the FOT, with 99% confidence level. For this reason, it is economically impossible to demonstrate the safety of automated driving systems with widespread usage statistically before introduction, defined as an approval trap.

While the critical traffic events are typically rare and not reproducible, early identification of functional deficiencies is essential for automated driving. Despite the difficulty of predicting all possible operating scenarios a priori, the coverage of critical driving scenarios needs to be adequately investigated. Recent research suggests the hypothesis of Poisson distribution to calculate the required validation distance with the following assumptions [Wac17, KP16]. The Poisson distribution is a discrete probability distribution that expresses the probability of a given number of events in a continuum of time or space. Here, the route used is representative, while the critical events occur independently of each other within a random process.

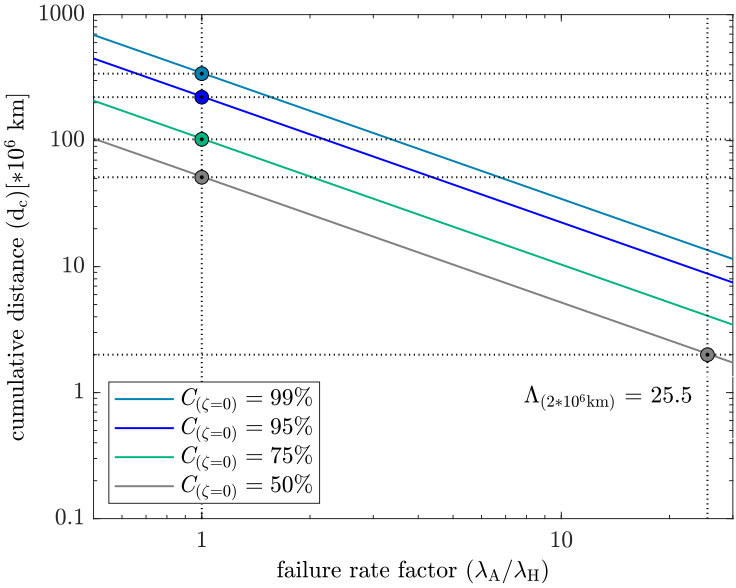

Figure 2.7: Prospective failure-free kilometers for a failure rate factor compared to human-driven CMV fatality rate of the year 2015 [ESO+19].

In the equation 2.9, m corresponds to the number of accident events and $\lambda[1/\text{km}]$ is the predicted distance at which this event occurs at a given confidence level.

$$C_{(\zeta=m)} = \frac{\lambda^{\zeta}}{\zeta!} e^{-\lambda};\ \zeta = 0, 1, 2, \cdots, m \tag{2.9}$$

The MTBF can be determined at a given confidence level using the hypothesis of the Chi-square distribution according to ISO 26262:2018. The Chi-square distribution is a probability density function that calculates the MTBF failure rate based on observed failures. Accordingly, an exponential failure distribution with a constant failure rate is assumed. Regarding the safeguarding of driver assistance systems, there are no legal requirements for the validation distance.

Since unintended reactions are rare events, a Chi-square distribution can be applied. If no critical event occurs at a sample distance with a required failure rate of one million kilometers each, the necessary validation requires around three million kilometers. In this case, no event should occur during the driven interval to argue the residual risk with a confidence level of 95%. The required mileage will increase if more events occur during validation (e.g. $d_c = 4.8 * 10^6$[km] at $\zeta = 1$, $d_c = 6.3 * 10^6$[km] at $\zeta = 2$, $d_c = 7.8 * 10^6$[km] at $\zeta = 3$, etc.), as illustrated in figure 2.8. In practice, the validation distance does not play the central role, but the variance of test conditions do, in order to cover maximum possible rate operating situations (e.g. different weather conditions, time of day, road conditions, traffic conditions, pedestrian conditions, etc.). Therefore, route diversity in physical road tests is a significant measure of the probability distribution.

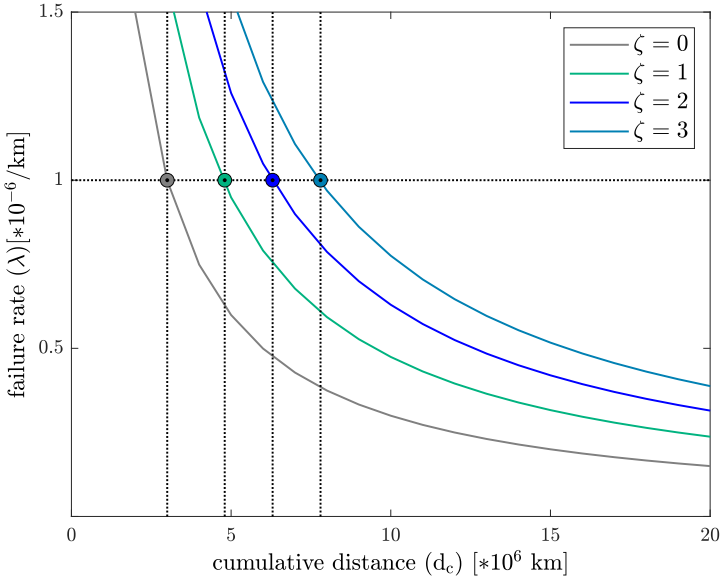

Figure 2.8: Required validation distance for various accident events using the Chi-square distribution with confidence level (C = 95%) [ESO+19].

However, statistical evidence of the accumulated road kilometers is potentially invalid with each software upgrade. Even if the FOT continues in the Spiral model of software development until no more errors are found, the safety case argument does not provide any proof that the ADF is absolutely safe due to the Pesticide paradox phenomenon [Bac18]. In software testing, the Pesticide paradox is an error detection phenomenon, where if the same test matrix is performed repeatedly, the same test matrix will eventually find no more errors [KW18]. It means that an automated driving algorithm that passes the same repetitive tests eventually builds up resistance to them. Consequently, it will not be viable to prove safety of the required level of system performance through driving test hours alone during the development phase. Furthermore, there is no complete public set of machine-interpretable traffic regulation with exception-handling rules.

The functional requirements for ADFs become thus implicit and incomplete by the learning process from machine learning data-sets, which are used to perform algorithmic operations. The reliance on data-driven mileage accumulation as the only credible safety argument points to an impractical safety validation strategy. Also, the real-world testing may not accumulate enough hours of exposure to observe critical scenarios that occur by chance. On the other hand, knowledge-based assessments can accelerate exposure to the known critical scenarios but suffer from the possibility of not verifying the unanticipated scenarios.

Alternative methods of safety assessment are therefore required, as the validation distance in the FOT will increase dramatically by using the current test concepts for automate driving without driver engagement. It is therefore obvious that a safety argument for these algorithms from SAE L3 onwards, based solely on the accumulation of road kilometers through endurance test campaigns, is no longer feasible. Therefore, knowledge-based test platforms can fulfill the associated test objectives in a time and cost efficient manner. However, these techniques alone cannot guarantee a sufficient confidence of safety for large-scale deployment without giving particular attention to data acquisition and analysis from the FOT [Z$^+$18b].

2.5 Measurable Safety Methodology

The functional safety is an absence of unreasonable risks due to hazards caused by malfunctioning behavior of E/E systems. Therefore, the ISO 26262:2018 norm is a risk-based approach for the development of safety-critical software systems in passenger cars, motorcycles and commercial vehicles. The ISO 26262:2018 includes a hazard analysis and risk assessment to determine the required Automotive Safety Integrity Level (ASIL) and to assess the potential risks of E/E malfunctions that may violate the safety goals. The risk-oriented approach classifies the risk $\mathcal{R}$ for each potentially hazardous driving situation, as shown in equation 2.10 [Hun18].

$$\mathcal{R} \approx \sum_{0}^{\mathrm{h} \in \mathcal{H}} (\mathcal{S}_\mathrm{h} * \mathcal{X}_\mathrm{h} * \mathcal{O}_\mathrm{h}) \tag{2.10}$$

The hazard classification defines each potentially hazardous driving situation with the following three impact factors, where $\mathcal{H}$ corresponds to the set of hazardous driving situations [HKM17]:

- Severity level $\mathcal{S}_\mathrm{h}$ of a hazardous driving situation h with the consequence of injuries or fatalities, where $\mathcal{S}_\mathrm{h} \in \{ \mathcal{S}_\mathrm{h}^0 , \mathcal{S}_\mathrm{h}^1 , \mathcal{S}_\mathrm{h}^2 , \mathcal{S}_\mathrm{h}^3 \}$.
- Exposure probability level $\mathcal{X}_\mathrm{h}$ of hazardous driving situations, where $\mathcal{X}_\mathrm{h} \in \{ \mathcal{X}_\mathrm{h}^0 , \mathcal{X}_\mathrm{h}^1 , \mathcal{X}_\mathrm{h}^2 , \mathcal{X}_\mathrm{h}^3 , \mathcal{X}_\mathrm{h}^4 \}$.
- Controllability probability level $\mathcal{O}_\mathrm{h}$ of not avoiding an accident in a harmful situation, where $\mathcal{O}_\mathrm{h} \in \{ \mathcal{O}_\mathrm{h}^0 , \mathcal{O}_\mathrm{h}^1 , \mathcal{O}_\mathrm{h}^2 , \mathcal{O}_\mathrm{h}^3 \}$.

In this context, the system follows one of five classes to define risk reduction requirements, where ASIL D is the highest and Quality Management (QM) the lowest risk reduction class (ASIL 26262-1:2018). For example, a system specified for the implementation of a truck platooning may exhibit undesirable behavior due to a misclassification of objects and require driver intervention. Therefore, controllability presents the probability of controlling driving situations within system limits and failures. The processes and methods for assessing the controllability of unintended driver assistance reactions are specified in the Code of Practice (CoP) for the design and evaluation of human assisted driving systems [KNB+09].

The ASIL determination can be assigned to the level O_h^3 of controllability for automated driving without driver intervention, where the intended function is difficult to control or uncontrollable, as illustrated in table 2.5 [KW16].

ASIL level	S_h^1	S_h^1	S_h^1	S_h^1	S_h^2	S_h^2	S_h^2	S_h^2	S_h^3	S_h^3	S_h^3	S_h^3
	X_h^1	X_h^2	X_h^3	X_h^4	X_h^1	X_h^2	X_h^3	X_h^4	X_h^1	X_h^2	X_h^3	X_h^4
QM	●	●	○	○	●	○	○	○	○	○	○	○
ASIL A	○	○	●	●	○	●	○	○	●	○	○	○
ASIL B	○	○	○	○	○	○	●	○	○	●	○	○
ASIL C	○	○	○	○	○	○	○	●	○	○	●	○
ASIL D	○	○	○	○	○	○	○	○	○	○	○	●

● : relevant
○ : irrelevant

Table 2.5: ASIL requirements for automated driving devices without driver monitoring using uncontrollable level O_h^3 according to ISO 26262.

Despite the updating of the scope of the ISO 26262:2018 norm for the inclusion of CMVs in Edition 2, its safety goals mainly address undetected random hardware failures of the system components and systematic software failures [SH19c]. Assuming that the E/E system malfunctions are managed using ISO 26262:2018, the safety violations, which may be caused by the environmental perception sensors, remain outside the scope [SH20]. Redundancy, diversity and functional restrictions can compensate system limitations [SH19a].

The ISO/PAS 21448:2019 serves as an extension scheme to specify the intended function in such a way that it is robust and safe enough to take the variations in sensor inputs and the different environmental conditions into account [fS$^+$19b]. The OEDR examines whether the vehicle can correctly detect objects and events and execute an appropriate response. Therefore, the context of the OEDR is similar to the case in SOTIF, but with a different designation given by the NHTSA. However, new verification and validation measures are needed to assess unintended system behavior due to technological and systemic deficiencies. At the same time, the ISO/PAS 21448:2019 activities complement the ISO 26262:2018 norm with its focus on driver assistance rather than automated driving without driver engagement.

On 22. Mai 2022, the German Federal Council (*Bundesrat*) approved an ordinance that regulates the operation of motor vehicles, in particular CMVs, with SAE L4 ADFs (*Verordnung zur Regelung des Betriebs von Kraftfahrzeugen mit automatisierter und autonomer Fahrfunktion und zur Änderung straßenverkehrsrechtlicher Vorschriften*) on public roads within a specified operating range [Bun22]. The ordinance supplements the German act on autonomous driving (*Gesetz zur Änderung des Straßenverkehrsgesetzes und des Pflichtversicherungsgesetzes – Gesetz zum autonomen Fahren*) and establishes a legal framework for the safe deployment of CMVs with a SAE L4 ADF for operation on public roads in Germany [KML22]. In addition, the ordinance introduces a uniform procedure for the approval of tests by the German Federal Motor Transport Authority (*Kraftfahrt-Bundesamt*) in order to standardize and centralize the tests pursuant to §1i of the German Road Traffic Act[2] (*Straßenverkehrsgesetz – Erprobung von automatisierten und autonomen Fahrfunktionen*), with the provision of a safety concept for functional safety.

[2] https://www.gesetze-im-internet.de/stvg/BJNR004370909.html

3 State-of-the-Art and Research Perspectives

The systems engineering process requires the state-of-the-art evaluation procedures to verify and validate ADFs [SZS+15]. The evaluation of software releases needs to be carried out in different phases up to the start of production to prove that the residual risk is below an acceptable level. Initially, the driving simulator tests are used to evaluate various system concepts of a CMV [Has14]. Testing programs are carried out in XiL, on test tracks and in environments under real traffic conditions, as already defined in subsection 2.1.7. Several tools are used in this context including the following: Hardware/ Software (HW/SW)-open-Loop reprocessing, HW/SW-in-the-closed-Loop simulation, customer studies in driving simulator as well as real-world test drive open-loop and closed-loop, as already defined in subsection 2.1.3. The scenario databases are then integrated into the data analysis and assessment tools. Therefore, the induced traffic situations with unintended reactions continuously extend the scenario databases. Finally, homologation is used to indicate that the approval requirements for the market-specific type are met, based on the evidence collected during the development process [R+18a]. Appropriate quality measures are essential to achieve sufficient reliability, safety and availability in the framework of the software quality management process.

3.1 Data-driven and Knowledge-based Test Platforms

The data-driven test methods use empirical data to obtain new insights into the system behavior under specific traffic situations or field conditions, while the knowledge-based test methods convert the implicit information into explicit ones.

Data-driven test methods require many prerequisites such as fleet vehicles equipped with additional high performance measurement systems and data processing pipelines. Meanwhile, knowledge-based test methods use abstract information to create functional, logical or even directly concrete scenarios for the database. The information can be in the form of abstract knowledge from experts, standards and guidelines. Table 3.1 illustrates various data-driven and knowledge-based test methods (e.g. New Car Assessment Programme (NCAP) tests, requirements test metrics, FOT and scenario databases) for safety and reliability assessments [ZKK+16]. Although ISO 26262:2018 and its V-framework reflect generally accepted practices to ensure functional safety, ADFs present unique challenges in mapping the technical and functional requirements to the classical V&V methods [KW17]. Therefore, the validation procedures offer a range of activities to generate confidence that an ADF can achieve its intended purpose and goals.

Scope	**Inference**	
	Data-driven testing (induction)	Knowledge-based testing (deduction)
Case-by-case analysis	Empirical data (e.g. NCAP tests)	System use cases (e.g. requirements test matrix)
Traffic-based evaluation	Route profile (e.g. on-road field tests)	Scenario meta-model (e.g. scenario databases)

Table 3.1: Classification of data-driven and knowledge-based test methods for safety and reliability assessment [ESm+19]

3.1.1 HW/SW-in-the-open-Loop Simulation

Figure 3.1 depicts the general data flow of a regression test using the example of an eliminated software error within the object detection of a monocular camera sensor. The relevant situations for the camera ECU are recorded by the Ego-vehicle with an appropriate measurement equipment and collected within a data ingestion process. A software update occurs within a software development process. While the environmental perception sensors react sensitively to the target hardware constraints, the monocular camera sensor without camera optics is integrated for regression tests with recorded sequences.

This is followed by the Hardware-in-the-open-Loop (HoL) test bench stimulating the optical interfaces of the monocular camera sensor with the recorded data to verify the functionality of the new software release [WRM+19].

Definition 3.1 (Open-Loop Reprocessing): It verifies the error correction by reprocessing of field measurements with new software releases. In the case of software reprocessing, the target software is executed on prototypical hardware, whereby the software decisions have no influence on the stimulus. In the case of hardware processing, the target software is executed on the target hardware, while the hardware outputs have no influence on the hardware inputs.

Definition 3.2 (HoL Test Platform): A Hardware-open-Loop reprocessing platform to enable reprocessing of the target software on the target hardware, while the hardware outputs have no effect on the hardware inputs. The test platform is typically conducted within a validated environment for components and subsystems of the environmental perception.

Therefore, the HoL test bench stimulates the external interfaces in real-time by utilizing a recorded sequence of on-road tests with an original detection result (v) to generate a new detection result ($\tilde{v}$) after updating the application software of the target hardware, as illustrated in figure 3.1. Thus, original and new results can be compared to decide whether the error is fixed according to defined pass/fail criteria. The hardware open-loop reprocessing generates driving situations from the test-case description and evaluates the behavioral response without feeding it back into future situations.

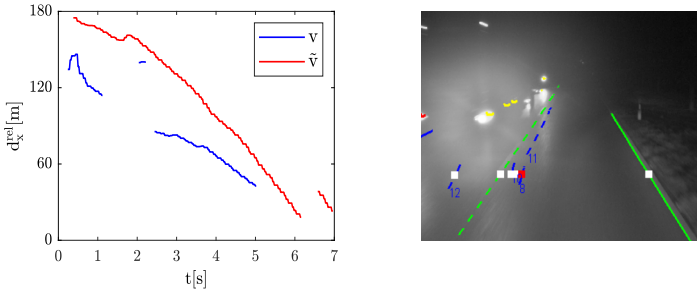

Figure 3.1: Regression test process with the HoL test bench using the example of detection of oncoming objects d_x^{rel} (left) with a monocular camera sensor (right) [E+16].

3.1.2 HW/SW-in-the-closed-Loop Simulation

The HW/SW-in-the-closed-Loop simulation verifies the behavior in a closed-loop to prove functional correctness of an artifact against its functional specification. A scenario-based testing requires various technical requirements for the simulation environment of roads, traffic objects, environmental perception sensors, the driver of the Ego-vehicle, commercial vehicle dynamics and actuators. For a camera HiL test bench, the traffic environment shall be animated in 3D perspective to display the traffic scenes in a virtual world in the form of a video sequence recorded by a monocular camera sensor, as illustrated in figure 3.2. Therefore, closed-loop testing specifies an entire scenario in a test case that contains a sequence of scenes, actions, events and goals for the ADF. The behavioral reactions are used to influence future scenes and thus also future situations. The HiL test method integrates the ADF into the traffic environment and vehicle dynamics simulations by combining the simulation models and the ECU hardware into a real-time ECU test bench. However, the Software-in-the-Loop (SiL) test platform integrates the executable software code generated from the same source as for the automotive ECU.

Definition 3.3 (HiL **Test Platform):** It refers to a test bench to enable processing of the target software on the target hardware, with the hardware outputs influencing the hardware inputs. The test platform can be used either at component, subsystem or system level.

The monocular camera ECU is connected to the HiL via a Controller Area Network (CAN) bus for rest-bus simulation purposes [TH13]. By the means of a monitor positioned in front of the camera, the camera is stimulated to pretend reality by processing the displayed realistic images under real-time conditions. Figure 3.2 illustrates an example with a lane change scenario, which is applied to a LDW function. As an objective of benchmarking, the reference values are compared with the detected values by the camera projection of the monocular camera ECU within the real-time simulation environment; e.g. d_y^{lt} [m] and v_y^{lt} [m/s] for the left lane and d_y^{rt} [m] and v_y^{rt} [m/s] for the right lane. In chapter 4, a detailed approach to HiL techniques, with their implementation is discussed.

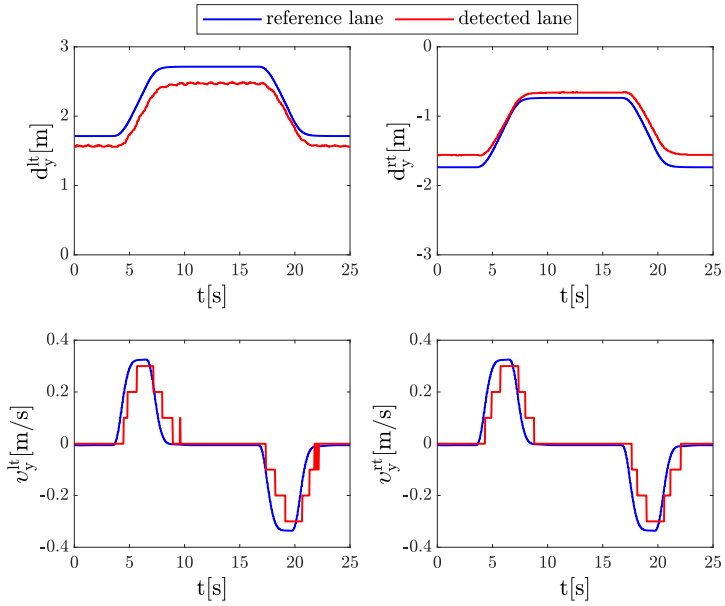

Figure 3.2: Verification process with the closed-loop HiL test bench using the example of the detection of left lane markings (left) and right lane markings (right) on the basis of a monocular camera sensor [ESW+16].

3.1.3 Model-based Back-to-Back Testing

Model-based software development in the automotive industry uses tools such as Simulink® or dSPACE TargetLink to implement a software module within the ADF. Simulink® is a graphical programming environment for modeling and simulating of automotive dynamical systems. The dSPACE TargetLink is a tool for automatic C and C++ code generation based on a subset of Simulink®/Stateflow® models. Tools such as Simulink Coder™ or dSPACE TargetLink are then used to automatically generate C source code from the resulting models. Simulink Coder™ is a tool for automatic code generation from Simulink® diagrams, Stateflow® charts and MATLAB® functions. Automated test data generation can be applied if a test oracle can be defined automatically with its reference information.

Definition 3.4 (**Test Oracle**)**:** A mechanism in software testing to determine whether a test case has passed or failed. For example, structural testing is typically employed to generate test data based on the internal structure of the test object. Therefore, the identification of input values depends on a selected path or branch that is executed within the test object.

The ISO 26262:2018 demands that coverage metrics shall be taken into account when testing at the model and software code level, such as MC/DC. A major cause of such semantic differences is the application of scaling to variables during the software code generation to optimize code efficiency and value precision. In the figure 3.3, back-to-back tests generate a collection of structural test cases to compare the software generated with the behavior of the underlying model.

Definition 3.5 (**Back-to-back Consistency**)**:** It is a type of software testing. Two or more variants of a software module are generally tested with the same stimulus inputs. Their corresponding outputs are compared and evaluated if there are discrepancies in the software.

Figure 3.3 illustrates an example for defining a test oracle to evaluate the automatically generated test cases according to the accepted time and value tolerances. Therefore, both, the model and software are executed with the same input data and then the corresponding output data entries are compared. The value tolerance is determined by the difference between the v_{MiL} and v_{SiL} of the longitudinal acceleration a_x^{ego} [m/s^2] within an ACC function, as illustrated in the left side of figure 3.3. In contrast, the time tolerance is determined by the difference between the v_{MiL} and v_{SiL} of the required mode $roll_{mode}$ within an ACC function, as depicted in the right side of figure 3.3. The model runs with the Model-in-the-Loop (MiL) test suite and is verified back-to-back with the C code running on the SiL test suite. Wilmes introduces a hybrid test data generation approach to combine static analysis and dynamic test data generation [Wil15]. The test data finding problem is converted into an optimization problem by defining a cost function. As a result, the generated test data is evaluated to distinguish between relevant and irrelevant test data and to generate new test data in each iterative cycle [Wil16]. In addition, the static analysis serves to accelerate the automatic search by identifying unreachable model states.

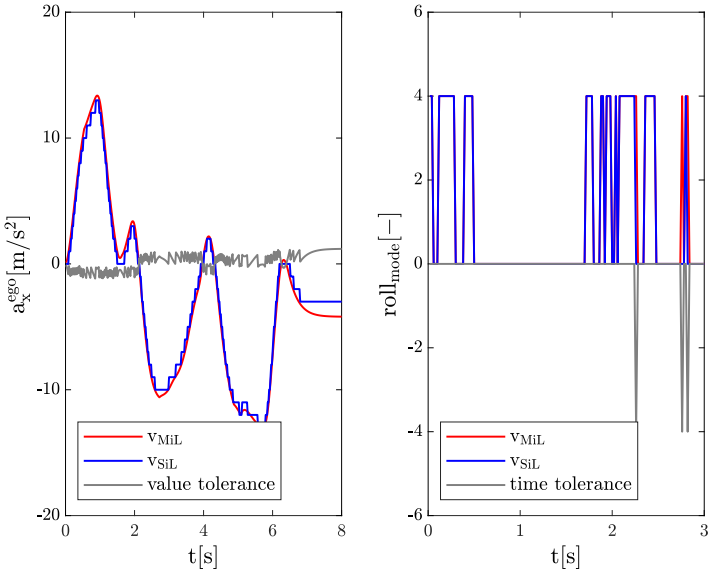

Figure 3.3: Back-to-back test process with the model-based code generator (dSPACE TargetLink) using the example of an ACC function based on value tolerance (left) and time tolerance (right) [ESO+19].

3.2 Goal-based Safety Case Assessments

The ISO 26262:2018 standard represents the state-of-the-art with respect to functional safety for safety-critical E/E systems in road vehicles. The CoP for the design of human assisted driving includes the evaluation of the human assisted driving functionality and is based on the driver's controllability to maneuver the CMV's reliability in road traffic [KNB+09]. These practices assume that the human driver remains responsible for CMV behavior to override or deactivate the system at any time. If the driver is no longer responsible for the behavior of the CMV, which is already the case with intervening emergency functions, the driver's controllability test catalogs are no longer sufficient. The evidence shall also be provided that the Type I and Type II rates are reasonably low [AW17].

The approaches of safety validation for ADFs beyond the mileage accumulation are in high demand. Thereby, a falsification approach shall be coupled with concrete, verifiable safety objectives and requirements. In parallel, the verification procedures according to the ISO 26262:2018 V-shaped model assume that high-quality requirements for implementation are further developed. Therefore, the traditional V-shaped model engineering process can pose a challenge in articulating the functional requirements of machine learning algorithms [KW16]. With the V-shaped model, the training set is more related to the functional requirements and the validation set to a test plan. The verification arguments with sufficient training and validation data leads to the need to develop the data ingestion system according to safety-critical software standards [Hil12].

The American National Standards Institute (ANSI)/ Underwriters Laboratories (UL) 4600 is a standardization activity to address the ability to autonomous products to perform the intended function without human intervention. This is based on their current state and sensing the operating environment. The standard intends to apply a goal-based approach that specifies tasks that need to be addressed in creating a safety case [UL22]. The safety case shall argue that relevant objects from existing sensors can be successfully detected and classified within the intended ODD, rather than determining whether a system design with a new sensor setup is required. Meanwhilst, the safety case should provide an argumentation to ensure an appropriate safety level through a robust combination of analysis, simulation and testing, rather than deciding how many kilometers have to be accumulated for safety demonstration purposes [KW17].

The Goal Structuring Notation (GSN) presents a safety case to highlight the verification methods for automated truck driving. The GSN is a graphical structure notation for structuring an assurance case in connection with argument, context, assumptions and evidence. The safety case is a reasoned and verifiable artifact that supports the contention that its top level claim is satisfied, including systematic reasoning, its underlying evidence and explicit assumptions that support the claim according to ISO/IEC 15026-2:2011. Therefore, the assurance case ensures that sufficient evidence is systematically gathered to argue a tolerable residual risk through adaptive verification for ADFs, as illustrated in figure 3.4. Each hypothesis identifies the residual risks for a test or simulation environment. The assumptions that are covered by other verification approaches are a part of the safety argumentation chain [BGH17].

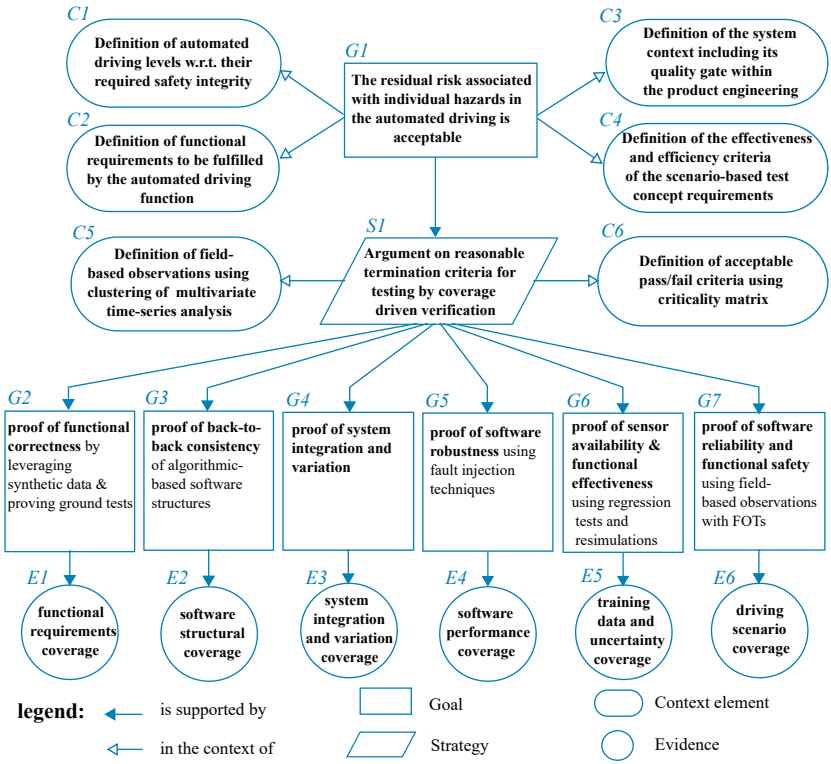

Figure 3.4: Argument structure of context-driven test concept based on the ODD coverage using GSN [ESO+19].

The goal-based safety case approach follows a decomposition approach on functional, logical and concrete levels. Therefore, the test scenarios can be described either as functional with NL description without values, logical with an assignment of value ranges or concrete with an association of fixed values. A major challenge in achieving the decomposition approach is to classify the test objectives and the coverage criteria according to their respective test environments [Ame20]. Therefore, the test objectives imply measurable quality criteria for the V&V strategy.

Strategy (*S1*)**:** It describes the verification case strategy to define the required termination criteria for testing through adaptive test coverage with the context elements {*C5*,*C6*}. Subsequently, the following sub-goals {*G2*, ··· , *G7*} provide the evidence arguments {*E1*, ··· , *E6*} of test coverage within the verification case [GMB18].

Goal (*G1*)**:** It is the main goal to constitute the top-level claim of the verification case scope with the context elements {*C1*, ··· , *C4*}. The *G1* argues a sufficiently low level of residual risk associated with individual hazards in automated driving, although all possible driving situations might not be verified during the development phase, detailed in section 6.1.

Goal (*G2*)**:** The proof of functional correctness verifies that the test object fulfils the required functionality according to its specifications by leveraging synthetic and real-world data, discussed in subsections 3.1.2 and 5.3.3.

Goal (*G3*)**:** The proof of back-to-back consistency is the verification of the required consistency between the various execution platforms (e.g. model and code) within the permissible discrepancies by back-to-back tests for algorithmic-based software structures, illustrated in subsection 3.1.3.

Goal (*G4*)**:** The proof of system integration and variation is the evidence to the validity of the system maturity to cover the system variations that may include static and dynamic tolerances of truck-trailer combinations, expound in subsections 4.2.1 and 4.2.2.

Goal (*G5*)**:** The proof of software robustness shall provide a sufficient probability of coping with system boundaries and faults using fault injection techniques, as explained in subsections 3.3.2 and 3.3.3.

Goal (*G6*)**:** The proof of sensor availability and functional effectiveness focuses on the presence of the environment perception sensor within the defined deviation and tolerance limits of the specified time and range using regression tests and re-simulations, as focused within subsection 3.1.1.

Goal (*G7*)**:** The proof of the software reliability and functional safety is the proof that the test object is reliable enough with respect to functionality and safety mechanisms by leveraging field observations with FOTs and synthetic data. The proof of functional safety refers to the functional safety requirements of ISO 26262:2018 in order to avoid systematic software and random hardware failures, as portrayed in sections 6.2, 7.1, 7.2 and 7.3.

Table 3.2 explains the assignment of different possible test environments with the respective test objectives. The selection of suitable test environments from XiL to FOT depends on the effectiveness and efficiency criteria of the test conditions and their validity. The effectiveness criteria indicate the intended results such as representative valid and observable interfaces. However, the efficiency criteria reflect the desired performance in comparison with the resources used to achieve the economic use, reproducibility and promptness. The proving ground is an area dedicated to putting a vehicle's performance to the test. In the driving simulator, the driver is in an artificial environment designed to replace one or more aspects of real driving behavior.

Test environments	**Test objectives**					
	G2	*G3*	*G4*	*G5*	*G6*	*G7*
MiL tests	●	●	○	●	○	○
SiL tests	●	●	○	●	○	○
Component HiL tests	●	○	○	●	○	●
Subsystem HiL tests	●	○	○	●	○	●
System HiL tests	●	○	●	○	○	●
Proving ground tests	●	○	●	○	●	●
Driving simulator studies	○	○	●	○	○	●
Naturalistic FOTs	●	○	●	○	●	●
HW/SW reprocessing & regression tests	●	○	●	○	●	●

● : recommended test objective
○ : not useful

Table 3.2: Assignment of potential test environments to the corresponding test objectives — driven from figure 3.4 — [ESO+19].

Table 3.3 explains the assignment of different possible test suites with the respective test coverage criteria. The test coverage criteria provide an indicator of the software testing effort during a test run within the V&V strategy.

Test environments	**Test coverage criteria**					
	E1	*E2*	*E3*	*E4*	*E5*	*E6*
MiL tests	●	●	○	○	○	○
SiL tests	●	●	○	○	○	○
Component HiL tests	●	○	○	●	○	●
Subsystem HiL tests	●	○	●	●	○	●
System HiL tests	●	○	●	○	○	○
Test tracks & proving grounds	●	○	●	○	●	○
Driving simulator studies	○	○	●	○	○	●
Naturalistic FOTs	○	○	●	●	●	●
HW/SW reprocessing & regression tests	○	○	○	●	●	○

● : recommended test coverage
○ : not useful

Table 3.3: Assignment of possible test environments to the corresponding test coverage criteria — driven from figure 3.4 — [ESO+19].

Evidence (*E1*)**:** The functional requirements coverage defines a relationship between the functional requirements and the executed test cases, whereby at least one test case is defined for each requirement.

Evidence (*E2*)**:** The software structure coverage provides the code coverage of model-based software structure components, such as MC/DC.

Evidence (*E3*)**:** The system integration coverage includes detected failures in the interfaces and interactions between integrated components, subsystems or systems. The system variation coverage defines the robustness against variations in the system context. For example, when automated driving software modules are developed, a variety of system variants which can include static and dynamic tolerances within the Ego-vehicle, are put in.

Evidence (*E4*)**:** The software performance coverage determines the robustness of the software using fault injection techniques.

Evidence (*E5*)**:** The training data coverage specifies which training data is required for a particular application and which data leads to the most accurate results, such as training of neural networks for image processing. The uncertainty coverage quantifies the Aleatoric and Epistemic uncertainties of machine learning algorithms.

Evidence (*E6*)**:** The driving scenarios coverage identifies the known critical scenarios, which should exhibit similar behavior, and minimizes unknown critical scenarios.

3.3 Standardization Activities and Research Projects

There are various standards that determine the assessment of SAE L0-L2 functions for commercial vehicles and buses based on the physical characteristics of the vehicle. The ISO 19377:2017 describes a test method for determining the path deviation of the braking maneuvers induced by an AEB of CMVs from a predefined desired trajectory. If the AEB function is utilized in further stages of the automated driving systems, the deviation needs to be compensated.

The ISO 15622:2018 presents the minimum requirements for failure reactions and performance test procedures for ACC systems. The ISO 22839:2013 applies criticality assessment metrics on an AEB system to provide a certain ratio of Type I and Type II errors within the receiver operating characteristic curve space. The ISO 22839:2013 is a test method standard for defining the required behaviors and test criteria of an AEB system. Shladover et al. [SN19], Takács et al. [TDG+18] and Junietz et al. [JWKW18] provide a comprehensive overview of the activities concerning regulations and standards for the type approval of automated vehicles. These standardization activities can be considered as development guidelines for manufacturers. In recent years, many research projects have been completed dealing with the V&V activities for SAE L3 automated vehicles. The following is a summary of representative completed projects:

- The research project PEGASUS (Project for Establishing Generally Accepted good quality criteria, tools, methods, Scenarios and Situations for the approval of highly ADFs) aimed at developing methods for en-

suring the safety of ADFs [N+20]. The aim of the research project was to develop uniform technical standards for the verification of conditional ADFs, and to define thresholds for a sufficiently high controllability level of SAE L3 systems [Sch15]. The PEGASUS[1] project has seeked to define risk scenarios, assessment criteria for sufficient safety in various traffic situations on the basis of a generally accepted method for evaluating the safety of conditional ADFs [WLFM19]. The project results were demonstrated in the middle of 2019 on the basis of a highway chauffeur as a SAE L3 system for passenger cars [M+16].

- The research project L3Pilot[2] (large-scale Piloting of SAE L3 functions on European roads) was concerned with large-scale field tests of SAE L3 functions under variable conditions with 1000 drivers and 100 vehicles [HSK+19]. The field tests served to evaluate the technical aspects, user acceptance, driving behavior and safety impacts for SAE L3 applications. With the comprehensive piloting of ADFs in test vehicles, L3Pilot has paved the way for large-scale field tests of series cars on public roads.

- The research project AdaptIVe[3] (Automated driving applications and technologies for Intelligent Vehicles) was an European research project dealing with safety validation, technical system limits and legal aspects of the release of automated driving applications. The project results have presented best practices in systems engineering and safety validation to establish a CoP for the design and evaluation of ADFs [Ete17].

- The research project SaLsA (Safe autonomous Logistics and transportation vehicles in outdoor Areas) focused on developing automated trucks that can operate safely outside warehouses in a shared work environment with conventional human-driven vehicles and pedestrians. The cooperative scanning of the environment through mobile and stationary sensors enables safe operation even at higher speeds [Wie17].

- The research project ATLaS[4] (AuTomated and networked driving in Logistics - opportunities for more added value) investigated the influ-

[1] https://www.pegasusprojekt.de/en/home
[2] https://www.l3pilot.eu/
[3] https://www.adaptive-ip.eu/
[4] https://www.tib.eu/de/suchen/id/TIBKAT:1697842658/

ences of automated and connected driving on the logistics chain in order to identify deployment scenarios accepted by stakeholders [FL+20].

- The research project ENABLE-S3[5] (European initiative to ENABLE validation for highly automated safe and secure systems) was a research project to enable an accelerated assessment of highly automated and autonomous systems in the mobility domains (e.g. automotive, avionics, rail and maritime) [LAH+19]. The project objective is to establish cost-efficient cross-domain virtual and semi-virtual V&V platforms and methods for autonomous cyber-physical systems [Lei20].

In parallel, numerous research projects have been launched to contribute to developing a general consensus or uniform framework for safety validation and reliability assessment of SAE L4-L5 automated driving. The representative ongoing projects in industry and science to safeguard automated driving are summarized as follows:

- The test field TAF-BW[6] (Test Area Autonomous Driving Baden-Württemberg) offers a testing environment under real traffic conditions for automated and connected driving applications. The test area is funded by the state of Baden-Württemberg and includes all relevant road types. The aim of the test area is to promote the development of future-oriented solutions for individual traffic and local public transport. The ground truth data is derived from the infrastructure sensors for real-time recording of traffic and its influencing factors as well as from 3D HD maps of the road network in everyday road traffic [FDW+18].

- The research project VVM[7] (Verification and Validation Methods for SAE L4 and L5 automated vehicles) aims to develop procedures to combine the virtual and the real-life tests. The developed procedures can be used for legally compliant and efficient homologation of automated vehicles [ZRBE20].

- The research project SET LEVEL 4to5[8] (Simulative dEvelopment and Testing of LEVEL 4 and 5 systems) addresses the simulation-based

[5] https://enable-s3.eu/
[6] https://taf-bw.de/en/
[7] https://www.vvm-projekt.de/
[8] https://setlevel.de/

development and testing of SAE L4 and L5 vehicles for urban areas. The project intends to provide tool-oriented techniques for efficiently designed simulation-based test and release procedures [HTR+20].

- The research project KI-Absicherung[9] (Methods and measures to safeguard artificial intelligence-based perception functions for automated driving) seeks to develop measures and methods to validate neural network based perception and multi-model sensor fusion. The pedestrian detection is identified as a representative function for the project in terms of multi-sensory perception [GGSB19].

In addition, a significant number of publications on safety validation of automated vehicles have been published in recent years in response to the strong interest in a rapid market introduction of automated vehicles. Table 3.4 refers to a comparison with a scale from not applicable to optimal using the following evaluation criteria. According to the current state of science and technology, the diverse approaches of safety assessment for automated driving can be divided into four categories: shadow-mode approach, formal safety verification, traffic simulation-based approach and scenario-based approach, as described in the following subsections 3.3.1, 3.3.2, 3.3.3 and 3.3.4.

First, the representativeness of scenarios reflects how realistic road traffic conditions can be used. Second, efficiency in the identification of corner cases is of crucial importance for system developers. Third, the scenario space coverage represents how the permutation coverage is achieved within the physically possible parameter space. Fourth, the safety assessment approaches can be distinguished based on the system applicability for perception, prediction and planning modules. Fifth, the computational feasibility shows the applicability of the assessment approach for ADFs, e.g. truck platooning. Sixth, the reliability of statements indicates which evidence exists for the safety arguments. Then, the extrapolation of risk metrics illustrates the scalability by transferring the microscopic assessment results to macroscopic assessment.

Next, the modeling dependencies define the model validation efforts to argue the safety statement. After that, the dependence on safety drivers points to the challenges to launch ADFs under supervision of safety drivers. Finally, the closed-loop interactions define the approach validity for predication and

[9] https://www.ki-absicherung-projekt.de/

planning modules in which future trajectories of other road users are influenced by the automated vehicle's behavior.

Evaluation criteria	Safety assessment approaches			
	Shadow-mode approach	Formal safety verification	Traffic simulation-based approach	Scenario-based approach
Representativeness of scenarios	○	○	◒	●
Identification of corner cases	●	◓	◒	◔
Scenario space coverage	◒	●	◔	◒
System applicability	◒	◔	●	●
Computational feasibility	◒	○	◓	◓
Reliability of statements	◔	●	◒	◓
Extrapolation of risk metrics	◒	○	◔	●
Modeling dependencies	◔	◔	◓	◓
Dependence on safety drivers	●	○	◓	◒
Closed-Loop interactions	○	●	◒	◒

● : optimal ◓ : fairly optimal
◒ : natural ◔ : fairly poor
○ : not applicable

Table 3.4: Comparable evaluation according to the state-of-the-art of categorized safety assessment approaches [SWB+20, Jun19].

3.3.1 Shadow Mode Approach

Wachenfeld proposes the use of stochastic methods for introduction of automated driving without driver supervision. Here, the safety of these automated driving levels cannot be proven statistically using accumulated kilometers of physical onroad testing under consideration of an estimated uncertainty

[Wac17]. Wang and Winner [WW19] describe an approach in which the ADF in end-customer vehicles is operated passively and is known as shadow mode. The passive function is equipped with the relevant sensory perception and logging devices, but does not have an access to the actuators [Wan21]. Therefore, a large fleet of human-driven vehicles can collect raw sensor data for HW/SW-open-Loop reprocessing. In addition, these open-Loop recordings are retrieved to validate prediction and planning algorithms, so that the recorded data can be used in a HW/SW-in-the-closed-Loop simulation environment. In this way, the simulation can be used to evaluate the decisions of the ADF in passive mode and thus determine the required safety level [Kar20]. One of the major advantages is the absence of safety risks during the data acquisition phase. However, one of the main disadvantages is that the behavior of the objects in the simulation does not correspond to reality because other road users also plan and execute their actions based on the actions of the active driving function [Wag21]. Therefore, the results of the simulation need to be argued. The car manufacturer (Tesla) announces to use the shadow mode testing to validate new ADFs and new software versions of existing systems [FBK17, Hul16].

3.3.2 Formal Safety Verification

Althoff et al. [ASB07] introduced a reachability analysis which argues the safety of automated vehicles. Reachability analysis seeks to determine the states that a system can reach from given initial states and possible inputs and parameters [Alt10]. Through formal verification, a mathematical model is used to formally demonstrate the safety of trajectory planning across the entire ODD [AKM17]. The approach distinguishes between perceptual and planning concerns. Using a safety envelope, the trajectory planner model is defined as intrinsically safe [AL18]. The valid and explainable set of traffic rules is aligned with the employed model that restricts the actions of the actual trajectory planner and increases the transparency of the behavior planning.

The IEEE 2846 is a standardization activity to define a formal rules-based mathematical model for automated vehicle decision making using discrete mathematics and logic. Arechiga [Are19] defines a set of rules for automated driving and other road users that are formally verified using a formal language called Signal Temporal Logic (STL). The argument is that the entire traffic system

will be collision-free if all road users follow these rules. Loos et al. [LPN11] used a formal proof calculus for safety verification by proofing safety separately for adaptive cruise control cases applied in a distributed car-control system. Nilsson et al. [N+15] verified an AEB system using closed-form expressions for robust avoidance scenarios. The closed-form expressions are derived based on worst-case performance as an optimization problem between FP and TP interventions (elucidated in chapter 2). Since the model and its parameterization take a central role, the modeling is a key challenge for this approach. The concept of safety envelopes can be illustrated by a cut-out test scenario. In the cut-out scenario, a vehicle in front suddenly leaves the lane to avoid a stopped vehicle in its lane.

The situation gives the AEB function a short time to detect and react accordingly. Figure 3.5 depicts a cut-out scenario in which the moving vehicle [obj2] is triggered at a cut-out distance $d_x^{co} = 45[m]$ between the stationary vehicle [obj1] and the moving vehicle [obj2] to change to the adjacent lane after a cut-out delay time $t_{co} = 2[s]$.

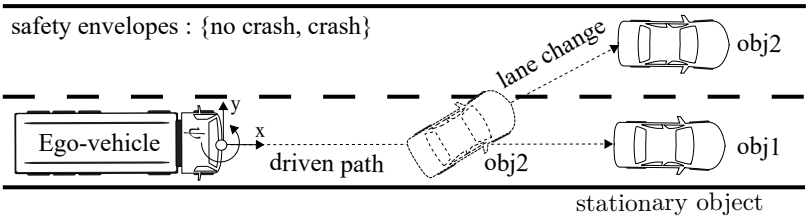

Figure 3.5: Schematic diagram of a cut-out driving scenario with the safety envelopes.

Figure 3.6 shows the pass/fail envelopes either as crash or no crash with ($\mathcal{E}_s^1$, $\mathcal{E}_s^2$ and $\mathcal{E}_s^3$) as escalation levels of the AEB function. The driving scenario with the different vehicles (ego, obj1 and obj2) is carried out on the HiL test bench and controlled by a test script based on a track co-ordinate system.

The HiL platform includes perception sensor models for the RADAR and monocular camera sensors that are used within the AEB function. Therefore, the stationary vehicle [obj1] is occluded through the moving vehicle [obj2], where the Ego-vehicle and the moving vehicle have the same constant velocities on the same lane $v_x^{ego} = v_x^{obj2} = \{30, 35, 40, 45\}[km/h]$. In the case of idealized

sensor models without occlusion effects, the cut-out scenario cannot be realized in which the RADAR sensor model provides the object list to the AEB function for the moving as well as stationary objects. The idealized sensor models have no occlusion effects, represented on the left hand of the figure 3.6. The cut-out delay time t_{co} [s] and distance d_x^{co} [m] parameters of cut-out scenarios have no influence on present crash events, while the idealized sensor model constantly recognizes the object [obj1] as a relevant stationary object.

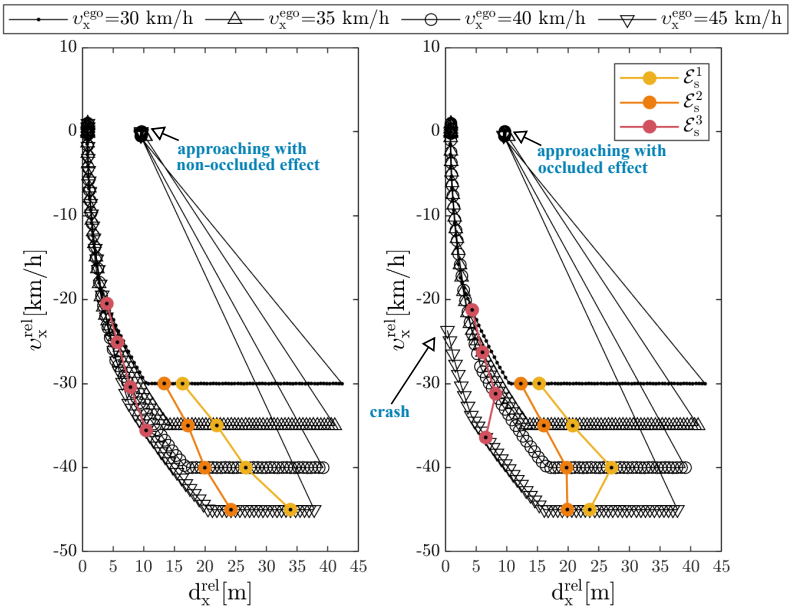

Figure 3.6: Safety envelopes based on comparison between simulation results of idealized sensor models (left) and high-fidelity sensor models (right) in the example of cut-out test scenarios.

The left side of the figure 3.6 shows the three types of escalations ($\mathcal{E}_s^1$, $\mathcal{E}_s^2$ and $\mathcal{E}_s^3$) that occur at different Ego-vehicle velocities $v_x^{ego} = \{30, 35, 40, 45\}$ [km/h], respectively, while a collision with the non-occluded object [obj1] does not occur. Due to the margin scale of the figure 3.6, though it might seem that the Ego-vehicle collides with the object [obj1] after stopping, this is not the case. A positive distance between the Ego-vehicle and the object [obj1] remains even

after the Ego-vehicle stops due to the emergency braking of the AEB function. The same test scenarios are repeated with high-fidelity sensor models to show the influence of cut-out scenario parameterization with a crash event in the case of $v_x^{ego} = 45$ [km/h], as illustrated on the right side of figure 3.6. Meanwhile, an occluded object [obj1] is simulated to less idealize the driving scenario, it is evident that the three types of escalations ($\mathcal{E}_s^1$, $\mathcal{E}_s^2$ and $\mathcal{E}_s^3$) happen at $d_x^{rel} = \{24, 20, 7\}$ [m], respectively, resulting in a crash (with $v_x^{ego} = 45$ [km/h]) due to insufficient time to respond. In chapter 4, the developed HiL test platform is elaborately discussed with its implementation.

3.3.3 Traffic Simulation-based Approach

The introduction of automated vehicles will change the flow of road traffic. Therefore, a macroscopic statement about the safety of automated vehicles can be obtained by the traffic simulation based approach [PSS19]. The concept of traffic simulation is to simulate not only a single driving scenario but a whole road network with numerous road users (so-called Agents). Kitajima et al. [K+19a] developed a multi-agent simulation for estimating the impact of automated vehicles on road traffic. Rösener et al. [R+18b] investigate the change in the occurrence frequency of scenarios to assess the safety performance of ADFs. Saraoglu et al. [SMJ19] present a framework called MOdel-Based Autonomous Traffic Simulation Framework (MOBATSim) for the analysis of traffic safety including automated vehicles with a focus on the fault-error-failure chain. In traffic simulation, the entire ODD can be simulated to increase the efficiency of the staged introduction of automated vehicles [Hal20]. Bach et al. introduce a model-based specification of driving scenarios with the example use case of an ACC system based on the abstraction of temporal and spatial information [BOS16, BHOS17]. Otten et al. extend the model-based scenario specification with an automated assessment and evaluation concept for stochastic digital test drives [OBW+18].

To speed up the required FOT of the automated vehicles in car-following scenarios, Zhao et al. propose an accelerated evaluation method using stochastic optimization and importance sampling methods [Zha16]. Although the simulation-based falsification is relatively similar to testing, a search is conducted within a logical scenario for a parameter set, where the ADF violates the requirements [AGZ18]. Koren et al. [K+18b] use a reinforcement learning

formulation to identify the trajectories that are likely to be critical or lead to collisions. Thus, the Adaptive Stress Testing (AST) framework searches for the most likely path to a failure event [CDDCm19, L+15, L+18b]. Gangopadhyay et al. [GKD+19] use a Bayesian optimization to learn parameter values by observing the system's output. Nabhan et al. [NSTH19] used a Random Forest model to detect the maximum amount of faulty scenarios in the search space.

The falsification concept can be demonstrated using a driving scenario with crossing pedestrians for the verification of a multi-sensor fusion unit utilizing the implemented HiL test platform; Chapter 4 discusses such a concept in detail. The driving scenario distinguishes between the own lane and the adjacent lane with two classification areas. The pedestrian [obj1] crosses from the path of the moving Ego-vehicle to the adjacent lane and another pedestrian [obj2] comes from the opposite direction at the same velocity $v_x^{ego} = v_x^{obj1} = v_x^{obj2} = 5[\mathrm{km/h}]$, as seen in figure 3.7.

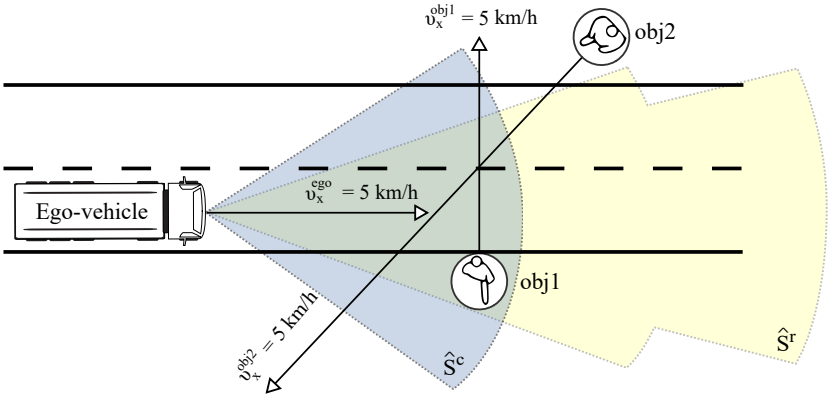

Figure 3.7: Schematic diagram of a pedestrian crossing scenario with simulation-based falsification [ESW+16].

The falsification is implemented using the Lemniscate of Bernoulli's equation in the track co-ordinate system to obtain the lying-eight as a noise offset to the camera sensor model for object detection. The parametric equations for the

Lemniscate with a half-width η[m] are $\tilde{d}_x^{rel}$[m] and $\tilde{d}_y^{rel}$[m] of the camera-based detection, as shown in equation 3.1.

$$\tilde{d}_x^{rel} = \left(\frac{\cos(t)}{\sin^2(t)+1}\right) * \eta, \quad \tilde{d}_y^{rel} = \left(\frac{\cos(t) * \sin(t)}{\sin^2(t)+1}\right) * \eta \tag{3.1}$$

Figure 3.8 depicts the object detection from the RADAR sensor model ($\hat{S}^r_{obj1}$ and $\hat{S}^r_{obj2}$) and the camera sensor model ($\hat{S}^c_{obj1}$ and $\hat{S}^c_{obj2}$). The sensor models assign the track IDs based on the criticality of the two pedestrians [obj1] and [obj2]. While their walking paths cross each other, their track IDs are swapped.

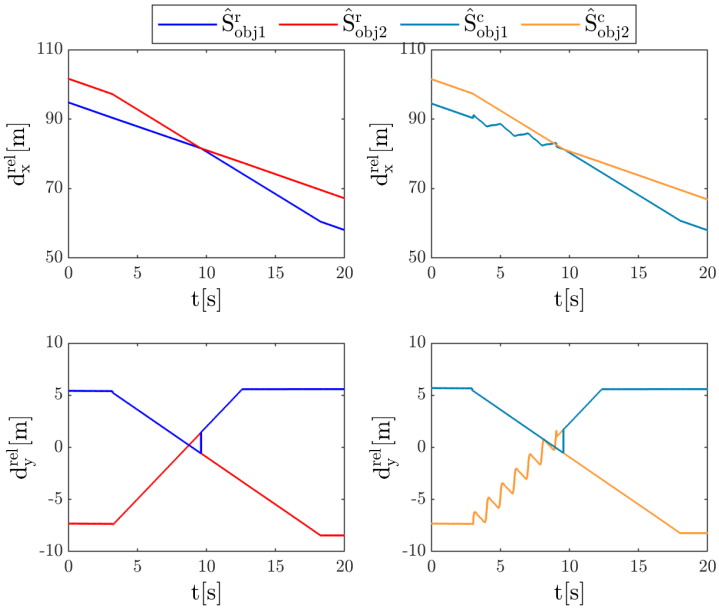

Figure 3.8: Falsification of object detection (right) for verification of multi-sensor fusion using the example of a pedestrian crossing scenario compared to object detection without falsification (left) [ESS+16].

Thereby, the Lemniscate noise model is applied to the object motion from the camera sensor model, so that the covariance matrix of the Mahalanobis distance can be identified. The covariance matrix represents the uncertainty of

the state estimation of longitudinal and lateral information. The Mahalanobis distance of the object is manipulated using the Lemniscate noise model until the fusion algorithm no longer associates the tracked object. Therefore, the covariance matrix can be determined over the entire ellipse.

3.3.4 Scenario-based Approach

In Japan, Antona-Makoshi et al. [J+19] explains the impact of the scenario-based approach on the safety assurance process for autonomous vehicles. Krishna [Kri19] extracts the corner case scenarios that occur on the roads of Singapore for automated vehicles. The corner-case scenario is one in which two or more parameter values are each within the capabilities of the system, but together constitute a rare condition that challenges its capabilities [fS+20], [Pon21]. In the Netherlands, the StreetWise methodology provides scenario-based safety validation of automated vehicles using a database of real-world scenarios [EPG+18].

In Germany, the PEGASUS project relies on scenario-based testing to derive scenarios from system knowledge, domain modeling and field observation [WLFM19, PEG19]. The scenarios are applicable as test cases, which are executed and evaluated in the simulation or on test tracks [DG18]. The corresponding criticality metrics are employed to determine the automated driving capabilities [BKB+19, Hau21]. Schuldt [Sch17b] proposes the use of equivalence classes, boundary value analysis and combinational methods for identifying the representative driving scenarios. The proposed approach of Schuldt provides a systematic generation of test cases, but lacks a method to determine a meaningful test coverage [S+18a]. Schuldt motivates a scenario-based test process and presents a systematic test case generation by use of a four layer abstraction model. The concept of criticality analysis can be explained using the example of an off-tracking driving scenario.

Figure 3.9 illustrates the rear axle path prediction of the off-tracking phenomenon based on Tractrix motion [ESS+19b]. The off-tracking phenomenon occurs when a vehicle turns and its rear wheels do not follow the same path as its front wheels [RA12].

The low-speed transient off-tracking describes the lateral offset between the turning paths of the front and rear axle before steady-state off-tracking is

reached. While commercial vehicles traverse shorter curves or curves of smaller radius, non-steady-state off-tracking increases gradually at low-speed vehicle turning scenarios. Thereby, the swept path width is the difference in paths between the outside front tractor tire and the inside rear trailer tire.

Definition 3.6 (**Tractrix Motion**)**:** The vehicle off-tracking behavior at low speeds is approximated by a Tractrix motion, where the rear axle of a tractor-semitrailer combination truck follows a given steering curve for low speed turning scenarios. The rear axle is always moving in the direction of the front axle based on a given wheelbase d_w [m] [RA12]. The velocity of the rear axle depends on the direction of the velocity vector of the front axle.

Definition 3.7 (TTC)**:** It refers to the time required for two objects to collide; if they are moving at their current velocity and following the same path [W+16].

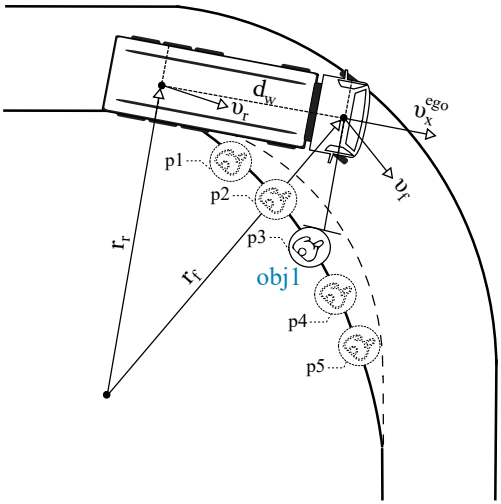

Figure 3.9: Tractrix motion for low-speed transient off-tracking within a cornering scenario [ESS+19b].

Figure 3.10 shows the criticality analysis with a stationary pedestrian at different positions (p1, p2, p3, p4 and p5) and velocities of v_x^{ego} [km/h]. The

off-tracking scenario is executed on the HiL test bench and managed by a test script for the right-hand turns with the SGA function.

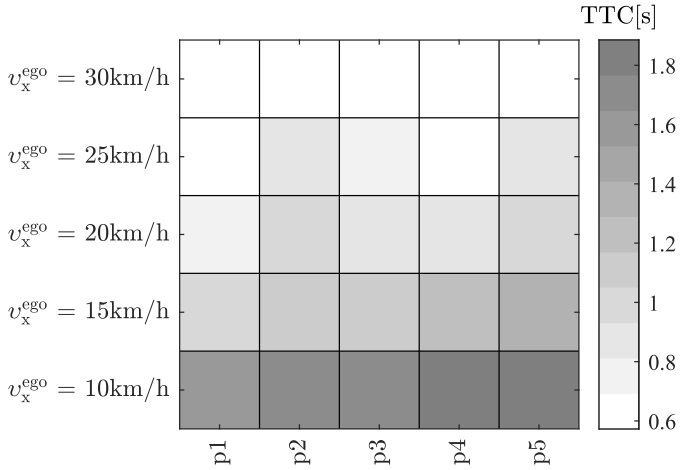

Figure 3.10: TTC criticality analysis with a stationary pedestrian in the Tractrix zone using the SGA function within a cornering scenario [ESS+19b].

4 Hardware-in-the-Loop Simulation

In the automotive industry, HiL-based test methods are commonly used and they provide a significant advantage for the functional verification of software components on the ECU in the laboratory [Düs10]. The integration capability of perception sensors in HiL test benches reduces the dependence on simulation models and their validation in contrast to those in SiL environments. Accordingly, the Device under Test (DuT) can be evaluated at the hardware level and also its influence on the system, such as poor performance due to limited computing power or memory space [FBS09]. In addition, various effects, such as message loss, time delays and limited signal value ranges due to communication between hardware components, can be investigated. There are three main architecture levels of the HiL test system (component, subsystem and system). At the component level, each individual ECU can be verified with respect to its functional correctness. While ADFs commonly include several ECUs, the data flow can be verified at the subsystem or system level. Although the integration of diverse ECUs in a virtual test environment increases realism in simulation, the test system complexity can be unmanageable [NGB13]. Consequently, the design of the HiL test system needs to meet the requirements derived from the desired test objective and not for every degree of realism or fidelity [Tel14].

4.1 Simulation Co-ordinate Systems

The entire simulation relies on the right-handed inertial co-ordinate system $\mathcal{R}_I$, where the x indicates the east direction, the y the north direction and the z the elevation direction. The rotation matrix $\mathbf{R}_{\mathrm{E}}^{\mathrm{I}}$ is generated from the intrinsic Euler angles by multiplying the three matrices generated by rotations around the axes.

According to the ISO 8855:2011 for right-handed vehicle co-ordination systems [fS+11c], the rotation matrix $\mathbf{R}_{\mathrm{E}}^{\mathrm{I}}$ follows a z-, y- and x- rotation sequence, as depicted in equation 4.1.

$$\mathbf{R}_{\mathrm{E}}^{\mathrm{I}} = \begin{bmatrix} \cos\psi & -\sin\psi & 0 \\ \sin\psi & \cos\psi & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} \cos\theta & 0 & \sin\theta \\ 0 & 1 & 0 \\ -\sin\theta & 0 & \cos\theta \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ 0 & \cos\phi & -\sin\phi \\ 0 & \sin\phi & \cos\phi \end{bmatrix} \tag{4.1}$$

4.1.1 Ego-vehicle Co-ordinate System

The origin of the Ego-vehicle co-ordinate system $\mathcal{R}_{\mathcal{E}}$ in neutral loading conditions is on road level at the center of the truck rear axle. The Ego-vehicle co-ordinate system determines the position of environmental perception sensors and detected objects. The rotation matrix $\mathbf{R}_{\mathrm{S}}^{\mathrm{E}}$ is an identity matrix, where the sensor co-ordinate orientation corresponds to the Ego-vehicle orientation, as illustrated in equation 4.2. The origin of the RADAR sensor co-ordinate system $\mathcal{R}_{\mathcal{S}}$ is at the installation position of the RADAR, mounted on the front of the Ego-vehicle.

$$\mathbf{R}_{\mathrm{S}}^{\mathrm{E}} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \tag{4.2}$$

The rotation matrix $\mathbf{R}_{\mathrm{I}}^{\mathrm{S}}$ describes the rotation from the inertial co-ordinates to the RADAR sensor co-ordinates, as shown in equation 4.3.

$$\mathbf{R}_{\mathrm{I}}^{\mathrm{S}} = \left(\mathbf{R}_{\mathrm{E}}^{\mathrm{I}}.\mathbf{R}_{\mathrm{S}}^{\mathrm{E}}\right)^{-1} \tag{4.3}$$

The transformation matrix $\mathbf{T}_{\mathrm{I}}^{\mathrm{S}}$ is calculated from the inversion matrix of the sub-transformations from Ego-vehicle into inertial co-ordinate systems $\mathbf{T}_{\mathrm{E}}^{\mathrm{I}}$ and from sensor into Ego-vehicle co-ordinate systems $\mathbf{T}_{\mathrm{S}}^{\mathrm{E}}$, as represented in equation 4.4. The distance vector $\mathbf{t}_{\mathrm{E}}^{\mathrm{I}}$ specifies the distance between the origin points of Ego-vehicle and inertial co-ordinate systems, while the distance vector $\mathbf{t}_{\mathrm{S}}^{\mathrm{E}}$ describes the distance between origin points of sensor and Ego-vehicle co-ordinate systems.

$$\mathbf{T}_{\mathrm{I}}^{\mathrm{S}} = \left(\mathbf{T}_{\mathrm{E}}^{\mathrm{I}} . \mathbf{T}_{\mathrm{S}}^{\mathrm{E}}\right)^{-1} = \left(\begin{bmatrix} \mathbf{R}_{\mathrm{E}}^{\mathrm{I}} & \mathbf{t}_{\mathrm{E}}^{\mathrm{I}} \\ \mathbf{0}^{\mathrm{T}} & 1 \end{bmatrix} \cdot \begin{bmatrix} \mathbf{R}_{\mathrm{S}}^{\mathrm{E}} & \mathbf{t}_{\mathrm{S}}^{\mathrm{E}} \\ \mathbf{0}^{\mathrm{T}} & 1 \end{bmatrix}\right)^{-1} \tag{4.4}$$

4.1.2 Object Co-ordinate System

The origin of the object co-ordinate system $\mathcal{R}_{\mathcal{O}}$ corresponds to the origin point of the detected object in front. Equation 4.5 represents the transformation matrix $\mathbf{T}_{\mathrm{O}}^{\mathrm{I}}$ from the object to inertial co-ordinate systems, where the distance vector $\mathbf{t}_{\mathrm{O}}^{\mathrm{I}}$ describes the distance between origin points of object and inertial co-ordinate systems.

$$\mathbf{T}_{\mathrm{O}}^{\mathrm{I}} = \begin{bmatrix} \mathbf{R}_{\mathrm{O}}^{\mathrm{I}} & \mathbf{t}_{\mathrm{O}}^{\mathrm{I}} \\ \mathbf{0}^{\mathrm{T}} & 1 \end{bmatrix} \tag{4.5}$$

The origin of the detection co-ordinate system $\mathcal{R}_{\mathcal{D}}$ allocates on road level at the center of the rear bumper of the detected preceding object. Figure 4.1 depicts the various simulation reference co-ordinate systems of environmental perception sensors [Amm15].

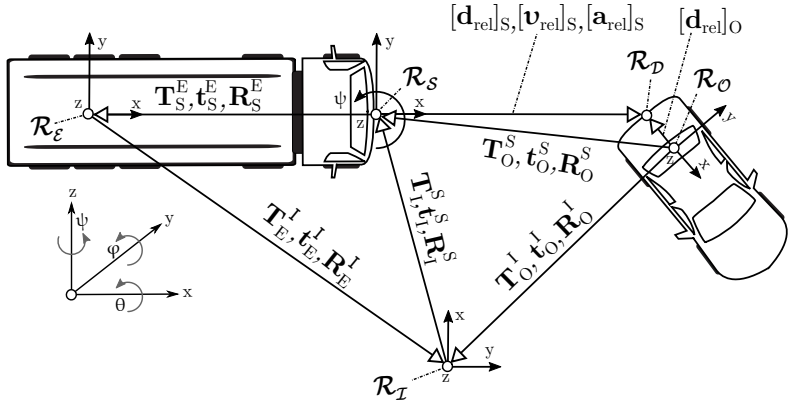

Figure 4.1: Simulation reference co-ordinate systems of environmental perception sensors in the example of detection of a vehicle ahead.

The equation 4.6 describes the Ego-vehicle velocity and acceleration vector components at the sensor co-ordinate system $\mathcal{R}_S$ as follows:

$$[\boldsymbol{\upsilon}^{\text{ego}}]_S = \begin{bmatrix} \upsilon_x^{\text{ego}} \\ \upsilon_y^{\text{ego}} \\ \upsilon_z^{\text{ego}} \end{bmatrix}_S = \mathbf{R}_I^S . [\boldsymbol{\upsilon}^{\text{ego}}]_I , [\mathbf{a}^{\text{ego}}]_S = \begin{bmatrix} a_x^{\text{ego}} \\ a_y^{\text{ego}} \\ a_z^{\text{ego}} \end{bmatrix}_S = \mathbf{R}_I^S . [\mathbf{a}^{\text{ego}}]_I \tag{4.6}$$

The equations 4.7, 4.8 and 4.9 depict the relative vector movement components at the sensor co-ordinate system $\mathcal{R}_S$ as follows:

$$[\mathbf{d}^{\text{rel}}]_S = \begin{bmatrix} d_x^{\text{rel}} \\ d_y^{\text{rel}} \\ d_z^{\text{rel}} \end{bmatrix}_S = \left(\mathbf{T}_I^S . \mathbf{T}_O^I\right) . [\mathbf{d}^{\text{rel}}]_O \tag{4.7}$$

$$[\boldsymbol{\upsilon}^{\text{rel}}]_S = \begin{bmatrix} \upsilon_x^{\text{rel}} \\ \upsilon_y^{\text{rel}} \\ \upsilon_z^{\text{rel}} \end{bmatrix}_S = \left(\mathbf{R}_I^S . \mathbf{R}_O^I\right) . [\boldsymbol{\upsilon}^{\text{obj}}]_O - [\boldsymbol{\upsilon}^{\text{ego}}]_S \tag{4.8}$$

$$[\mathbf{a}^{\text{rel}}]_S = \begin{bmatrix} a_x^{\text{rel}} \\ a_y^{\text{rel}} \\ a_z^{\text{rel}} \end{bmatrix}_S = \left(\mathbf{R}_I^S . \mathbf{R}_O^I\right) . [\mathbf{a}^{\text{obj}}]_O - [\mathbf{a}^{\text{ego}}]_S \tag{4.9}$$

The relative longitudinal movement components are represented by $d_x^{\text{rel}}[\text{m}]$, $\upsilon_x^{\text{rel}}[\text{m/s}]$ and $a_x^{\text{rel}}[\text{m/s}^2]$, while the relative lateral movement components are defined by $d_y^{\text{rel}}[\text{m}]$, $\upsilon_y^{\text{rel}}[\text{m/s}]$ and $a_y^{\text{rel}}[\text{m/s}^2]$. Also, the variables $d_z^{\text{rel}}[\text{m}]$, $\upsilon_z^{\text{rel}}[\text{m/s}]$ and $a_z^{\text{rel}}[\text{m/s}^2]$ represent the relative vertical movement components. The relative start reference co-ordinate system $\mathcal{R}_R$ is a special co-ordinate system to allow the zero initialization of the external vehicle dynamics simulation at its own origin points, when the Ego-vehicle needs to be initialized at a non-zero location or orientation within the driving scenario. In addition, the track co-ordinate system $\mathcal{R}_T$ is used to allocate and control objects through the reactive test automation. The co-ordinate axes are defined on the road, which are applied along the reference road center.

Equation 4.10 depicts the curvature of the driven distance of the Ego-vehicle in the sensor co-ordinate system.

$$\kappa_{\text{ego}} = \frac{\psi}{v_{\text{x}}^{\text{ego}}} \tag{4.10}$$

In case of the simulation of an eHorizon sensor via an OpenDRIVE[ASA21] database, the geographic co-ordinate system should be applied to the geographical data based on World Geodetic System 1984 (WGS84) co-ordinates. The OpenDRIVE provides an open file format for the logical description of road networks[ASA21]. The camera co-ordinate system $\mathcal{R}_C$ represents the co-ordinate axes of the physical camera ECU. If the camera sensor is stimulated via a monitor, the monitor co-ordinate system $\mathcal{R}_M$ is used, which represents a 2D co-ordinate system. The origin of the monitor is located in the lower left corner [Nen14]. The co-ordinates of the monitor are normalized and valid in a range between 0.0 and 1.0.

4.2 Vehicle Dynamics Simulation

The dynamic behavior of heavy-duty trucks differs considerably from that of passenger cars due to the geometry and dimensions of the truck-trailer combinations. Also, the loading conditions have a considerable influence on the longitudinal and vertical positions of the vehicle center of gravity. The wheel configuration variants influence the off-tracking behavior of the trailer in cornering situations with different wheelbases, drawbars and kick angles. Pitch and roll movements need to be taken into account, which are strongly influenced by the aforementioned factors. Therefore, vehicle dynamics is an important topic to discuss. Moreover, a CMV consists of a driver's cabin suspended from the vehicle frame in order to reduce mechanical road excitation. As a result, the environmental perception sensors have to cope with static and dynamic tolerances of truck-trailer combinations. The static tolerances relate mainly to sensor mounting position, vehicle type, tire pressure, type of cab suspension and load condition [SMAN08]. Additionally, the dynamic tolerances refer to the tractor cab movements. The vertical movement shifts the sensor height in relation to the road surface. The cabin pitch and roll angles change the optical axis position of the forward-facing camera mounted behind the windshield in relation to the horizontal axis.

Moreover, the steady-state behavior of a truck is determined in accordance with ISO 14792:2011 by the steady-state circular tests [fS+11b]. The change in the required steering as a function of lateral acceleration is represented by the under-steer gradient, which shall be positive for heavy vehicles. The negative under-steer gradient leads to instability at a certain critical velocity [MP17].

4.2.1 Truck Cabin Simulation

Figure 4.2 shows the pitch movements when braking in straight ahead direction using a parameterized simulation model of the Mercedes-Benz Actros [Tru18] tractor with 4x2 axle configuration and a weight of 18 tonnes [Mar04]. The truck's braking and driving dynamics differ from those of a passenger car. The pneumatic brake system has a time delay and slower response than the hydraulic brake system in passenger cars. Therefore, the load condition influences both dynamic stability and braking dynamics, whereby the difference between full and empty weight is very large.

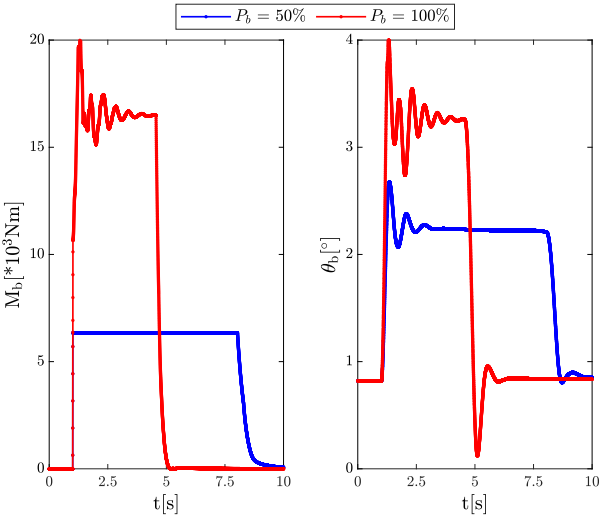

Figure 4.2: Step-shaped half and full braking in a straight line with Mercedes-Benz Actros 4x2 tractor simulation model at a constant longitudinal velocity of 80 [km/h] including braking torques (left) and pitch angles (right).

Moreover, the process of stopping a CMV in an emergency requires a complex interaction between the braking system, the CMV tires, the CMV's dimensions, the loading conditions and the road surface characteristics. The realistic simulation of CMV dynamics enables the virtual movements of the environmental perception sensors to investigate the effects of cabin dynamics in braking situations. Therefore, the multi-body vehicle dynamics need to simulate the roll and pitch movements of the truck's tractor cabin. Meanwhile, the pitching movements during braking affect the detection performance of the perception sensors. If the brake pedals are set in a step-wise manner to P_b = 50% and 100%, braking torques are applied at half and full braking with $M_b \approx 6.3$ [kNm] and 16.5 [kNm] respectively. The pitch angles of the truck's tractor cabin are shifted from $\theta_b \approx 2.2°$ to 3.2° with an initial offset of 1° at a constant longitudinal velocity v_x^{ego} = 80 [km/h]. Figure 4.3 represents the roll movements when driving around a tight curve using the same parameterization of the employed simulation model.

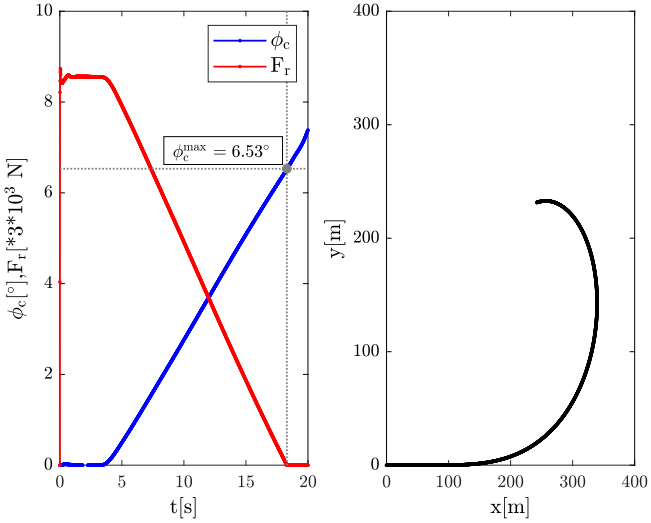

Figure 4.3: Driving around a tight curve at lateral acceleration of 6 [m/s^2] and longitudinal velocity of 80 [km/h] including roll movement compared to weight force on the rear axle (left) and Ego-vehicle trajectory (right).

The roll-over angle of the truck's tractor cabin can be estimated on the basis of the tight curve maneuvers according to ISO 16333:2011 [fS+11a]. The calculation of the tire lift-off and roll-over limits specifies the stability limit of the vehicle when the rear axle tires are lifted off the road. Thus, ISO 16333:2011 presents a test method for estimating the maximum lateral acceleration that a CMV could withstand in steady-state turning maneuvers without rolling over [fS+11a]. From the right side of the figure 4.3, the Ego-vehicle accelerates in a tight curve maneuver, which is expressed in the left side of the figure 4.3 in terms of a corresponding roll angle of 6.53° with a weight of 0 [kN], which means that the rear axle of the Ego-vehicle does not exert any force. The roll movement of the truck's tractor cabin is shifted to $\phi_c^{max} = 6.53°$ at actual weight on rear axle with $F_r = 0$ [kN], lateral acceleration with $a_y^{ego} = 6$ [m/s^2] and longitudinal velocity with $\upsilon_x^{ego} = 80$ [km/h].

4.2.2 Run-time Analysis of Vehicle Dynamics

The applied vehicle dynamics simulation uses an explicit Euler integration method based on Simulink® with a sampling rate of 1000[Hz]. While the turn-around time represents the computing time required to calculate the executed tasks, the sampling time indicates the time interval for the integration step of the simulation. Therefore, the turn-around time should be shorter than the sampling time to ensure that the simulation is executed in real-time. The explicit Euler integration method utilizes the equations 4.11 and 4.12 for each Simulink® block at each simulation time step t_s= 1[ms]. The [mdlOutputs] task comprises the output part and the [mdlUpdate] task contains the update part of the discrete state-space equation. The [mdlUpdate] task updates the block inputs u_s with the discrete state x_{d+1} for each Simulink® block that has a discrete state x_d, while the [mdlOutputs] task calculates the block outputs y_o

$$y_o = \text{mdlOutputs}(u_s, x_d, t_s) \tag{4.11}$$

$$x_{d+1} = \text{mdlUpdate}(u_s, x_d, t_s) \tag{4.12}$$

As distinct from a general-purpose operating system, a Real Time Operating System (RTOS) is expected to meet computational time constraints when executing software applications. The implemented HiL simulator uses a 2.1 GHz Intel Quad-Core Ivy-Bridge processor with a VxWorks® RTOS.

Definition 4.1 (VxWorks®): An embedded RTOS developed as proprietary software by Wind River Systems. The VxWorks® kernel utilizes a shared memory model to control the communication between task states based on preemptive scheduling, where a first-in-first-out policy is used to schedule all tasks. The binary semaphores of VxWorks® provide a synchronization between the tasks controlled by the functions [semTake] or [semGive]. The semaphore is a binary variable for multi-process access control of a common resource within a RTOS, where a binary semaphore acts as a flag that can be blocked or released [HVJ14].

The test cases are graphically modeled with special state machines on the host computer and transformed as byte code to the virtual machine task [JavaVM]. The compiled test cases thus run on the real-time HiL simulator using the [JavaVM] task and reactively control the synthetic driving scenario. Figure 4.4 shows the turn-around time t_ϱ[s] of each task at the HiL simulator, where the total turn-around time does not exceed the sampling time $t_s = 0.001$[s].

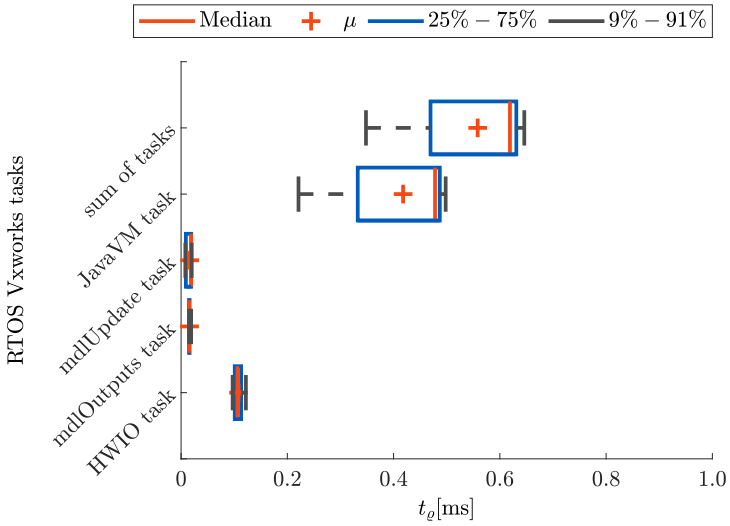

Figure 4.4: Turn-around time of task of vehicle dynamics, hardware in-/output and test automation under hard real-time conditions.

In this way, the [HWIO] task sets up the communication between the HiL simulation and the DuT. The entire execution time at the HiL is represented by the turn-around time and can be summarized in terms of [mdlOutputs], [mdlUpdate], [JavaVM] and [HWIO] tasks. The highest processing usage is provided by the [JavaVM] task, and then by the [HWIO] task, as illustrated in figure 4.4. For this reason, an optimization is required to use the available time and hardware resources better for further necessary simulation components. Consequently, the employed HiL test bench utilizes a distributed heterogeneous co-simulation environment in order to realize a real-time capable system architecture. The co-simulation environment provides the core functionality of task and data management for the simulation of road traffic as well as Ego-vehicle driver and environmental perception sensors at a different sampling time.

4.3 Road Traffic Simulation

The fidelity of road traffic simulations can be divided into four categories: Macroscopic, mesoscopic, microscopic and nanoscopic. First, macroscopic simulations characterize the traffic flow, velocity and density of traffic (e.g. the number of road users who travel a certain distance per unit of time). Therefore, macroscopic fidelity is well suited for the analysis of large or complex road networks. Second, traffic units can be grouped in the mesoscopic models, where each group is treated as a single traffic unit (e.g. queues of vehicles). Third, the microscopic models simulate the behavior and interactions of each simulated traffic unit individually with specific state variables such as position, velocity and acceleration. The rules of vehicle behavior such as speed and lane changes are also taken into account. Fourth, an additional level of detail in nanoscopic models is achieved by dividing each vehicle into a number of sub-units. The nanoscopic model thus allows an extension of the vehicle dynamics modeling and the driver behavior fidelity. In this way, vehicle dynamics, complex driver decision-making processes or interaction with the vehicle environment can be modeled in more detail. In general, the computing time for traffic simulation increases considerably with increasing fidelity. Consequently, the HiL test system requires a compromise between fidelity level and computing effort for real-time traffic simulation under the given real-time conditions.

4.3.1 Multi-rate Co-simulation Setting

The developed real-time co-simulation platform is utilized to evaluate the ADFs in a closed-loop HiL configuration. The Functional Mock-up Interface (FMI) co-simulation manages the scheduling and data exchange between multiple simulation units with independent solvers using VxWorks® callback routines [Sch16b], [Sch17]. As a result, the FMI integrates the heterogeneous modules to make one real-time-capable application for the HiL simulation. Figure 4.5 illustrates the co-simulation integration scheme to provide the Ego-vehicle dynamics simulation with contact point information at a sampling rate of 1000[Hz]. In order to place the Ego-vehicle correctly in the 3D space, the computation of the road contact is necessary.

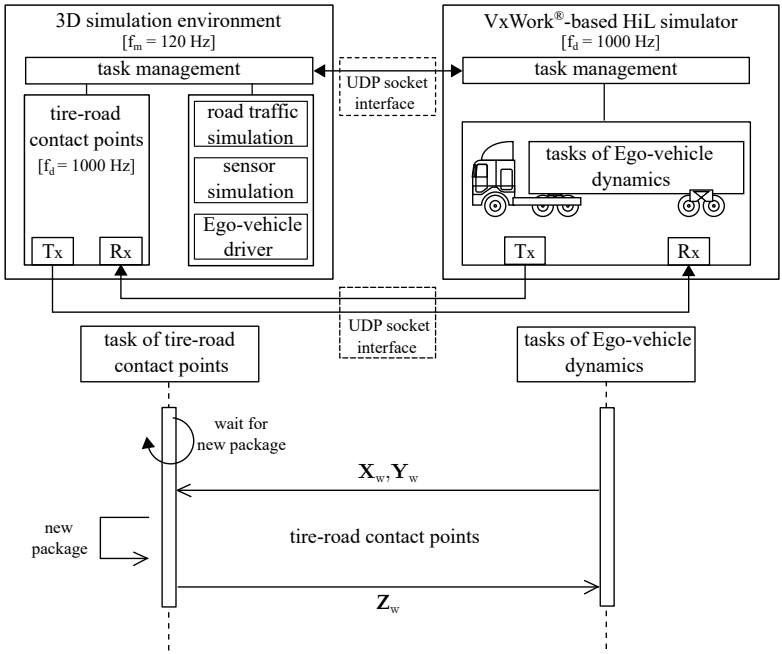

Figure 4.5: Sequence and schematic diagrams of the road traffic co-simulation with exchange of contact point information using asynchronous real-time simulation.

The number of contact points between the tires of the truck-trailer combination and the road surface is 10 points (one contact point per tire) as the maximum number of contact points for the applied vehicle dynamics model, described in section 4.2. The real-time vehicle dynamics simulation needs contact point information at a higher frequency $f_d = 1000[Hz]$ than the frequency of the simulation task management $f_m = 120[Hz]$. Real-time operation in asynchronous mode utilizes a common time domain between simulation components to run them in real-time. The gradients are computed by each simulation component to handle and align asynchronously computed results. The data packages of tire-road contact points introduce the road profile in correspondence of the tire contact points. The task of Ego-vehicle dynamics interacts with the task of tire-road contact points to provide the tire longitudinal and vertical positions $\mathbf{X}_w$, $\mathbf{Y}_w$ respectively. The vertical displacement $\mathbf{Z}_w$ of all road-tire contact points is determined and sent back to the task of Ego-vehicle dynamics, especially in uphill/downhill driving situations. The vector of road's longitudinal gradient $\boldsymbol{\alpha}_w^{long}$, lateral slope $\boldsymbol{\alpha}_w^{lat}$ and friction coefficient $\boldsymbol{\mu}_w$ are assumed with constant values.

4.3.2 Integration of HiL Simulation Modules

The data management through equivalent FMI implementation provides the co-simulation scheme with high portability between various HiL simulation platforms and enables the run-time fault injection scheme triggered by the test execution, where as, the FMI standardization of interfaces are set in table 4.1. The road traffic simulation supports the traffic scenarios with visual and logical databases stored in an eXtensible Markup Language (XML) format [vNC14]. The scenario database is based on de-facto standard formats OpenDRIVE [ASA21] and OpenSCENARIO [ASA22] for specifying road networks and dynamic contents respectively [vNCDW09]. While OpenDRIVE road networks assign static features to the driving scenario (e.g. lane, traffic sign, intersection, junction, crossing, etc.) [ASA21], OpenSCENARIO databases provide the synthetic scenario with the detailed dynamic contents of objects (e.g. actor, object, trigger, action, etc.) [ASA22]. Due to performance reasons, the change between a non-animated and an animated pedestrian depends on the distance between the Ego-vehicle position and the pedestrian position. If the distance to the pedestrian is less than a certain minimum distance, the animated pedestrian is used.

FMI category	**Description**
Actor	vehicle: • vehicle category (e.g. car, truck, motorcyclist and bicyclist) • position (x[m], y[m], z[m]) and orientation (ϕ[°], θ[°], ψ[°]) • driver behaviors (e.g. distraction and drowsiness) • vehicle behaviors (e.g. overtaking and rear-ending) • wheel information (e.g. steering angle, radius and forces) • power-train information (e.g. speed and torque) pedestrian: • pedestrian category (e.g. single and group) • position (x[m], y[m], z[m]) and orientation (ϕ[°], θ[°], ψ[°]) • pedestrian behaviors (e.g. crossing, inattentiveness and failure to obey traffic laws) • gesture and motion pattern (e.g. run and walk)
Road traffic	road information: • road geometry (e.g. curvature information, intersections, roundabouts and rural areas) • road condition (e.g. road damage, uneven surfaces and road construction) • road type (e.g. urban, rural and highway) • traffic condition (e.g. traffic sign, traffic light, high speed limit, heavy flows and potential accidents) • lane marking (e.g. type, color, broken/missing markings and irregular lane/road shapes) • drive lane information (e.g. width and ID) • static objects (e.g. fence, pole, vegetation and curbstone)
Camera ECU	image generator: • image height and width • depth and camera information ambient conditions: • time of day, sky state and visiblity • illumination (e.g. shadow, night, dawn/dusk and directly facing the sun) • road conditions (e.g. dry and wet) • weather conditions (e.g. fog, rain and snow)
RADAR ECU	sensor stimulation: • Over-The-Air movable antennas for the horizontal positions • distance, velocity, and size of the RADAR objects • radial distance, Doppler velocity, azimuth angle and elevation angle of the RADAR detections.
Virtual sensor	sensor simulation: • data communication (e.g. detection, feature and object level) • sensor position and orientation • relative movement of recognized objects (d_x^{rel}[m], d_y^{rel}[m], v_x^{rel}[m/s], v_y^{rel}[m/s], a_x^{rel}[m/s^2], a_y^{rel}[m/s^2])

Table 4.1: List of information categories of data management with equivalent FMI modules via well-defined in-/output interfaces.

The developed real-time co-simulation platform is utilized to evaluate the automated driving ECUs in a HiL configuration, as demonstrated in figure 4.6.

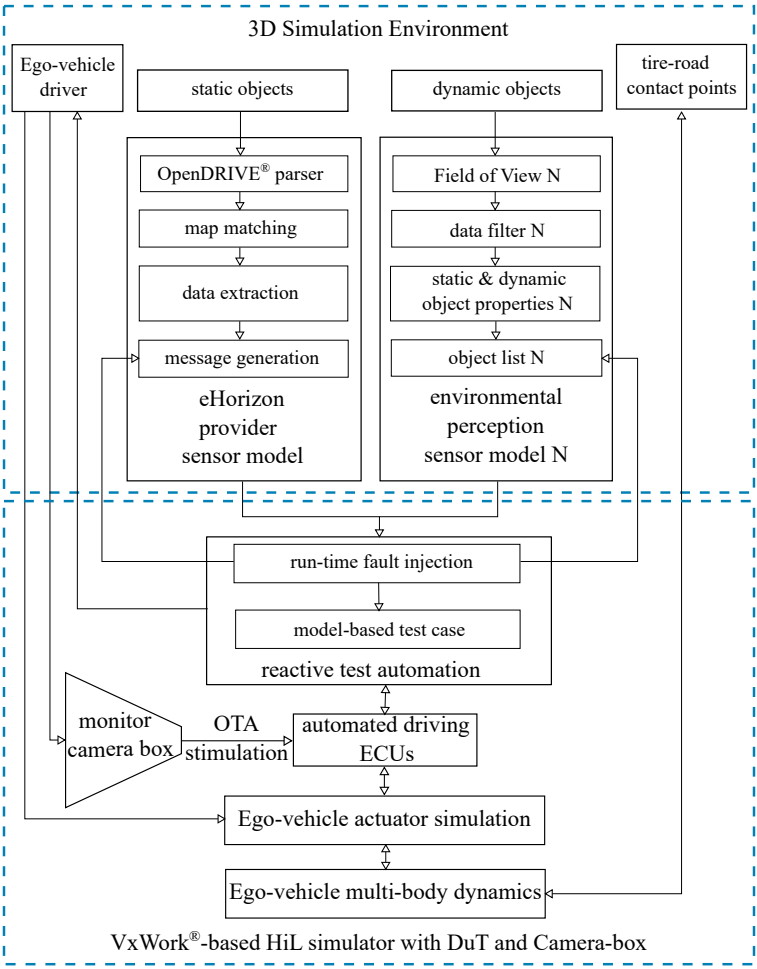

Figure 4.6: Schematic diagram of the dynamic behavior testability of ADFs within a HiL framework. Partially mentioned in chapter 3, and is integrated within chapter 4.

The multi-rate numerical simulation enables the integration of heterogeneous simulation modules (sensor models, vehicle dynamics and test automation) using the User Datagram Protocol (UDP) socket interface mechanism. The real-time HiL simulator uses two Ethernet connectors to access both the simulation environment via Transmission Control Protocol (TCP)/UDP ports and the host computer via TCP ports. In addition, the monitor camera box is connected to the HiL simulator for the Over-The-Air (OTA) stimulation of the camera ECU. The monitor is connected to the simulation environment via a Digital Visual Interface (DVI) connector [Elg12].

The ECU in-/outputs can be plugged into the break out box for measurement and hardware fault injection purposes. The Automotive Data and Time-triggered Framework (ADTF) tool provides a measurement framework for image recognition algorithms based on received data. The test automation program manages the actions of all actors, such as lane changes or speed changes by the traffic module, to control the Ego-vehicle and other traffic objects. The target acceleration and steering are translated with the vehicle dynamics simulation into throttle, brake and steering wheel positions within the applicable limits.

The model-based test methods support reactive tests, where the execution of a test case depends on what the DuT is doing while being tested [ESW+16]. Therefore, the model-based testing is integrated into the closed-loop HiL platform, as illustrated in figure 4.6. Furthermore, the robustness tests are integrated by run-time fault injection to assess the degree of correct functionality under invalid inputs or in stressful environmental conditions. For the minimization of human effort in test execution, test automation provides computer-aided execution of dynamic verification of a function's behavior on a limited number of test cases against the specified expected behavior.

4.3.3 Round-Trip Time Estimation

The delta time of each simulation step is identical (fixed-step solver) but the real-world time between the steps differs and, therefore, influences the correlation between simulation time and real-world time. A system that accumulates simulation time in-sync with the progress of real-world time is called a real-time system.

There is a distinction between hard and soft real-time systems based on the consequences in case of a violation of the real-time requirements. While the hard real-time systems must reliably deliver the correct result within the required response time, the soft real-time systems can statistically meet the required response time within a tolerance margin [LSE12]. Consequently, the co-simulation framework allows the execution of heterogeneous real-time simulation components, whereby the real-time data exchange must not exceed the defined tolerance margin. The real-time requirements have to be fulfilled for the HiL test bench for ADFs. Hence, the real-time behavior validation of the distributed HiL components is typically as important as their functional correctness. Accordingly, a case study is conducted to investigate the time constraints using a multi-sensor data fusion module, a camera sensor model $\hat{S}^c$ and a RADAR sensor model $\hat{S}^r$, as demonstrated in figure 4.7.

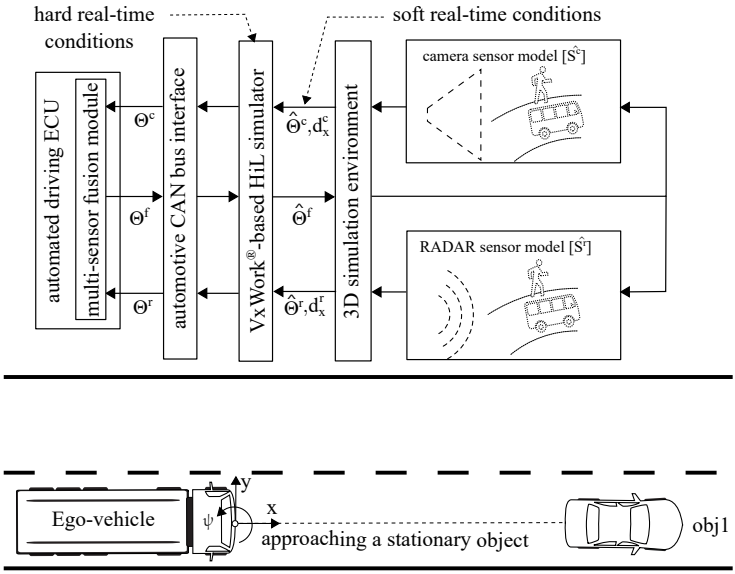

Figure 4.7: Driving scenario setup with approaching a stationary object for validation of the timing constraints within the HiL co-simulation framework.

The camera sensor model calculates the relative longitudinal distance d_x^c between the Ego-vehicle and the stationary object. Similarly, the RADAR sensor model computes the relative longitudinal distance d_x^r between these two vehicles. The sent timestamp message $\hat{\Theta}^f$ is employed for the fusion of non-synchronized measurements from heterogeneous sensor models using received timestamp messages $\hat{\Theta}^c$ and $\hat{\Theta}^r$.

The above-mentioned case study applies a driving scenario in which the Ego-vehicle approaches a stationary object. The driving scenario integrates an AEB function with reaction to the stationary object within the HiL co-simulation framework. Consequently, approaching a stationary object is carried out at different velocities $v_x^{ego} = \{5, 10, \cdots, 105\}$ [km/h] to evaluate the time constraints. If the Ego-vehicle velocity is increased, the temporal longitudinal distance to the stationary object is shortened. The number of corresponding timestamps decreases accordingly. In order to determine the Round Trip Time (RTT) at the soft real-time communication, the difference between d_x^c[m] and d_x^r[m] is calculated. The RTT can be calculated at the start time of the same sent and received timestamp messages, as described in equation 4.13. Thus, the RTT describes the time interval to send a message from a starting point to a destination and return to the same starting point.

$$\mathrm{RTT} = \frac{d_x^c - d_x^r}{v_x^{ego}}, \forall\ \hat{\Theta}^f = \hat{\Theta}^c = \hat{\Theta}^r \tag{4.13}$$

The co-simulation interface introduces latency that determines the limits of the real-time capability for the entire framework based on the configuration frequencies, where $f_d = 1000$[Hz] and $f_m = 120$[Hz]. Figure 4.8 presents the latency between timestamp measurements from the camera and RADAR sensor models without taking the additional delay, caused by the CAN bus interface, into account. Although the Ego-vehicle velocity increases, the RTT remains at $f_m = 120$[Hz] with a variation between ±2 simulation cycles with a reduced number of matching timestamps. This variation is considered as an accepted tolerance for the soft real-time conditions.

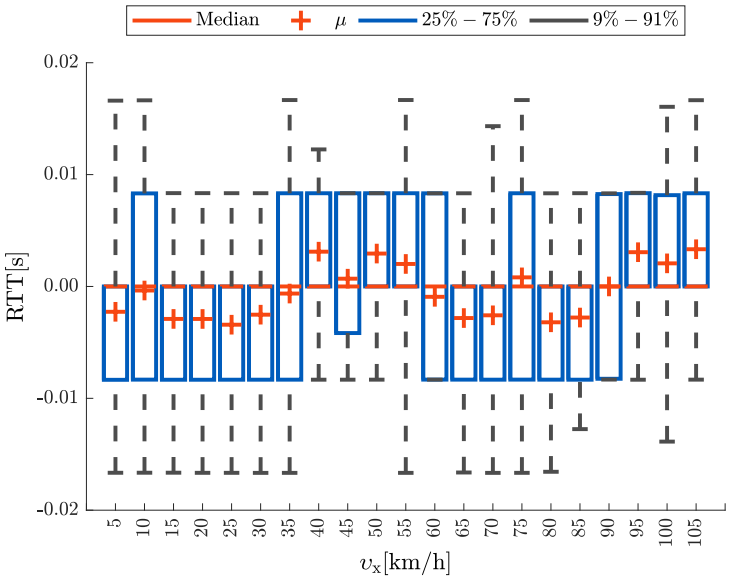

Figure 4.8: RTT calculation at start time of the same timestamp messages for sensor fusion using different velocities of the Ego-vehicle in the scenario with approach to a stationary object.

4.4 Perception Sensor Simulation

The HiL-based test approaches can provide an efficient functional evaluation of environmental perception sensors under reproducible conditions. There are different methods, which can be practiced at three logical interface levels (detection, feature and object level) [fS+21] to stimulate the sensor behavior in a HiL environment. Therefore, the quality of the environment simulation depends on the layer of injection into the DuT. While injecting the data via the physical sensor layer or the unprocessed raw data, the simulated data has to represent the real world at a high fidelity level [NS11].

4.4.1 Sensor-in-the-Loop Testing

Diewald et al. presented an antenna coupling approach for an OTA RADAR sensor stimulation, which considers the three subsequent processing steps (measurement, processing and observation) [Die15]. The stimulation describes the act of manipulating an entity in which its state corresponds to a driving scenario. The antenna coupling simulator allows dynamic simulation of moving objects using opposite RADAR antennas. Gowdu et al. improved the antenna coupling approach to stimulate automotive RADAR sensors in a virtual electromagnetic environment using an OTA interface with wideband horn antennas at 77 GHz [GA+18a]. The employed RADAR target simulator generates back-scattered radar signals to emulate the DuT by synthetic RADAR target signatures with specific attributes like RCS, range and velocity.

Weiskopf et al. described various approaches of digital RADAR signal injection-based integration for RADAR ECUs in a HiL setup [W+15]. Furthermore, Hanke et al. proposed a way to design a realistic description of the automotive RADAR system with the help of sophisticated sensor models, whereby the data is generated synthetically at a phenomenological level [H+15, H+12]. The same integration approaches can also be employed to verify the camera based functions in a 3D synthetic testing environment. Nentwig and Stamminger worked on the applicability of real-time generated computer graphics for camera-based detection algorithms [NS11]. Tan and Hassan presented a projection-based approach which is used to stimulate the camera ECU using synthetic traffic scenes [TH13]. The 3D scenes are visualized in real-time on a flat monitor. The 3D orientation of the camera is estimated within a global frame of reference, whereby the video is displayed in a 3D virtual environment [HH15].

The 3D orientation is converted into the Camera Reference Frame (CRF), which is compared to the FOV of the camera ECU [SS13]. In order to match each monitor pixel with a point in the CRF, the camera ECU calibration with respect to the monitor is to be performed [NMS12]. In particular, the camera ECU is characterized in terms of its intrinsic parameters, e.g. focal length, skew coefficient and distortion coefficient [HM+14]. The extrinsic parameters of the camera ECU with respect to the 2D screen are the rotation matrix and the translation vector [Kan16]. The screen co-ordinate system is a 2D system, where the x-axis corresponds to the right hand direction and the y-axis corresponds to the upward direction.

The camera ECU records the artificial scene and provides the image processing unit with the raw image data in a digital form. The image processing unit generates the object list from the resulting image data to the CAN vehicle bus interface and tracks the objects based on the simulated Ego-vehicle state data. Pfeffer and Haselhoff illustrated the injection-based approach, in which the synthetic raw image data is injected directly into the image processing unit via the bypass of the image sensor module [PH16]. The standardization of sensor injection interfaces play a crucial role, while the current stimuli injection interfaces are highly product-specific.

Table 4.2 — terminologies are discussed in section 2.3 — summarizes the manifold ways of injecting synthetically generated stimuli into the DuT. Because of the technical limitations of covering all the sensor physical characteristics, the suitability of each approach should be determined according to the required test objective of the environmental perception sensor.

Stimulus injection	**Over-The-Air stimulation**	**Raw-data simulation**	**Target-list simulation**	**Object-list simulation**
DuT hardware modification	not required	sensor specific	not required	not required
DuT software modification	not required	software bypass	software bypass	not required
Additional hardware	sensor dependent	hardware adapter	not required	not required
Simulation quality	sensor dependent	high	simplified	simplified
Sensor ECU pipeline	complete	excluding measurement	excluding measurement & processing	excluding measurement, processing & observation

Table 4.2: Overview of different Sensor-in-the-Loop approaches and their advantages and drawbacks [FHW16, Fei18].

4.4.2 Object-list based Sensor Models

The environmental perception sensor simulation is positioned on the Ego-vehicle and detects the objects within its pyramid-shaped cone. The sensor position is computed relative to the carrying CMV's reference point (typically the center of the rear axle on ground level).

The sensor uses a purely geometrical approach, which calculates the nearest point of extended object within the applied FOV. The modular architecture enables an iterative development of the phenomenological sensor model to achieve increased realism, as presented in figure 4.9.

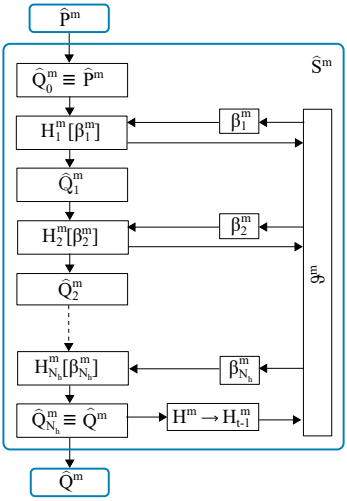

Figure 4.9: Modular architecture of the phenomenological sensor simulation [ESFG19b].

Equations 4.14 and 4.15 describe the sensor model components. Therefore, the sensor model $\hat{S}^m$ maps a set of detected objects $\hat{P}^m$ from the Ego-vehicle's environment to a set of tracked objects $\hat{Q}^m$ into the object-level data fusion algorithm. The characteristics of the sensor mapping H are divided into H_υ^m modules with $\upsilon \in [1, N_h]$. The set of configuration parameters for module H_υ^m is denoted by β_υ^m and comprises the relevant subset of sensor properties.

$$\hat{S}^m : \hat{P}^m \rightarrow \hat{Q}^m \tag{4.14}$$

$$H = H_{N_h}^m \left[\beta_{N_h}^m\right] \circ \cdots \circ H_2^m \left[\beta_2^m\right] \circ H_1^m \left[\beta_1^m\right] \tag{4.15}$$

Therefore, an incremental development process of simulation models is proposed to reproduce the physical relationship of reality with proper modeling depth. The procedure consists of the analysis step, the implementation step, the verification step, and the accreditation step.

The analysis step obtains a conceptual model from reality through analyzing the physical principle of the real component. The goal of this action is to construct specified instructions for the implementation of an executable model. The executable model is a computer program, which should be verified based on its specifications. The accreditation process compares the test results with its component test bench to determine whether the executable model's accuracy is adequate to accomplish its intended function. The influence of sensor modeling is investigated on the simulation results using typical driving scenarios from the scenario database.

The ADASIS interface is applied as a proper interface for information exchange to enable access to relevant map data via the CAN vehicle bus interface for advanced assistance features. Therefore, the eHorizon can be perceived as a virtual sensor to anticipate the driving path based on a non-physical measuring principle [BMBL06, KP+14]. The eHorizon is implemented as a sensor model with a predefined declaration of input information and defined output signals. The sensor model has an access to the map provided by the environment simulation and the position of the Ego-vehicle in this map. Using this information, the relevant section of the map is parsed and the position of the Ego-vehicle is obtained in path co-ordinates [HS15]. Along the aforementioned MPP, all the information is extracted and ADASIS v2.0 conform messages are generated. The configuration of the sensor model can be set up from the test automation, allowing the test cases to be linked to specific sensor configurations.

The modular architecture enables iterative development of the eHorizon sensor model to achieve increased realism. The MPP consists of data structures that are provided through an OpenDRIVE parser [ASA21]. In reality, discrepancy between visual and map data is omnipresent due to map errors, old map data or optical detection failure. Through fault injection, defined inconsistencies can be produced within the HiL simulation environment. Amongst others, the fault injection covers the placement and value of traffic signs. These failures can be used for robustness testing of the fusion algorithms. The implementation of the eHorizon sensor model focuses on the basic road geometry and attributes (e.g. speed limits, overtaking signs) along the MPP, as shown in table 4.3. The ADASIS v2.0 is based on a path offset model, where the path is the first entity that must be obtained to retrieve the required information identified by a path identifier [ESA+17].

Message	Description
META DATA	semi-permanent data (e.g. country code)
POSITION	vehicle positioning in geographic co-ordinates
SEGMENT	road ahead data (e.g. speed limits, tunnel)
STUB	branch points (e.g. turn angle, intersection)
PROFILE SHORT	road's course data (e.g. curvature, slope)
PROFILE LONG	road specific data (e.g. longitude, latitude)

Table 4.3: List of eHorizon sensor messages and their descriptions [Are16].

The [META DATA] message contains country-specific information, where implicit speed limits are included implicitly for each relevant country. The [POSITION] message is determined by a path identifier and an offset along the path. The start of each path is defined to be at offset 0. The [SEGMENT] summarizes the most important attributes for a part of path. The [STUB] message defines the relationship between the paths. While the [PROFILE SHORT] message contains 10-bit variables for road's course data, the [PROFILE LONG] message holds 32-bit variables for geographical co-ordinate information. The data of interest are either on the same path or on one of the sub-paths ahead of the Ego-vehicle. In order to fulfill the highly automated driving requirements, the ADASIS v3.0 is responsible for transferring high precision and up-to-date map data from the cloud to the ECU. Up-to-date information about traffic and road conditions are not provided by the ADASIS v2.0 protocol.

The capability to integrate vehicle-independent information sources by extending the environment simulation makes the proposed framework adaptable for further applications, such as Vehicle to X(everything) (V2X) communication[1]. The eHorizon data contains vehicle position data as well as road segment attributes, such as road geometry, road class, number of lanes, speed limits, etc. The iterative development process enables the re-design of the modular simulation models based on the claimed modeling depth ranging from ideal to phenomenological sensor models.

[1] The V2X communication summarizes various vehicle technologies that enable a vehicle to communicate, e.g. with other vehicles, with systems integrated into the infrastructure, with pedestrains' mobile devices, or to a database.

4.5 Ontology-based Scenario Management

The ontology-based NL notations are meta-data representations of the data elements and their semantic relationships in a structure that is understandable for humans and machines [Z+18a]. An ontology is equivalent to a Description Logic (DL) knowledge base [HPSVH03]. While knowledge representation in an ontology is based on DL, the Ontology Web Language (OWL) is a common popular file format for storing ontologies based on the Resource Description Framework (RDF) data graphs [M+15]. Therefore, the ontology comprises Terminological Box (TBox) and Assertional Box (ABox) statements [EH16]. The TBox statements describe object-oriented classes within a knowledge base as a schema or data model. In contrast, the ABox statements are associated with instances of these classes [CK18].

Automated driving involves the use of ontologies in various applications, especially in situation assessment, scene understanding and behavioral planning [BMM18]. Armand et al. describes the application of ontologies to model interactions based on spatial-temporal relationships between road users and infrastructure [AIGZ17]. The sensor data is employed as ABoxs of an ontology to develop a human-like understanding of scenes. The scene understanding relies on object tracking, map data and the dynamic states of the subject vehicle. Behavior rules are stored in the semantic web rule language to infer knowledge from the TBoxs to a given ABox from the sensor data [KLN+18].

Ulbrich et al. proposes an environmental model derived from a knowledge base with hierarchical classes and relations between the entities [UNMH14]. The environmental model is updated by sensor data and utilized for online decisions. Geyer et al. proposes nomenclatures of a unified ontology for generating test cases and scenario catalogs [GBF+13]. However, Geyer et al. describes that each scenario catalog shall have its own nomenclature and concepts for knowledge organization.

Figure 4.10 shows an ontology-based test scenario synthesis based on knowledge discovery from triggered FOT events. The systematic test case generation leads to concrete scenarios and test cases based on a generic model consisting of four layers of scenario description. The first layer belongs to road geometry, the second layer to static objects, the third layer to dynamic objects and finally the fourth layer to weather conditions.

The test cases are also deduced from the functional specification, which results from the top-level requirements and the use cases, whereby the use cases are also inferred from the top-level requirements. A category of adequate and relevant scenarios for existing field tests is extracted using the ontology-based scenario management. The semantic representation of worst-case scenarios can be obtained by using data mining techniques and systematically processed in requirements for ODD coverage.

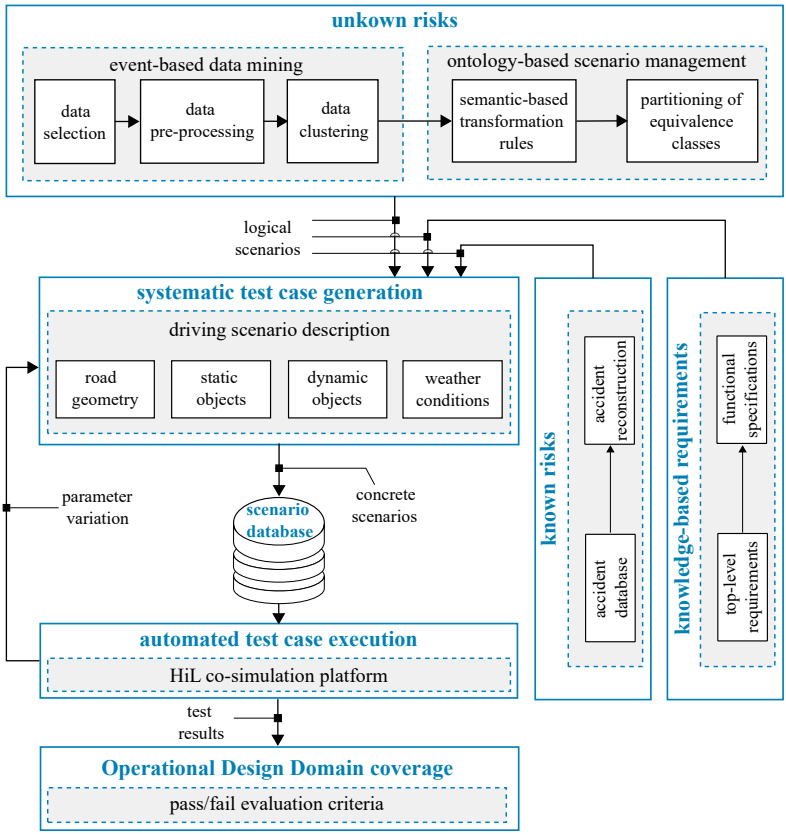

Figure 4.10: Coverage-driven test concept with systematic test case generation based on field-based observation.

Definition 4.2 (Data Mining): It is an integrated Knowledge Discovery in Databases (KDD) process for the automatic identification of useful information in big data repositories [FPSS96]. Therefore, the data mining process comprises several transformation steps, from data pre-processing tasks to post-processing of the data mining results. The knowledge management process serves to obtain representative driving situations that are recorded by FOTs in the form of time-series of environmental perception sensor data.

The implemented framework facilitates a functional verification of ADFs precisely and more efficiently on the target ECU in the laboratory. The test cases are executed on a HiL co-simulation platform. The ontology combines relevant entities in natural language and assigns a formal order through conceptualization. Irrelevant entities or relations are excluded for the traceability of parameter changes associated with an ontology-based scenario synthesis. The ontology provides a method to derive possible observations of ABoxs from modeled knowledge of TBoxs [LTW20]. Subsequently, the ontology-based scenario synthesis can be converted into de facto simulation data formats (e.g. OpenDRIVE [ASA21], OpenSCENARIO [ASA22], etc.).

The identification of situations from FOTs using data mining techniques offers a valuable solution for the generation of simulated test scenarios to extend the validity of test coverage of ADFs more cost-effectively [EUA+19]. Although the corner case scenarios occur rarely in real traffic, the ontology-based approach provides the relevant parameter space from open-loop sensor signals to ensure the validity of the generated closed-loop test cases. The ontology-based transformation rules provide the syntheses of relevant driving scenarios based on the parameterized characteristic waveforms for the HiL scenarios. In case of further induced traffic situations, the classification maps them into one of the predefined categorical classes. Subsequently, the ADFs can be verified effectively and efficiently through an ODD coverage using test oracles based on envelope components of pass and fail criteria [L+19]. The proposed test concept complements the traditional knowledge-based text matrix and enriches the test cases with a maximum test coverage. Furthermore, the context-driven verification approach determines how such an argument can be formed by decomposing the test coverage and objectives. Therefore, the proposed approach is complemented by the systematic provision of various evidence of ODD coverage to meet the required test termination criteria.

5 Data-driven Scenario Extraction

Safety in automated truck driving can be maximized if human-like errors of automated CMVs can be avoided. Hence, the sense-plan-act robot control procedure is an essential part of the automated CMV's response to the dynamic driving environment. Functions of the environmental perception and situation analysis are responsible for the recognition of the vehicle environment and the associated situational awareness. While the planning function is responsible for determining the driving trajectory, the motion control function operates the throttle or brakes, turns the wheels and otherwise actuates the vehicle to follow the plan precisely [Smi17]. To validate these functions, the field tests are carried out after verifying the absence of logical errors and determining the required total mileage and geographical variation [Ebn14]. Accordingly, the automated CMV fleet is equipped with data-logging devices for recording automotive communication buses, sensor raw data, additional context cameras and inertial measurement systems (e.g. Inertial Measurement Unit (IMU) and Global Navigation Satellite System (GNSS) sensors) for precise vehicle positioning. Then, data mining techniques are usually used to retrieve novel and useful patterns from large databases [FPSS96].

5.1 Measurement Methods for Scenario Mining

Scenario-based testing relies on the measurement data from real-world scenarios to derive the required knowledge within the scenario elicitation process. Consequently, the measurement method requirements include the acquisition of data with reasonable effort, while ensuring an adequate quality of the dynamic scenario description and the naturalistic behavior of the road users [Kri19]. Common measurement methods for obtaining data in scenario-based testing are vehicle data loggers, drones equipped with cameras and roadside infrastructure sensors.

Various publications identify the requirements of automated driving systems based on the German In-Depth Accident Study (GIDAS) accident database to prevent as many such human-like accidents as possible [S+19a, S+19b, FWP+19]. According to the National Motor Vehicle Crash Causation Survey (NMVCCS) database conducted by NHTSA for U.S. police-reported passenger vehicle crashes, driver-related contributing factors can be divided into five categories: First, the sensing and perception factors contribute with 24% to accidents caused by unrecognized hazards. Second, the incapacitation factors represent 10% of accidents due to drivers who are alcohol-impaired or otherwise incapacitated drivers. Third, the prediction factors due to a misjudgment of the other vehicle behavior account for 17%. Fourth, the factors of planning due to illegal maneuvers or poor decision making behind compliance with traffic rules and defensive driving cause 39% of the accidents. Fifth, the factors of execution and performance share with 23% due to inappropriate vehicle control. While a crash event can result in multiple common factors, the total percentage is more than 100% [M+20b].

5.1.1 In-vehicle Data Loggers

The development of automated driving algorithms requires extensive tests on real-world driving scenarios, which are collected and aggregated from the FOT [L+18a]. The FOT represents a study undertaken to evaluate an ADF under normal operating conditions in road traffic environments [Mas19]. Consequently, data loggers are installed in test fleet CMVs to record raw data from the perception sensors and data from vehicle networks in a time-synchronous way [AADN+16]. Thus, the data loggers for test fleets allow an in-depth analysis of the entire data flow from sensing, perception, prediction and planning to motion control software modules [ZWZ18, GKS18, KYB18]. In addition, a number of public data-sets, such as the Karsruhe Institute of Technology and Toyota technological Institute data-set (KITTI) [GLSU13], the Cityscapes data-set [COR+16] and the truck-specific TuSimple data-set [N+18], were published, which can be utilized for scenario-based testing. Section 5.2 discusses in-vehicle data loggers describing their implementation.

5.1.2 Drones Equipped with Cameras

Trajectories of individual road users can be extracted from an aerial perspective using drones equipped with high-resolution cameras. While the traffic is not affected by the measurement, computer vision algorithms are used to process the extracted naturalistic images [K+18c]. Several public data-sets provide vehicle trajectory data with the bounding-box based annotation of the detected vehicles. The highway Drone data-set (highD) provides a large-scale data-set of naturalistic trajectories of vehicles such as cars, trucks and buses on German highways using unmanned aerial vehicles [KBKE18]. The Stanford drone data-set consists trajectories of VRUs, such as pedestrians and bicyclists, extracted from drone video recordings of the university campus [R+16]. The intersection Drone data-set (inD) contains trajectories of naturalistic road users at German urban intersections [BKM+19]. The road user trajectory data-sets can be used for scenario-based testing of prediction and path-planning software modules.

5.1.3 Roadside Infrastructure Sensors

The infrastructure can be equipped with sensors installed on roadside masts to detect traffic or weather conditions. Ground truth bounding boxes can be extracted by permanently monitoring a certain road segment. If both vehicle and infrastructure data are collected separately, the two data-sets need to be time-synchronized with a GNSS clock for scenario reconstruction. The TAF-BW test field provides a distributed intelligent infrastructure that can handle traffic light states, road topology and data about monitored road users [FDW+18]. The test field in Lower Saxony covers sections of motorway with a variety of traffic environments and situations [R+15].

5.2 In-vehicle Data Logging System

The Automated Driving Data Recorder (ADDR) gathers data from the ADFs and from sensors mounted on the truck. Figure 5.1 shows an in-vehicle data logging system and data analysis concept for the worldwide validation of ADFs.

In accordance with the circular buffer concept, representative driving situations in FOT at object list level can be separated from permanent recording in the form of time-series data from environmental perception sensors. The triggering events can be individually defined in the data logger depending on the DuT reaction, so that the event-based recordings are transported to the databases [BBLF19]. The recordings include vehicle bus data, sensor object lists and reference video streams for data analysis purposes.

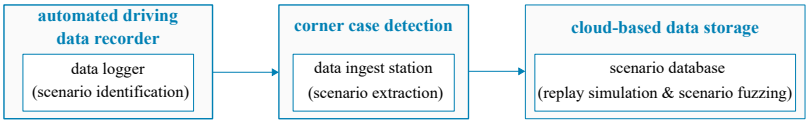

Figure 5.1: Data handling structure of the in-vehicle data logging system with its main elements.

The geographically distributed data sources constitute a challenge for existing development processes and information technology infrastructures. Therefore, a data ingest station is necessary for data storage and retrieval, where the corner cases can be identified for replay purposes. Also, the cloud-based data storage contains big data tools, architectures and analytics, that provide the database infrastructure for ingesting, storing, accessing and processing of logged data simulations [SPW+19]. For an objective assessment of the driving function used, suitable evaluation criteria are integrated into the database.

5.2.1 Automated Driving Data Recorder

According to the California Code of Regulations (CCRs) §228.06, ADDR is a mechanism installed in a CMV under test. This test CMV is equipped with an ADF to record technical information about the status and operation of the environmental perception sensors either continuously or for 30 seconds preceding a critical event (e.g. disengagement, crash, etc.) [LS19]. The data captured by the ADDR and stored in a read-only format shall be accessible and retrievable with a commercially available tool [DMV18]. According to SAE J3197, the ADDR does not interfere with the ability of the ADF to perform the DDT [Int20]. While the automated driving technology is still being developed and is not yet commercially deployed, the data included in the ADDR is used for validation and replay purposes.

Meanwhile, the SAE J1698 Event Data Recorder (EDR) is mainly used for traditional accident reconstruction analysis [Int17]. Since 2017, the German law requires data processing of disengagements with storage of the position and time data captured by a GPS antenna for vehicles with highly or fully ADFs [BMV17]. Therefore, data has to be timestamped in a way that allows perfect synchronization of multiple data streams within the ADDR using a Precision Time Protocol (PTP) grand-master. According to the IEEE 1588-2019, PTP is a protocol used to synchronize clocks in a networked measurement system [IEC+19]. Figure 5.2 illustrates an ADDR with real-time streaming of data source and sink components [BES+21].

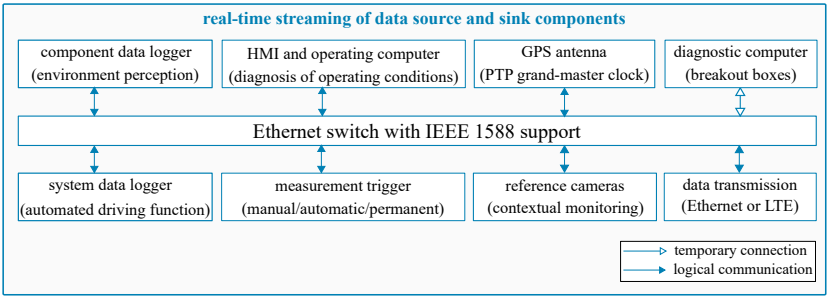

Figure 5.2: Schematic diagram of an ADDR with its main components.

The component data logger records the raw data of the environmental perception sensors, which can be regarded as a data source. For the component data logger, the Plug On Device (POD) interface allows access to the internal ECU software of environmental perception sensors as a standardized hardware interface [ASA17a]. Additionally, the system data logger collects the ADFs data as a data sink, which is communicated via an Ethernet switch capable of supporting IEEE 1588. The CMV driver utilizes the Human-Machine Interface (HMI) and operating computer to start, stop and monitor the measurements. Moreover, the HMI acquires the health status of the various measuring components via diagnostics interfaces, such as Universal (X) Measurement and Calibration Protocol (XCP) [ASA17b] and REpresentational State Transfer - Application Programming Interface (REST-API), to control the start/stop of the measurements. The ADDR provides contextual monitoring to gather information about the surrounding traffic.

Therefore, the reference cameras assist the analysis process to understand the underlying context in relation to the surroundings. For commissioning, the diagnostic computer is temporarily connected to check the prerequisites of the measuring components and ECUs via breakout boxes. The triggered measurements typically use a ring buffer to record data before and after the relevant event. The triggering events serve as a supplementary source of information to identify relevant driving scenarios occurring in the FOT [K+18a].

In practice, two types of recording have become established: Continuous and triggered measurements. If the recording is continuous, a reduced set is recorded with the most important measurement signals for statistical statements in order to keep the data volume manageable. If certain situations are triggered, the entire vehicle bus can be recorded. For each event, a recorded video snippet of the traffic situation around the vehicle is generated, supplemented by information on system states and signal characteristics. The triggered measurements cover a shorter period of time, which is concentrated on certain driving situations and is automatically triggered when predefined conditions are met [dGP17]. In case of false negative events, a manual trigger button provides the trigger to record situations in which the system intervention is missing.

As mentioned in section 2.5 to the ordinance implementing the act amending the road traffic act and the compulsory insurance act, the CMV with a SAE L4 ADF needs to be equipped with a digital data storage system [Bun22], as follows:

Ordinance (Event-based Data Storage):
Ä data storage system must be integrated in the motor vehicle with autonomous driving function that collects, uses and stores data concerning the motor vehicle with autonomous driving function on an event basis and during operation in accordance with §9(5) and §15 only for the purpose of improving road safety. The data to be collected is conclusively laid down in §1g(1) of the road traffic act in conjunction with Annex 2 to this ordinance."
Ïm Kraftfahrzeug mit autonomer Fahrfunktion muss ein entsprechender Datenspeicher integriert sein, der ereignisbasiert und während des Betriebs nach §Absatz 5 und §15 Daten des Kraftfahrzeugs mit autonomer Fahrfunktion ausschließlich zu dem Zweck der Verbesserung der Verkehrssicherheit erfasst, speichert und verwendet. Die zu erfassenden Daten sind in §1g Absatz 1 des Straßenverkehrsgesetzes in Verbindung mit Anlage 2 zu dieser Verordnung abschließend geregelt."

5.2.2 Corner Case Detection

The data retrieval procedures are needed to make sure that all collected data is backed-up and stored in a safe place. The FESTA handbook compares the different data transfer modes, either manual data ingestion via external hard disks and Network Attached Storage (NAS) devices or data transfer via Ethernet cable or Long Term Evolution (LTE) connection [FC11]. In the first step, the measurements of test CMVs are transferred to an ingest station. For this purpose, the geographically distributed data sources are collected according to a predefined ingest process [TM20]. The second step is the extraction of corner cases from the recorded measurements using data pre-processing methods. The data logistics process can either be stored on an on-premise data storage, or a cloud-based one. The optimization is executed via pre-processing of the data snippets, that are extracted from the computing services (e.g. Amazon Web Services (AWS)[1]), such a technique is illustrated in section 5.3.1.

Numerous literature on time-series representations is available to facilitate the tasks of data search and knowledge discovery [AFS93, ANR74, DTS+08]. In addition, the data is converted from the automotive data formats such as Measurement Data Format (MDF) [ASA19] and ADTF data format into big data file formats such as Comma-Separated Values (CSV) and Hadoop Distributed File System (HDFS) [AJK+17].

5.2.3 Cloud-based Data Storage

Statistical statements can be derived from the recorded naturalistic data and the real customer behavior can be identified. Meanwhilst, the simulation of logged data is a well-known technique for generating virtual open-loop test drives based on sensor data collected from field tests. Figure 5.3 shows the worldwide FOTs in CMVs over different periods of time. Therefore, the common workflow for the development of automated driving algorithms includes many iterations of simulations.

[1] https://aws.amazon.com/what-is-cloud-computing/

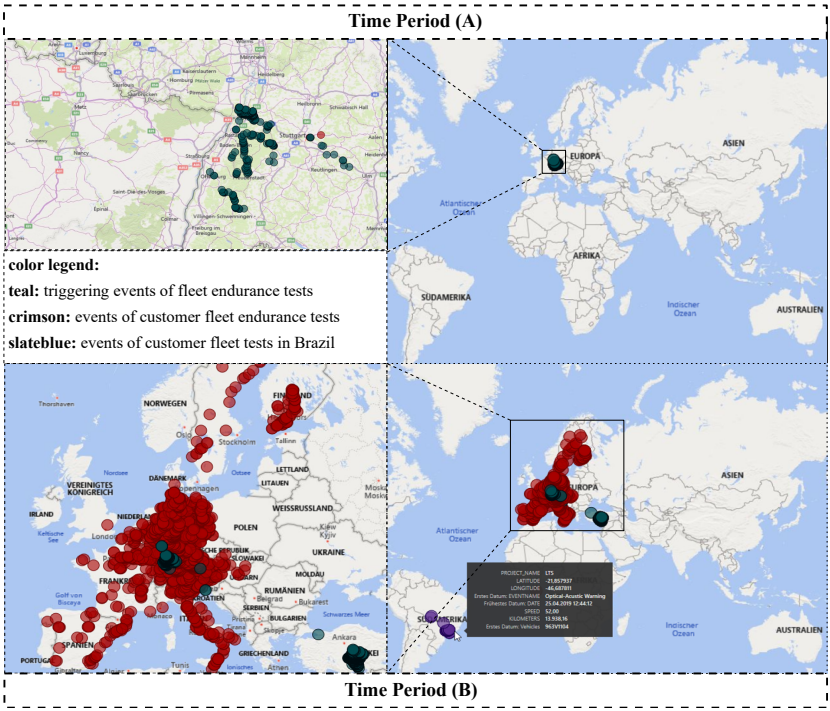

Figure 5.3: Web-based ODD coverage for worldwide FOTs of an AEB function over various periods of time in CMVs [BES+21].

Each iteration consists of three steps. First, the relevant input data is selected. Second, the playback simulation is performed. Third, the output of the logged data simulation is post-processed to evaluate the functional performance of the algorithm in terms of KPIs. In such a way, the measurement data of the test vehicles is automatically processed and stored on central servers, while the meta-information is stored in a database. An adaptable user interface is used to visualize, operate and monitor the database and the measurement data procedures. Consequently, logged data simulations with different sensor datasets are performed with modified versions of the algorithm to select the most suitable version for release in automated CMVs.

5.3 Analysis and Triage of Time-Series Data

The time-series analysis allows the extraction of representative situations observed during on-road test drives [ESFG18]. The data consists of processed object lists of environmental perception sensors that change over time. The cluster analysis refers to an unsupervised classification to group the time-series data, based on the information retrieved in the data that describes the time-series signals and their relationships [W+19]. Consequently, the cluster analysis is performed to extract homogeneous groups (clusters) from data-sets according to a defined similarity metrics [Pfe20].

The time-series within a cluster should be similar to each other and different from the time-series in other clusters. In prototype-based clustering, a cluster consists of a set of time-series in which each signal is similar to the prototype describing the cluster compared to the prototype of another cluster. The objective of clustering is to extract clusters from a data-set, where the distance between members of a cluster is minimized and the distance between different clusters is maximized. The efficiency of the cluster analysis is determined by the selection of the time-series to be analyzed, the distance measures to be used and the clustering algorithm to be employed.

5.3.1 Time-Series Data Pre-processing

The time-series data refers to a specific type of sequential data, where each data-set represents a time-series, i.e. a sequence of values that change over time. The outliers represent values of time-series data that have some characteristics that differ from most other time-series in the data-set. After filtering out the outliers from the time-series signals, the rolling standard deviation method is used to quantify the degree of variation of each value in the time-series and to select the optimized time interval around the triggering event. The low standard deviation indicates that the time-series tends to be closed to the mean, while high standard deviation implies that the time-series tends to be spread over a wider range of values. The determination of the optimized time interval depends on the assumption that time intervals with high variations of the sensor values are more representative for an efficient cluster analysis of time-series signals. Therefore, the interval around the trigger event is selected with a higher rolling standard deviation than other intervals.

Equation 5.1 represents the normalized rolling standard deviation of each point $u(a_i)$ of the time-series $\mathtt{A}= \{a_i\}_{i=1}^{m}$ to determine the corresponding time intervals of the measured sensor variables, where $\bar{a}_{i+j} = \frac{1}{w} \sum_{j=1}^{w} a_{i+j}$ indicates the rolling mean and w indicates the rolling window.

$$u(a_i) = \frac{\sum_{j=1}^{w} (a_{i+j} - \bar{a}_{i+j})^2}{w * \sqrt{\frac{1}{w^2} \sum_{j=1}^{w} a_{i+j}^2}} \tag{5.1}$$

The proximity between the two time-series signals is a numerical measure of the degree to which the two signals are equal. Proximities are usually not negative and are often between 0 for no similarity and 1 for full similarity. The time-series proximity can be performed with the Minkowski distance metric, as in the equation 5.2 to compare the original time-series with each other, where $r = 1$ for the distance from Manhattan (L_1 norm), $r = 2$ for Euclidean distance (L_2 norm) and $r = \infty$ for Supremum distance (L_∞ norm) [Els18].

$$\mathcal{D}_r(\mathtt{A}, \mathtt{B}) = \left(\sum_{i=1}^{m} |a_i - b_i|^r \right)^{\frac{1}{r}} \tag{5.2}$$

Although the Normalized Cross Correlation (NCC) is often used in image and signal processing for template matching, similar traffic situations can be effectively described by similarities in sensor measurements with time shifts and sliding windows. Therefore, the NCC provides as a suitable measure for the proximity in the time-series analysis of the measurements from the environmental perception sensors. Equation 5.3 refers to the NCC calculation between two different time-series $\mathtt{A} = \{a_i\}_{i=1}^{m}$ and $\mathtt{B} = \{b_i\}_{i=1}^{m}$ with a time shift $s_j \in \{s_{-m}, \cdots, s_m\}$ and a time interval $j \in \{-m, \cdots, m\}$.

$$\mathcal{D}_{\mathrm{NCC}}^{s_j}(\mathtt{A}, \mathtt{B}) = \frac{1}{m} \sum_{i=1}^{m} \frac{(a_i - \bar{a})(b_{i+j} - \bar{b})}{\sigma_a \sigma_b}, \forall\, i + j \in [1,\ m] \tag{5.3}$$

The mean and standard deviation of the time-series $\mathtt{A}$ indicates $\bar{a} = \frac{1}{m} \sum_{i=1}^{m} a_i$ and $\sigma_a = \sqrt{\frac{1}{m} \sum_{i}^{m} (a_i - \bar{a})^2}$. The mean and standard deviation of the time-series $\mathtt{B}$ gives $\bar{b} = \frac{1}{m} \sum_{i=1}^{m} b_i$ and $\sigma_b = \sqrt{\frac{1}{m} \sum_{i=1}^{m} (b_i - \bar{b})^2}$ respectively.

Based on the equation 5.3, the distance measure $\mathcal{D}_{\mathrm{NCC}}^{s_j}$ shows a value between $[-1,\ 1]$ for a certain time shift $j \in \{-m, \cdots, m\}$, where -1 indicates a complete dissimilarity and 1 denotes a perfect match. Equation 5.4 searches for a shift with the maximum proximity between the two given time-series over all possible shifts and converts the selected shift into a distance measure $\mathcal{D}_{\mathrm{NCC}}$. The computed distance is then defined in the range $[0,\ 1]$, where 0 indicates a minimum distance for perfect match and 1 indicates a maximum distance for complete dissimilarity.

$$\mathcal{D}_{\mathrm{NCC}}\left(\mathtt{A}, \mathtt{B}\right) = \frac{1}{2}\left(1 - \max_{j} \mathcal{D}_{\mathrm{NCC}}^{S_j}(\mathtt{A}, \mathtt{B})\right), \forall\, j \in [-m,\ m] \tag{5.4}$$

5.3.2 Principal Component Analysis

Data-driven test development necessitates continuous extraction of knowledge from the recorded sequences. In this context, reducing complexity of data is essential in order to perform meaningful analysis [BSMH18]. To this end, PCA has been widely used as a simple, non-parametric method of extracting useful information from complex data-sets. This principal was defined in chapter 2, through subsection 5.3.2, a discussion providing its integration is unraveled.

The PCA aims to reduce dimensionality variance of data using principle components. This is achieved by orthogonal transformation of the multivariate data-set into a new basis which best expresses the data-set. As a case study, the PCA is applied to a LDW function in CMVs during the FOTs. The cluster analysis of the right side lane departures is based on the lateral distance to the right lane $\mathrm{d}_{\mathrm{y}}^{\mathrm{rt}}[\mathrm{m}]$ and takes into account a total of 250 triggered events from FOTs, as depicted in figure 5.4. The time-series patterns of $\mathrm{d}_{\mathrm{y}}^{\mathrm{rt}}[\mathrm{m}]$ are divided into three clusters, as illustrated in 5.5. The cluster C_{rt}^{1} shows 236 events for different driving situations with a deviation to the detected right lane marking and a return from this deviation. The cluster C_{rt}^{2} illustrates 16 events for a number of driving situations with a sudden jump in the distance to lane indicating the detection of new lane lines. The cluster C_{rt}^{3} shows 3 events for driving situations with a deviation and a temporary sudden change of distance to the right line and back to normal deviation due to painted islands.

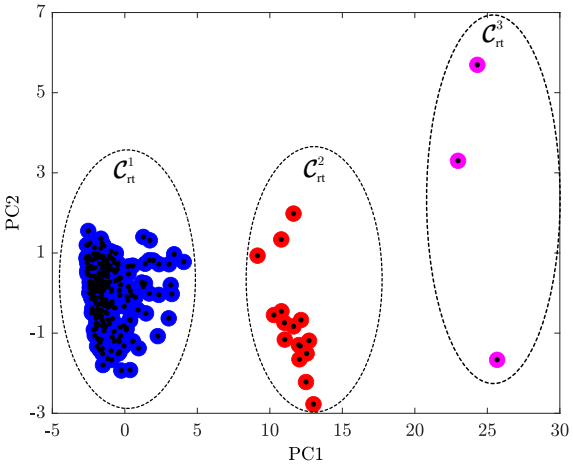

Figure 5.4: Cluster analysis of 250 events with right lane departures using the PCA of time-series data from $d_y^{rt}[m]$ [ESF19].

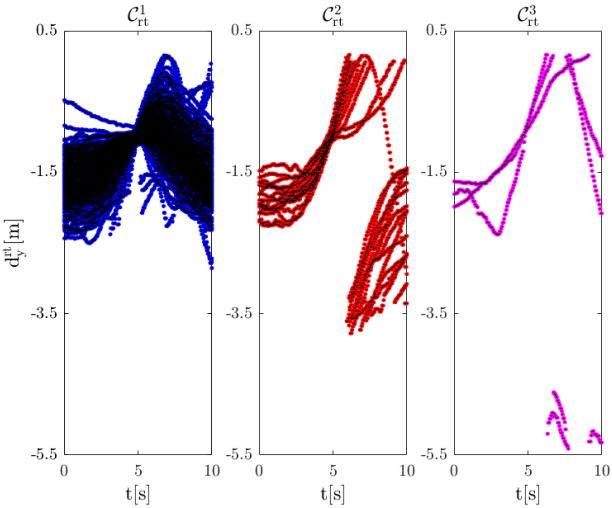

Figure 5.5: Time-series data of $d_y^{rt}[m]$ for each observed cluster from 250 events with right lane departures using a PCA algorithm [ESF19].

5.3.3 Hierarchical Agglomerative Clustering

The hierarchical clustering technique is one of the popular clustering techniques in multivariate time-series analysis and is often represented graphically in a cluster-map consisting of a dendrogram and a heat map [E+18]. The dendrogram is a tree-like diagram that shows the cluster-subcluster relationships and the order in which clusters are merged. The heat map represents the proximity matrix after the merge operations. The Hierarchical Agglomerative Clustering (HAC) produces a hierarchy of nested clusters. The cluster analysis method follows a bottom-up approach in performing clustering, starting at the lowest level where each time-series is a cluster and performing merge operations in sequence until one cluster including all time-series is left.

In addition, the inputs of the HAC algorithm consist of a pairwise distance matrix and a linkage criterion to update the matrix during the merge operations. Three of the most popular linkage criteria are: Single, average and complete. First, the single linkage defines new distances between clusters as the minimum distance between any two series in the different clusters when merging two clusters. Next, the average linkage defines the distance between two clusters as the average pairwise distance between all time-series in both clusters. Finally, the complete linkage defines new distances as the maximum distance between any two time-series in the different clusters. The choice of the appropriate matrix and the corresponding linkage criterion play a major role for the HAC algorithm to effectively update the proximity matrix during the merge process. Therefore, a complete linkage for the algorithm f_{HAC} is chosen to be less susceptible to noise and outliers, as shown in algorithm 1.

Algorithm 1 HAC algorithm with NCC-based distance measure and complete linkage.

procedure $f_{\mathrm{HAC}}(\mathcal{G} = \left\{g^{(\mathrm{t})}\right\}_{\mathrm{t}=1}^{\mathrm{T}}, \mathcal{D}_{\mathrm{NCC}}$, complete linkage)
  while length$(\mathcal{G}) > 1$ **do**
    $g^{(1)}, g^{(2)} \leftarrow \underset{g^{(i)}, g^{(j)} \in \mathcal{G}}{\max} \mathcal{D}_{\mathrm{NCC}}(g^{(\mathrm{i})}, g^{(\mathrm{j})}) \; \forall \; g^{(i)} \neq g^{(j)}$
    $\mathcal{G} = (\mathcal{G} \setminus \left\{g^{(1)}, g^{(2)}\right\}) \cup \left\{g^{(1)} \cup g^{(2)}\right\}$
    return $\mathcal{G}$
  end while
end procedure

The hierarchical clustering is executed by starting with each element as a singleton cluster $\mathcal{G}$ and then recursively merging the two nearest clusters $g^{(1)}$ and $g^{(2)}$ until a single cluster remains. As a further case study on cluster analysis, the HAC is applied to an AEB function in response to stationary objects. The emergency braking operation is initiated to achieve a predetermined target safety distance between the Ego-vehicle and the preceding object. Two object list variables measured by the RADAR sensor are selected to characterize the event-based driving situations, namely the lateral deviation $\mathrm{d}_{\mathrm{y}}^{\mathrm{rel}}[\mathrm{m}]$ of the relevant stationary object and the predicted road curvature $\kappa_{\mathrm{ego}}[1/\mathrm{km}]$.

The three types of escalation levels where abstractly mentioned in 3, as type 1, 2 and 3; in which, the first escalation level $\mathcal{E}_{\mathrm{s}}^{1}$, visual and audible alarms are emitted as a warning for the driver as long as a relevant object is within a predefined distance and closing speed. If the truck driver does not react to the haptic and automatic partial braking of the second escalation level $\mathcal{E}_{\mathrm{s}}^{2}$, and if the threat worsens further, then emergency braking occurs in the third escalation level $\mathcal{E}_{\mathrm{s}}^{3}$ at a point where the collision is imminent. Figure 5.6 shows the signal prototype of selected sensor data variables. The positive curvature $\kappa_{\mathrm{ego}}[1/\mathrm{km}]$ values indicate driving in a left turn, while the negative ones indicate right turn driving. A positive lateral distance $\mathrm{d}_{\mathrm{y}}^{\mathrm{rel}}[\mathrm{m}]$ indicates that the object is on the left side of the Ego vehicle, while a negative lateral distance indicates that the object is on the right side.

Figure 5.7 shows a clustermap to visualize the hierarchical clustering of time-series events using complete linkage within four clusters of unlabeled trigger-events caused by stationary objects with only the escalation levels $\mathcal{E}_{\mathrm{s}}^{1}$ and $\mathcal{E}_{\mathrm{s}}^{2}$. No triggering events were recorded for the escalation level $\mathcal{E}_{\mathrm{s}}^{3}$ of the entire FOT campaign. The well-separated clusters show a very strong, block-diagonal pattern in the reordered proximity matrix.

The cluster C_{s}^{1} displays 179 driving situations with 53% in total of 337 triggered events in which the Ego-vehicle triggers a false alarm in a left turn due to an irrelevant obstacle on the right lane. The characteristic waveforms of cluster C_{s}^{1} represent driving in a left turn. This cluster indicates an increasing negative lateral distance $\mathrm{d}_{\mathrm{y}}^{\mathrm{rel}}[\mathrm{m}]$ to the object in front, which indicates that the truck is moving to the left. The driving in a left-hand curve is also evident from the increasing value of road curvature positively.

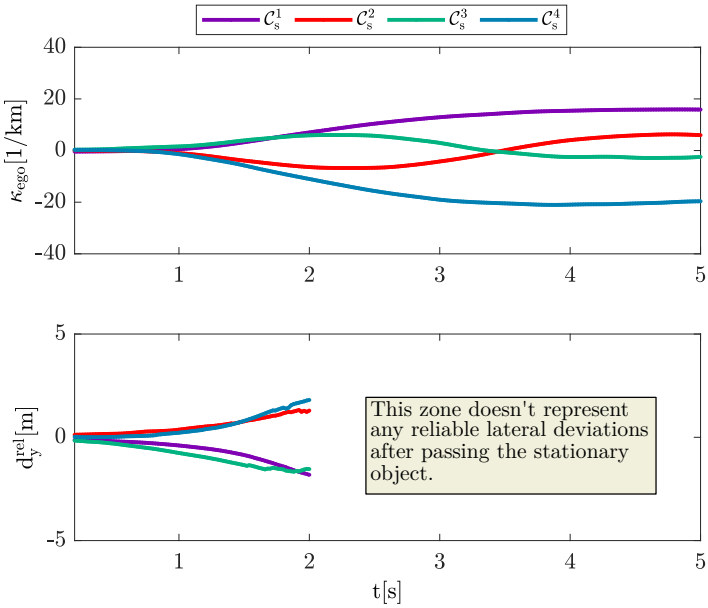

Figure 5.6: Characteristic waveforms of the trajectories based on the road curvature of the Ego-vehicle $\kappa_{ego}\,[1/\mathrm{km}]$ and the lateral distance of the stationary object in front of the Ego-vehicle $d_y^{rel}\,[\mathrm{m}]$ for each cluster [ESFG20].

The cluster $\mathcal{C}_s^2$ indicates 58 driving situations with 17% in which the Ego-vehicle erroneously triggers a warning with an unrelated obstacle located on a left-hand traffic island. This cluster shows events where the truck has driven to the right and subsequently to the left, basically driving around the object from the right-hand side. As a result, characteristic signal courses of the cluster $\mathcal{C}_s^2$ show driving around an object from the right-hand side. Furthermore, the cluster $\mathcal{C}_s^3$ collects 32 driving situations with 10% in which the Ego-vehicle triggers a false-triggered event with an unrelated obstacle located on a right-hand traffic island. The cluster $\mathcal{C}_s^3$ demonstrates that trucks drove around an object from the left-hand side. Eventually, the cluster $\mathcal{C}_s^4$ represents 68 driving situations with 20% in which the Ego-vehicle triggers an inappropriate warning in a right turn due to an irrelevant obstacle on the left lane. The cluster $\mathcal{C}_s^4$ indicates events where trucks are driving in a right curve.

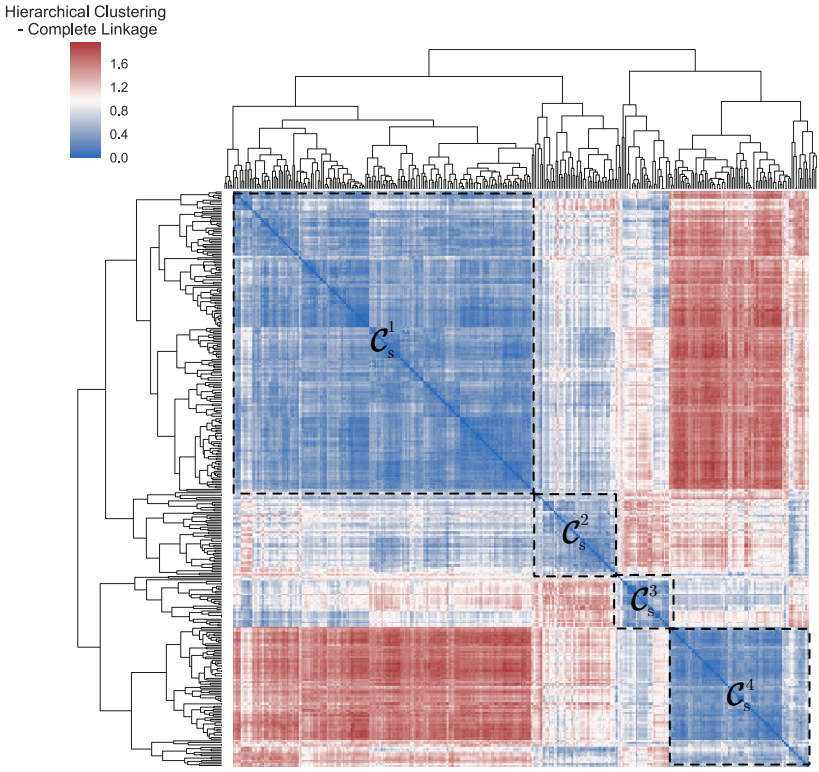

Figure 5.7: Clustermap of hierarchical clustering with complete linkage from 337 driving situations using the HAC of time-series data from $\kappa_{ego}\,[1/km]$ and $d_y^{rel}\,[m]$ [ESS$^+$19c].

Figure 5.8 illustrates the four clusters, which cause false warnings of the AEB function by detecting bounding objects of the road as relevant objects, and which can be identified by trigger conditions. In the highway engineering, the transition curve is an essential design element in the horizontal projection of the road design besides the straight line and the circular arc. The transition curve is a curve in which curvature $\kappa_{\text{road}}\,[1/\text{km}]$ varies uniformly with respect to its length. It allows a gradual change from one radius to another or from a straight line to a circular curve, since a straight line is merely a curve of infinite radius. Therefore, a circle has a constant curvature and a straight line has a curvature of 0. The most common transition curve is the clothoid, which ensures a smooth transition between the horizontal alignment elements with a constant curvature. As a transition curve type, the combined curve follows the sequence (clothoid - circular arc – clothoid). As a further type of the transition curve, the S-shaped clothoid, also called inflection line, consists of two clothoid branches with curvature in opposite directions [Küh13].

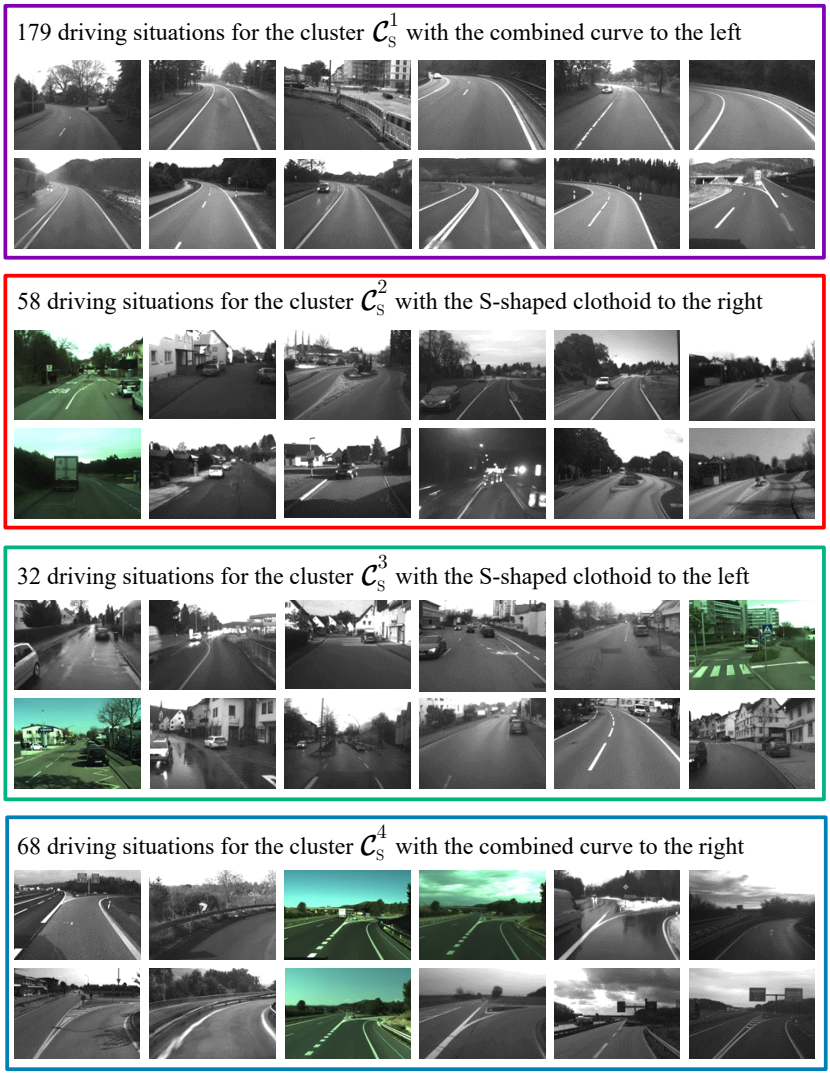

Figure 5.8: Characterization of 337 driving situations on the basis of recorded sensor signals for environment perception and their system reactions during the FOTs [BES+21]. In real-world footage (Source: documentation camera).

6 Operational Design Domain Coverage

As expound in chapters 2,3 and 4, the vision of accident-free driving accelerates the development of intelligent, connected and complex driving functions in CMVs. The conventional test methods pose two major challenges for the testing of ADFs. The numerous combinations of relevant driving scenarios present a challenge to meet a complete set of functional requirements. On the other hand, ADFs employ machine learning algorithms with opaque functional requirements. As discussed in chapters 3 and 4, the release of automated CMVs requires not only comprehensive testing against realistic driving scenarios, but also the KPIs to evaluate the ADFs with respect to dealing with uncertainties [Rös20]. Consequently, the further development and market introduction of ADFs have led to a need for new methods of software development and evaluation [HS10].

The road traffic expresses an open parameter space in which an infinite number of different traffic situations can occur. But it is also not possible to guarantee absolute safety for automated CMVs. Due to the complexity of the development of ADFs, SOTIF recommends continuous improvement iteratively, so that the residual risk can be accepted [SH19b]. Although a human driver doesn't drive perfectly at times, especially during the first few driving lessons, the development of his predictive mental model enables him to expand his accumulated driving skills. Therefore, the software product lines have established knowledge-based and data-driven test methods to ensure the functionality of their products in terms of robustness, reliability and safety. Moreover, a functional decomposition is necessary to present appropriate arguments by measuring and addressing the residual risks caused by deficiencies in the environmental perception sensors [Hül18]. Accordingly, the measurable safety framework provides a data-driven test method that complements the knowledge-based test method and adaptively enriches test cases with an ODD coverage.

6.1 Measurable Safety Framework

There are numerous steps in the data mining process to ensure continuous monitoring and learning from field observations. For this purpose, the measurable safety framework aims at bridging the gap between the knowledge-based and the data-driven test methods by continuously expanding knowledge in an ODD coverage. In addition, the framework provides a meaningful termination criteria for testing the ADFs based on the performance evaluation of individual components as well as that of the entire system. Figure 6.1 schematically illustrates a test method using the measurable safety framework.

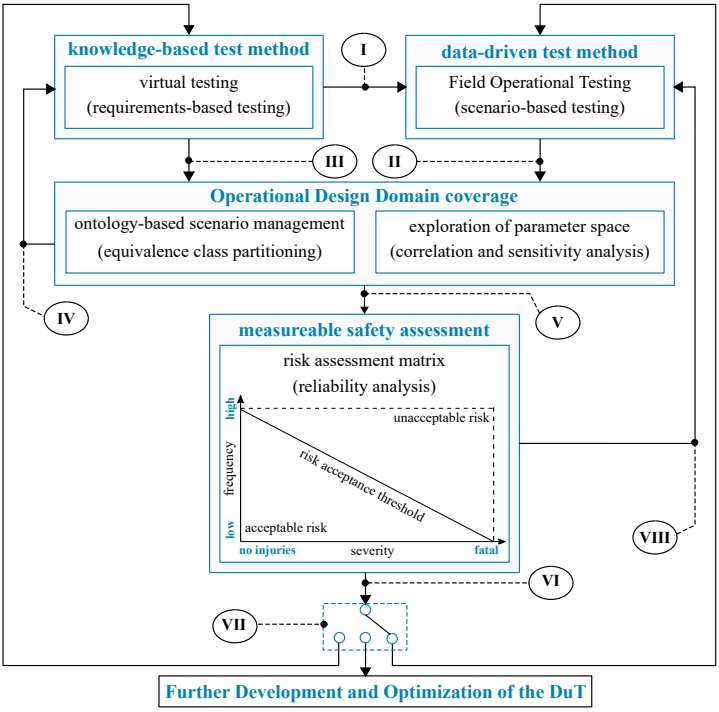

Figure 6.1: Overview of the measurable safety framework with its main elements by means of test termination criteria and ODD coverage [ESS+19c].

All ADFs are required to be compliant with the ISO 26262:2018 for functional safety and shall not adversely affect each other. Consequently, the ADFs need to be adequately safeguarded to ensure that the functional safety requirements of ISO 26262:2018 are met, in particular with regard to unintended interventions. The following steps are obligatory steps performed in the test method to evaluate the optimization quality of ADFs within the ODD of the vehicle. They do not need to always be in same sequence as mentioned below, their order can vary as per requirement.

(I) Verification of the absence of logical errors for the ADF by leveraging the HiL test bench according to the knowledge-based requirements, as presented in section 6.2.

(II) Identification of statistical errors within the ADF by scenario-based testing using cluster analysis of time-series data in a cloud database, as shown in section 5.3.

(III) Determination of criticality metrics for the ADF through correlation analysis between the criticality threshold of synthetic driving scenarios and events from the various clustered naturalistic driving situations, as displayed in section 6.4.

(IV) Execution of test cases on the HiL test bench for the extracted logical scenarios with the ontology-based scenario management after converting the open-loop detection data to closed-loop control data within the ODD coverage, as presented in section 6.3

(V) Exploration of the parameter space for the observed logical scenarios using sensitivity analysis to generate an approximation model of the parameters under consideration, as demonstrated in section 6.5.

(VI) Prediction of the risk acceptance threshold as a confidence level using reliability analysis to estimate the probability of exceeding the safety margin using various sampling methods, as exposed in section 7.2.

(VII) Definition of the test termination criteria on the basis of the generated tolerable risk curve as reference safety threshold for the further development of the ADFs and the decision whether more kilometers are needed or more simulations have to be carried out, as shown in section 7.1.

Another approach to account for additional field data is specified as an optional method step:

(VIII) Comparison of the additional field data with observed logical scenarios, possible extension of the number of clusters, repetition and application of the steps (II) to (VII).

6.2 Demonstration Case Study

The AEB function comprises a detection unit to determine the distance and speed of the Ego-vehicle relative to the leading object or obstacle. The emergency braking procedure is initiated in steps to achieve a predetermined target safety distance between the Ego-vehicle and the object or obstacle in front. The target safety distance corresponds to the latest time at which full braking must be initiated to avoid a collision. As per seen in chapters 3 and 4, the process of bringing a CMV to standstill requires a complex interaction between the brake actuator, the CMV tires, the CMV dimensions, the load conditions and the road surface.

False interventions can be avoided or at least significantly reduced by the detection of edge structures and boundary objects of the road (e.g. guideposts, crash barriers, traffic signs, etc.). Such constructions and objects depend on the road type and are taken into account when triggering the escalation levels $\mathcal{E}_s^1$, $\mathcal{E}_s^2$ and $\mathcal{E}_s^3$ by adapting the triggering conditions to the road classification. The ISO 22839:2013 specifies a test method to define the minimum functionality requirements of the AEB system with respect to FP and FN events. The scenario-based test concept emphasizes the testing of particularly critical traffic scenarios. Most typical traffic situations are not regarded as particularly dangerous and therefore contribute relatively less to the PoS. The identification of critical scenarios therefore requires indicators that quantify the criticality of traffic situations [Sch17a]. The deterministic TTR indicators are used, which describe the remaining time until a critical event occurs, such as Time HeadWay (THW) and TTC. The THW describes the required time for the Ego-vehicle to reach the current position of the relevant object, as illustrated in equation 6.1.

$$\mathrm{THW} = \frac{\mathrm{d}_\mathrm{x}^\mathrm{rel}}{\upsilon_\mathrm{x}^\mathrm{ego}} \quad , \forall \ \upsilon_\mathrm{x}^\mathrm{ego} > 0 \tag{6.1}$$

The TTC describes the time at which a collision occurs, if Ego-vehicle and detected object don't change their speed and direction. Nevertheless, TTC is used as a generic KPI for the functional evaluation even with constant relative speeds and accelerations. Equation 6.2 presents the mathematical description of the TTC at a constant relative velocity between the Ego-vehicle and the stationary object for each escalation level ($\mathcal{E}_s^1$, $\mathcal{E}_s^2$ and $\mathcal{E}_s^3$) with emergency braking of the Ego-vehicle until standstill.

$$\mathrm{TTC} = \frac{\mathrm{d}_x^{\mathrm{rel}}}{\upsilon_x^{\mathrm{rel}}} \quad ,\forall \ \mathrm{a}_x^{\mathrm{rel}} = 0, \quad \upsilon_x^{\mathrm{rel}} > 0, \quad \mathrm{d}_x^{\mathrm{rel}} > 0 \tag{6.2}$$

In case of a constant relative acceleration between the Ego-vehicle and the preceding object for each escalation level ($\mathcal{E}_s^1$, $\mathcal{E}_s^2$ and $\mathcal{E}_s^3$) with emergency braking of the Ego-vehicle until standstill, the equation 6.3 gives the mathematical description of the TTC, as follows:

$$\mathrm{TTC} = \frac{-\upsilon_x^{\mathrm{rel}} + \sqrt{\left(\upsilon_x^{\mathrm{rel}}\right)^2 + 2 * \mathrm{a}_x^{\mathrm{rel}} * \mathrm{d}_x^{\mathrm{rel}}}}{\mathrm{a}_x^{\mathrm{rel}}} \quad ,\forall \ \mathrm{a}_x^{\mathrm{rel}} < 0, \quad \upsilon_x^{\mathrm{rel}} > 0, \quad \mathrm{d}_x^{\mathrm{rel}} > 0 \tag{6.3}$$

The developed test method executes step (I) by elimination of logical errors through verification of the AEB function using the HiL platform — briefed in chapter 3, and detailed in chapter 4 — according to the knowledge-based requirements, as depicted in figure 6.2. The logical errors represent a subset of the discrepancies between the specified and implemented behavior of the AEB algorithm. The simulation results rely on synthetic data from a camera ECU and a RADAR sensor model, which are applied to the HiL test bench.

The TTC parameters are determined on the basis of the temporal progression of the Ego-vehicle's braking when approaching a stationary object at different longitudinal velocities $\upsilon_x^{\mathrm{rel}}$[km/h] at the various escalation levels ($\mathcal{E}_s^1$, $\mathcal{E}_s^2$ and $\mathcal{E}_s^3$). If the braking cascade of AEB controller is triggered when approaching a stationary object, the minimum TTC can be evaluated accordingly. Therefore, the pass/fail criteria use criticality parameters such as THW and TTC for the criticality assessment of the test results [JBKW18]. Other deterministic indicators use vehicle dynamics parameters and physical capabilities of the CMV to assess criticality, such as the required braking acceleration.

The deficit of these indicators is their reliance on the assumption of proper trajectory prediction. Consequently, small changes in motion prediction can

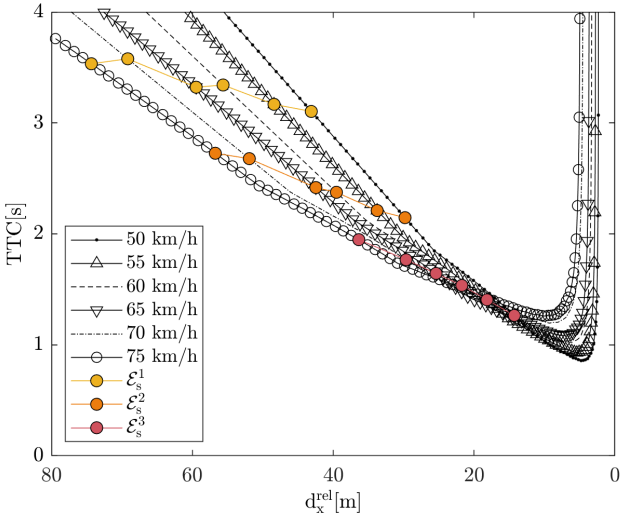

Figure 6.2: Identification of the TTC parameters using the HiL platform by scenarios approaching a stationary object with a straight road at different longitudinal velocities of the Ego-vehicle [ESS+19c].

lead to significantly different results [W+16]. In step (II), the measurement data from country-specific field tests are recorded and stored centrally in a cloud database. Then, the cluster analysis takes place to identify the FP and FN errors (elucidated in chapter 2) from the field observation. These error types occur due to environmental influences during operation and are statistical in nature for the discrepancy between the intended and specified behavior. The criticality thresholds are derived in step (III) from the simulation results using an appropriate regression function. Subsequently, a correlation analysis is applied between the proposed criticality threshold from synthetic driving scenarios and events from the individual grouped naturalistic driving situations. In step (IV), the HiL platform employs systematic test case generation to extend the test coverage within the applied ODD using ontologies and equivalence classes. The conversion rules derive the control data of concrete scenarios from the characteristic waveforms of the acquired sensor signals from the FOT.

6.3 Test Case Generation

Figure 6.3 illustrates exemplary time-series of the detected sensor signals of the cluster C_s^1. The cluster C_s^1 presents driving situations in a combined curve to the left, where the predicted road curvature κ_{ego}[1/km] is increased. At the same time, the lateral deviation to the stationary object in front d_y^{rel}[m] is decreased.

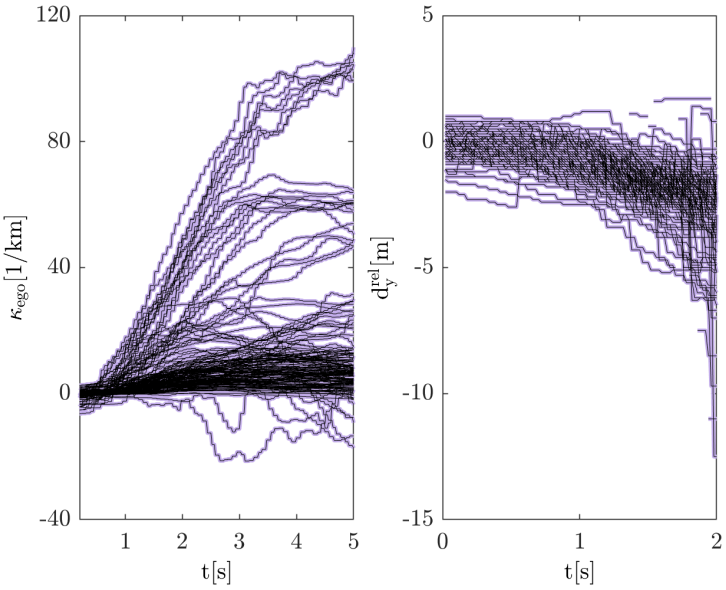

Figure 6.3: Event-based data acquisition of the curvature of the Ego-vehicle κ_{ego}[1/km] (left) and lateral offset of the stationary object in front d_y^{rel}[m] (right) for the cluster C_s^1 [ESD20].

On the basis of the stochastic variation within the parameter space of the logical scenarios, the process of generating the concrete parameter sets for the ontologies is ensured. Thereby, the identification of a concrete set of parameters within the ODD coverage follows the search for critical parameter sets within the parameter space. Accordingly, each concrete parameter set corresponds to a concrete scenario and vice versa.

Figure 6.4 depicts a logical scenario from the cluster C_s^1. The characteristic waveforms of cluster C_s^1 represent driving in a left turn, as seen in figures 5.6 and 5.8. The ontology uses a [consists_of] statement to model the elements of a road network layout with two lane classes and one class for a hard shoulder. The statements [has_right_neighbour] and [has_left_neighbour] are used to arrange road elements to each other. The position instances are generated on the basis of a relation [offers_position] for the ontology road elements. The statements [left_of], [right_of], [front_of] and [rear_of] are utilized to arrange the position instances with logic reasoning. The statements [driving_on] and [located_on] are employed to control the dynamic objects with different position instances. The object [obj1] is defined as a stationary object on the hard shoulder. The ontology specifies axioms that define constraints on attributes and relationships for specific concepts.

The semantic transformation rules utilize the open-loop sensor signals acquired by the ADDR to generate the respective closed-loop control data within the scenario synthesis process. These rules are applied, for example, to express the obtained curvature data $\kappa_{ego}[1/km]$ as Ego-vehicle trajectory data within Cartesian co-ordinates for the concrete scenarios. The direction angle $\omega_i[°]$ is measured between the tangent in the current position (i) of the clothoid and the initial direction with ($\kappa_{road} \approx 0$). The formula [$\omega_i = 0.5 * L_i * \kappa_i$] calculates the direction angle $\omega_i[°]$ based on the length of the clothoid arc to the current point L_i[km] and the current curvature κ_i[1/km] [Kol10]. Equations 6.4 and 6.5 calculate the x_i[m] and y_i[m] in Cartesian co-ordinates as follows:

$$x_i = \sqrt{\frac{2L_i}{\kappa_i}} * \sqrt{\omega_i} * \left\{ 1 - \frac{\omega_i^2}{(5)2!} + \frac{\omega_i^4}{(9)4!} - \cdots \right\} \tag{6.4}$$

$$y_i = \sqrt{\frac{2L_i}{\kappa_i}} * \sqrt{\omega_i} * \left\{ \frac{\omega_i}{3} - \frac{\omega_i^3}{(7)3!} + \frac{\omega_i^5}{(11)5!} - \cdots \right\} \tag{6.5}$$

The OWL is implemented in the ontology for each logical scenario, that allows the Semantic Web Rule Language (SWRL)[1] to combine logic operators into rules. Therefore, invalid or forbidden combinations can be eliminated from the scenario catalog.

[1] https://www.w3.org/Submission/SWRL/

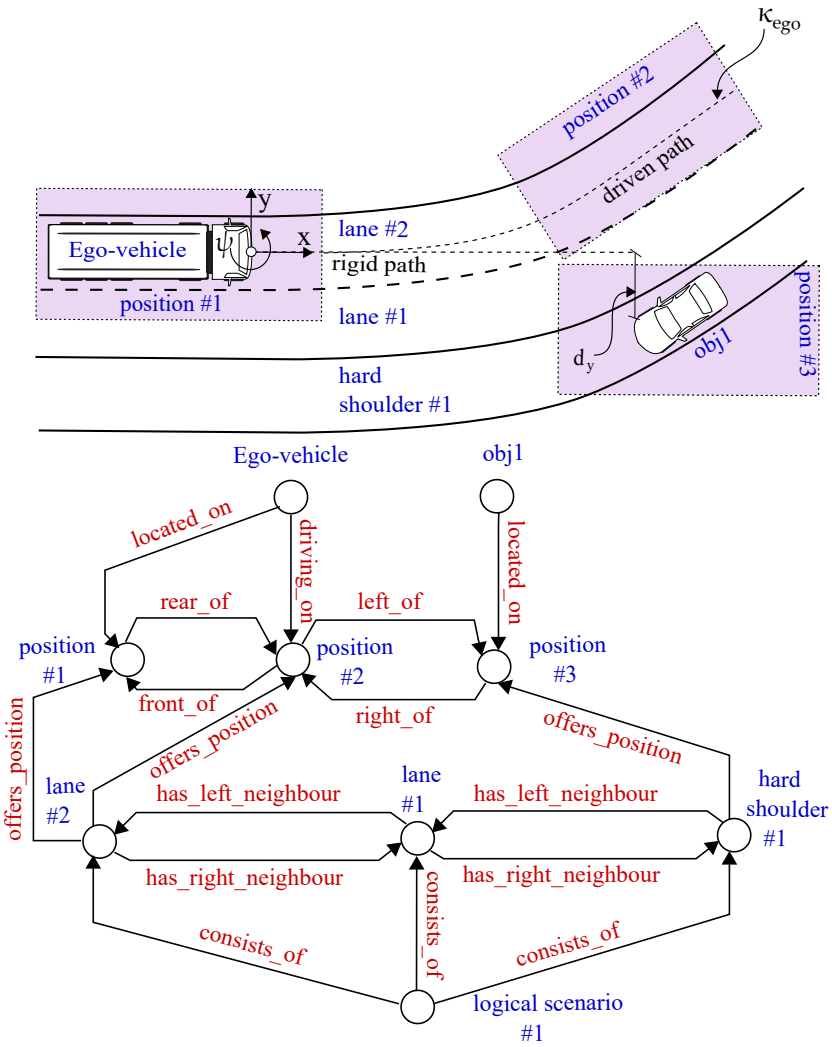

Figure 6.4: Logical scenario synthesis of cluster C_s^1 when driving on a left curve during coming close to a stationary object with 179 events [ESD20].

The SWRL rules are included in the ontology file[2] to extend the axioms of the logical scenario by inferring knowledge [UNGO21]. The context-driven test concept requires the overall probability P_T that an AEB function is faulty and executes a set of scenarios $\mathbf{F}_T$ in which the AEB function performed incorrectly. Since the logical scenario $\mathbf{F}_C$ for combined curves presents clusters C_s^1 and C_s^4 as a set of scenarios in which the Ego-vehicle drives in a combined curve with a stationary object on the side, its conditional probability is determined by $P(\mathbf{F}_T|\mathbf{F}_C)$ as probability that the system operates erroneously when driving in a curve with a stationary object on the side. $P(\mathbf{F}_C)$ represents the probability, whereby driving in curves with the object standing sideways occurs in all traffic scenarios. If the clusters C_s^2 and C_s^3 have to be considered, the ontology has to be expanded with the functional scenario $\mathbf{F}_S$ for S-shaped clothoid curves. The combined curve consists of a sequence of a clothoid, then a circular, and then a clothoid again. Meanwhile, the S-shaped clothoid consists two clothoid branches with curvature in the opposite direction and abutting at their zero point [Küh13].

Since the logical scenario $\mathbf{F}_S$ describes the set of scenarios in which the vehicle goes around a stationary object, its conditional probability is determined by $P(\mathbf{F}_T|\mathbf{F}_S)$ as probability that the system operates erroneously when driving around a stationary object. $P(\mathbf{F}_S)$ represents the probability, whereby driving around a stationary object occurs in all traffic scenarios. Equation 6.6 presents the overall probability P_T to fulfill the functional scenario catalogs extracted from cluster analysis of triggered events for an AEB function.

$$P_T = P(\mathbf{F}_T) = P(\mathbf{F}_T|\mathbf{F}_C) * P(\mathbf{F}_C) + P(\mathbf{F}_T|\mathbf{F}_S) * P(\mathbf{F}_S) \tag{6.6}$$

In general, a functional scenario can be defined as a combination of logical scenarios $\mathcal{U} = [\mathcal{U}_1, \mathcal{U}_2, \cdots, \mathcal{U}_n]$ of n-dimension. The overall probability of functional scenarios can be generally formulated as in equation 6.7.

$$P_T = P(\Psi) = \sum_{i=1}^{n} P(\Psi|\mathcal{U}_i) * P(\mathcal{U}_i) \tag{6.7}$$

[2] https://protege.stanford.edu/

6.4 Correlation Analysis

The curve fitting process uses an appropriate regression function to generate an ideal criticality threshold from the simulation calculations applied to the HiL test bench. While the variable slope sigmoidal equation is often used for a regression analysis of pharmacological dose-response curves, the four-parameter logistic regression is generally also employed for curve fitting analysis. The dose-response curve takes the form of a sigmoid curve that is shaped like the letter S.

Equation 6.8 uses the four-parameter logistic nonlinear regression to fit the curve of the $\mathrm{TTC}_{\mathrm{ref}}^{\mathcal{E}1}$ [s] measurements, where P1 denotes the baseline response, P2 the maximum response, P3 the turning point of $\upsilon_{\mathrm{x}}^{\mathcal{E}1}$ [km/h] giving a halfway response between the baseline and maximum, and P4 the curve slope. The $\mathrm{TTC}_{\mathrm{ref}}^{\mathcal{E}1}$ [s] refers to the ideal criticality threshold of the TTC for the escalation level $\mathcal{E}_{\mathrm{s}}^{1}$ based on synthetic data using driving scenarios approaching of a stationary object with a straight road at different longitudinal velocities of the Ego-vehicle.

$$\mathrm{TTC}_{\mathrm{ref}}^{\mathcal{E}1} = \mathrm{P1} + \frac{\mathrm{P2} - \mathrm{P1}}{1 + \left[\frac{10^{\mathrm{P3}}}{10^{\upsilon_{\mathrm{x}}^{\mathcal{E}1}}}\right]^{\mathrm{P4}}} = \mathrm{P1} + \frac{\mathrm{P2} - \mathrm{P1}}{1 + 10^{(\mathrm{P3} - \upsilon_{\mathrm{x}}^{\mathcal{E}1}) * \mathrm{P4}}} \quad , \forall \ \upsilon_{\mathrm{x}}^{\mathcal{E}1} > 0 \tag{6.8}$$

The sigmoidal dose-response equation extracts the logistic curve parameters based on the $\mathrm{TTC}_{\mathrm{ref}}^{\mathcal{E}1}$ [s] measurements from the HiL test bench, where (P1= 0.52, P2= 4.24, P3= 37.6 and P4= 0.02). Subsequently, the correlation analysis measures the statistical relationship between the FOT data and the ideal criticality threshold to a standardized covariance ranging between −1 and 1. The direction and strength of the linear dependence between the two time-series are quantified by that coefficient. The estimated Pearson Product-Moment Correlation Coefficient (PPMCC) becomes more inaccurate, as its value is closer to zero. If both variables have a strong positive correlation, the PPMCC is close to one, for a strong negative correlation is close to minus one. The linear dependence between $\mathrm{TTC}_{\mathrm{ref}}^{\mathcal{E}1}$ [s] and $\mathrm{TTC}_{\mathrm{fot}}^{\mathcal{E}1}$ [s] can be expressed as ratio between the covariance and the product of standard deviations.

The $\mathrm{TTC}_{\mathrm{fot}}^{\mathcal{E}1}$ [s] designates the TTC of the escalation level $\mathcal{E}_{\mathrm{s}}^{1}$ for the cluster C_{s}^{1} events with the combined curve to the left. Equation 6.9 calculates the PPMCC = 0.8055, where $\mathrm{cov}(\mathrm{TTC}_{\mathrm{ref}}^{\mathcal{E}1}, \mathrm{TTC}_{\mathrm{fot}}^{\mathcal{E}1})$ refers to the covariance, $\sigma_{\mathrm{TTC}_{\mathrm{ref}}^{\mathcal{E}1}}$ is the standard deviation of $\mathrm{TTC}_{\mathrm{ref}}^{\mathcal{E}1}$ [s] and $\sigma_{\mathrm{TTC}_{\mathrm{fot}}^{\mathcal{E}1}}$ is the standard deviation of $\mathrm{TTC}_{\mathrm{fot}}^{\mathcal{E}1}$ [s].

$$\mathrm{PPMCC} = \rho_{(\mathrm{TTC}_{\mathrm{ref}}^{\mathcal{E}1}, \mathrm{TTC}_{\mathrm{fot}}^{\mathcal{E}1})} = \frac{\mathrm{cov}(\mathrm{TTC}_{\mathrm{ref}}^{\mathcal{E}1}, \mathrm{TTC}_{\mathrm{fot}}^{\mathcal{E}1})}{\sigma_{\mathrm{TTC}_{\mathrm{ref}}^{\mathcal{E}1}} * \sigma_{\mathrm{TTC}_{\mathrm{fot}}^{\mathcal{E}1}}} \tag{6.9}$$

Equations 6.10 and 6.11 calculate the covariance and the standard deviations for $\mathrm{TTC}_{\mathrm{ref}}^{\mathcal{E}1}$ [s] and $\mathrm{TTC}_{\mathrm{fot}}^{\mathcal{E}1}$ [s], where $\mu_{\mathrm{TTC}_{\mathrm{ref}}^{\mathcal{E}1}}$ and $\mu_{\mathrm{TTC}_{\mathrm{fot}}^{\mathcal{E}1}}$ are the estimates of the mean values, respectively.

$$\mathrm{cov}(\mathrm{TTC}_{\mathrm{ref}}^{\mathcal{E}1}, \mathrm{TTC}_{\mathrm{fot}}^{\mathcal{E}1}) = \sum_{\mathrm{i}=1}^{\mathrm{n}} (\mathrm{TTC}_{\mathrm{ref}}^{\mathcal{E}1} - \mu_{\mathrm{TTC}_{\mathrm{ref}}^{\mathcal{E}1}})(\mathrm{TTC}_{\mathrm{fot}}^{\mathcal{E}1} - \mu_{\mathrm{TTC}_{\mathrm{fot}}^{\mathcal{E}1}}) \tag{6.10}$$

$$\sigma_{\mathrm{TTC}_{\mathrm{ref}}^{\mathcal{E}1}} = \sqrt{\sum_{\mathrm{i}=1}^{\mathrm{n}} \left(\mathrm{TTC}_{\mathrm{ref}}^{\mathcal{E}1} - \mu_{\mathrm{TTC}_{\mathrm{ref}}^{\mathcal{E}1}}\right)^2}, \sigma_{\mathrm{TTC}_{\mathrm{fot}}^{\mathcal{E}1}} = \sqrt{\sum_{\mathrm{i}=1}^{\mathrm{n}} \left(\mathrm{TTC}_{\mathrm{fot}}^{\mathcal{E}1} - \mu_{\mathrm{TTC}_{\mathrm{fot}}^{\mathcal{E}1}}\right)^2} \tag{6.11}$$

The coefficient of determination, denoted R^2, is the square of PPMCC with (R^2= 0.6488) and evaluates the strength of the linear relationship between the two time-series $\mathrm{TTC}_{\mathrm{ref}}^{\mathcal{E}1}$ [s] and $\mathrm{TTC}_{\mathrm{fot}}^{\mathcal{E}1}$ [s]. While the PPMCC and R^2 values show a strong positive correlation between $\mathrm{TTC}_{\mathrm{ref}}^{\mathcal{E}1}$ [s] and $\mathrm{TTC}_{\mathrm{fot}}^{\mathcal{E}1}$ [s], the equation 6.12 calculates the $\mathrm{d}_{\mathrm{TTC}}^{\mathcal{E}1}$ [s], which is assumed as a quantifiable error indicator of the environment perception.

$$\mathrm{d}_{\mathrm{TTC}}^{\mathcal{E}1} = \mathrm{TTC}_{\mathrm{fot}}^{\mathcal{E}1} - \mathrm{TTC}_{\mathrm{ref}}^{\mathcal{E}1} \tag{6.12}$$

The $\mathrm{TTC}_{\mathrm{fot}}^{\mathcal{E}1}$ parameters of triggered events from FOT can be correlated with the pass/fail criteria obtained from the HiL to identify the criticality of each triggered event. Figure 6.5 shows the parameter identification of $\mathrm{d}_{\mathrm{TTC}}^{\mathcal{E}1}$ [s] based on the correlation and regression analysis between $\mathrm{TTC}_{\mathrm{ref}}^{\mathcal{E}1}$ criticality threshold and the $\mathrm{TTC}_{\mathrm{fot}}^{\mathcal{E}1}$ for 179 events of cluster C_{s}^{1} at the escalation level $\mathcal{E}_{\mathrm{s}}^{1}$.

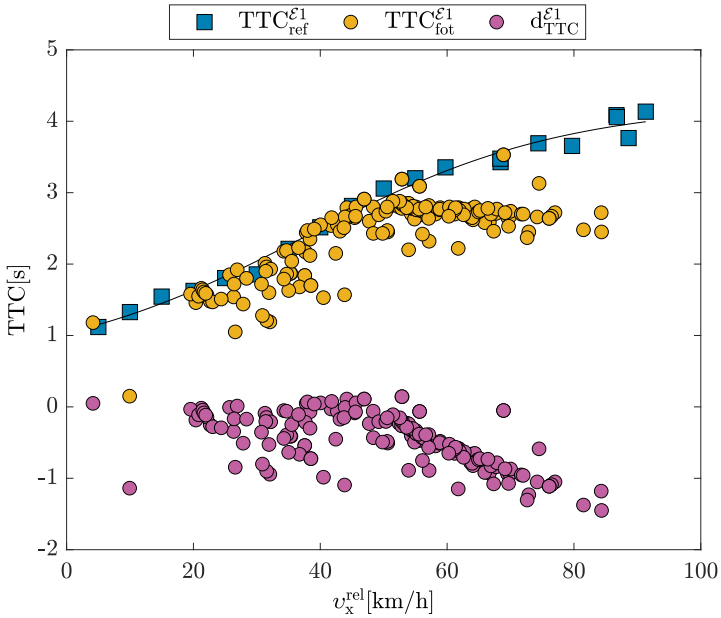

Figure 6.5: Correlation and regression analysis between digital and physical tests with the example of the TTC criticality assessment [ESD20].

To avoid accidents, CMVs need greater distance and time to come to a complete stop compared to passenger cars. The braking distance of a truck is even longer when it is hauling a heavy load and/or there are adverse road conditions such as snow, ice or rain, as demonstrated in chapter 4. At the same time, the AEB function requires to operate in the robust detection ranges, so the $\text{TTC}^{\mathcal{E}1}_{\text{fot}}$ curve saturates at higher velocities of $v_{\text{x}}^{\mathcal{E}1}$ above 50[km/h] to keep the reaction distance increasing linearly with the $v_{\text{x}}^{\text{rel}}$ and thus optimize the sensitivity and specificity of the AEB. The $\text{TTC}^{\mathcal{E}1}_{\text{ref}}$ is calculated by the HiL test platform based on synthetic sensor models, whereby the simulation environment is communicating with the embedded ECU for AEB function under real-time conditions. The $\text{d}^{\mathcal{E}1}_{\text{TTC}}$[s] represents the subtraction between $\text{TTC}^{\mathcal{E}1}_{\text{ref}}$[s] and $\text{TTC}^{\mathcal{E}1}_{\text{fot}}$[s] values and is expressed as a numerical value denoting a quantifiable error indicator.

6.5 Sensitivity Analysis

Saltelli et al. [S+08] defines the sensitivity analysis as a study of how the uncertainty in the output of a model or system can be apportioned and allocated to different sources of variation in its inputs. After defining the input variables and model responses, the sensitivity analysis then scans the parameter space with Design of Experiments (DoE) or sampling methods. Ebert et al. [EGK+20] applies the sensitivity analysis to reduce the dimension of the DoE to the most important factors as a part of the engine calibration process within the Real Driving Emissions (RDE) test cycle. Accordingly, the sensitivity analysis can be used to analyze the contribution of each input variable to the scatter of each model response. Each scattering input specifies a distribution type including mean value and variance, whereby the input variables are defined as random variables. The dependencies between the input variables can be formulated in terms of linear correlations. The meta-models are also used to represent the model responses by surrogate functions in terms of the model inputs [BSH17]. Meta-model approaches such as Kriging approximation, Metamodel of Optimal Prognosis (MOP) approximation, Support Vector Regression (SVR) and Artifical Neural Network (ANN) provide an automatic reduction of the variable space. The meta-model applied in this work is based on the anisotropic kriging approach using the Metamodel of Optimal Prognosis (MOP) solver [MW08].

Most et al. defines the coefficient of prognosis as a sensitivity measure to determine the optimal input variable subspace together with the optimal meta-model approach. The sensitivity analysis is applied in the example of the cluster C_s^1 at the escalation level $\mathcal{E}_s^1$. The sensitivity measures are determined in step (V) using the sensitivity analysis, whereby the input variables of the approximation model are ($v_x^{\mathcal{E}1}$[km/h],$d_x^{\mathcal{E}1}$[m], $d_y^{\mathcal{E}1}$[m], $\kappa_{ego}^{\mathcal{E}1}$[1/km] and $d_{TTC}^{\mathcal{E}1}$[s]) and the approximation model output is $TTC_{min}^{\mathcal{E}1}$[s]. The scalar random variables can be used to model the scattering inputs. The MOP solver reduced the input variables from [$v_x^{\mathcal{E}1}$,$d_x^{\mathcal{E}1}$,$d_y^{\mathcal{E}1}$,$\kappa_{ego}^{\mathcal{E}1}$,$d_{TTC}^{\mathcal{E}1}$] into [$v_x^{\mathcal{E}1}$,$d_{TTC}^{\mathcal{E}1}$]. While it is important to note that the variable $\kappa_{ego}^{\mathcal{E}1}$ was eliminated by the MOP solver, although curvature could play a role in determining $TTC_{min}^{\mathcal{E}1}$ for such FP events — as shown in the figure 6.4 — , it did not, arguing that correlation does not imply causation.

Definition 6.1 (minimum TTC): It describes the lowest time-to-collision in which an environmental perception sensor reports a tracked object list with objects classified as relevant for the ADF. The minimum TTC relates to the error-object classification in an inverse-proportional sense; the lower the value of $\mathrm{TTC}_{\mathrm{min}}^{\mathcal{E}1}$, the more prominent the error in classifying the object for such FP events in the cluster C_{s}^{1} [BES+21].

6.6 Exploration of Parameter Space

In general, the mean value and the standard deviation are used to describe the variation of a random variable. However, the shape of the distribution function can have a significant influence on the results of a stochastic analysis. Accordingly, the parameter space exploration requires the definition of distribution types including mean value and standard deviation for all random input variables. The truncated normal distribution is chosen as a suitable distribution type for the measurements of Ego-vehicle curvature $\kappa_{\mathrm{ego}}^{\mathcal{E}1}$[1/km], lateral deviation to the stationary object $\mathrm{d}_{\mathrm{y}}^{\mathcal{E}1}$[m] and error indicator of environment perception $\mathrm{d}_{\mathrm{TTC}}^{\mathcal{E}1}$[s] at the escalation level $\mathcal{E}_{\mathrm{s}}^{1}$ of the cluster C_{s}^{1}. The truncated normal distribution is a normal Gaussian distribution that is restricted to lie within a finite range, i.e. $[\kappa_{\mathrm{ego}}^{\mathrm{min}} \leq \kappa_{\mathrm{ego}}^{\mathcal{E}1} \leq \kappa_{\mathrm{ego}}^{\mathrm{max}}]$. The truncated normal distribution can be expressed in terms of the normal Gaussian distribution as follows:

$$f_{\mathrm{PDF}}(\kappa_{\mathrm{ego}}^{\mathcal{E}1}; \kappa_{\mathrm{ego}}^{\mathrm{min}}, \kappa_{\mathrm{ego}}^{\mathrm{max}}, \mu, \sigma) = \frac{1}{\sigma} \frac{\Omega\left(\frac{\kappa_{\mathrm{ego}}^{\mathcal{E}1}-\mu}{\sigma}\right)}{\Phi\left(\frac{\kappa_{\mathrm{ego}}^{\mathrm{max}}-\mu}{\sigma}\right) - \Phi\left(\frac{\kappa_{\mathrm{ego}}^{\mathrm{min}}-\mu}{\sigma}\right)}, \kappa_{\mathrm{ego}}^{\mathrm{min}} \leq \kappa_{\mathrm{ego}}^{\mathcal{E}1} \leq \kappa_{\mathrm{ego}}^{\mathrm{max}} \tag{6.13}$$

$$f_{\mathrm{CDF}}(\kappa_{\mathrm{ego}}^{\mathcal{E}1}; \kappa_{\mathrm{ego}}^{\mathrm{min}}, \kappa_{\mathrm{ego}}^{\mathrm{max}}, \mu, \sigma) = \frac{\Phi(\frac{\kappa_{\mathrm{ego}}^{\mathcal{E}1}-\mu}{\sigma}) - \Phi(\frac{\kappa_{\mathrm{ego}}^{\mathrm{min}}-\mu}{\sigma})}{\Phi(\frac{\kappa_{\mathrm{ego}}^{\mathrm{max}}-\mu}{\sigma}) - \Phi(\frac{\kappa_{\mathrm{ego}}^{\mathrm{min}}-\mu}{\sigma})}, \kappa_{\mathrm{ego}}^{\mathrm{min}} \leq \kappa_{\mathrm{ego}}^{\mathcal{E}1} \leq \kappa_{\mathrm{ego}}^{\mathrm{max}} \tag{6.14}$$

where μ and σ denote the mean and standard deviation of the parent normal distribution and $\kappa_{\mathrm{ego}}^{\mathrm{min}} = -4.5$[1/km] and $\kappa_{\mathrm{ego}}^{\mathrm{max}} = 3.9$[1/km] as the lower and upper truncation points respectively.

The Ω and Φ designate the Probability Distribution Function (PDF) and the Cumulative Distribution Function (CDF) for the normal Gaussian distribution respectively. Figure 6.6 illustrates the PDF and CDF estimations using the truncated normal distribution. Equations 6.15 and 6.16 estimate the mean and standard deviation of the truncated normal distribution.

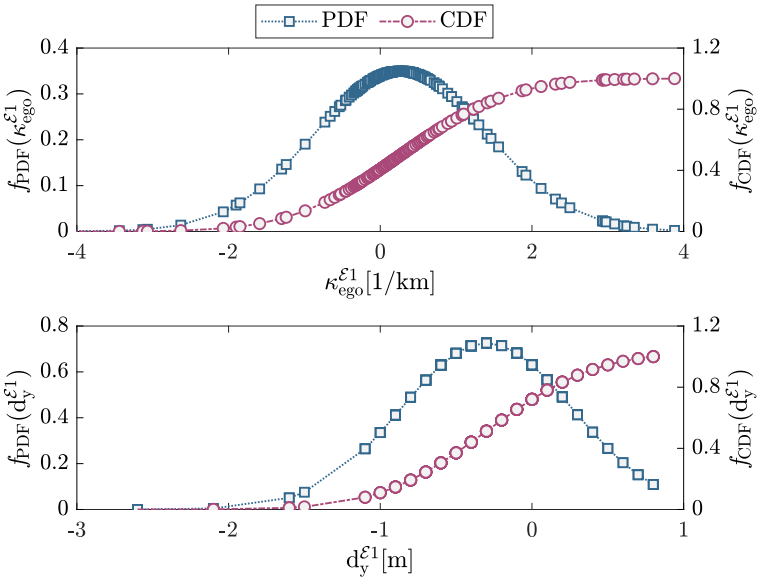

Figure 6.6: Estimation of the PDF and CDF of the Ego-vehicle curvature $\kappa_{\mathrm{ego}}^{\mathcal{E}1}$[1/km] (top) and the lateral deviation to the stationary object $\mathrm{d}_{\mathrm{y}}^{\mathcal{E}1}$[m] (bottom) at the escalation level $\mathcal{E}_{\mathrm{s}}^{1}$ of the cluster C_{s}^{1}.

$$\mu_{\kappa} = \mu + \sigma * \frac{\Omega(\frac{\kappa_{\mathrm{ego}}^{\mathrm{min}} - \mu}{\sigma}) - \Omega(\frac{\kappa_{\mathrm{ego}}^{\mathrm{max}} - \mu}{\sigma})}{\Phi(\frac{\kappa_{\mathrm{ego}}^{\mathrm{max}} - \mu}{\sigma}) - \Phi(\frac{\kappa_{\mathrm{ego}}^{\mathrm{min}} - \mu}{\sigma})} \tag{6.15}$$

$$\sigma_\kappa = \sigma * \sqrt{1 + \frac{(\frac{\kappa_{ego}^{min}-\mu}{\sigma})\Omega(\frac{\kappa_{ego}^{min}-\mu}{\sigma}) - (\frac{\kappa_{ego}^{max}-\mu}{\sigma})\Omega(\frac{\kappa_{ego}^{max}-\mu}{\sigma})}{\Phi(\frac{\kappa_{ego}^{max}-\mu}{\sigma}) - \Phi(\frac{\kappa_{ego}^{min}-\mu}{\sigma})} - \left[\frac{\Omega(\frac{\kappa_{ego}^{min}-\mu}{\sigma}) - \Omega(\frac{\kappa_{ego}^{max}-\mu}{\sigma})}{\Phi(\frac{\kappa_{ego}^{max}-\mu}{\sigma}) - \Phi(\frac{\kappa_{ego}^{min}-\mu}{\sigma})}\right]^2} \tag{6.16}$$

The same calculations are applied for the lateral deviation to the stationary object $d_y^{\mathcal{E}1}$[m], where $d_y^{min} = -2.6$[m] and $d_y^{max} = 0.8$[m] as the lower and upper truncation points respectively. The false warnings from the cluster C_s^1 are presented as an example for the sensitivity analysis. Accordingly, the sensitivity measures are analyzed to define the important input variables in relation to the response variable $TTC_{min}^{\mathcal{E}1}$[s].

The triangular distribution is chosen as a suitable distribution type for the measurements of the Ego-vehicle longitudinal velocity $v_x^{\mathcal{E}1}$[km/h] and the minimum time to collision $TTC_{min}^{\mathcal{E}1}$[s] at the escalation level $\mathcal{E}_s^1$ of the cluster C_s^1. The following equations calculate the PDF, CDF, mean and standard deviation of the Ego-vehicle longitudinal velocity $v_x^{\mathcal{E}1}$[km/h] based on the triangular distribution respectively, where $v_x^{min} = 7.9$[km/h], $v_x^{max} = 83.4$[km/h] and $v_x^{peak} = 59.6$[km/h] as the lower limit, upper limit and mode respectively. Equations 6.19 and 6.20 estimate the mean and standard deviation of the triangular distribution of the $v_x^{\mathcal{E}1}$[km/h].

$$f_{PDF}(v_x^{\mathcal{E}1}; v_x^{min}, v_x^{max}, v_x^{peak}) = \begin{cases} \frac{2[v_x^{\mathcal{E}1} - v_x^{min}]}{[v_x^{max} - v_x^{min}][v_x^{peak} - v_x^{min}]}, & \forall v_x^{min} \leq v_x^{\mathcal{E}1} \leq v_x^{peak} \\ \frac{2[v_x^{max} - v_x^{\mathcal{E}1}]}{[v_x^{max} - v_x^{min}][v_x^{max} - v_x^{peak}]}, & \forall v_x^{peak} < v_x^{\mathcal{E}1} \leq v_x^{max} \end{cases} \tag{6.17}$$

$$f_{CDF}(v_x^{\mathcal{E}1}; v_x^{min}, v_x^{max}, v_x^{peak}) = \begin{cases} \frac{[v_x^{\mathcal{E}1} - v_x^{min}]^2}{[v_x^{max} - v_x^{min}][v_x^{peak} - v_x^{min}]}, & \forall v_x^{min} \leq v_x^{\mathcal{E}1} \leq v_x^{peak} \\ 1 - \frac{[v_x^{max} - v_x^{\mathcal{E}1}]^2}{[v_x^{max} - v_x^{min}][v_x^{max} - v_x^{peak}]}, & \forall v_x^{peak} < v_x^{\mathcal{E}1} \leq v_x^{max} \end{cases} \tag{6.18}$$

$$\mu_v = \frac{1}{3}[v_x^{min} + v_x^{max} + v_x^{peak}] \tag{6.19}$$

$$\sigma_{\upsilon} = \sqrt{\frac{[\upsilon_{\mathrm{x}}^{\mathrm{min}}]^2 + [\upsilon_{\mathrm{x}}^{\mathrm{max}}]^2 + [\upsilon_{\mathrm{x}}^{\mathrm{peak}}]^2 - [\upsilon_{\mathrm{x}}^{\mathrm{min}} * \upsilon_{\mathrm{x}}^{\mathrm{max}}] - [\upsilon_{\mathrm{x}}^{\mathrm{min}} * \upsilon_{\mathrm{x}}^{\mathrm{peak}}] - [\upsilon_{\mathrm{x}}^{\mathrm{max}} * \upsilon_{\mathrm{x}}^{\mathrm{peak}}]}{18}} \tag{6.20}$$

The same calculations are applied for the minimum time to collision $\mathrm{TTC}_{\mathrm{min}}^{\mathcal{E}1}$[s] on the basis of the triangular distribution function. Figure 6.7 shows the PDF and CDF estimations of the $\upsilon_{\mathrm{x}}^{\mathcal{E}1}$[km/h] using the triangular distribution and the $\mathrm{d}_{\mathrm{TTC}}^{\mathcal{E}1}$[s] using the truncated normal distribution. The calculations are applied for the error indicator of environment perception $\mathrm{d}_{\mathrm{TTC}}^{\mathcal{E}1}$[s], where the minimum value of $\mathrm{d}_{\mathrm{TTC}}^{\mathcal{E}1} = -1.5$[s] and the maximum value of $\mathrm{d}_{\mathrm{TTC}}^{\mathcal{E}1} = 0.14$[s] are the lower and upper truncation points respectively.

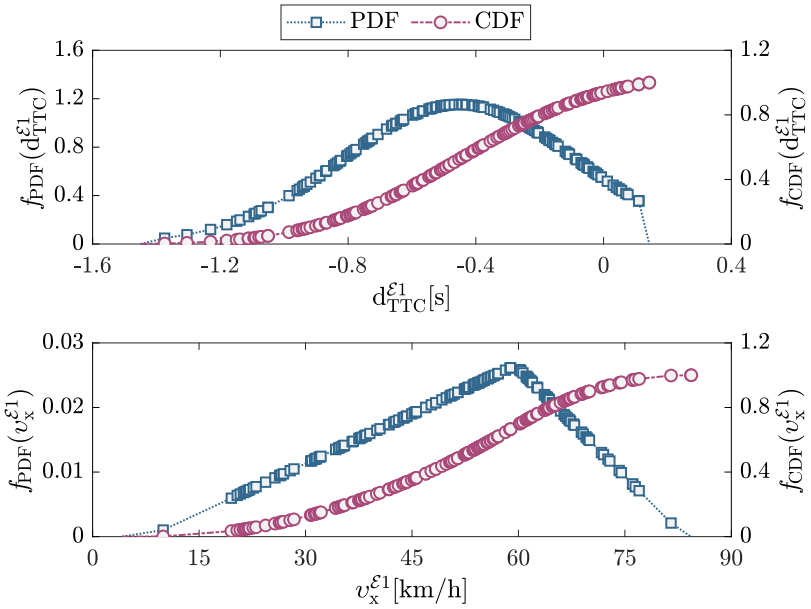

Figure 6.7: Estimation of the PDF and CDF of the error indicator of environment perception $\mathrm{d}_{\mathrm{TTC}}^{\mathcal{E}1}$[s] (top) and the Ego-vehicle longitudinal velocity $\upsilon_{\mathrm{x}}^{\mathcal{E}1}$[m] (bottom) at the escalation level $\mathcal{E}_{\mathrm{s}}^{1}$ of the cluster C_{s}^{1}.

7 Reliability Analysis Using Sampling Methods

The ADFs in a vehicle have to comply with the ISO 26262:2018 standard for functional safety and are not allowed to influence each other negatively. The Regulation No. 79 applies to the steering equipment of CMVs for lane change operations. Meanwhilst, the ACSF of Category E requires minimum monitoring ranges for the front, rear and side detection of automated vehicles [EST+17]. If a motorcycle approaches the Ego-vehicle from behind, the rear detection has to take place early enough to prevent the motorcycle from braking abruptly when the Ego-vehicle changes lanes. If the distance is below a critical range, the lane change request must be suppressed. At the same time, the CMVs include a variety of series and models with tractors, semi-tractors or trailer combinations. Consequently, the tractor-mounted environmental perception sensors cannot provide a full 360° surround-view when a trailer is coupled to the tractor. In Germany, according to Section 4 of the German road traffic regulations, the Ego-vehicle in front must not brake suddenly without a compelling reason. Therefore, compliance with ISO 26262:2018 is achieved if an ADF represents an acceptable residual risk for the subsequent traffic. A probabilistic safety analysis is needed to quantify the uncertainties and to enable a prospective risk assessment for the testing and further development of ADFs [BRW+17].

7.1 Probabilistic Safety Assessment

Following the application of cluster and sensitivity analysis to the real traffic data-set — interpreted in chapters 5 and 6 — , the data is then used to determine the probability distributions of the input variables and the probability of occurrence of the response variables [Mul18].

In accordance with ISO 26262:2018, the reference values of the tolerable risk curve are derived in step (VI) as a confidence level for the optimization quality of ADFs. The measurable safety methodology aims at evaluating the SOTIF- or OEDR-related capabilities of ADFs. Accordingly, the reliability analysis aims at identifying the probability of safety violation for each logical scenario [AKK+19]. The parameter space is searched with a suitable sampling method to determine the probability that a predefined safety margin is exceeded [Roo02]. The scattering input variables are considered as random variables within the framework of probabilistic safety analysis. Therefore, the probability of failure P_f is defined as the probability of the event that a random vector Z falls into the failure domain. The failure mode involves defining a failure criterion known as the Limit State Function (LSF) or safety margin. A direct formulation of the LSF in dependence of the input random variables is often not possible, but is implicitly given by the simulation results of the investigated system [Bay99]. The LSF forms the basis for the reliability calculations and denoted by $g(Z)$. The function expresses Resistance-Load as a function of Z, where Z is a vector of all uncertainty variables describing the loads and resistances. The failure criterion is consequently defined as $g(Z) \leq \gamma$, where γ denotes a predefined safety margin. For a given LSF with $g(Z)$, the probability of failure P_f is defined as follows:

$$P_f = P[g(Z) \leq \gamma] = \int \cdots_{\{z \in \mathbb{R}^n | g(z) \leq \gamma\}} \int f_Z(z) dz \tag{7.1}$$

where $\{z \in \mathbb{R}^n | g(z) \leq \gamma\}$ represents the failure domain, (z) designates the realization of Z in the variable space and $f_Z(z)$ denotes the joint density function for Z. Accordingly, the probability of failure is the integral of the joint PDF over the failure domain. While the probability of failure is the complementary of reliability, the failure criteria do not have to indicate a system breakdown. Such failure criteria can be defined by the violation of quality or safety requirements. The probability integral can be expressed as the expectation of the indicator function $I(g(z))$, where $I(g(z)) = 1$ if $g(z) \leq \gamma$ and $I(g(z)) = 0$ otherwise. The probability integral can be interpreted as the expected value $\mathbb{E}$ of the indicator $I(g(z))$ as follows:

$$P_f = \mathbb{E}[I(g(z))] = \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} I(g(z)) f_Z(z) dz \tag{7.2}$$

While the space of all random variables can be separated into a safe domain and a failure domain, the indicator's expected value $\mathbb{E}[I(\mathrm{g}(z))]$ calculates the probability of severity violations within the prospective risk assessment. The reliability analysis identifies the failure region in the parameter space to predict the probability of failure for each logical scenario. While the analytical calculation of expected values $\mathbb{E}[I(\mathrm{g}(z))]$ with the dimensional increase of input random variables is no longer suitable for practical use, the Monte Carlo integration approximates the indicator's expectation $\mathbb{E}[I(\mathrm{g}(z))]$ by simulation-based approximation. The simulation of rare events approximates the integral solution using Monte Carlo Sampling (MCS) and Adaptive Importance Sampling (AIS) methods [WPZ19]. In compliance with the HARA of the ISO 26262:2018, the failure criteria present the different levels of severity [LPP11], whereby the $\mathrm{TTC}_{\mathrm{min}}^{\mathcal{E}1}$[s] defines the different severity levels in an exemplary way.

The $\mathrm{TTC}_{\mathrm{min}}^{\mathcal{E}1}$ describes the lowest TTC in which an environmental perception sensor reports a tracked object list with objects classified as relevant for the AEB function of the cluster $\mathcal{C}_{\mathrm{s}}^{1}$ at the escalation level $\mathcal{E}_{\mathrm{s}}^{1}$. The sensitivity analysis defines the variables $v_{\mathrm{x}}^{\mathcal{E}1}$[km/h] and $\mathrm{d}_{\mathrm{TTC}}^{\mathcal{E}1}$[s] as the most important factors influencing the variable $\mathrm{TTC}_{\mathrm{min}}^{\mathcal{E}1}$[s] as a response within the probabilistic safety assessment. The selection is determined by the variables with the greatest impact on the model output $\mathrm{TTC}_{\mathrm{min}}^{\mathcal{E}1}$[s]. Accordingly, the reliability analysis is applied to the inputs and response of the cluster $\mathcal{C}_{\mathrm{s}}^{1}$ at the escalation level $\mathcal{E}_{\mathrm{s}}^{1}$. As discussed in section 5.3.3, the $\mathcal{E}_{\mathrm{s}}^{1}$ events represent phantom warnings, where the situation was safe and a warning was erroneously signaled. The prospective risk assessment $\mathcal{R}\left(\mathcal{C}_{\mathrm{s}}^{1}\right)$ is defined as the combination of the probability of occurrence $P(\mathcal{C}_{\mathrm{s}}^{1})$ of harm and the severity of that harm. Table 7.1 shows the severity levels, which are divided into four regions as follows:

Severity level condition	**Hypothetical severity level**
$\mathcal{S}_{\mathrm{s}}^{3} := [\ \mathrm{TTC}_{\mathrm{min}}^{\mathcal{E}1} \leq 0.5\mathrm{s}]$	critical severity
$\mathcal{S}_{\mathrm{s}}^{2} := [0.5\mathrm{s} < \mathrm{TTC}_{\mathrm{min}}^{\mathcal{E}1} \leq 1\mathrm{s}]$	high severity
$\mathcal{S}_{\mathrm{s}}^{1} := [1\mathrm{s} < \mathrm{TTC}_{\mathrm{min}}^{\mathcal{E}1} \leq 2\mathrm{s}]$	medium severity
$\mathcal{S}_{\mathrm{s}}^{0} := [\ \mathrm{TTC}_{\mathrm{min}}^{\mathcal{E}1} > 2\mathrm{s}]$	low severity

Table 7.1: Hypothetical severity levels for object classification error based on the criticality indicator $\mathrm{TTC}_{\mathrm{min}}^{\mathcal{E}1}$[s].

Equation 7.3 represents the prospective risk for the cluster C_s^1 with the combined curve to the left. In step (VII), the tolerable risk curve is generated as a reference safety threshold with respect to the logical scenarios. Thereby, the logical scenario with combined curve to the left is shown as an example in order to prospectively estimate the risk regarding false alarms on subsequent traffic.

$$\mathcal{R}\left(C_s^1\right) \approx \left[\mathrm{P_f}(\mathcal{S}_s^3 \mid C_s^1) * P(C_s^1)\right] + \left[\mathrm{P_f}(\mathcal{S}_s^2 \mid C_s^1) * P(C_s^1)\right] + \left[\mathrm{P_f}(\mathcal{S}_s^1 \mid C_s^1) * P(C_s^1)\right] + \left[\mathrm{P_f}(\mathcal{S}_s^0 \mid C_s^1) * P(C_s^1)\right] \tag{7.3}$$

Figure 7.1 represents the PDF and CDF estimations of the minimum time to collision $\mathrm{TTC}_{\mathrm{min}}^{\mathcal{E}1}$[s] using the triangular distribution function with the corresponding severity levels. The $\mathrm{TTC}_{\mathrm{min}}^{\mathcal{E}1}$[s] is used as a prototypical safety measure for the reliability analysis. The original measurements of $\mathrm{TTC}_{\mathrm{min}}^{\mathcal{E}1}$ range between 0.03[s] and 2.27[s]. The triangular distribution is selected to approximate the $\mathrm{TTC}_{\mathrm{min}}^{\mathcal{E}1}$ distribution, where the lower bound of $\mathrm{TTC}_{\mathrm{min}}^{\mathcal{E}1} = -0.3$[s] and the upper bound and mode of $\mathrm{TTC}_{\mathrm{min}}^{\mathcal{E}1} = 1.91$[s].

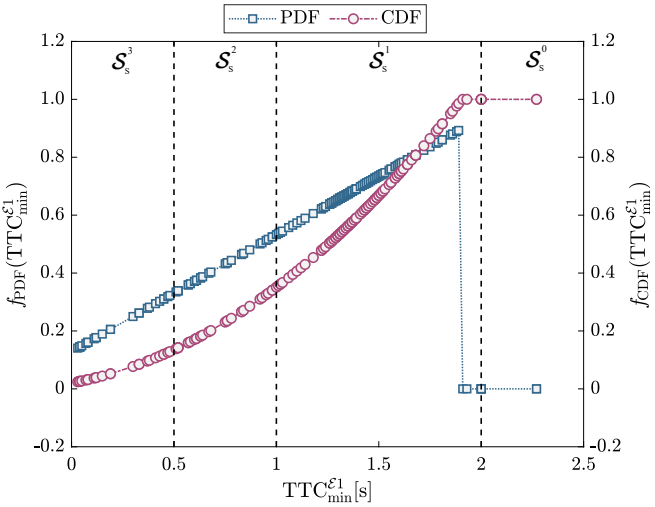

Figure 7.1: Estimation of the PDF and CDF of the response variable $\mathrm{TTC}_{\mathrm{min}}^{\mathcal{E}1}$[s] of the cluster C_s^1 at the escalation level $\mathcal{E}_s^1$.

7.2 Monte Carlo Simulation

The MCS method estimates P_f by generating independent samples from the PDF of $f_Z(z)$ and taking the average of the sample indicator values as an unbiased estimator of the expected probability of failure [Zio13]. While the joint density function $f_Z(z)$ depends on random input variables, the estimator of its probability failure $\bar{P}_f$ is also a random variable [Buc09]. Equation 7.4 represents the probability failure estimator $\bar{P}_f$ as follows:

$$\bar{P}_f = \bar{\mathbb{E}}[I(g(z))] = \frac{1}{N_{mc}} \sum_{k=1}^{N_{mc}} I(g(z_k)) \tag{7.4}$$

The estimator variance $\mathrm{Var}(\bar{P}_f)$ can be computed approximately from the generated samples. In accordance with the central limit theorem, the standard deviation converges towards 0 for a large number of samples $N_{mc} \rightarrow \infty$. If the estimator's variance converges towards 0, the confidence interval of the estimator becomes narrower and thus the estimator gets more confident. The number of samples is used to normalize the standard deviation of the unbiased estimator $\bar{P}_f$ of the target failure probability P_f [P+14]. The normalized standard deviation is used as measure for the completion of the required number of generated samples and is known as standard error $e_{\bar{P}_f}$.

$$e_{\bar{P}_f} = \sqrt{\frac{\mathrm{Var}(\bar{P}_f)}{N_{mc}}} \tag{7.5}$$

The MCS method evaluates the probability of failure by determining whether the LSFs are exceeded. The trial is repeated many times to ensure the convergence of the statistical results. In each trail, a sample value is generated and evaluated as a Hit or Miss (relevant or irrelevant) according to the LSF definition. The MCS requires a large number of samples to accurately predict the probability of failure, especially if the expected value is small. Figure 7.2 illustrates the result of MCS method used to obtain the failure probability of the severity level $\mathcal{S}_s^3$. As shown in figure 7.2, the failure probability $\bar{P}_f(\mathcal{S}_s^3|\mathcal{C}_s^1)$ and the standard error $e_{\bar{P}_f}(\mathcal{S}_s^3|\mathcal{C}_s^1)$ are plotted against the number of samples N_{mc}.

The MCS sampling is a Hit-or-Miss sampling method. Here, the blue dots indicate the target Hit samples at which the corresponding LSF condition for each severity level is met. Meanwhilst, the yellow spots show the Miss condition, that means they don't fulfill the conditions required for the severity level. The Hit-or-Miss MCS method generates N_{mc} samples with 1000 (assumed budget of samples), where the number of Hits is 99 and the number of Misses is 901. The failure probability $\bar{P}_f(\mathcal{S}_s^3|\mathcal{C}_s^1)$ is calculated to be $\frac{99}{1000} = 0.1$. For this severity level, the Hit samples are very little compared to Miss samples and are located for the lower values of $v_x^{\mathcal{E}1}$[km/h] towards the left end of the plot. The target is missed for the values of $v_x^{\mathcal{E}1}$ higher than 50[km/h] (In which $v_x^{\mathcal{E}1}$ indicates the velocity at escalation level 1). The footage representing such a scenario — cluster 1 — is depicted in chapter 5. The stopping criteria of simulation depend on the comparison between the standard error of the exceedance probability[1] $e_{\bar{P}_f}(\mathcal{S}_s^3|\mathcal{C}_s^1) = 0.0094$ and the target standard error $\delta_{\bar{P}_f} = 0.05$.

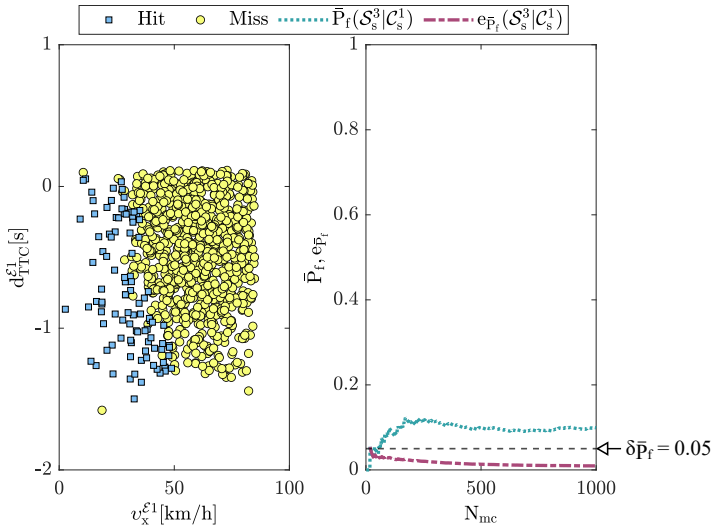

Figure 7.2: Estimation of the failure probability for the severity level $\mathcal{S}_s^3$ using the MCS method.

[1] The exceedance probability represents a probability of exceeding a certain limit that can be quantified and demonstrated to be less than an acceptable value.

Figure 7.3 shows the failure probability of the severity level $\mathcal{S}_s^2$ using the MCS method with a number of samples $N_{mc} = 1800$. The number of Hits is 335 and the number of Misses is 1465. Considering the graph for severity level $\mathcal{S}_s^2$ as shown in figure 7.3, here, the condition seems to be fulfilled for more numbers of samples. This means that the number of Hit samples is increased and also the value of $\upsilon_x^{\mathcal{E}1}$[km/h] for which the condition seems to be fulfilled. After the first few samplings, this failure probability value falls and then becomes stable at a level $\bar{P}_f(\mathcal{S}_s^2|\mathcal{C}_s^1) = \frac{335}{1800} = 0.19$. In addition, the standard error becomes stable gradually at the value $e_{\bar{P}_f}(\mathcal{S}_s^2|\mathcal{C}_s^1) = 0.0092$ to be less than the required standard error $\delta_{\bar{P}_f}$ with 0.05. Ultimately, the standard error does not seem to be entirely nullified, but is really low. It makes a cloud of the Hit samples slightly shifted towards the right side. It is important to note that the figure 7.3 shows the ratio of Hits and their corresponding velocity $\upsilon_x^{\mathcal{E}1}$[km/h]. It is obvious that as the velocity $\upsilon_x^{\mathcal{E}1}$[km/h] increases, the sensor error $d_{TTC}^{\mathcal{E}1}$[s] also increases for the specified criticality range $\mathcal{S}_s^2$.

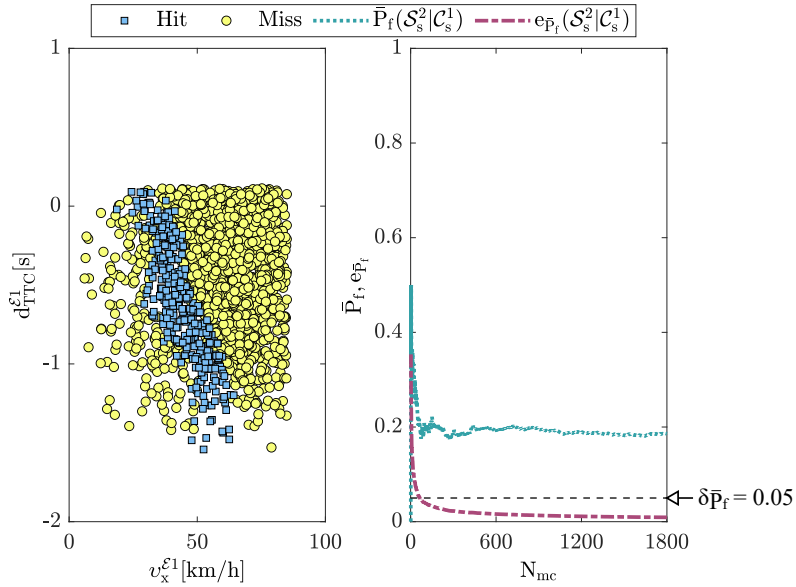

Figure 7.3: Estimation of the failure probability for the severity level $\mathcal{S}_s^2$ using the MCS method.

Figure 7.4 illustrates the failure probability of the severity level $\mathcal{S}_s^1$ using the MCS method, with a number of samples N_{mc}= 500. The number of Hits is 257 and the number of Misses is 243. Here, the Hit condition is improved as compared to the previous severity level. For the severity level $\mathcal{S}_s^1$, the values of $\upsilon_x^{\mathcal{E}1}$[km/h] have also increased and the cloud of Hit samples moves rightwards again. According to the results of this sampling, the failure probability of this level of severity is high at the beginning of sampling. It falls immediately after some samples and then stabilizes at a certain level of $\bar{P}_f(\mathcal{S}_s^1|\mathcal{C}_s^1) = \frac{257}{500} = 0.51$ with infinitesimal difference in each sampled value. The standard error $e_{\bar{P}_f}(\mathcal{S}_s^1|\mathcal{C}_s^1)$ gives the value 0.022 as abort criterion of the simulation, which is smaller than the target standard error $\delta_{\bar{P}_f}$.

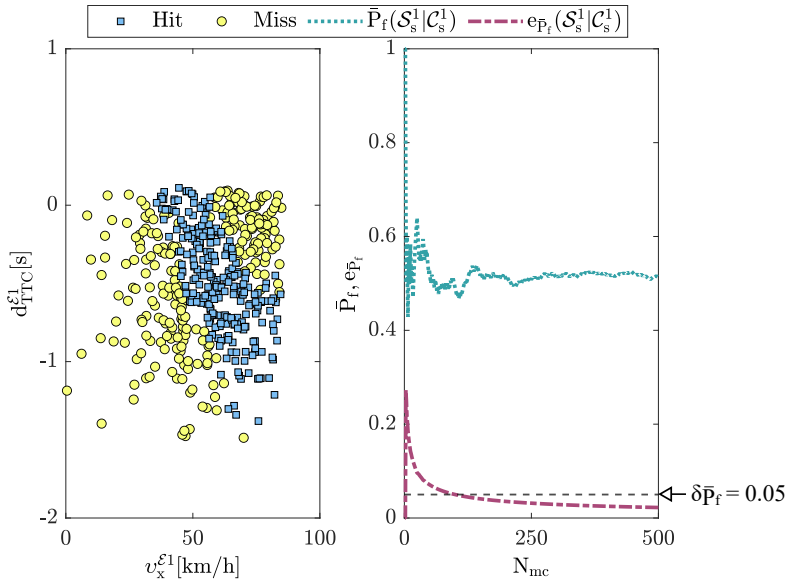

Figure 7.4: Estimation of the failure probability for the severity level $\mathcal{S}_s^1$ using the MCS method.

Figure 7.5 describes the failure probability for the severity level $\mathcal{S}_s^0$ using the MCS with a number of samples N_{mc} = 1600. The number of Hits is 321 and the number of Misses is 1279. This sampling does show the stability in the failure probability with $\bar{P}_f(\mathcal{S}_s^0|\mathcal{C}_s^1) = \frac{321}{1600} = 0.2$. The value of the standard error is less than the previous case with $e_{\bar{P}_f}(\mathcal{S}_s^0|\mathcal{C}_s^1) = 0.01$.

Finally switching to the figure 7.5, the cloud of Hit samples lies to the right side of the graph for the severity level $\mathcal{S}_s^0$. Here, the target Hits are a little less in number compared to severity levels $\mathcal{S}_s^1$ and $\mathcal{S}_s^2$. It is significant to note, that with correspondence to figure 7.1 the studied range contains only a small number of events. Thus, rendering this statistic improbable. However, the indication from the graph follow accordingly to the expected flow.

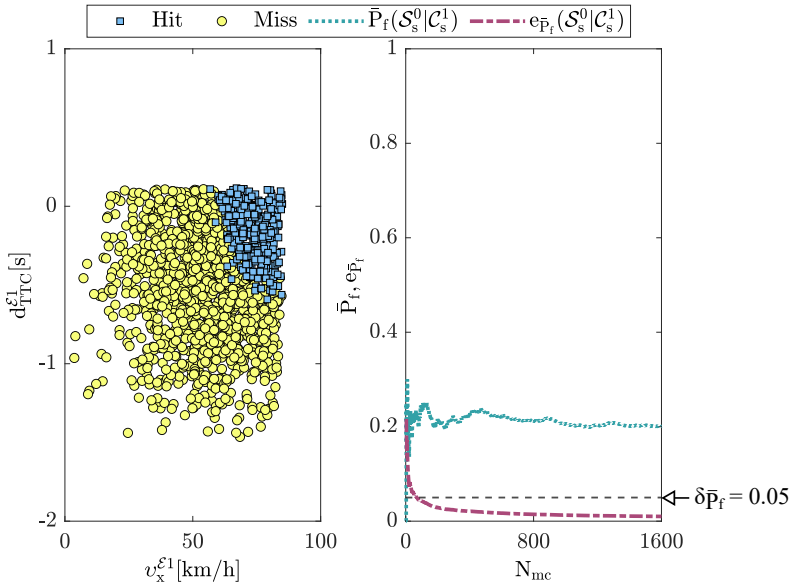

Figure 7.5: Estimation of the failure probability for the severity level $\mathcal{S}_s^0$ using the MCS method.

The figures 7.2 to 7.5 illustrate that there is a movement in the position of the target Hits starting from the bottom left moving towards the top right through the severity levels, looking like a cloud of Hits.

7.3 Adaptive Importance Sampling

Since the estimator variance corresponds to the confidence, variance reduction techniques aim at influencing the sampling in such a way that the estimator variance becomes smaller.

The Importance Sampling is a variance reduction method to guide the sampling using biased normal distribution for variance reduction. The estimator includes the Importance Sampling weight to warrant unbiasedness of the estimator. The samples are not generated following the prescribed joint PDF of $f_Z(z)$, but with an importance sampling density denoted as $h_Y(y)$. Setting $h_Y(y)/h_Y(y)$ into the probability integral does not change its value, but the integral provides the expected value.

$$\mathrm{P_f} = \mathbb{E}\left[I(\mathrm{g}(y))\frac{f_Z(y)}{h_Y(y)}\right] = \int_{-\infty}^{\infty}\cdots\int_{-\infty}^{\infty} I(\mathrm{g}(y))\frac{f_Z(y)}{h_Y(y)}h_Y(y)\mathrm{d}y \tag{7.6}$$

$$\bar{\mathrm{P}}_\mathrm{f} = \bar{\mathbb{E}}\left[I(\mathrm{g}(y))\frac{f_Z(y)}{h_Y(y)}\right] = \frac{1}{\mathrm{N_{as}}}\sum_{\mathrm{k}=1}^{\mathrm{N_{as}}} I(\mathrm{g}(y_\mathrm{k}))\frac{f_Z(y_\mathrm{k})}{h_Y(y_\mathrm{k})} \tag{7.7}$$

The AIS reduces the estimator variance and involves several simulation runs [O+18]. The importance sampling density $h_Y(y)$ may have a larger scatter. In the first iteration, the samples falling into the failure domain $\{z \in \mathbb{R}^\mathrm{n} | \mathrm{g}(z) < \gamma\} : Z|\mathrm{g}(z) \leq \gamma$ are statistically evaluated. The result is used to define the parameter of normal Gaussian distribution type for the subsequent Importance Sampling iterations. To increase the number of Hits, the AIS method is managed through the use of information about the LSF domain. Each sample is weighted according to the ratio of the original density function $f_Z(z)$ to the importance density function $h_Y(y)$ to ensure correct statistics [KS13]. Furthermore, the AIS techniques can be applied to efficiently search for critical scenarios by speeding up the parameter space exploration [O+18]. Equations 7.8 and 7.9 show the expected values of the importance sampling density $h_Y(y)$ according to the estimated mean and covariance of the samples in the failure domain in iterative steps [Li07].

$$\mathbb{E}[Y] = \mathbb{E}[Z|\mathrm{g}(z) \leq \gamma] \tag{7.8}$$

$$\mathbb{E}[YY^T] = \mathbb{E}[ZZ^T|\mathrm{g}(z) \leq \gamma] \tag{7.9}$$

The third iteration should be performed to prove the stability of the results in an iterative adaption scheme. Accordingly, the goal of the importance sampling density $h_Y(y)$ is to converge the estimator standard error to zero by adjusting the importance sampling density $h_Y(y)$ iteratively [BMC15]. Typically, the target standard error for the termination criteria is set to 0.1, but in order to conduct more samples, the target standard error was set to 0.05. (The value = 0.05 was previously utilized in MCS sampling).

Figure 7.6 shows the estimation of failure probability of the severity level $\mathcal{S}_s^3$ using the AIS method with number of samples N_{as} = 800. The AIS algorithm applies four iterations, where each iteration consists of 200 samples. The number of Hits is 512 and the number of Misses is 288. The standard error is calculated to be $e_{\bar{P}_f}(\mathcal{S}_s^3|\mathcal{C}_s^1) = 0.0088$. The standard error of the estimator doesn't appear to be entirely eliminated, but is extremely low. The failure probability is computed to be $\bar{P}_f(\mathcal{S}_s^3|\mathcal{C}_s^1) = 0.1$ as confirmation of the MCS result for the same severity level.

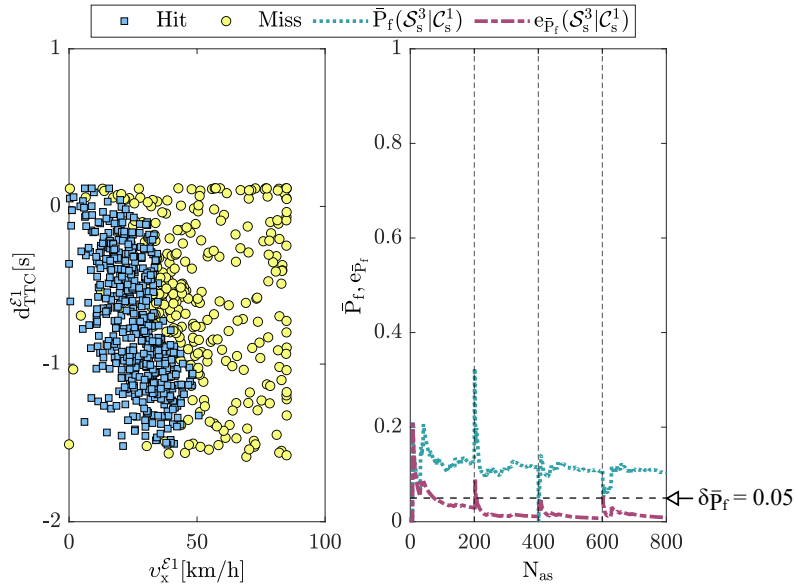

Figure 7.6: Estimation of failure probability for the severity level $\mathcal{S}_s^3$ using the AIS method. For comparative purposes, these calculations in MCS are shown in figure 7.2.

The figure 7.7 describes the behavior of the failure probability and standard error after sampling using number of samples $N_{as} = 900$. The sampling is performed in five iterations, with the first three iterations having a budget of 100 samples, the fourth with 200 samples, and the fifth with 400 samples. The number of Hits is 434 and the number of Misses is 446. The standard error is computed to be $e_{\bar{P}_f}(\mathcal{S}_s^2|\mathcal{C}_s^1) = 0.0088$. This shows that with AIS, the variance is successfully eliminated. This proves that the AIS method is effective in eliminating the variance with less numbers of samples compared to the MCS method. The failure probability is calculated to be $\bar{P}_f(\mathcal{S}_s^2|\mathcal{C}_s^1) = 0.18$ as confirmation of the MCS result for the severity level $\mathcal{S}_s^2$.

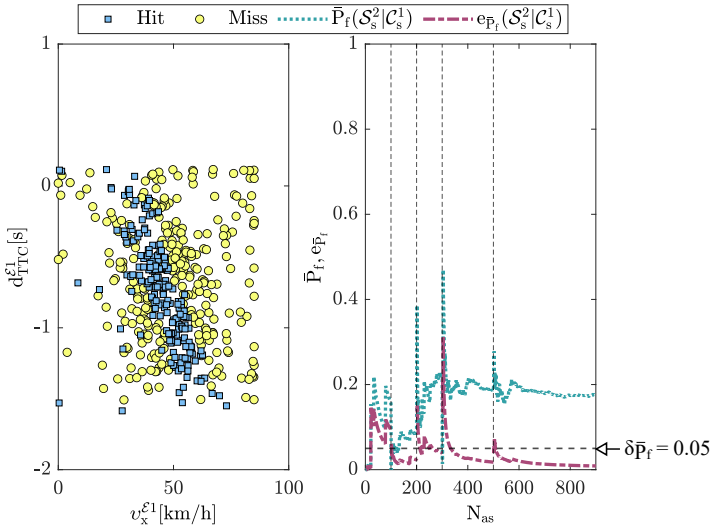

Figure 7.7: Estimation of failure probability for the severity level $\mathcal{S}_s^2$ using the AIS method. For comparative purposes, these calculations in MCS are shown in figure 7.3.

As shown in figure 7.8, the probability of failure $\bar{P}_f(\mathcal{S}_s^1|\mathcal{C}_s^1)$ with 0.52 using the AIS method confirms the result extracted from the MCS method for the same severity level. The convergence of the AIS method occurs with the required standard error of 0.05 using a number of samples $N_{as} = 800$. Accordingly, the standard error $e_{\bar{P}_f}(\mathcal{S}^1|\mathcal{C}_s^1)$ of 0.036 is used as abort criteria for the simulation runs.

The sampling is performed in four iterations, with the first three iterations having a budget of 100 samples and the fourth with 500 samples. The number of Hits is 616 and the number of Misses is 184.

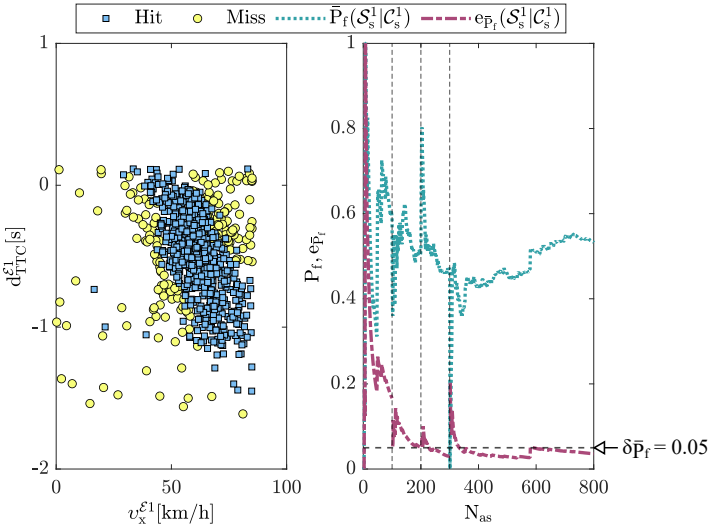

Figure 7.8: Estimation of failure probability for the severity level $\mathcal{S}_s^1$ using the AIS method. For comparative purposes, these calculations in MCS are shown in figure 7.4.

Figure 7.9 shows that the failure probability at severity level $\mathcal{S}_s^0$ using the AIS method and number of samples $N_{as} = 800$. The sampling is executed in four iterations, with the first three iterations having a budget of 100 samples and the fourth with 500 samples. The number of Hits is 581 and the number of Misses is 219. Because of the very few available $TTC_{min}^{\mathcal{E}1}$ events of the cluster $\mathcal{C}_s^1$ at the severity level $\mathcal{S}_s^0$ as shown in figure 7.1, the AIS method generates a statistical uncertainty with the result of the failure probability. As aforementioned, due to poor sampling in the region of $TTC_{min}^{\mathcal{E}1} > 2[s]$, there were not enough sample to explorate the values within the specific used budget of samples, resulting in an non-coherent probability failure not achieving $e_{\bar{P}_f}(\mathcal{S}^0|\mathcal{C}_s^1)$ less than 0.05. Figure 7.9 shows that AIS method is the least suitable for these error ranges of $TTC_{min}^{\mathcal{E}1}$ and MCS method ought to be used instead.

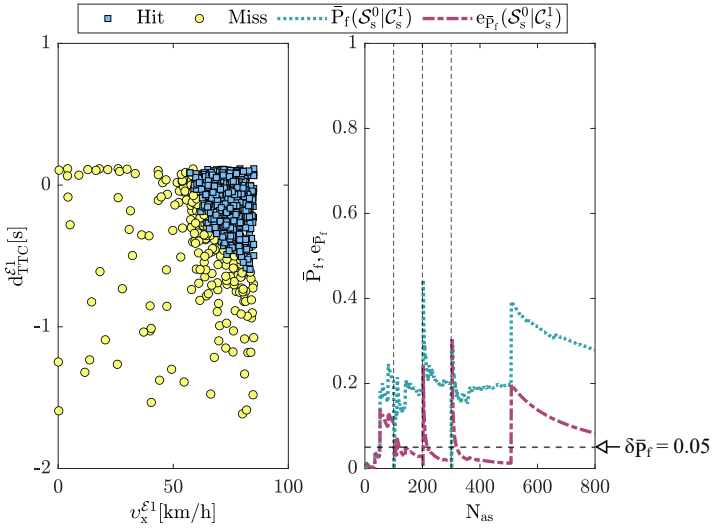

Figure 7.9: Estimation of failure probability for the severity level $\mathcal{S}_s^0$ using the AIS method. For comparative purposes, these calculations in MCS are shown in figure 7.5.

Figures 7.6 to 7.9 show a similar type of movement of the Hit cloud from the bottom left to the top right in comparison with figures 7.2 to 7.5. In many cases, the AIS method requires fewer samples to achieve the required standard error. For several iterations, the parameters of the importance sampling density function $h_Y(y)$ can be determined adaptively. The probability of failure decreases from $\mathcal{S}_s^1$ to $\mathcal{S}_s^3$, while $\mathcal{S}_s^0$ occurs with a bigger probability than $\mathcal{S}_s^1$. The safety statement can be obtained in that the failure probability for each traffic scenario is approximated by estimating the exceedance region in the parameter space. At the beginning of the AIS method, the required standard error $\delta_{\bar{P}_f}$ of 0.1 is chosen for the investigations, which generally represents a good confidence in the result. While the AIS fulfills its required standard error of 0.1 after a few simulation runs, but the iteration results differ strongly, the required standard error is reduced to $\delta_{\bar{P}_f} = 0.05$ to force more iterations for the AIS method. For comparability reasons, $\delta_{\bar{P}_f} = 0.05$ is also given for the MCS algorithm in all severity levels. The MCS method is used as a reference, where the results of MCS and AIS methods confirm each other.

8 Conclusions

The incomplete and implicit requirements are a part of the verification challenges of ADFs. There is no complete public computer-aided traffic simulator that contains the rules for dealing with exceptions. Moreover, testing can detect the presence of errors, but not their absence. In addition, the Pesticide paradox phenomenon occurs after fixing the software errors found by a scenario catalog. As a consequence, the same catalog may no longer be suitable for the residual errors. The incompleteness of the functional requirements can thus be covered by requirements extraction from the field observation. The on-road testing plays a decisive role to identify the missing functional requirements and transform these requirements adaptively into the deployed scenario catalogs. The scenario-based testing assumes that the majority of the mileage on the motorway is accomplished without particular events, while critical scenarios in real traffic are rather rare and randomly distributed. The identification and reproducibility of critical scenarios from the KDD of on-road testing in laboratory simulations or proving grounds leads to a significant reduction of the FOT distances required for statistical approval. The scenario-based functional decomposition, therefore, splits complex functions into sub-functions in order to identify relevant driving scenarios that reduce the approval effort for ADFs.

8.1 Executive Summary

The status quo evaluation refers to large-scale verification as one of the decisive challenges for the economical, reliable and safe use of ADFs in CMV product engineering. Therefore, the systems engineering has established data-based and knowledge-driven test methods to assure the required dependability of their products.

The required test distance is approximately 220 million kilometers with error-free driving to show that the automated CMV operates with a confidence level of 95% as safely as a human driver, as indicated in section 2.4.

RQ1: **How can the knowledge-based and data-driven test platforms be combined in a complementary and collaborative manner?**

Due the rare nature of critical events, the FOTs will require a prohibitive number of driving hours to prove the statistical superiority. At the same time, the required total mileage is increased if an error occurs during the validation process. Therefore, the sole reliance on FOTs is inadequate and, in particular, time and cost-intensive when applied to the next generation of ADFs, e.g. truck platooning and hub-to-hub transportation. Owing to the complexity of the ADFs, a test oracle for safety certification is unlikely to be provided by a single test platform. Therefore, innovative approaches are required to enable systematic testing under reproducible conditions with regard to robustness, reliability and safety, as discussed in chapters 3 and 5. In addition, functional decomposition is also necessary to support the argumentation for a reasonably low residual risk resulting from imperfections of the environmental perception sensors, as elucidated in chapter 4.

RQ2: **How can the ADFs be effectively and efficiently tested?**

The context-driven approach — as indicated in section 3.2 — employs a grey-box test strategy that combines the insights from the observed road data with functional requirements within the ODD coverage. In figure 3.4, the argument structure of the context-driven test concept is shown with the help of the GSN. The required test termination criteria can be met by quantifying the conflicts between the test concept requirements, and thus, achieving a trade-off between efficiency and effectiveness criteria of the test procedures. The correlation between the various test coverage criteria intends to support controlling the decision of choosing whether more test kilometers need to be collected or more simulations require to be carried out. The optimized test strategy requires a selection of the necessary test methods for various scenarios and their interaction with other test methods. Thereby, new and innovative approaches need to be designed, especially in simulation and in the laboratory. The knowledge-based test platforms are executed at component and system levels to generate the corresponding test evidence, as explained in chapter 4.

Moreover, the test results of the data-driven test platforms are evaluated using statistical extrapolation techniques and field-based observations within the ODD coverage, as discussed in chapter 5.

RQ3: **How can the prospective risk be measured in order to achieve reasonable termination criteria for the testing of the ADFs?**

As indicated in chapter 5, the measurable safety approach envisages the identification of failure types to break down the functional complexity. Furthermore, the presented framework structure utilizes a back-end database filled with catalogs of corner driving scenarios from different sources of field-based observations. The extensive data-sets are collected world-wide under realistic driving conditions with test CMV fleets. The evaluation of algorithms requires an efficient search and interpretation of relevant traffic situations with help of RCA to modify the algorithm accordingly. In this scheme, the processing chain includes clustering of multivariate time-series data-sets and finding critical driving situations to identify and allocate the necessary test cases for various suitable test environments. The platform-independent mechanism is intended to offer a consistent scenario description format for the various test environments. Further, these new test cases complement the existing test cases developed from expert knowledge in an adaptive ODD coverage manner.

A case study is provided to illustrate the measurable safety framework, in which the behavior of CMVs equipped with an AEB system with respect to stationary objects is discussed. The AEB can also unexpectedly issue warnings and brake the vehicle if it detects stationary objects next to the vehicle's own lane, e.g. broken-down vehicles, signs, bridges and traffic islands. The extracted clusters and their parameter space define the probability distributions of the associated parameters of each logical scenario. The minimum TTC is used as a prototypical safety measure. Thereby, the sensitivity analysis reduces the number of input variables to the most important ones — as discussed in chapter 6 — , which influence the safety measure. Subsequently, the reliability analysis identifies the failure region in the parameter space to predict the probability of failure for each logical scenario using MCS and AIS methods, as indicated in chapter 7. The generated tolerable risk curve is then used as a reference safety threshold — as portrayed in figure 6.1 — for the test termination criteria.

Meanwhilst, the real PoS can only be provided after the market introduction of automated CMVs in which the probability of collisions can be estimated based on the relevant parameter space within the ODD. However, the probability of failure can be approximated by estimating the exceedance probability of parameters with higher criticality while validating the performance of existing levels of automation. Accordingly, there is a hypothesis based on the assumption that erroneous triggering events of the present ADFs during the field observation can be considered as disengagement events for automated CMVs. Therefore, the prospective risk assessment of an existing ADF provides reference values for the risk acceptance threshold. Such reference values act as a benchmark for the further development and optimization of a similar ADF at the next level of automation.

8.2 Technical and Scientific Contributions

The proposed framework enables functional verification of ADFs on the embedded ECUs in complex automotive sensor networks. The framework utilizes a real-time capable system architecture in a distributed heterogeneous co-simulation environment. In this architecture, object-list-based sensor models are designed to simulate realistic sensor behavior. The real-time interaction between the HiL test bench and the DuT imposes timing constraints to ensure traceability and reproducibility of the test results. The described architecture contains both the structural design and time-efficient integration into the test bench and is applied to RADAR and camera sensors. An important feature of the architecture is the high portability between various driving simulation frameworks due to well-defined in-/output interfaces [BMK+16]. Furthermore, a run-time fault injection has been introduced to simulate sensor failures, such as latency, detection failure and false one-to-many object labeling. These failures are used for testing the robustness of the ADFs. Accordingly, a phenomenological sensor model can be developed iteratively. Sensor failures can be added in a continuous manner in order to achieve increasing degrees of realism. For efficient verification, an astute selection of relevant test scenarios is conducted to avoid repetitive test scenarios. The research results of this thesis have been published in internationally well-recognized journals and peer-reviewed conferences.

Six invention disclosures focusing on test procedures have been filed and used as references in this work. In the *DE 102017009971A1*, the invention relates to a method for testing a lane assistance system for a vehicle by extracting an ontology-based category of adequate and relevant scenarios for existing field tests [ESF19]. In the *DE 102018004429A1*, the invention deals with a cluster analytical characterization of driving situations based on detected sensor signals for the detection of surroundings and their system reactions in the driving operation of the vehicle [ESS+19c]. In the *DE 102018005864A1*, the invention relates to a method for testing a blind spot assistance system for a vehicle, particularly for a truck. On the basis of the Tractrix zone of the Ego-vehicle, objects are detected by sensors and their distance and relative speed to the Ego-vehicle are measured. The cluster analysis is applied to determine the criticality of the detected ambient objects [ESS+19b].

In the *DE 102018005865A1*, the invention relates to a method for testing a map-based system for speed limit in a vehicle. The system detects traffic signs based on maps and fused image, wherein acquired information to be tested are classified by means of a cluster analysis for scenarios into equivalence classes [ESS+19a]. In the *DE 102020005507A1*, the invention refers to a method for testing an emergency braking function of a vehicle based on a confidence for existing field tests. Thereby, the sensitivity and reliability analysis identifies a failure range in the parameter space to predict a failure probability for each extracted logical scenario using sampling methods [ESD20].

In the *DE 102020006644A1*, the invention relates to a method for evaluating systematic and statistical errors of multi-radar-based recognition systems of a vehicle. According to the invention, exceptions from field-based observations and knowledge experience are identified by means of processes of knowledge development by clustering acquired data records and determining the exceptions in order to identify and assign necessary test data records for various suitable test environments. By means of the determined test data-sets, test data-sets developed on the basis of expert knowledge are completed in an adaptive test coverage [EDA20].

8.3 Future Research Directions

The presented work raises several issues that require substantial future research activities. The quality of ADFs depends primarily on the environmental sensors providing the vehicle's environmental perception as the basis for situation analysis and decision making. An acceptable level of maturity of these sensors must be accomplished as a prerequisite for an adequate field validation. The logical and statistical errors or functional deficiencies are assessed by objective and subjective evaluation criteria.

Since it is not possible to guarantee complete safety for automated CMVs, one of the major challenges in automated truck driving is to argue for a reasonably low residual risk resulting from limitations of the environmental perception sensor. Currently, relevant safety norms do not support such arguments. While the Threat Assessment and Remediation Analysis (TARA) is one of the key activities defined in the ISO/SAE 21434 for automotive cyber-security, a SOTIF similar standard is also required for security issues.

Moreover, the standards for safety (ISO 26262 and ISO/PAS 21448) and security (ISO/SAE 21434) need to be more harmonized. The cyber-security is a condition in which assets are adequately protected against threat scenarios to electrical or electronic components of road vehicles and their functions [fS$^+$20]. Therefore, further research will include the application of the test method for security issues. These activities have to be integrated into an ASE approach that supports the structure of the context-driven test concept. This research work needs to be complemented by activities with standard organizations to form a consensus on risk evaluation and acceptable argumentation structures that would feed into future standards and CoP guidelines for the safeguarding of ADFs.

9 Glossary

9.1 List of Acronyms

Preliminaries

ABA	Active Brake Assist
ACC	Adaptive Cruise Control
ACSF	Automatically Commanded Steering Function
ADA	Active Drive Assist
ADAS	Advanced Driver Assistance System
ADASIS	ADAS Interface Specification
ADF	Automated Driving Function
AEB	Autonomous Emergency Braking
ANSI	American National Standards Institute
ASE	Automotive Systems Engineering
ASIL	Automotive Safety Integrity Level
ASPICE	Automotive Software Process Improvement and Capability dEtermination
AST	Adaptive Stress Testing
ATA	American Trucking Associations
BMWi	German Federal Ministry for Economic Affairs and Energy
CAN	Controller Area Network
CFAR	Constant False Alarm Rate
CMMI	Capability Maturity Model Integration
CMV	Commercial Motor Vehicle
CoP	Code of Practice
DBSCAN	Density-Based Spatial Clustering of Applications with Noise
DDT	Dynamic Driving Task

DESTATIS	Federal STATIStical Office of Germany
DIN	German Insitute for Standardization
DIS	Draft International Standard
EC	European Commission
ECU	Electronic Control Unit
E/E	Electrical and/or Electronic
eHorizon	electronic Horizon
FMI	Functional Mock-up Interface
FN	False Negative
FOT	Field Operational Test
FOV	Field of View
FP	False Positive
GPS	Global Positioning System
GSN	Goal Structuring Notation
HARA	Hazard Analysis and Risk Assessment
HD	High Definition
HiL	Hardware-in-the-Loop
HoL	Hardware-in-the-open-Loop
HW/SW	Hardware/ Software
IEC	International Electrotechnical Commission
IEEE	Institute of Electrical and Electronics Engineers
ISO	International Organization for Standardization
KPI	Key Performance Indicator
LDP	Lane Departure Protection
LDW	Lane Departure Warning
LiDAR	Light Detection And Ranging
MC/DC	Modified Condition/Decision Coverage
MIIT	Ministry of Industry and Information Technology
MiL	Model-in-the-Loop
MOBATSim	MOdel-Based Autonomous Traffic Simulation Framework
MOP	Metamodel of Optimal Prognosis
MPP	Most Probable Path

MRC	Minimal Risk Condition
MTBF	Mean Time Between Failures
MVP	Minimum Viable Product
NCAP	New Car Assessment Programme
NDS	Navigation Data Standard
NHTSA	National Highway Traffic Safety Administration
NL	Natural Language
ODD	Operational Design Domain
OEDR	Object and Event Detection and Response
PAS	Publicly Available Specification
PCA	Principal Component Analysis
PoS	Proof of Safety
PPV	Positive Predictive Value
QG	Quality Gate
QM	Quality Management
RADAR	RAdio Detection And Ranging
RCA	Root Cause Analysis
RCS	Radar Cross Section
ROC	Receiver Operating Characteristic
SAE	Society of Automotive Engineers
SDLC	Software Development Life Cycle
SWE	Software Engineering Process Group
SENSORIS	SENSOR Interface Specification
SGA	Side Guard Assist
SiL	Software-in-the-Loop
SOTIF	Safety Of the Intended Functionality
SPI	Safety Performance Indicator
STL	Signal Temporal Logic
SVM	Support Vector Machine
TaaS	Transport-as-a-Service
TN	True Negative
TP	True Positive

TPEG	Transport Protocol Experts Group
TRL	Technology Readiness Level
TSR	Traffic Sign Recognition
TTC	Time To Collision
TTR	Time To React
UL	Underwriters Laboratories
UML	Unified Modeling Language
UN/ECE	United Nations Economic Commission for Europe
UNGA	United Nations General Assembly
UNR	United Nations Regulation
U.S.	United States
USDOT	United States Department of Transportation
VCRT	Vienna Convention on Road Traffic
VRU	Vulnerable Road User
VSSA	Voluntary Safety Self Assessment
V&V	Verification and Validation
VVA	Verification, Validation and Accreditation
WHO	World Health Organization
XiL	X(something)-in-the-Loop

Framework Conception

ABox	Assertional Box
ADDR	Automated Driving Data Recorder
ADTF	Automotive Data and Time-triggered Framework
AWS	Amazon Web Services
CCR	California Code of Regulation
CRF	Camera Reference Frame
CSV	Comma-Separated Values
DL	Description Logic
DuT	Device under Test
DVI	Digital Visual Interface
EDR	Event Data Recorder

GIDAS	German In-Depth Accident Study
GNSS	Global Navigation Satellite System
HAC	Hierarchical Agglomerative Clustering
HDFS	Hadoop Distributed File System
highD	highway Drone data-set
HMI	Human-Machine Interface
KDD	Knowledge Discovery in Databases
IMU	Inertial Measurement Unit
inD	intersection Drone data-set
KITTI	Karsruhe Institute of Technology and Toyota technological Institute data-set
LTE	Long Term Evolution
MDF	Measurement Data Format
NAS	Network Attached Storage
NCC	Normalized Cross Correlation
NMVCCS	National Motor Vehicle Crash Causation Survey
OTA	Over-The-Air
OWL	Ontology Web Language
POD	Plug On Device
PTP	Precision Time Protocol
RDF	Resource Description Framework
REST-API	REpresentational State Transfer - Application Programming Interface
RTT	Round Trip Time
RTOS	Real Time Operating System
TBox	Terminological Box
TCP	Transmission Control Protocol
UDP	User Datagram Protocol
V2X	Vehicle to X(everything)
WGS84	World Geodetic System 1984
XCP	Universal (X) Measurement and Calibration Protocol
XML	eXtensible Markup Language

Framework Implementation

AIS	Adaptive Importance Sampling
ANN	Artifical Neural Network
CDF	Cumulative Distribution Function
DoE	Design of Experiments
LSF	Limit State Function
MCS	Monte Carlo Sampling
MLS	Moving Least Squares
MOP	Metamodel of Optimal Prognosis
PDF	Probability Distribution Function
PPMCC	Pearson Product-Moment Correlation Coefficient
RDE	Real Driving Emissions
SVR	Support Vector Regression
SWRL	Semantic Web Rule Language
TARA	Threat Assessment and Remediation Analysis
THW	Time HeadWay

9.2 List of Variables and Constants

Preliminaries

a_x^{ego}	longitundinal acceleration of the Ego-vehicle.
C	confidence level of the Binomial distribution.
d_c	accumulated driving distance for the statistical PoS.
$\tilde{d}_x^{rel}$	falsified relative longitudinal distance of the preceding pedestrian.
$\tilde{d}_y^{rel}$	falsified relative lateral distance of the preceding pedestrian.
d_y^{lt}	relative lateral distance to the left lane marking.
d_y^{rt}	relative lateral distance to the right lane marking.
d_x^{co}	cut-out distance for lane change of the following vehicle.
d_w	wheelbase of the Ego-vehicle tractor.
h	index of the hazardous driving situations.
$\mathcal{H}$	set of the hazardous driving situations.
i	index of the number of detected objects.
j	index of the number of tracked objects.
k	index of the sensor node.
m	number of the hazardous events along the cumulative driving distance for the statistical PoS.
N_s	number of the sensor nodes for environmental perception within the vehicle-mounted sensor setup module.
λ	failure rate of the hazardous events.
λ_A	failure rate of an ADF without driver supervision.
λ_H	benchmark failure rate of a human driver.
Λ	fatality factor of the ADF without driver supervision in relation to the benchmark failure rate of a human driver.
γ	reliability as a complementary of the failure rate λ.
ζ	number of the hazardous events observed by field operational tests.
$roll_{mode}$	required mode of rolling within an ACC function.
t_{co}	cut-out delay time for lane change of the following vehicle.
v_x^{ego}	longitudinal velocity of the Ego-vehicle.

υ_x^{obj1}	longitudinal velocity of the object with the identification number 1.
υ_x^{obj2}	longitudinal velocity of the preceding object with the identification number 2.
υ_y^{lt}	relative lateral velocity to the left lane marking.
υ_y^{rt}	relative lateral velocity to the right lane marking.
η	distance from crossing point at the origin to a horizontal extremity of a Lemniscate elliptic function.

Framework Conception

a_y^{ego}	lateral acceleration of the Ego-vehicle.
a_x^{rel}	longitudinal acceleration between the Ego-vehicle and the object ahead.
a_y^{rel}	lateral acceleration between the Ego-vehicle and the object ahead.
a_z^{rel}	vertical acceleration between the Ego-vehicle and the object ahead.
d_x^{rel}	longitudinal distance between the Ego-vehicle and the object ahead.
d_y^{rel}	lateral distance between the Ego-vehicle and the object ahead.
d_z^{rel}	vertical distance between the Ego-vehicle and the object ahead.
d_x^{c}	longitudinal distance between the Ego-vehicle and the stationary vehicle in front, which is detected by the camera sensor model.
d_x^{r}	longitudinal distance between the Ego-vehicle and the stationary object in front, which is detected by the RADAR sensor model.
f_d	sampling rate of the VxWorks® RTOS-based HiL simulator.
f_m	sampling rate of the 3D simulation environment.
F_r	weight force on the rear axle.
κ_{ego}	RADAR-based estimation of curvature for the Ego-vehicle trajectory.
M_b	braking torque of the Ego-vehicle.
P_b	foot placement on the brake pedal.
r	proximity distance index between two time series vectors.
t_s	sampling period of the vehicle dynamics model.

t_{ϱ}	turn-around time.
υ	number of subsets for features of sensor modelling.
v_x^{ego}	longitudinal velocity of the Ego-vehicle.
v_x^{rel}	longitudinal velocity between the Ego-vehicle and the object ahead.
v_y^{rel}	lateral velocity between the Ego-vehicle and the object ahead.
v_z^{rel}	vertical velocity between the Ego-vehicle and the object ahead.
x	longitudinal position of the Ego-vehicle in Cartesian co-ordinates.
y	lateral position of the Ego-vehicle in Cartesian co-ordinates.
z	vertical position of the Ego-vehicle in Cartesian co-ordinates.
ϕ	roll angle of the Ego-vehicle.
θ	pitch angle of the Ego-vehicle.
ψ	heading (yaw) angle of the Ego-vehicle.
ϕ_c^{max}	maximum roll angle with zero weight force acting on the rear axle.
θ_b	pitch angle while braking in the straight-ahead direction.

Framework Implementation

$d_x^{\mathcal{E}1}$	longitudinal distance between the Ego-vehicle and the object ahead at the time of triggering the first escalation event.
$d_y^{\mathcal{E}1}$	lateral distance between the Ego-vehicle and the object ahead at the time of triggering the first escalation event.
$d_{TTC}^{\mathcal{E}1}$	distance between the TTC of FOT and the theoretical ideal TTC curve from HiL testing.
d_y^{min}	lower bound of the truncated normal distribution of the lateral distance between the Ego-vehicle and the object ahead at the time of triggering the first escalation event.
d_y^{max}	upper bound of the truncation normal distribution of the lateral distance between the Ego-vehicle and the object ahead at the time of triggering the first escalation event.
$e_{\bar{P}_f}(\mathcal{S}_s^3 \vert C_s^1)$	standard error of $\mathcal{S}_s^3$ of the cluster C_s^1 at the escalation level $\mathcal{E}_s^1$.
$e_{\bar{P}_f}(\mathcal{S}_s^2 \vert C_s^1)$	standard error of $\mathcal{S}_s^2$ of the cluster C_s^1 at the escalation level $\mathcal{E}_s^1$.

$e_{\bar{P}_f}(\mathcal{S}_s^1 \mid C_s^1)$	standard error of $\mathcal{S}_s^1$ of the cluster C_s^1 at the escalation level $\mathcal{E}_s^1$.
$e_{\bar{P}_f}(\mathcal{S}_s^0 \mid C_s^1)$	standard error of $\mathcal{S}_s^0$ of the cluster C_s^1 at the escalation level $\mathcal{E}_s^1$.
κ_{road}	actual road curvature.
κ_i	road curvature at the position (i) of the clothoid.
$\kappa_{ego}^{\mathcal{E}1}$	curvature of the Ego-vehicle at the time of triggering the first escalation event.
κ_{ego}^{min}	lower bound of the truncated normal distribution for the road curvature at the time of triggering the first escalation event.
κ_{ego}^{max}	upper bound of the truncated normal distribution for the road curvature at the time of triggering the first escalation event.
L_i	length of the clothoid arc to the point (i).
$\bar{P}_f(\mathcal{S}_s^3 \mid C_s^1)$	failure probability of $\mathcal{S}_s^3$ of the cluster C_s^1 at the escalation level $\mathcal{E}_s^1$.
$\bar{P}_f(\mathcal{S}_s^2 \mid C_s^1)$	failure probability of $\mathcal{S}_s^2$ of the cluster C_s^1 at the escalation level $\mathcal{E}_s^1$.
$\bar{P}_f(\mathcal{S}_s^1 \mid C_s^1)$	failure probability of $\mathcal{S}_s^1$ of the cluster C_s^1 at the escalation level $\mathcal{E}_s^1$.
$\bar{P}_f(\mathcal{S}_s^0 \mid C_s^1)$	failure probability of $\mathcal{S}_s^0$ of the cluster C_s^1 at the escalation level $\mathcal{E}_s^1$.
$TTC_{ref}^{\mathcal{E}1}$	Time to Collision from the HiL testing.
$TTC_{fot}^{\mathcal{E}1}$	Time to Collision from the on-road testing.
$TTC_{min}^{\mathcal{E}1}$	minimum Time to Collision at the last moment when the object is classified as relevant.
$v_x^{\mathcal{E}1}$	longitudinal velocity of the Ego-vehicle at the time of triggering the first escalation event.
v_x^{min}	lower bound of the triangular distribution of the longitudinal velocity at the time of triggering the first escalation event.
v_x^{max}	upper bound of the triangular distribution of the longitudinal velocity at the time of triggering the first escalation event.
v_x^{peak}	mode of the triangular distribution of the longitudinal velocity at the time of triggering the first escalation event.
x_i	longitudinal position at the position (i) of the clothoid.
y_i	lateral position at the position (i) of the clothoid.

ω_i	direction angle between the tangent of the current position (i) of the clothoid and the initial direction of the straight line.

9.3 List of Quantities and States

Preliminaries

$\mathcal{A}$	set of missing specifications for the intended behavior within the three-circle Venn diagram.
$\mathcal{B}$	set of robust unspecified behaviors within the three-circle Venn diagram.
$\mathcal{C}$	intersection of the three Venn diagram sets of the specified, implemented and intended behaviors.
$\mathcal{D}$	Venn diagram set of the missing implementations for correct specifications.
$\mathcal{J}$	set of unexpected wrong behaviors within the three-circle Venn diagram.
$\mathcal{K}$	set of wrong specifications or technical limitations within the three-circle Venn diagram.
$\mathcal{L}$	set of missing implementations for wrong specifications within the three-circle Venn diagram.
$\mathcal{M}$	set of implemented behaviors within the three-circle Venn diagram.
$\mathcal{N}$	set of intended behaviors within the three-circle Venn diagram.
$\mathcal{S}$	set of specified behaviors within the three-circle Venn diagram.
$\mathcal{E}_{\mathrm{s}}^{1}$	first escalation event, where visual and acoustic alarms are issued to warn the driver about a possible collision with a stationary object.
$\mathcal{E}_{\mathrm{s}}^{2}$	second escalation event, where haptic and automatic particle braking is applied to assist the driver in braking in order to prevent a collision with a stationary object.
$\mathcal{E}_{\mathrm{s}}^{3}$	third escalation event, where emergency braking occurs at a point where the collision with a stationary object is imminent.
p1	first position of the stationary object within the Tractrix area of the Ego-vehicle defined in the Cartesian co-ordinates.
p2	second position of the stationary object within the Tractrix area of the Ego-vehicle defined in the Cartesian co-ordinates.
p3	third position of the stationary object within the Tractrix area of the Ego-vehicle defined in the Cartesian co-ordinates.

p4	fourth position of the stationary object within the Tractrix area of the Ego-vehicle defined in the Cartesian co-ordinates.
p5	fifth position of the stationary object within the Tractrix area of the Ego-vehicle defined in the Cartesian co-ordinates.
v	recorded sequence during on-road testing by reporting an event of measurement and detection errors.
$\tilde{\text{v}}$	newly produced sequence after updating of the ECU software with bug fixing of the faulty object detection during regression and progression tests.
v_{MiL}	model result of a back-to-back test within a MiL test setup.
v_{SiL}	generated code result of a back-to-back test within a SiL test setup.

Framework Conception

$\bar{\text{a}}$	mean of the time-series vector of the variable A.
$\bar{\text{b}}$	mean of the time-series vector of the variable B.
σ_{a}	standard deviation of the time-series vector of the variable A.
σ_{b}	standard deviation of the time-series vector of the variable B.
$\mathcal{C}_{\text{s}}^{1}$	group of driving situations with a combined curve to the left.
$\mathcal{C}_{\text{s}}^{2}$	group of driving situations with a S-shaped clothoid to the right.
$\mathcal{C}_{\text{s}}^{3}$	group of driving situations with a S-shaped clothoid to the left.
$\mathcal{C}_{\text{s}}^{4}$	group of driving situations with a combined curve to the right.
$\mathcal{C}_{\text{rt}}^{1}$	group of driving situations with a deviation to the detected right lane marking and a return afterwards from this deviation.
$\mathcal{C}_{\text{rt}}^{2}$	group of driving situations with a sudden jump in the distance to the lane, which indicates the detection of new lane lines.
$\mathcal{C}_{\text{rt}}^{3}$	group of driving situations with a deviation and a temporary sudden change of the distance to the right line and back to normal deviations due to painted islands.
$\mathcal{D}_{\text{NCC}}$	distance measure of the NCC algorithm.
$\mathcal{D}_{\text{NCC}}^{s_{\text{j}}}$	distance measure of the NCC algorithm with a time shift s_{j}.
f_{HAC}	HAC function.
$\mathcal{G}$	resulting cluster of the HAC algorithm.

$g^{(1)}$	merged cluster according to the complete-linkage criterion.
$g^{(2)}$	merged cluster according to the complete-linkage criterion.
$\hat{\mathrm{S}}^{\mathrm{c}}$	radar sensor model.
$\hat{\mathrm{S}}^{\mathrm{m}}$	3D cone model of a perception sensor.
$\hat{\mathrm{S}}^{\mathrm{r}}$	camera sensor model.
$\hat{\mathrm{S}}^{\mathrm{c}}_{\mathrm{obj1}}$	first relevant object detected by the camera sensor model.
$\hat{\mathrm{S}}^{\mathrm{c}}_{\mathrm{obj2}}$	second relevant object detected by the camera sensor model.
$\hat{\mathrm{S}}^{\mathrm{r}}_{\mathrm{obj1}}$	first relevant object detected by the radar sensor model.
$\hat{\mathrm{S}}^{\mathrm{r}}_{\mathrm{obj2}}$	second relevant object detected by the radar sensor model.
$\hat{\Theta}^{\mathrm{c}}$	timestamp message of the camera sensor model.
$\hat{\Theta}^{\mathrm{f}}$	timestamp message of the fusion module.
$\hat{\Theta}^{\mathrm{r}}$	timestamp message of the radar sensor model.
$\mathrm{u}(\mathrm{a_i})$	point of the normalized rolling standard deviation.
w	rolling window of the NCC algorithm.
$\mathrm{u_s}$	input of discrete state.
$\mathrm{x_{d+1}}$	input update of discrete state.
$\mathrm{x_d}$	input of discrete state.
$\mathrm{y_o}$	output of continuous state.

Framework Implementation

$\mathbb{E}$	expectation of a target indicator function.
$\mathrm{e}_{\bar{\mathrm{P}}_{\mathrm{f}}}$	standard error of the failure probability estimator.
$f_Z(z)$	joint density function for the random vector Z.
$I(\mathrm{g}(z))$	indicator function.
$h_Y(y)$	importance sampling density.
$\mathrm{g}(Z)$	limit state function.
$\mathrm{N_{mc}}$	number of samples according to the MCS method.
$\mathrm{N_{as}}$	number of samples according to the AIS method.
$\mathrm{P_f}$	probability of failure.
$\bar{\mathrm{P}}_{\mathrm{f}}$	estimator of the failure probability.

$P(C_s^1)$	prbability of occurrence for the group of driving situations with a combined curve to the left.
P_T	overall probability of occurrence for a functional scenario.
$\mathrm{Var}(\bar{P}_f)$	variance of the failure probability estimator.
R^2	coefficient of determination.
$\mathcal{R}\left(C_s^1\right)$	prospective risk assessment for the group of driving situations with a combined curve to the left.
$\mathcal{S}_s^0$	hypothesis for low severity.
$\mathcal{S}_s^1$	hypothesis for medium severity.
$\mathcal{S}_s^2$	hypothesis for high severity.
$\mathcal{S}_s^3$	hypothesis for critical severity.
γ	predefined safety margin.
$\delta_{\bar{P}_f}$	required standard error for the failure probability estimator.
μ	mean of the parent normal distribution.
$\mu_{\mathrm{TTC}_{\mathrm{ref}}^{\mathcal{E}1}}$	mean of the variable $\mathrm{TTC}_{\mathrm{ref}}^{\mathcal{E}1}$.
$\mu_{\mathrm{TTC}_{\mathrm{fot}}^{\mathcal{E}1}}$	mean of the variable $\mathrm{TTC}_{\mathrm{fot}}^{\mathcal{E}1}$.
σ	standard deviation of the parent normal distribution.
$\sigma_{\mathrm{TTC}_{\mathrm{ref}}^{\mathcal{E}1}}$	standard deviation of the variable $\mathrm{TTC}_{\mathrm{ref}}^{\mathcal{E}1}$.
$\sigma_{\mathrm{TTC}_{\mathrm{fot}}^{\mathcal{E}1}}$	standard deviation of the variable $\mathrm{TTC}_{\mathrm{fot}}^{\mathcal{E}1}$.
Ω	PDF of the normal Gaussian distribution.
Φ	CDF of the normal Gaussian distribution.
z	realization of Z in the variable space.
Z	random vector.

9.4 List of Matrices and Vectors

Preliminaries

O_h	controllability of hazardous events.
O_h^0	controllable hazardous event in general.
O_h^1	simply controllable event, where 99% or more of all drivers or other traffic participants are usually able to avoid hazardous event.
O_h^2	normally controllable event, where 90% or more of all drivers or other traffic participants are usually to avoid hazardous event.
O_h^3	difficult to control or uncontrollable event, where less than 90% of all drivers or other traffic participants are usually able, or barely able, to avoid the hazardous event.
$P^{(k)}$	set of sensor properties at detection level for each sensor node.
$Q^{(k)}$	set of sensor properties at object level for each sensor node.
$p_i^{(k)}$	subset of the sensor properties at detection level for each sensor node.
$q_j^{(k)}$	subset of the sensor properties at object level for each sensor node.
$\mathcal{R}$	scenario-based risk assessment of ADFs.
$S^{(k)}$	number of sensor nodes for environmental perception within the vehicle-mounted sensor setup module.
$\mathcal{S}_h$	severity of the hazardous events.
$\mathcal{S}_h^0$	no injuries.
$\mathcal{S}_h^1$	light and moderate injuries.
$\mathcal{S}_h^2$	severe and life-threatening injuries (survival probable).
$\mathcal{S}_h^3$	life-threatening injuries (survival uncertain or fatal injuries).
$\mathcal{X}_h$	exposure of the hazardous events.
$\mathcal{X}_h^0$	probability of unlikely occurrence of hazardous events.
$\mathcal{X}_h^1$	very low probability of hazardous events.
$\mathcal{X}_h^2$	low probability of hazardous events.
$\mathcal{X}_h^3$	medium probability of hazardous events.
$\mathcal{X}_h^4$	high probability of hazardous events.

Framework Conception

$\mathbf{A}$	time series vector of the variable A.
$\mathbf{B}$	time series vector of the variable B.
H	overall set of the implemented features of a sensor model.
$\mathrm{H}_{\upsilon}^{\mathrm{m}}$	subset of the implemented features of a sensor model.
$\beta_{\upsilon}^{\mathrm{m}}$	set of configuration parameters for each subset of the implemented features of a sensor model.
$\hat{\mathrm{P}}^{\mathrm{m}}$	set of detection-level sensor properties for a sensor model.
$\hat{\mathrm{Q}}^{\mathrm{m}}$	set of object-level sensor properties for a sensor model.
$\mathbf{R}_{\mathrm{E}}^{\mathrm{I}}$	rotation matrix from Ego-vehicle co-ordinate system to inertial co-ordinate system.
$\mathbf{R}_{\mathrm{S}}^{\mathrm{E}}$	rotation matrix from sensor co-ordinate system to Ego-vehicle co-ordinate system.
$\mathbf{R}_{\mathrm{I}}^{\mathrm{S}}$	rotation matrix from inertial co-ordinate system to sensor co-ordinate system.
$\mathbf{t}_{\mathrm{E}}^{\mathrm{I}}$	translation vector from Ego-vehicle to initial co-ordinate system.
$\mathbf{t}_{\mathrm{S}}^{\mathrm{E}}$	translation vector from sensor co-ordinate system to Ego-vehicle co-ordinate system.
$\mathbf{t}_{\mathrm{O}}^{\mathrm{I}}$	translation vector from object co-ordinate system to initial co-ordinate system.
$\mathbf{T}_{\mathrm{I}}^{\mathrm{S}}$	roto-translation matrix from initial co-ordinate system to sensor co-ordinate system.
$\mathbf{T}_{\mathrm{E}}^{\mathrm{I}}$	roto-translation matrix from Ego-vehicle co-ordinate system to initial co-ordinate system.
$\mathbf{T}_{\mathrm{S}}^{\mathrm{E}}$	roto-translation matrix from sensor co-ordinate system to Ego-vehicle co-ordinate system.
$\mathbf{T}_{\mathrm{O}}^{\mathrm{I}}$	roto-translation matrix from object co-ordinate system to initial co-ordinate system.
s_{j}	time shift vector.
$\mathbf{X}_{\mathrm{w}}$	longitudinal contact point vector from Ego-vehicle dynamics model to road model.
$\mathbf{Y}_{\mathrm{w}}$	lateral contact point vector from Ego-vehicle dynamics model to road model.

$\mathbf{Z}_{\mathrm{w}}$	vertical contact point vector from road model to Ego-vehicle dynamics model.
$\boldsymbol{\mu}_{\mathrm{w}}$	friction coefficient of road surface in contact with the Ego-vehicle wheels from road model to Ego-vehicle model.
$\boldsymbol{\alpha}_{\mathrm{w}}^{\mathrm{long}}$	longitudinal gradient of road surface in contact with the Ego-vehicle wheels from road model to Ego-vehicle model.
$\boldsymbol{\alpha}_{\mathrm{w}}^{\mathrm{lat}}$	lateral gradient vector of road surface in contact with the Ego-vehicle wheels from road model to Ego-vehicle model.

Framework Implementation

$\mathbf{F}_{\mathrm{C}}$	logical scenario for driving situations with a stationary object on the side.
$\mathbf{F}_{\mathrm{S}}$	logical scenario for driving situations with the Ego-vehicle driving around a stationary object.
$\mathbf{F}_{\mathrm{T}}$	functional scenario catalog.
$\mathcal{U}$	set of logical scenarios with (n) dimension.

9.5 List of Notations and Co-ordinate Systems

Preliminaries

$A^{(k)}$	measuring principle block that generates a detection list for each environmental perception sensor.
$B^{(k)}$	processing block that generates a feature list for each environmental perception sensor.
$C^{(k)}$	observation block that generates an object list for each environmental perception sensor.
C1	context element to define the safety requirements according to the various automated driving levels.
C4	context element to use the test platforms according to the effectiveness and efficiency criteria.
C5	context element to use a measurement system for field-based observations.
C6	context element to define the criticality matrices for safety envelopes.
E1	sub-evidence to define the coverage of functional requirements.
E2	sub-evidence to define the coverage of value and time tolerances within software structures.
E3	sub-evidence to define the coverage of system integration and variation.
E4	sub-evidence to define the coverage of system performance.
E5	sub-evidence to define the coverage of training data and sensor uncertainties.
E6	sub-evidence to define the coverage of driving scenarios.
G1	main goal to provide a safety argumentation based on an accepted residual risk of an ADF.
G2	sub-goal to track functional correctness for an ADF.
G3	sub-goal to ensure the back-to-back consistency for a model-based code generator employed to software development.
G4	sub-goal to specify the system integration and variation for deployment of an automated driving function.
G5	sub-goal to present the software robustness.

G6	sub-goal to deliver the sensor availability and functional effectiveness of an automated driving function.
G7	sub-goal to supply the software reliability and functional safety of an automated driving function.
S1	strategy of the GSN safety case based on the test termination criteria.

Framework Conception

$\mathcal{R}_{\mathcal{C}}$	camera reference co-ordinate system.
$\mathcal{R}_{\mathcal{D}}$	detected object point co-ordinate system.
$\mathcal{R}_{\mathcal{E}}$	subject vehicle reference co-ordinate system.
$\mathcal{R}_{\mathcal{I}}$	inertial (global) reference co-ordinate system.
$\mathcal{R}_{\mathcal{M}}$	2D monitor reference co-ordinate system.
$\mathcal{R}_{\mathcal{S}}$	sensor reference co-ordinate system.
$\mathcal{R}_{\mathcal{O}}$	target object reference co-ordinate system.
$\mathcal{R}_{\mathcal{R}}$	relative start reference co-ordinate system.
$\mathcal{R}_{\mathcal{T}}$	track co-ordinate system.

Framework Implementation

P1	baseline response of the logistic regression function.
P2	maximum response of the logistic regression function.
P3	turning point giving a halfway response between the baseline and maximum of the logistic regression function.
P4	curve slope of the logistic regression function.

List of Figures

List of Tables

Bibliography

[AADN+16] G. Andria, F. Attivissimo, A. Di Nisio, A.M.L. Lanzolla, and A. Pellegrino. Development of an automotive data acquisition platform for analysis of driving behavior. In *Measurement*. Elsevier Ltd., 2016. DOI: 10.1016/j.measurement.2016.07.035.

[AB15] M. Aeberhard and T. Bertram. Object Classification in a High-Level Sensor Data Fusion Architecture for Advanced Driver Assistance Systems. In *18th International Conference on Intelligent Transportation Systems*. IEEE, Gran Canaria, Spain, 2015. DOI: 10.1109/ITSC.2015.76.

[Abe08] S. Abendroth. Automatisiertes Testen im Nutzfahrzeugbereich. In *Automatisiertes Testen Eingebetteter Systeme in der Automobilindustrie*. Carl Hanser Verlag, 2008. DOI: 10.3139/9783446419018.008.

[AD11] M. Althoff and J.M. Dolan. Set-based computation of vehicle behaviors for the online verification of autonomous vehicles. In *14th International IEEE Conference on Intelligent Transportation Systems (ITSC)*, pages 1162–1167. IEEE, Washington, DC, USA, 2011. DOI: 10.1109/ITSC.2011.6083052.

[AD12] M. Althoff and J.M. Dolan. Reachability computation of low-order models for the safety verification of high-order road vehicle models. In *2012 American Control Conference (ACC)*, pages 3559–3566. IEEE, Montreal, QC, Canada, 2012. DOI: 10.1109/ACC.2012.6314777.

[AD14] M. Althoff and J.M. Dolan. Online Verification of Automated Road Vehicles Using Reachability Analysis. In *IEEE Transactions on Robotics*, volume 30, pages 903–918. IEEE, 2014. DOI: 10.1109/TRO.2014.2312453.

[ADPZ19] S. Abadpour, A. Diewald, M. Pauli, and T. Zwick. Extraction of Scattering Centers Using a 77 GHz FMCW Radar. In *2019 12th German Microwave Conference (GeMiC)*, pages 79–82. IEEE, Stuttgart, Germany, 2019. ISBN: 978-1-7281-0242-9.

[AFS93] R. Agrawal, C. Faloutsos, and A. Swami. Efficient similarity search in sequence databases. In *International Conference on Foundations of Data Organization and Algorithms*, volume 730, pages 69–84. Springer Verlag, Berlin, Heidelberg, 1993. DOI: 10.1007/3-540-57301-1_5.

[AG19] Daimler Truck AG. The new Actros - all new features in detail. In *Actros*. Article - (Retrieval date: 01.06.2022), Barcelona, Spain, 2019.

[AGZ18] M. Arief, P. Glynn, and D. Zhao. An Accelerated Approach to Safely and Efficiently Test Pre-Production Autonomous Vehicles on Public Streets. In *21st International Conference on Intelligent Transportation Systems*, pages 2006–2011. IEEE, Maui, HI, USA, 2018. DOI: 10.1109/ITSC.2018.8569371.

[AIGZ17] A. Armand, J. Ibanez-Guzman, and C. Zinoune. Digital Maps for Driving Assistance Systems and Autonomous Driving. In *Automated Driving*, pages 201–244. Springer International Publishing, Cham, Switzerland, 2017. DOI: 10.1007/978-3-319-31895-0_9.

[AJK+17] T. Abthoff, N. John, K. Kreplin, J. Lorenz, A. Pawlik, T. Piffl, J. Villaveces, and A. Villwock. Big data re-simulations for autonomous driving using DaSense. In *Fahrerassistenzsysteme 2017*, pages 359–372. Springer Vieweg, Wiesbaden, Germany, 2017. DOI: 10.1007/978-3-658-19059-0_21.

[AKK+19] Y. Akagi, R. Kato, S. Kitajima, J. Antona-Makoshi, and N. Uchida. A Risk-index based Sampling Method to Generate Scenarios for the Evaluation of Automated Driving Vehicle Safety. In *Intelligent Transportation Systems Conference (ITSC)*, pages 667–672. IEEE, Auckland, New Zealand, 2019. DOI: 10.1109/ITSC.2019.8917311.

[AKM17] M. Althoff, M. Koschi, and S. Manzinger. CommonRoad: Composable benchmarks for motion planning on roads. In *Intelligent Vehicles Symposium (IV)*, pages 719–726. IEEE, Los Angeles, CA, USA, 2017. DOI: 10.1109/IVS.2017.7995802.

[AL18] M. Althoff and S. Lutz. Automatic Generation of Safety-Critical Test Scenarios for Collision Avoidance of Road Vehicles. In *Intelligent Vehicles Symposium (IV)*, pages 1326–1333. IEEE, Changshu, China, 2018. DOI: 10.1109/IVS.2018.8500374.

[Alt10] M. Althoff. Reachability Analysis and its Application to the Safety Assessment of Autonomous Cars. In *Technische Universität München*. Dissertation - (Retrieval date: 01.06.2022), 2010. Lehrstuhl für Steuerungs- und Regelungstechnik.

[Ame20] C. Amersbach. Functional Decomposition Approach - Reducing the Safety Validation Effort for Highly Automated Driving. In *Technische Universität Darmstadt*. Dissertation - (Retrieval date: 01.06.2022), 2020. Fachgebiet Fahrzeugtechnik.

[ANR74] N. Ahmed, T. Natarajan, and K.R. Rao. Discrete Cosine Transform. In *IEEE Transactions on Computers*, volume C-23, pages 90–93. IEEE, 1974. DOI: 10.1109/T-C.1974.223784.

[Are19] N. Arechiga. Specifying Safety of Autonomous Vehicles in Signal Temporal Logic. In *Intelligent Vehicles Symposium (IV)*, pages 58–63. IEEE, Paris, France, 2019. DOI: 10.1109/IVS.2019.8813875.

[ASA17a] ASAM. Plug-on Device Interface. In *ASAM MCD-1 POD*. Datasheet - (Retrieval date: 01.06.2022), 2017.

[ASA17b] ASAM. Universal Measurement and Calibration Protocol. In *ASAM MCD-1 XCP*. Datasheet - (Retrieval date: 01.06.2022), 2017.

[ASA19] ASAM. Measurement Data Format. In *ASAM MDF*. Datasheet - (Retrieval date: 01.06.2022), 2019.

[ASA21] ASAM. Open Dynamic Road Information for Vehicle Environment. In *ASAM OpenDRIVE*. Datasheet - (Retrieval date: 01.06.2022), 2021.

[ASA22] ASAM. Open Dynamic Content of Driving and Traffic Simulators. In *ASAM OpenSCENARIO*. Datasheet - (Retrieval date: 01.06.2022), 2022.

[ASB07] M. Althoff, O. Stursberg, and M. Buss. Reachability analysis of linear systems with uncertain parameters and inputs. In *2007 46th IEEE Conference on Decision and Control (CDC)*, pages 726–732. IEEE, New Orleans, LA, USA, 2007. DOI: 10.1109/CDC.2007.4434084.

[Ass20] Zehl & Associates. Driver Fatigue Causes Nearly 4,000 Truck Crash Deaths Each Year. In *Premier Truck Accident Lawyers*. Article - (Retrieval date: 01.06.2022), 2020.

[Aur21] Aurora. Aurora voluntary safety self-assessment. In *The New Era of Mobility*. Report - (Retrieval date: 01.06.2022), 2021.

[AW17] C. Amersbach and H. Winner. Functional Decomposition: An Approach to Reduce the Approval Effort for Highly Automated Driving. In *8. Tagung Fahrerassistenz*. Article - (Retrieval date: 01.06.2022), 2017.

[BA92] R. Billinton and R.N. Allan. System reliability evaluation using probability distributions. In *Reliability Evaluation of Engineering Systems*. DOI: 10.1007/978-1-4899-0685-4_7, 1992.

[Bac18] J. Bach. Methoden und Ansätze für die Entwicklung und den Test prädiktiver Fahrzeugregelungsfunktionen. In *Karlsruher Institut für Technologie*. Dissertation - (Retrieval date: 01.06.2022), 2018. Institut für Technik der Informationsverarbeitung.

[Bär08] T. Bäro. Analyse, Bewertung und Verbesserung von Testprozessen. In *Automatisiertes Testen Eingebetteter Systeme in der Automobilindustrie*. Carl Hanser Verlag, 2008. DOI: 10.3139/9783446419018.003.

[Bay99] V. Bayer. Zur Zuverlässigkeitsbeurteilung von Baukonstruktionen unter dynamischen Einwirkungen. In *Bauhaus-Universität Weimar*. Dissertation - (Retrieval date: 01.06.2022), 1999. Fakultät Bauingenieurwesen.

[BBLF19] J. Bolte, A. Bar, D. Lipinski, and T. Fingscheidt. Towards Corner Case Detection for Autonomous Driving. In *Intelligent Vehicles Symposium*. IEEE, 2019. ISBN: 978-1-7281-0561-1.

[BDF+14] K. Bengler, K. Dietmayer, B. Farber, M. Maurer, C. Stiller, and H. Winner. Three decades of driver assistance systems: Review and future perspectives. In *IEEE Intelligent Transportation Systems Magazine*, volume 6, pages 6–22. IEEE, 2014. DOI: 10.1109/MITS.2014.2336271.

[Ber19] M. Berk. Safety Assessment of Environment Perception in Automated Driving Vehicles. In *Technische Universität München*. Dissertation - (Retrieval date: 01.06.2022), 2019. Fachgebiet für Risikoanalyse und Zuverlässigkeit.

[BGH17] S. Burton, L. Gauerhof, and C. Heinzemann. Making the Case for Safety of Machine Learning in Highly Automated Driving. In *International Conference on Computer Safety, Reliability, and Security*, pages 5–16. Springer International Publishing, Cham, Switzerland, 2017. DOI: 10.1007/978-3-319-66284-8_-1.

[BHOS17] J. Bach, M. Holzäpfel, S. Otten, and E. Sax. Reactive-Replay Approach for Verification and Validation of Closed-Loop Control Systems in Early Development. In *SAE International in United States*. SAE Mobilus, 2017. DOI: 10.4271/2017-01-1671.

[BHS17] O. Bartels, A. Hellmann, and P. Seiniger. Driving tests for the approval of automatically commanded steering functions. In *International Technical Conference on the Enhanced Safety of Vehicles*. Article - (Retrieval date: 01.06.2022), 2017.

[BKB+19] M. Büker, B. Kramer, E. Böde, S. Vander Maelen, and M. Fränzle. Identifikation von Automationsrisiken hochautomatisierter Fahrfunktionen in PEGASUS. In *AAET Automatisiertes und vernetztes Fahren*, pages 315–329. ITS mobility e.V., 2019. Article - (Retrieval date: 01.06.2022).

[BKM+19] J. Bock, R. Krajewski, T. Moers, S. Runde, L. Vater, and L. Eckstein. The inD Dataset: A Drone Dataset of Naturalistic Road User Trajectories at German Intersections. In *Computer Vision and Pattern Recognition*. arXiv - (Retrieval date: 01.06.2022), 2019.

[BLO+17] J. Bach, J. Langner, S. Otten, E. Sax, and M. Holzäpfel. Test scenario selection for system-level verification and validation of geolocation-dependent automotive control systems. In *International Conference on Engineering, Technology and Innovation*, pages 203–210. IEEE, Madeira, Portugal, 2017. DOI: 10.1109/ICE.2017.8279890.

[BLS22] BLS. Heavy and Tractor-trailer Truck Drivers. In *U.S. Bureau of Labor Statistics*. Article - (Retrieval date: 01.06.2022), 2022.

[BMBL06] V. Blervaque, K. Mezger, L. Beuk, and J. Loewenau. ADAS Horizon - How Digital Maps can contribute to Road Safety. In *Advanced Microsystems for Automotive Applications*, pages 427–436. Springer, Berlin, Heidelberg, 2006. DOI: 10.1007/3-540-33410-6_30.

[BMC15] M.F. Bugallo, L. Martino, and J. Corander. Adaptive importance sampling in signal processing. In *Digital Signal Processing*, volume 47, pages 36–49. Elsevier Inc., 2015. DOI: 10.1016/j.dsp.2015.05.014.

[BMK+16] R.L. Bücs, L. Murillo, E. Korotcenko, G. Dugge, R. Leupers, G. Ascheid, A. Ropers, M. Wedler, and A. Hoffmann. Virtual Hardware-in-the-Loop co-simulation for multi-domain automotive systems via the functional mock-up interface. In *Languages, Design Methods, and Tools for Electronic System Design*, pages 3–28. Springer International Publishing, Cham, Switzerland, 2016. DOI: 10.1007/978-3-319-31723-6_1.

[BMM18] G. Bagschik, T. Menzel, and M. Maurer. Ontology based Scene Creation for the Development of Automated Vehicles. In *Intelligent Vehicles Symposium (IV)*, pages 1813–1820. IEEE, Changshu, China, 2018. DOI: 10.1109/IVS.2018.8500632.

[BMV17] BMVI. Eight Act amending the Road Traffic Act. In *Road Traffic Act*. Report - (Retrieval date: 01.06.2022), 2017.

[BOS16] J. Bach, S. Otten, and E. Sax. Model based scenario specification for development and test of automated driving functions. In *Intelligent Vehicles Symposium (IV)*. IEEE, Gothenburg, Sweden, 2016. DOI: 10.1109/IVS.2016.7535534.

[BOS17] J. Bach, S. Otten, and E. Sax. A Taxonomy and Systematic Approach for Automotive System Architectures - From Functional Chains to Functional Networks. In *Proceedings of the 3rd International Conference on Vehicle Technology and Intelligent Transport Systems*. SciTePress, 2017. ISBN: 978-989-758-242-4.

[BOS19] J. Bach, S. Otten, and E. Sax. Classification of Automotive Electric/Electronic Features and the Consequent Hierarchization of the Logical System Architecture. In *Smart Cities, Green Technologies, and Intelligent Transport Systems*. Springer, 2019. DOI: 10.1007/978-3-030-02907-4_12.

[Bow01] J.P. Bowen. Z: A Formal Specification Notation. In *Formal Approaches to Computing and Information Technology*, pages 3–19. Springer, London, UK, 2001. DOI: 10.1007/978-1-4471-0701-9_1.

[Bra13] M. Brahmi. Reference Systems for Environmental Perception. In *Automotive Systems Engineering*. Springer Verlag, Berlin, Heidelberg, 2013. DOI: 10.1007/978-3-642-36455-6_9.

[BRW+17] H. Beglerovic, A. Ravi, N. Wikström, H.M. Koegeler, A. Leitner, and J. Holzinger. Model-based safety validation of the automated driving function highway pilot. In *International Munich Chassis Symposium*. Springer, 2017. DOI: 10.1007/978-3-658-18459-9_21.

[BSH17] H. Beglerovic, M. Stolz, and M. Horn. Testing of autonomous vehicles using surrogate models and stochastic optimization. In *International Conference on Intelligent Transportation Systems*. IEEE, Yokohama, Japan, 2017. DOI: 10.1109/ITSC.2017.8317768.

[BSMH18] H. Beglerovic, T. Schloemicher, S. Metzner, and M. Horn. Deep Learning Applied to Scenario Classification for Lane-Keep-Assist Systems. In *Applied Sciences*. MDPI, 2018. DOI: 10.3390/app8122590.

[Buc09] C. Bucher. Computation of Failure Probabilities. In *Computational Analysis of Randomness in Structural Mechanics*, volume 32, pages 56–67. CRC Press, London, UK, 2009. DOI: 10.1201/9780203876534.

[Bun22] Bundesrat. Verordnung zur Regelung des Betriebs von Kraftfahrzeugen mit automatisierter und autonomer Fahrfunktion und zur änderung strassenverkehrsrechtlicher Vorschriften vom 24. Juni 2022. In *Bundesgesetzblatt Teil I*. Standard - (Retrieval date: 01.07.2022), Bonn, Germany, 2022.

[CAR15] V. Cantillo, J. Arellana, and M. Rolong. Modelling pedestrian crossing behaviour in urban roads: A latent variable approach. In *Transportation Research Part F: Traffic Psychology and Behaviour*, volume 32, pages 56–67. Elsevier Ltd., 2015. DOI: 10.1016/j.trf.2015.04.008.

[CDDCm19] A. Corso, P. Du, K. Driggs-Campbell, and M.J. Kochenderfer. Adaptive Stress Testing with Reward Augmentation for Autonomous Vehicle Validation. In *Intelligent Transportation Systems Conference (ITSC)*, pages 163–168. IEEE, Auckland, New Zealand, 2019. DOI: 10.1109/ITSC.2019.8917242.

[CFHL07] X. Chen, J. Feng, M. Hiller, and V. Lauer. Application of Software Watchdog as a Dependability Software Service for Automotive Safety Relevant Systems. In *International Conference on Dependable Systems and Networks*. IEEE, 2007. DOI: 10.1109/DSN.2007.14.

[Cho16] J.B. Choi. Environmental Perception for Automated Vehicles: Localization, Mapping and Tracking. In *Technische Universität Braunschweig*. Dissertation - (Retrieval date: 01.06.2022), 2016. Institut für Regelungstechnik.

[CI14] K. Cooper and M. Ito. Formalizing a Structured Natural Language Requirements Specification Notation. In *INCOSE International Symposium*. Wiley Online Library, 2014. DOI: 10.1002/j.2334-5837.2002.tb02569.x.

[CK18] W. Chen and L. Kloul. An Ontology-based Approach to Generate the Advanced Driver Assistance Use Cases of Highway Traffic. In *10th International Joint Conference on Knowledge Discovery*. Report - (Retrieval date: 01.06.2022), 2018.

[COR+16] M. Cordts, M. Omran, S. Ramos, T. Rehfeld, M. Enzweiler, R. Benenson, U. Franke, S. Roth, and B. Schiele. The Cityscapes Dataset for Semantic Urban Scene Understanding. In *Computer Vision and Pattern Recognition (CVPR)*. IEEE, Las Vegas, NV, USA, 2016. DOI: 10.1109/CVPR.2016.350.

[CS15] B. Costello and R. Suarez. Truck driver shortage analysis 2015. In *American Trucking Associations*. Article - (Retrieval date: 01.06.2022), 2015.

[Cza18] K. Czarnecki. Operational Design Domain for Automated Driving Systems: Taxonomy of Basic Terms. In *Waterloo Intelligent Systems Engineering Lab*. WISE Lab, 2018. DOI: 10.13140/RG.2.2.18037.88803.

[DES16] DESTATIS. Unfallentwicklung auf deutschen straßen 2015. In *Statistisches Bundesamt 2016*. DESTATIS - (Retrieval date: 01.06.2022), Wiesbaden, Germany, 2016.

[DFS+10] M. Darms, F. Foelster, J. Schmidt, D. Froehlich, and A. Eckert. Data Fusion Strategies in Advanced Driver Assistance Systems. In *SAE International Journal of Passenger Cars - Electronic and Electrical Systems*. SAE Mobilus, 2010. DOI: 10.4271/2010-01-2337.

[DG18] W. Damm and R. Galbas. Exploiting learning and scenario-based specification languages for the verification and validation of highly automated driving. In *1st International Workshop on Software Engineering for AI in Autonomous Systems*. ACM, 2018. DOI: 10.1145/3194085.3194086.

[dGP17] E. de Gelder and JP. Paardekooper. Assessment of Automated Driving Systems using real-life scenarios. In *IEEE Intelligent Vehicles Symposium (IV)*, pages 589–594. IEEE, Los Angeles, CA, USA, 2017. DOI: 10.1109/IVS.2017.7995782.

[DH17] W. Damm and P. Heidl. SafeTRANS Working Group "Highly Automated Systems: Test, Safety, and Development Processes". In *Recommendations on Actions and Research Challenges*. Article - (Retrieval date: 01.06.2022), 2017.

[Die15] A.R. Diewald. Antenna coupling for computational radar simulation. In *Loughborough Antennas & Propagation Conference (LAPC)*, pages 1–5. IEEE, Loughborough, UK, 2015. DOI: 10.1109/LAPC.2015.7366079.

[DMV18] California DMV. Testing of Autonomous Vehicles. In *Order to Adopt. Title 13, Division 1, Chapter 1. Article 3.7*. Article - (Retrieval date: 01.06.2022), 2018.

[DS11] S. Durekovic and N. Smith. Architectures of Map-Supported ADAS. In *Intelligent Vehicles Symposium (IV)*, pages 207–211. IEEE, Baden-Baden, Germany, 2011. DOI: 10.1109/IVS.2011.5940402.

[DS15] B. Duraisamy and T. Schwarz. On Track-to-Track Data Association for Automotive Sensor Fusion. In *International Conference on Information Fusion*. IEEE, Washington, DC, USA, 2015. Electronic ISBN: 978-0-9824-4386-6.

[DTS+08] H. Ding, G. Trajcevski, P. Scheuermann, X. Wang, and E. Keogh. Querying and mining of time series data: Experimental comparison of representations and distance measures. In *Proceedings of the VLDB Endowment*. Association for Computing Machinery, 2008. DOI: 10.14778/1454159.1454226.

[Düs10] T. Düser. X-in-the-Loop-ein durchgängiges Validierungsframework für die Fahrzeugentwicklung am Beispiel von Antriebsstrangfunktionen und Fahrerassistenzsystemen. In *Karlsruher Institut für Technologie*. Dissertation - (Retrieval date: 01.06.2022), 2010. Institut für Produktentwicklung.

[Ebn14] A. Ebner. Referenzszenarien als Grundlage für die Entwicklung und Bewertung von Systemen der aktiven Sicherheit. In *Technische Universität Berlin*. Dissertation - (Retrieval date: 01.06.2022), 2014. Fakultät V - Verkehrs- und Maschinensysteme.

[EGK+20] N. Ebert, JC. Goos, F. Kirschbaum, E. Yildiz, and T. Koch. Methods of sensitivity analysis in model-based calibration. In *Automotive and Engine Technology*, volume 5, pages 45–56. Springer Nature, 2020. DOI: 10.1007/s41104-020-00058-x.

[EH16] M. El-Haji. Ontologie-basierte Definition von Anforderungen an Validierungswerkzeuge in der Fahrzeugtechnik. In *Karlsruher Institut für Technologie*. Dissertation - (Retrieval date: 01.06.2022), 2016. Institut für Fahrzeugsystemtechnik - Institutsteil Fahrzeugtechnik.

[Elg12] M. Elgharbawy. Further development and optimization of a real-time Hardware-in-the-Loop test bench for camera-based driver assistance systems. In *Master's Thesis*. Paderborn University, 2012. Heinz Nixdorf Institut.

[EPG+18] H. Elrofai, JP. Paardekooper, E. Gelder, S. Kalisvaart, and O. Op den Camp. StreetWise: Scenario-based safety validation of connected automated driving. In *TNO innovation for life*. Report - (Retrieval date: 01.06.2022), Helmond, Netherlands, 2018.

[EST+17] M. Edwards, M. Seidl, M. Tress, A. Pressley, and S. Mohan. Study on the Assessment and Certification of Automated Vehicles. In *European Commission Directorate General for Internal Market, Industry, Entrepreneurship and SMEs*. Report - (Retrieval date: 01.06.2022), 2017.

[Ete17] A. Etemad. AdaptIVe: Automated Driving Applications and Technologies for Intelligent Vehicles. In *Automated Driving: Safer and More Efficient Future Driving*, pages 535–540. Springer International Publishing, Cham, Switzerland, 2017. DOI: 10.1007/978-3-319-31895-0_23.

[EUA+19] A. Erdogan, B. Ugranli, E. Adali, A. Sentas, E. Mungan, E. Kaplan, and A. Leitner. Real-World Maneuver Extraction for Autonomous Vehicle Validation: A Comparative Study. In *Intelligent Vehicles Symposium (IV)*, pages 267–272. IEEE, Paris, France, 2019. DOI: 10.1109/IVS.2019.8814254.

[FBK17] L. Fraade-Blanar and N. Kalra. Autonomous Vehicles and Federal Safety Standards: An Exemption to the Rule? In *RAND Cooperation*. Article - (Retrieval date: 01.06.2022), 2017.

[FBS09] S. Fuchs, B. Butting, and E. Sax. Automated synthesis and configuration of Hardware-in-the-Loop test equipment. In *Applied Electronics*. IEEE, Pilsen, Czech Republic, 2009. Print ISBN: 978-807043781-0.

[FC11] FESTA-Consortium. FESTA Handbook Version 4. In *Field Operational Tests Networking and Methodology Promotion*. FOT-NET - (Retrieval date: 01.06.2022), 2011.

[FDW+18] T. Fleck, K. Daaboul, M. Weber, Schörner, et al. Towards Large Scale Urban Traffic Reference Data: Smart Infrastructure in the Test Area Autonomous Driving Baden-Württemberg. In *International Conference on Intelligent Autonomous Systems*. Springer, 2018. DOI: https://doi.org/10.1007/978-3-030-01370-7_75.

[Fei18] M. Feilhauer. Simulationsgestützte Absicherung von Fahrerassistenzsystemen. In *Universität Stuttgart*. Dissertation - (Retrieval date: 01.06.2022), 2018. Institut für Höchstleistungsrechnen.

[FHW16] M. Feilhauer, J. Haering, and S. Wyatt. Current Approaches in HiL-Based ADAS Testing. In *SAE International Journal of Commercial Vehicles*. SAE Mobilus, 2016. DOI: 10.4271/2016-01-8013.

[Fis17] T. Fischer. Eine Technologie für das durchgängige und automatisierte Testen eingebetteter Software. In *Karlsruher Institut für Technologie*. Dissertation - (Retrieval date: 01.06.2022), 2017. KIT-Fakultät für Elektrotechnik und Informationstechnik.

[FL+20] H. Flämig, S. Lunkeit, et al. ATLaS - Automatisiertes und vernetztes Fahren in der Logistik - Chancen für mehr Wertschöpfung. In *Projektbericht des Verbundvorhabens*. Report - (Retrieval date: 01.06.2022), 2020.

[Flä16] H. Flämig. Autonomous Vehicles and Autonomous Driving in Freight Transport. In *Autonomous Driving: Technical, Legal and Social Aspects*, pages 365–385. Springer, Berlin, Heidelberg, 2016. DOI: 10.1007/978-3-662-48847-8_18.

[FPSS96] U. Fayyad, G. Piatetsky-Shapiro, and P. Smyth. The KDD Process for Extracting Useful Knowledge from Volumes of Data. In *Communications of the ACM*, pages 27–34. Association for Computing Machinery, 1996. DOI: 10.1145/240455.240464.

[FRD18] D. Feng, L. Rosenbaum, and K. Dietmayer. Towards Safe Autonomous Driving: Capture Uncertainty in the Deep Neural Network for Lidar 3D Vehicle Detection. In *21st International Conference on Intelligent Transportation Systems*, pages 3266–3273. IEEE, Maui, HI, USA, 2018. DOI: 10.1109/ITSC.2018.8569814.

[fS+11a] International Organization for Standardization et al. Heavy commercial vehicles and buses - Steady-state rollover threshold - Tilt-table test method. In *ISO 16333:2011*. Standard - (Retrieval date: 01.06.2022), 2011.

[fS+11b] International Organization for Standardization et al. Road vehicles - Heavy commercial vehicles and buses - Steady-state circular tests. In *ISO 14792:2011*. Standard - (Retrieval date: 01.06.2022), 2011.

[fS+11c] International Organization for Standardization et al. Road vehicles - Vehicle dynamics and road-holding ability - Vocabulary. In *ISO 8855:2011*. Standard, 2011.

[fS+15] International Organization for Standardization et al. Systems and software engineering - System life cycle processes. In *ISO/IEC/IEEE 15288:2015*. Standard - (Retrieval date: 01.06.2022), 2015.

[fS+19a] International Organization for Standardization et al. Information technology - Process assessment - Process measurement framework for assessment of process capability. In *ISO/IEC 33020*. Standard - (Retrieval date: 01.06.2022), 2019.

[fS+19b] International Organization for Standardization et al. Road vehicles - Safety of the intended functionality. In *ISO/PAS 21448*. Standard - (Retrieval date: 01.06.2022), 2019.

[fS+20] International Organization for Standardization et al. Road vehicles - Safety and cybersecurity for automated driving systems - Design, verification and validation. In *ISO/TR 4804:2011*. Standard - (Retrieval date: 01.06.2022), 2020.

[fS+21] International Organization for Standardization et al. Road vehicles - Data communication between sensors and data fusion unit for automated driving functions - Logical interface. In *ISO/DIS 23150*. ISO 23150 - (Retrieval date: 01.06.2022), 2021.

[fS+22a] International Organization for Standardization et al. Road vehicles - Test scenarios for automated driving systems - Scenario based safety evaluation framework. In *ISO/DIS 34502*. Standard - (Retrieval date: 01.06.2022), 2022.

[fS+22b] International Organization for Standardization et al. Road vehicles - Test scenarios for automated driving systems - Vocabulary. In *ISO/DIS 34501*. Standard - (Retrieval date: 01.06.2022), 2022.

[fSIRTA22] Association for Safe International Road Travel (ASIRT). Annual Global Road Crash Statistics. In *Road Safety Facts*. Article - (Retrieval date: 01.06.2022), 2022.

[FSLL19] P. Feig, J. Schatz, V. Labenski, and T. Leonhardt. Assessment of Technical Requirements for Level 3 and Beyond Automated Driving Systems Based on Naturalistic Driving and Accident Data Analysis. In *Proceedings of ESV Conference*. Article - (Retrieval date: 01.06.2022), 2019.

[FWP+19] F. Fahrenkrog, L. Wang, T. Platzer, A. Fries, F. Raisch, and K. Kompaß. Prospective Safety Effectiveness Assessment of

Automated Driving Functions – from the Methods to the Results. In *Proceedings of ESV Conference*. Article - (Retrieval date: 01.06.2022), 2019.

[GA+18a] S.B.J. Gowdu, M.E. Asghar, et al. System architecture for installed-performance testing of automotive radars over-the-air. In *International Conference on Microwaves for Intelligent Mobility*, pages 1–4. IEEE, Munich, Germany, 2018. DOI: 10.1109/ICMIM.2018.8443490.

[GA18b] F. Gruber and M. Althoff. Anytime Safety Verification of Autonomous Vehicles. In *21st International Conference on Intelligent Transportation Systems (ITSC)*, pages 1708–1714. IEEE, Maui, HI, USA, 2018. DOI: 10.1109/ITSC.2018.8569950.

[GB12] B. Geller and T. Bradley. Quantifying Uncertainty in Vehicle Simulation Studies. In *SAE International Journal of Passenger Cars-Mechanical Systems*. SAE Mobilus, 2012. DOI: 10.4271/2012-01-0506.

[GBF+13] S. Geyer, M. Baltzer, B. Franz, S. Hakuli, M. Kauer, M. Kienle, S. Meier, et al. Concept and development of a unified ontology for generating test and use-case catalogues for assisted and automated vehicle guidance. In *Intelligent Transport Systems*. Transportation Research Board, 2013. ISSN: 1751-956X.

[GGSB19] J. Günther, O. Grau, I. Sharma, and B. Brücher. Advantages of Physically Based Rendering for Autonomous Driving Validation. In *3rd ACM Computer Science in Cars Symposium*. Article - (Retrieval date: 01.06.2022), Kaiserslautern, Germany, 2019.

[GKD+19] B. Gangopadhyay, S. Khastgir, S. Dey, P. Dasgupta, et al. Identification of Test Cases for Automated Driving Systems Using Bayesian Optimization. In *Intelligent Transportation Systems Conference*, pages 1961–1967. IEEE, Auckland, New Zealand, 2019. DOI: 10.1109/ITSC.2019.8917103.

[GKS18] J. Guo, U. Kurup, and M. Shah. Is It Safe to Drive? An Overview of Factors, Challenges, and Datasets for Driveability Assessment in Autonomous Driving. In *Artificial Intelligence*. arXiv - (Retrieval date: 01.06.2022), 2018.

[GLSU13] A. Geiger, P. Lenz, C. Stiller, and R. Urtasun. Vision meets Robotics: The KITTI dataset. In *International Journal of Robotics Research*. Sage Publications, London, UK, 2013. DOI: 10.1177/0278364913491297.

[GMB18] L. Gauerhof, P. Munk, and S. Burton. Structuring Validation Targets of a Machine Learning Function Applied to Automated Driving. In *International Conference on Computer Safety, Reliability, and Security*. Springer Nature, Cham, Switzerland, 2018. DOI: 10.1007/978-3-319-99130-6_4.

[Gol18] K. Golowko. Absicherung der ADAS-Funktionen schwerer Nutzfahrzeuge. In *ATZ-Automobiltechnische Zeitschrift*. Springer, 2018. DOI: 10.1007/s35148-018-0117-1.

[H+12] L. Hammarstrand et al. Adaptive Radar Sensor Model for Tracking Structured Extended Objects. In *IEEE Transactions on Aerospace and Electronic Systems*, volume 48, pages 1975–1995. IEEE, 2012. DOI: 10.1109/TAES.2012.6237574.

[H+15] T. Hanke et al. Generic architecture for simulation of ADAS sensors. In *16th International Radar Symposium (IRS)*, pages 125–130. IEEE, Dresden, Germany, 2015. DOI: 10.1109/IRS.2015.7226306.

[Hal20] S. Hallerbach. Simulation-based testing of cooperative and automated vehicles. In *Carl von Ossietzky Universität Oldenburg*. Dissertation - (Retrieval date: 01.06.2022), 2020. Fakultät II - Informatik, Wirtschafts- und Rechtswissenschaften.

[HALZ19] Z. Huang, M. Arief, H. Lam, and D. Zhao. Evaluation Uncertainty in Data-Driven Self-Driving Testing. In *Intelligent Transportation Systems Conference*. IEEE, Auckland, New Zealand, 2019. DOI: 10.1109/ITSC.2019.8917406.

[Har03] D.W. Harwood. Review of Truck Characteristics as Factors in Roadway Design. In *National Cooperative Highway Research Program*. Transportation Research Board, 2003. DOI: 10.17226/23379.

[Has14] B. Hassan. A Design Framework for Developing a Reconfigurable Driving Simulator. In *Universität Paderborn*. Dissertation - (Retrieval date: 01.06.2022), 2014. Fakultät für Maschinenbau.

[Hau21] F. Hauer. On Scenario-based Testing of Automated and Autonomous Driving Systems. In *Technische Universität München*. Dissertation - (Retrieval date: 01.06.2022), 2021. Fakultät für Informatik.

[Hel14] T. Helmer. Development of a methodology for the evaluation of active safety using the example of preventive pedestrian protection. In *Technische Universität Berlin*. Dissertation - (Retrieval date: 01.06.2022), 2014. Fakultät V Institut für Land- und Seeverkehr, Fachgebiet Kraftfahrzeuge.

[HH15] M. Haselhoff and S. Hakuli. ECUs für kamerabasierte Fahrerassistenzsysteme im Closed-Loop-Verfahren. In *ATZextra*, pages 30–33. Springer, 2015. DOI: 10.1007/s35778-015-0002-4.

[Hil12] M. Hillenbrand. Funktionale Sicherheit nach ISO 26262 in der Konzeptphase der Entwicklung von Elektrik/Elektronik Architekturen von Fahrzeugen. In *Karlsruher Institut für Technologie*. Dissertation - (Retrieval date: 01.06.2022), 2012. Institut für Technik der Informationsverarbeitung.

[HKM17] H. Hungar, F. Köster, and J. Mazzega. Test Specifications for Highly Automated Driving Functions: Highway Pilot. In *Vehicle Test & Development Symposium*. Article - (Retrieval date: 01.06.2022), 2017.

[HM+14] D. Hospach, S. Mueller, et al. Simulation and evaluation of sensor characteristics in vision based advanced driver assistance systems. In *17th International IEEE Conference on Intelligent Transportation Systems (ITSC)*, pages 2610–2615. IEEE, Qingdao, China, 2014. DOI: 10.1109/ITSC.2014.6958108.

[HPSVH03] I. Horrocks, P.F. Patel-Schneider, and F. Van Harmelen. From SHIQ and RDF to OWL: The making of a Web Ontology Language. In *Journal of Web Semantics*, volume 1, pages 7–26. Elsevier B.V., 2003. DOI: 10.1016/j.websem.2003.07.001.

[HPT10] F. Hendriks, R. Pelders, and M. Tideman. Future Testing of Active Safety Systems. In *SAE International Journal of Passenger Cars-Electronic and Electrical Systems*, pages 955–960. SAE Mobilus, 2010. DOI: 10.4271/2010-01-2334.

[HS10] M. Hörwick and KH. Siedersberger. Strategy and architecture of a safety concept for fully automatic and autonomous driving assistance systems. In *Intelligent Vehicles Symposium*, pages 955–960. IEEE, La Jolla, CA, USA, 2010. DOI: 10.1109/IVS.2010.5548115.

[HS15] Y. Horita and R.S. Schwartz. Extended electronic horizon for automated driving. In *14th International Conference on ITS Telecommunications (ITST)*, pages 32–36. IEEE, Copenhagen, Denmark, 2015. DOI: 10.1109/ITST.2015.7377396.

[HSK+19] J. Hiller, E. Svanberg, S. Koskinen, F. Bellotti, and N. Osman. The L3Pilot Common Data Format - Enabling efficient automated driving data analysis. In *Proceedings of ESV Conference*. Article - (Retrieval date: 01.06.2022), 2019.

[HSP19] J. Hathcock, M. Shahandashti, and P.B. Pudasaini. Autonomous Vehicles and Freight Transportation Analysis. In *Technical Report*. The North Central Texas Council of Governments, 2019. Department of Civil Engineering - The University of Texas at Arlington.

[HSSB14] J. Holzinger, P. Schöggl, M. Schrauf, and E. Bogner. Objective Assessment of Driveability while Automated Driving. In *ATZ worldwide*. Springer, 2014. DOI: 10.1007/s38311-014-0250-8.

[HTR+20] M.F. Holder, J.R. Thielmann, P. Rosenberger, C. Linnhoff, and H. Winner. How to evaluate synthetic radar data? Lessons learned from finding driveable space in virtual environments. In *FAS-Workshop*. Uni-DAS e.V., 2020. Article - (Retrieval date: 01.06.2022).

[Hul16] D. Hull. The Tesla Advantage: 1.3 Billion Miles of Data. In *Bloomberg*. Article - (Retrieval date: 01.06.2022), 2016.

[Hül18] D. Hülsebusch. Fahrerassistenzsysteme zur energieeffizienten Längsregelung - Analyse und Optimierung der Fahrsicherheit. In *Karlsruher Institut für Technologie*. Dissertation - (Retrieval date: 01.06.2022), 2018. Institut für Fahrzeugsystemtechnik.

[Hun18] H. Hungar. Scenario-Based Validation of Automated Driving Systems. In *International Symposium on Leveraging Applications of Formal Methods*, pages 449–460. Springer, 2018. ISBN: 978-3-030-03423-8.

[HVJ14] P. Hambarde, R. Varma, and S. Jha. The Survey of Real Time Operating System: RTOS. In *2014 International Conference on Electronic Systems, Signal Processing and Computing Technologies*, pages 34–39. IEEE Computer Society, Nagpur, India, 2014. DOI: 10.1109/ICESC.2014.15.

[IEC+90] International Electrotechnical Commission et al. Glossary of Software Engineering Terminology. In *Std 610.12-1990*, pages 1–84. DOI: 10.1109/IEEESTD.1990.101064, 1990.

[IEC+10] International Electrotechnical Commission et al. Functional safety of electrical/electronic/programmable electronic safety-related systems. In *IEC 61508:2010*. Standard - (Retrieval date: 01.06.2022), 2010.

[IEC+19] International Electrotechnical Commission et al. Precision Clock Synchronization Protocol for Networked Measurement and Control Systems. In *IEEE 1588*. DOI: 10.1109/IEEESTD.2020.9120376, 2019.

[Int17] SAE International. Event Data Recorder. In *SAE J1698*. SAE Mobilus, 2017. Standard - (Retrieval date: 01.06.2022).

[Int18] SAE International. Taxonomy and Definitions for Terms Related to Driving Automation Systems for On-Road Motor Vehicles. In *SAE J3016*. SAE Mobilus, 2018. Standard - (Retrieval date: 01.06.2022).

[Int20] SAE International. Automated Driving System Data Logger. In *SAE J3197*. SAE Mobilus, 2020. Standard - (Retrieval date: 01.06.2022).

[J+11] H. Jamshidi et al. Fusion of digital map traffic signs and camera-detected signs. In *5th International Conference on Signal Processing and Communication Systems*. IEEE, Honolulu, HI, USA, 2011. DOI: 10.1109/ICSPCS.2011.6140824.

[J+19] A.M. Jacobo et al. Development of a Safety Assurance Process for Autonomous Vehicles in Japan. In *Proceedings of ESV Conference*. Article - (Retrieval date: 01.06.2022), 2019.

[JBKW18] P. Junietz, F. Bonakdar, B. Klamann, and H. Winner. Criticality Metric for the Safety Validation of Automated Driving using Model Predictive Trajectory Optimization. In *21st International Conference on Intelligent Transportation Systems (ITSC)*. IEEE, Maui, HI, USA, 2018. DOI: 10.1109/ITSC.2018.8569326.

[JN04] H. Janssen and W. Niehsen. Vehicle surround sensing based on information fusion of monocular video and digital map. In *Intelligent Vehicles Symposium*, pages 244–249. IEEE, Parma, Italy, 2004. DOI: 10.1109/IVS.2004.1336389.

[JS14] K. Jo and M. Sunwoo. Generation of a Precise Roadway Map for Autonomous Cars. In *IEEE Transactions on Intelligent Transportation Systems*, volume 15, pages 925–937. IEEE, 2014. DOI: 10.1109/TITS.2013.2291395.

[JS+19a] S. Jesenski, J.E. Stellet, et al. Generation of Scenes in Intersections for the Validation of Highly Automated Driving Functions. In *Intelligent Vehicles Symposium*, pages 502–509. IEEE, Paris, France, 2019. DOI: 10.1109/IVS.2019.8813776.

[JS+19b] S. Jesenski, J.E. Stellet, et al. Simulation-Based Methods for Validation of Automated Driving: A Model-Based Analysis and an Overview about Methods for Implementation. In *Intelligent Transportation Systems*, pages 1914–1921. IEEE, Auckland, New Zealand, 2019. DOI: 10.1109/ITSC.2019.8917072.

[JSW19] P. Junietz, U. Steininger, and H. Winner. Macroscopic Safety Requirements for Highly Automated Driving. In *Transportation Research Record*. SAGE journals, 2019. DOI: 10.1177/0361198119827910.

[Jun19] P.M. Junietz. Microscopic and Macroscopic Risk Metrics for the Safety Validation of Automated Driving. In *Technische Universität Darmstadt*. Dissertation - (Retrieval date: 01.06.2022), 2019. DOI: 10.25534/tuprints-00009282.

[JWKW18] P. Junietz, W. Wachenfeld, K. Klonecki, and H. Winner. Evaluation of Different Approaches to Address Safety Validation of Automated Driving. In *21st International Conference on Intelligent Transportation Systems (ITSC)*, pages 491–496. IEEE, Maui, HI, USA, 2018. DOI: 10.1109/ITSC.2018.8569959.

[K+18a] P. Koopman et al. Cerification of Highly Automated Vehicles for Use on UK Roads. In *Creating An Industry-Wide Framework for Safety*. Report - (Retrieval date: 01.06.2022), 2018.

[K+18b] M. Koren et al. Adaptive Stress Testing for Autonomous Vehicles. In *Intelligent Vehicles Symposium (IV)*. IEEE, Changshu, China, 2018. DOI: 10.1109/IVS.2018.8500400.

[K+18c] R. Krajewski et al. Data-Driven Maneuver Modeling using Generative Adversarial Networks and Variational Autoencoders for Safety Validation of Highly Automated Vehicles. In *21st International Conference on Intelligent Transportation Systems (ITSC)*. IEEE, Maui, HI, USA, 2018. DOI: 10.1109/ITSC.2018.8569971.

[K+19a] S. Kitajima et al. Multi-agent traffic simulations to estimate the impact of automated technologies on safety. In *Traffic injury prevention*. Taylor & Francis, 2019. DOI: 10.1080/15389588.2019.1625335.

[K+19b] F. Kruber et al. Unsupervised and Supervised Learning with the Random Forest Algorithm for Traffic Scenario Clustering and Classification. In *Intelligent Vehicles Symposium*. IEEE, 2019. DOI: 10.1109/IVS.2019.8813994.

[Kar20] A. Karpathy. System and method for obtaining training data. In *United States Patent and Trademark Office*. TESLA INC, 2020. AN: 2019051121.

[KBKE18] R. Krajewski, J. Bock, L. Kloeker, and L. Eckstein. The highD dataset: A drone dataset of naturalistic vehicle trajectories on german highways for validation of highly automated driving systems. In *21st International Conference on Intelligent Transportation Systems (ITSC)*, pages 2118–2125. IEEE, Maui, HI, USA, 2018. DOI: 10.1109/ITSC.2018.8569552.

[KFFW19] P. Koopman, U. Ferrell, F. Fratrik, and M. Wagner. A Safety Standard Approach for Fully Autonomous Vehicles. In *International Conference on Computer Safety, Reliability, and Security*. Springer Nature, Cham, Switzerland, 2019. DOI: 10.1007/978-3-030-26250-1_26.

[Kir15] M. Kirschbaum. Highly automated driving for commercial vehicles. In *6th International Munich Chassis Symposium*. Springer Vieweg, Wiesbaden, Germany, 2015. DOI: 10.1007/978-3-658-09711-0_2.

[KLN+18] F. Klueck, Y. Li, M. Nica, J. Tao, and F. Wotawa. Using ontologies for test suites generation for automated and autonomous driving functions. In *International Symposium on Software Reliability Engineering Workshops (ISSREW)*, pages 118–123. IEEE, Memphis, TN, USA, 2018. DOI: 10.1109/ISSREW.2018.00-20.

[KM05] A.J. Ko and B.A. Myers. A framework and methodology for studying the causes of software errors in programming systems. In *Journal of Visual Languages & Computing*, volume 16, pages 41–84. Elsevier Ltd., 2005. DOI: 10.1016/j.jvlc.2004.08.003.

[KML22] A. Kriebitz, R. Max, and C. Lütge. The German Act on Autonomous Driving - Why Ethics Still Matters. In *Philosophy & Technology*. Springer, 2022. DOI: 10.1007/s13347-022-00526-2.

[KNB+09] A. Knapp, M. Neumann, M. Brockmann, R. Walz, and T. Winkle. Code of Practice for the Design and Evaluation of ADAS. In *Preventive and Active Safety Applications, eSafety for road and air transport*. Report - (Retrieval date: 01.06.2022), 2009.

[Kod20] Kodiak. Safety, first and always. In *Kodiak Safety Report 2020*. Report - (Retrieval date: 01.06.2022), 2020.

[Kol10] M. Kolditz. Entwicklung einer Spur-Objekt-Fusion auf Basis von Radar- und Mono-Kamera-Daten fuŕ Fahrerassistenzsysteme im Nutzfahrzeug. In *Karlsruher Institut für Technologie*. Diplomarbeit, 2010. Institut für Produktentwicklung.

[Kop22] A. Kopestinsky. 24 Disturbing Truck Accident Statistics. In *Policy Advice*. Article - (Retrieval date: 01.06.2022), 2022.

[KP+14] M. Kutila, P. Pyykonen, et al. The DESERVE project: Towards future ADAS functions. In *International Conference on Embedded Computer Systems: Architectures, Modeling, and Simulation*. IEEE, 2014. DOI: 10.1109/SAMOS.2014.6893226.

[KP16] N. Kalra and S.M. Paddock. Driving to safety: How many miles of driving would it take to demonstrate autonomous vehicle reliability? In *Transportation Research Part A: Policy and Practice*, volume 94, pages 182–193. Elsevier Ltd., 2016. DOI: 10.1016/j.tra.2016.09.010.

[Kri19] S. Krishna. Data driven extraction of challenging situation for autonomous vehicles. In *Nanyang Technological University*. Dissertation - (Retrieval date: 01.06.2022), 2019. Computer science.

[KRK+19] C. King, L. Ries, C. Kober, C. Wohlfahrt, and E. Sax. Automated Function Assessment in Driving Scenarios. In *12th IEEE Conference on Software Testing, Validation and Verification (ICST)*. IEEE, Xi'an, China, 2019. DOI: 10.1109/ICST.2019.00050.

[KS13] N. Kurtz and J. Song. Cross-entropy-based adaptive importance sampling using Gaussian mixture. In *Structural Safety*, volume 42, pages 35–44. Elsevier Ltd., 2013. DOI: 10.1016/j.strusafe.2013.01.006.

[Küh13] W. Kühn. Road Categorization. In *Fundamentals of Road Design*. WIT PRESS, 2013. ISBN: 978-1-84564-097-2.

[Kv07] J.T.B.A. Kessels and P.P.J. van den Bosch. Electronic Horizon: Energy Management using Telematics Information. In *IEEE Vehicle Power and Propulsion Conference*. IEEE, Arlington, TX, USA, 2007. DOI: 10.1109/VPPC.2007.4544190.

[KW16] P. Koopman and M. Wagner. Challenges in Autonomous Vehicle Testing and Validation. In *SAE International Journal of Transportation Safety*. SAE Mobilus, 2016. DOI: 10.4271/2016-01-0128.

[KW17] P. Koopman and M. Wagner. Autonomous Vehicle Safety: An Interdisciplinary Challenge. In *Intelligent Transportation Systems Magazine*, pages 90–96. IEEE, 2017. DOI: 10.1109/MITS.2016.2583491.

[KW18] P. Koopman and M. Wagner. Toward a Framework for Highly Automated Vehicle Safety Validation. In *SAE International*. SAE Mobilus, 2018. DOI: 10.4271/2018-01-1071.

[KWB18] F. Kruber, J. Wurst, and M. Botsch. An Unsupervised Random Forest Clustering Technique for Automatic Traffic Scenario Categorization. In *21st International Conference on Intelligent Transportation Systems*, pages 2811–2818. IEEE, Maui, HI, USA, 2018. DOI: 10.1109/ITSC.2018.8569682.

[KYB18] Y. Kang, H. Yin, and C. Berger. Test Your Self-Driving Algorithm: An Overview of Publicly Available Driving Datasets and Virtual Testing Environments. In *IEEE Transactions on Intelligent Vehicles*, pages 171–185. IEEE, 2018. DOI: 10.1109/TIV.2018.2886678.

[L+15] R. Lee et al. Adaptive stress testing of airborne collision avoidance systems. In *34th Digital Avionics Systems Conference*. IEEE, Prague, Czech Republic, 2015. DOI: 10.1109/DASC.2015.7311450.

[L+18a] J. Langner et al. Estimating the Uniqueness of Test Scenarios derived from Recorded Real-World-Driving-Data using Autoencoders. In *Intelligent Vehicles Symposium*, pages 1860–1866. IEEE, Changshu, China, 2018. DOI: 10.1109/IVS.2018.8500464.

[L+18b] R. Lee et al. Differential Adaptive Stress Testing of Airborne Collision Avoidance Systems. In *Modeling and Simulation Technologies Conference*. Aerospace Research Central, 2018. DOI: 10.2514/6.2018-1923.

[L+19] J. Langner et al. Logical Scenario Derivation by Clustering Dynamic-Length-Segments Extracted from Real-World-Driving-Data. In *Proceedings of the 5th International Conference on Vehicle Technology and Intelligent Transport Systems*. SciTePress, 2019. DOI: 10.5220/0007723304580467.

[LAH+19] A. Leitner, A. Akkermann, B.A. Hjøllo, B. Wirtz, D. Nickovic, et al. ENABLE-S3: Testing & Validation of Highly Automated Systems. In *Summary of Results*. Report - (Retrieval date: 01.06.2022), 2019.

[Lat08] F. Lattemann. Test-Operations. In *Automatisiertes Testen Eingebetteter Systeme in der Automobilindustrie*. Carl Hanser Verlag, 2008. DOI: 10.3139/9783446419018.007.

[LBS19] A. Lauber, N. Brenner, and E. Sax. Automated vehicle depots as an initial step for an automated public transportation. In *UITP Global Public Transport Summit*. Article - (Retrieval date: 01.06.2022), Stockholm, Schweden, 2019. Institut für Technik der Informationsverarbeitung.

[Lei20] A. Leitner. ENABLE-S3: Project Introduction. In *Validation and Verification of Automated Systems*. Springer Nature, Cham, Switzerland, 2020. DOI: 10.1007/978-3-030-14628-3_4.

[Li07] M. Li. Robust optimization and sensitivity analysis with multi-objective genetic algorithms: Single-and multi-disciplinary applications. In *University of Maryland*. Dissertation - (Retrieval date: 01.06.2022), 2007. Department of Mechanical Engineering.

[LPN11] S.M. Loos, A. Platzer, and L. Nistor. Adaptive Cruise Control: Hybrid, Distributed, and Now Formally Verified. In *International Symposium on Formal Methods*. Springer Verlag, Berlin, Heidelberg, 2011. DOI: 10.1007/978-3-642-21437-0_6.

[LPP11] P. Löw, R. Pabst, and E. Petry. Konzept, Anwendungsbereich, Risikoanalyse. In *Funktionale Sicherheit in der Praxis: Anwendung von DIN EN 61508 und ISO/DIS 26262 bei der Entwicklung von Serienprodukten*. Dpunkt, 2011. ISBN: 978-3-8986-4898-1.

[LS19] S.H. Leilabadi and S. Schmidt. In-depth Analysis of Autonomous Vehicle Collisions in California. In *Intelligent Transportation Systems Conference*. IEEE, Auckland, New Zealand, 2019. DOI: 10.1109/ITSC.2019.8916775.

[LSE12] J. Lee, I. Shin, and A. Easwaran. Convex optimization framework for intermediate deadline assignment in soft and hard real-time distributed systems. In *Journal of Systems and Software*. Elsevier Inc., 2012. DOI: 10.1016/j.jss.2012.04.050.

[LSW18] A. Lauber, E. Sax, and M. Wiedemann. Autonomes Fahren auf dem Busbetriebshof. In *ATZ-Automobiltechnische Zeitschrift*. Springer, 2018. DOI: 10.1007/s35148-018-0047-y.

[LTW20] Y. Li, J. Tao, and F. Wotawa. Ontology-based test generation for automated and autonomous driving functions. In *Information and Software Technology*. Elsevier B.V., 2020. DOI: 10.1016/j.infsof.2019.106200.

[M+15] M.A. Mohammad et al. Ontology-based Framework for Risk Assessment in Road Scenes Using Videos. In *Procedia Computer Science*. Elsevier B.V., 2015. DOI: 10.1016/j.procs.2015.08.300.

[M+16] J. Mazzega et al. Testing of highly automated driving functions. In *ATZ worldwide*. Springer, 2016. DOI: 10.1007/s38311-016-0101-x.

[M+17] SM.S. Mahmud et al. Application of proximal surrogate indicators for safety evaluation: A review of recent developments and research needs. In *International Association of Traffic and Safety Sciences Research*. Elsevier Ltd., 2017. DOI: 10.1016/j.iatssr.2017.02.001.

[M+18] T. Menzel et al. Scenarios for development, test and validation of automated vehicles. In *Intelligent Vehicles Symposium (IV)*. IEEE, Changshu, China, 2018. DOI: 10.1109/IVS.2018.8500406.

[M+20a] V. Matthias et al. Modelling road transport emissions in Germany - current situation and scenarios for 2040. In *Transportation Research Part D: Transport and Environment*. Elsevier Ltd., 2020. DOI: 10.1016/j.trd.2020.102536.

[M+20b] A.S. Müller et al. What human-like errors do autonomous vehicles need to avoid to maximize safety? In *booktitle of Safety Research*. Elsevier Ltd., 2020. DOI: 10.1016/j.jsr.2020.10.005.

[Mar04] A. Marquardt. Modellierung der Mehrkörperdynamik von Lastkraftwagen zur Echtzeitsimulation. In *Universität Stuttgart*. Diplomarbeit, 2004. Institut B für Mechanik.

[Mas19] J. Masino. Road Condition Estimation with Data Mining Methods using Vehicle Based Sensors. In *Karlsruher Institut für Technologie*. Dissertation - (Retrieval date: 01.06.2022), 2019. Institut für Fahrzeugsystemtechnik - Institutsteil Fahrzeugtechnik.

[McN18] S. McNally. Turnover Rate at Large Truckload Carriers Rises in First Quarter. In *American Trucking Associations*. Article - (Retrieval date: 01.06.2022), 2018.

[McN19] S. McNally. New Survey Data Reveals Increases in Driver Compensation. In *American Trucking Associations*. Article - (Retrieval date: 01.06.2022), 2019.

[Min17] P.M. Minnerup. An Efficient Method for Testing Autonomous Driving Software against Nondeterministic Influences. In *Technische Universität München*. Dissertation - (Retrieval date: 01.06.2022), 2017. Fakultät für Informatik.

[MO17] M. Michałowska and M. Ogłoziński. Autonomous Vehicles and Road Safety. In *International Conference on Transport Systems Telematics*. Springer International Publishing, Cham, Switzerland, 2017. DOI: 10.1007/978-3-319-66251-0_16.

[MP17] G. Mastinu and M. Plöchl. Vehicle Models and Equations of Motion. In *Road and Off-Road Vehicle System Dynamics Handbook*. CRC Press, 2017. DOI: 10.1201/b15560.

[MPS+15] A. Mehmed, S. Punnekkat, W. Steiner, G. Spampinato, and M. Lettner. Improving Dependability of Vision-Based Advanced Driver Assistance Systems Using Navigation Data and Checkpoint Recognition. In *Computer Safety, Reliability, and Security*. Springer, 2015. DOI: 10.1007/978-3-319-24255-2_6.

[MSH+19] I. Majzik, O. Semeráth, C. Hajdu, K. Marussy, Z. Szatmári, et al. Towards System-Level Testing with Coverage Guarantees for Autonomous Vehicles. In *ACM/IEEE 22nd International Conference on Model Driven Engineering Languages and Systems (MODELS)*. IEEE, 2019. DOI: 10.1109/MODELS.2019.00-12.

[Mul85] M. Mulazzani. Reliability versus Safety. In *IFAC Proceedings Volumes*. Elsevier Ltd., 1985. DOI: 10.1016/S1474-6670(17)60097-1.

[Mul18] G.E. Mullins. Adaptive Sampling Methods for Testing Autonomous Systems. In *University of Maryland*. Dissertation - (Retrieval date: 01.06.2022), 2018. Computer Science.

[MV19] S. Müller and F. Voigtländer. Automated Trucks in Road Freight Logistics: The User Perspective. In *Interdisciplinary Conference on Production, Logistics and Traffic*. Springer Nature, Cham, Switzerland, 2019. DOI: 10.1007/978-3-030-13535-5_-8.

[MW08] T. Most and J. Will. Metamodel of Optimal Prognosis - an automatic approach for variable reduction and optimal metamodel selection. In *Proc. Weimarer Optimierungs- und Stochastiktage*. Article - (Retrieval date: 01.06.2022), 2008.

[N+09] D. Nienhuser et al. A Situation context aware Dempster-Shafer fusion of digital maps and a road sign recognition system. In *Intelligent Vehicles Symposium*. IEEE, Xi'an, China, 2009. DOI: 10.1109/IVS.2009.5164490.

[N+15] J. Nilsson et al. Worst Case Analysis of Automotive Collision Avoidance Systems. In *IEEE Transactions on Vehicular Technology*. IEEE, 2015. DOI: 10.1109/TVT.2015.2419196.

[N+18] D. Neven et al. Towards End-to-End Lane Detection: An Instance Segmentation Approach. In *Intelligent Vehicles Symposium*. IEEE, Changshu, China, 2018. DOI: 10.1109/IVS.2018.8500547.

[N+20] C. Neurohr et al. Fundamental Considerations around Scenario-Based Testing for Automated Driving. In *Software Engineering*. arXiv - (Retrieval date: 01.06.2022), 2020.

[Nen14] M. Nentwig. Untersuchungen zur Anwendung von computergenerierten Kamerabildern für die Entwicklung und den Test von Fahrerassistenzsystemen. In *Friedrich-Alexander-Universität Erlangen-Nürnberg*. Dissertation - (Retrieval date: 01.06.2022), 2014. Technische Fakultät.

[NGB13] F. Netter, F. Gauterin, and B. Butterer. Real-data validation of simulation models in a function-based modular framework. In *Software Testing, Verification and Validation*. IEEE, Luxembourg, 2013. DOI: 10.1109/ICST.2013.36.

[NHT17a] NHTSA. A Vision for Safety. In *Automated Driving Systems*. Article - (Retrieval date: 01.06.2022), 2017.

[NHT17b] NHTSA. Voluntary Safety Self-Assessment. In *U.S. Department of Transportation*. Article - (Retrieval date: 01.06.2022), 2017.

[NMS12] M. Nentwig, M. Miegler, and M. Stamminger. Concerning the applicability of computer graphics for the evaluation of image processing algorithms. In *International Conference on Vehicular Electronics and Safety (ICVES)*. IEEE, Istanbul, Turkey, 2012. DOI: 10.1109/ICVES.2012.6294288.

[NS11] M. Nentwig and M. Stamminger. Hardware-in-the-Loop testing of computer vision based driver assistance systems. In *Intelligent Vehicles Symposium*. IEEE, 2011. DOI: 10.1109/IVS.2011.5940567.

[NSTH19] M. Nabhan, M. Schoenauer, Y. Tourbier, and H. Hage. Optimizing coverage of simulated driving scenarios for the autonomous vehicle. In *International Conference on Connected Vehicles and Expo (ICCVE)*. IEEE, Graz, Austria, 2019. DOI: 10.1109/ICCVE45908.2019.8965211.

[O+18] M. O'Kelly et al. Scalable End-to-End Autonomous Vehicle Testing via Rare-event Simulation. In *Advances in Neural Information Processing Systems*. arXiv - (Retrieval date: 01.06.2022), 2018.

[OBW+18] S. Otten, J. Bach, C. Wohlfahrt, C. King, J. Lier, H. Schmid, S. Schmerler, and E. Sax. Automated Assessment and Evaluation of Digital Test Drives. In *Advanced Microsystems for Automotive Applications*. Springer International Publishing, Cham, Switzerland, 2018. DOI: 10.1007/978-3-319-66972-4_16.

[ODR+02] W.L. Oberkampf, S.M. DeLand, B.M. Rutherford, K.V. Diegert, and K.F. Alvin. Error and uncertainty in modeling and simulation. In *Reliability Engineering & System Safety*. Elsevier Ltd., 2002. DOI: 10.1016/S0951-8320(01)00120-X.

[OPLS11] C. Otto, F. Puente León, and A. Schwarzhaupt. A strategy on situation evaluation for driver assistance systems in commercial vehicles considering pedestrians in urban traffic. In *Intelligent Vehicles Symposium*. IEEE, Baden-Baden, Germany, 2011. DOI: 10.1109/IVS.2011.5940421.

[Ott13] C. Otto. Fusion of data from heterogeneous sensors with distributed fields of view and situation evaluation for advanced driver assistance systems. In *Karlsruher Institut für Technologie*. Dissertation - (Retrieval date: 01.06.2022), 2013. Institut für Industrielle Informationstechnik.

[P+01] A. Pasquini et al. Analysis of Incidents Involving Interactive Systems. In *Computer Safety, Reliability and Security*. Springer Verlag, Berlin, Heidelberg, 2001. DOI: 10.1007/3-540-45416-0_11.

[P+14] I. Papaioannou et al. Sequential importance sampling for structural reliability. In *Conference on Reliability and Optimization of Structural Systems*. Elsevier Ltd., 2014. DOI: 10.1016/j.strusafe.2016.06.002.

[P+17a] A. Pütz et al. Database approach for the sign-off process of highly automated vehicles. In *Proceedings of ESV Conference*. Article - (Retrieval date: 01.06.2022), 2017.

[P+17b] A. Pütz et al. System validation of highly automated vehicles with a database of relevant traffic scenarios. In *12th ITS European Congress*. Article - (Retrieval date: 01.06.2022), 2017.

[P+19a] D. Paddeu et al. New Technology and Automation in Freight Transport and Handling Systems. In *University of the West of England*. Report - (Retrieval date: 01.06.2022), 2019.

[P+19b] R. Pfeffer et al. Automated Driving-Challenges for the Automotive Industry in Product Development with Focus on Process Models and Organizational Structure. In *International Systems Conference*. IEEE, Orlando, FL, USA, 2019. DOI: 10.1109/SYSCON.2019.8836779.

[PEG19] PEGASUS. PEGASUS METHOD - An Overview. In *PEGASUS Project Office*. Report - (Retrieval date: 01.06.2022), 2019.

[Pfe20] R. Pfeffer. Szenariobasierte simulationsgestützte funktionale Absicherung hochautomatisierter Fahrfunktionen durch Nutzung von Realdaten. In *Karlsruher Institut für Technologie*. Dissertation - (Retrieval date: 01.06.2022), 2020. DOI: 10.5445/IR/1000125413.

[PH16] R. Pfeffer and M. Haselhoff. Video Injection Methods in a Real-world Vehicle for Increasing Test Efficiency. In *Auto Tech Review*. Springer, 2016. DOI: 10.1365/s40112-016-1181-0.

[Pon21] T. Ponn. How to Define System-specific Corner Cases for the Type Approval of Automated Vehicles. In *Technische Universität München*. Dissertation - (Retrieval date: 01.06.2022), 2021. Lehrstuhl für Fahrzeugtechnik.

[PSS19] R. Pfeffer, E. Sax, and S. Schmidt. Real Data Based Validation of Highly Automated Driving Functions Using Simulation Methods. In *ATZelectronics worldwide*. Springer, 2019. DOI: 10.1007/s38314-019-0116-3.

[PZA17] C. Pek, P. Zahn, and M. Althoff. Verifying the safety of lane change maneuvers of self-driving vehicles based on formalized traffic rules. In *Intelligent Vehicles Symposium (IV)*. IEEE, Los Angeles, CA, USA, 2017. DOI: 10.1109/IVS.2017.7995918.

[R+15] A. Richter et al. Reducing the gap between simulated and real life environments by introducing high-precision data. In *Proceedings of the Driving Simulation Conference 2015 Europe VR*. Driving Simulation Association, 2015. ISBN: 978-3-9813099-3-5.

[R+16] A. Robicquet et al. Learning Social Etiquette: Human Trajectory Understanding in Crowded Scenes. In *European Conference on Computer Vision*. Springer International Publishing, Cham, Switzerland, 2016. DOI: 10.1007/978-3-319-46484-8_33.

[R+18a] S. Riedmaier et al. Validation of X-in-the-Loop Approaches for Virtual Homologation of Automated Driving Functions. In *11th Graz Symposium Virtual Vehicle*. Article - (Retrieval date: 01.06.2022), 2018. Technische Universität München - Lehrstuhl für Fahrzeugtechnik.

[R+18b] C. Roesener et al. Modelling Human Driver Performance for Safety Assessment of Road Vehicle Automation. In *21st International Conference on Intelligent Transportation Systems*. IEEE, Maui, HI, USA, 2018. DOI: 10.1109/ITSC.2018.8569669.

[R+20] S. Riedmaier et al. Unified Framework and Survey for Model Verification, Validation and Uncertainty Quantification. In *Archives of Computational Methods in Engineering*. Springer, 2020. DOI: 10.1007/s11831-020-09473-7.

[RA12] Road and Transportation Research Association. Urban road classification. In *Directives for the Design of Urban Roads*. Report - (Retrieval date: 01.06.2022), 2012.

[Rod19] A. Rodrigues. Disengagement Report Update. In *Embark*. Report - (Retrieval date: 01.06.2022), 2019.

[Roo02] D. Roos. Approximation und Interpolation von Grenzzustandsfunktionen zur Sicherheitsbewertung nichtlinearer Finite-Elemente-Strukturen. In *Bauhaus-Universität Weimar*. Dissertation - (Retrieval date: 01.06.2022), 2002. Fakultät Bauingenieurwesen.

[Rös20] C. Rösener. A traffic-based method for safety impact assessment of road vehicle automation. In *RWTH Aachen University*. Dissertation - (Retrieval date: 01.06.2022), 2020. DOI: 10.18154/RWTH-2020-08950.

[S+08] A. Saltelli et al. Sensitivity Analysis: From Theory to Practice. In *Global Sensitivity Analysis: The Primer*, pages 237–275. John Wiley & Sons, Ltd, 2008. DOI: 10.1002/9780470725184.ch6.

[S+09] M. Serfling et al. Camera and imaging radar feature level sensorfusion for night vision pedestrian recognition. In *Intelligent Vehicles Symposium*. IEEE, Xi'an, China, 2009. DOI: 10.1109/IVS.2009.5164345.

[S+18a] F. Schuldt et al. A Method for an Efficient, Systematic Test Case Generation for Advanced Driver Assistance Systems in Virtual Environments. In *Automotive Systems Engineering II*, pages 147–175. Springer International Publishing, Cham, Switzerland, 2018. DOI: 10.1007/978-3-319-61607-0_7.

[S+18b] S. Steinmeyer et al. Method and Device in a Motor Vehicle for Improved Data Fusion in an Environment Detection. In *United States Patent and Trademark Office*. Volkswagen AG, 2018. AN: 201615749994.

[S+19a] L. Stark et al. Quantifying Vision Zero: crash avoidance in rural and motorway accident scenarios by combination of ACC, AEB, and LKS projected to German accident occurrence. In *Traffic injury prevention*. Taylor & Francis, 2019. DOI: 10.1080/15389588.2019.1605167.

[S+19b] L. Stark et al. Towards Vision Zero: Addressing White Spots by Accident Data based ADAS Design and Evaluation. In *International Conference of Vehicular Electronics and Safety*. IEEE, 2019. DOI: 10.1109/ICVES.2019.8906409.

[Sas17] V. Sasse. Autonomous driving from individualism towards collectivism. In *ATZelektronik worldwide*. Springer, 2017. DOI: 10.1007/s38314-017-0007-4.

[Sax08] E. Sax. Bedeutung des Testens in der Automobilindustrie. In *Automatisiertes Testen eingebetteter Systeme in der Automobilindustrie*. Carl Hanser Verlag, 2008. DOI: 10.3139/9783446419018.fm.

[SB08] M. Stämpfle and W. Branz. Kollisionsvermeidung im Längsverkehr - die Vision vom unfallfreien Fahren rückt näher. In *3. Tagung Aktive Sicherheit durch Fahrerassistenz*. Robert Bosch GmbH, 2008. Article - (Retrieval date: 01.06.2022).

[SBP+19] J.E. Stellet, T. Brade, A. Poddey, S. Jesenski, and W. Branz. Formalisation and algorithmic approach to the automated driving validation problem. In *Intelligent Vehicles Symposium*. IEEE, 2019. DOI: 10.1109/IVS.2019.8813894.

[SBWE19] J. Sauerbier, J. Bock, H. Weber, and L. Eckstein. Definition of scenarios for safety validation of automated driving functions. In *ATZ Worldwide*. Springer, 2019. DOI: 10.1007/s38311-018-0197-2.

[Sch05] Peter Scholz. Echtzeit, Echtzeitsysteme, Echtzeitbetriebssysteme. In *Softwareentwicklung eingebetteter Systeme: Grundlagen, Modellierung, Qualitätssicherung*. Springer Verlag, Berlin, Heidelberg, 2005. DOI: 10.1007/3-540-27522-3_3.

[Sch15] HP. Schöner. Challenges and approaches for testing of highly automated vehicles. In *Energy Consumption and Autonomous Driving*. Springer International Publishing, Cham, Switzerland, 2015. DOI: 10.1007/978-3-319-19818-7_11.

[Sch17a] M. Schreier. Bayesian environment representation, prediction, and criticality assessment for driver assistance systems. In *Technische Universität Darmstadt*. Dissertation - (Retrieval date: 01.06.2022), 2017. Institut für Automatisierungstechnik und Mechatronik.

[Sch17b] F. Schuldt. Ein Beitrag für den methodischen Test von automatisierten Fahrfunktionen mit Hilfe von virtuellen Umgebungen. In *Technische Universität Braunschweig*. Dissertation - (Retrieval date: 01.06.2022), 2017. Fakultät für Elektrotechnik, Informationstechnik, Physik.

[SH19a] L. Schnieder and R.S. Hosse. Das SOTIF-Vorgehensmodell. In *Leitfaden Safety of the Intended Functionality*. Springer Vieweg, Wiesbaden, Germany, 2019. DOI: 10.1007/978-3-658-25023-2_4.

[SH19b] L. Schnieder and R.S. Hosse. Verfeinerung der Sicherheit der Sollfunktion auf dem Weg zum autonomen Fahren. In *Leitfaden Safety of the Intended Functionality*. Springer Vieweg, Wiesbaden, Germany, 2019. DOI: 10.1007/978-3-658-25023-2.

[SH19c] L. Schnieder and R.S. Hosse. Warum brauchen wir SOTIF? In *Leitfaden Safety of the Intended Functionality*. Springer Vieweg, Wiesbaden, Germany, 2019. DOI: 10.1007/978-3-658-25023-2_3.

[SH20] L. Schnieder and R.S. Hosse. Fallstudie zur Gestaltung von SOTIF. In *Leitfaden Safety of the Intended Functionality*. Springer Vieweg, Wiesbaden, Germany, 2020. DOI: 10.1007/978-3-658-25023-2_5.

[SHE+17] B. Schürmann, D. Heß, J. Eilbrecht, O. Stursberg, F. Köster, and M. Althoff. Ensuring drivability of planned motions using formal methods. In *20th International Conference on Intelligent Transportation Systems*. IEEE, Yokohama, Japan, 2017. DOI: 10.1109/ITSC.2017.8317647.

[SHK+18] Y. Sun, X. Huang, D. Kroening, J. Sharp, M. Hill, and R. Ashmore. Testing deep neural networks. In *Machine Learning*. arXiv - (Retrieval date: 01.06.2022), 2018.

[Sie14] K. Siegemund. Contributions to ontology-driven requirements engineering. In *Technischen Universität Dresden*. Dissertation - (Retrieval date: 01.06.2022), 2014. Fakultät Informatik - Lehrstuhl Softwaretechnologie.

[Sjo22] K. Sjoberg. Automated trucks overtake self-driving cars [connected and automated vehicles]. In *IEEE Vehicular Technology Magazine*. IEEE, 2022. DOI: 10.1109/MVT.2021.3133558.

[SKOK18] S. Shafaei, S. Kugele, M.H. Osman, and A. Knoll. Uncertainty in machine learning: A safety perspective on autonomous driving. In *International Conference on Computer Safety, Reliability, and Security*. Article - (Retrieval date: 01.06.2022), 2018.

[SMAN08] K. Schmitt, J. Madsen, M. Anitescu, and D. Negrut. A Gaussian process-based approach for handling uncertainty in vehicle dynamics simulation. In *International Mechanical Engineering Congress and Exposition*. ASME, 2008. DOI: 10.1115/IMECE2008-66664.

[Smi17] B.W. Smith. Automated Driving and Product Liability. In *Michigan State Law Review*. Michigan State University College of Law, 2017. DOI: 10.17613/3bj8-z207.

[SMJ19] M. Saraoglu, A. Morozov, and K. Janschek. MOBATSim: MOdel-based Autonomous Traffic Simulation Framework for Fault-Error-Failure Chain Analysis. In *International Federation of Automatic Control*. Elsevier Ltd., 2019. DOI: 10.1016/j.ifacol.2019.08.077.

[SN19] S.E. Shladover and C. Nowakowski. Regulatory challenges for road vehicle automation: Lessons from the california experience. In *Transportation research part A: policy and practice*. Elsevier Ltd., 2019. DOI: 10.1016/j.tra.2017.10.006.

[SPW+19] J.J. So, I. Park, J. Wee, S. Park, and I. Yun. Generating Traffic Safety Test Scenarios for Automated Vehicles using a Big Data Technique. In *KSCE booktitle of Civil Engineering*. Springer, 2019. DOI: 10.1007/s12205-019-1287-4.

[SS06] LP. Shiu and CY. Sin. Top-Down, Middle-Out, and Bottom-Up Processes: A Cognitive Perspective of Teaching and Learning Economics. In *International Review of Economics Education*, volume 5, pages 60–72. Elsevier Ltd., 2006. DOI: 10.1016/S1477-3880(15)30124-9.

[SS13] B. Schick and S. Schmidt. Evaluation of Video-Based Driver Assistance Systems with Sensor Data Fusion by Using Virtual Test Driving. In *Proceedings of the FISITA 2012 World Automotive Congress*, pages 1363–1375. Springer Verlag, Berlin, Heidelberg, 2013. DOI: 10.1007/978-3-642-33738-3_36.

[Ste14] S. Steinmeyer. Probabilistische Fahrzeugumfeldschätzung für Fahrerassistenzsysteme. In *Technische Universität Braunschweig*. Dissertation - (Retrieval date: 01.06.2022), 2014. Fakultät für Elektrotechnik, Informationstechnik, Physik.

[Ste16] J.E. Stellet. Statistical modelling of algorithms for signal processing in systems based on environment perception. In *Karlsruher Institute for Technology*. Dissertation - (Retrieval date: 01.06.2022), 2016. Institute for Anthropomatics and Robotics.

[STZ+11] K. Siegemund, E.J. Thomas, Y. Zhao, J. Pan, and U. Assmann. Towards ontology-driven requirements engineering. In *Workshop Semantic Web Enabled Software Engineering at 10th International Semantic Web Conference (ISWC)*. Article - (Retrieval date: 01.06.2022), Bonn, Germany, 2011.

[SWB+20] J.E. Stellet, M. Wöhrle, T. Brade, A. Poddey, and W. Branz. Validation of automated driving - A structured analysis and survey of approaches. In *FAS-Workshop*. Uni-DAS e.V., 2020. Article - (Retrieval date: 01.06.2022).

[SWZ12] I. Scherhaufer, A. Wingert, and Z. Zomotor. Active Brake Assist, Six Years of Experience since Market Launch. In *3rd International Munich Chassis Symposium*, volume 2, pages 705–720. ATZlive - (Retrieval date: 01.06.2022), Munich, Germany, 2012.

[SZS+15] J.E. Stellet, M.R. Zofka, J. Schumacher, T. Schamm, F. Niewels, and J.M. Zollner. Testing of Advanced Driver Assistance Towards Automated Driving: A Survey and Taxonomy on Existing Approaches and Open Questions. In *18th International Conference on Intelligent Transportation Systems*, pages 1455–1462. IEEE, Gran Canaria, Spain, 2015. DOI: 10.1109/ITSC.2015.236.

[T+17] A.S. Trigell et al. Advanced vehicle dynamics of heavy trucks with the perspective of road safety. In *Vehicle System Dynamics*, volume 55, pages 1572–1617. Taylor & Francis, 2017. DOI: 10.1080/00423114.2017.1319964.

[Tal10] N.N. Taleb. A Brief History of the Black Swan Problem. In *The Black Swan: The Impact of the Highly Improbable*, volume 25. Random House, 2010. DOI: 10.5465/amp.25.2.87.

[Tar12] A.P. Tarko. Use of crash surrogates and exceedance statistics to estimate road safety. In *Accident Analysis & Prevention*, volume 45, pages 230–240. Elsevier Ltd., 2012. DOI: 10.1016/j.aap.2011.07.008.

[TDG+18] Á. Takács, D.A. Drexler, P. Galambos, I.J. Rudas, and T. Haidegger. Assessment and Standardization of Autonomous Vehicles. In *22nd International Conference on Intelligent Engineering Systems*, pages 185–192. IEEE, Las Palmas de Gran Canaria, Spain, 2018. DOI: 10.1109/INES.2018.8523899.

[Tel14] D. Tellmann. Hardware-in-the-Loop-gestützte Entwicklungsplattform für Fahrerassistenzsysteme: Modelle der Umfeldsensorik und angepasste Fahrermodelle. In *Universität Kassel*. Dissertation - (Retrieval date: 01.06.2022), 2014. Fachbereich 16 Elektrotechnik/Informatik.

[TH13] Y. Tan and B. Hassan. A Concept of Camera Test-Bench for Testing Camera Based Advanced Driver Assistance Systems. In *33rd Computers and Information in Engineering Conference*. The American Society of Mechanical Engineers (ASME), 2013. DOI: 10.1115/DETC2013-12996.

[TM20] S. Tiedemann and A. Mank. Solving the validation challenge of automated driving with a holistic test center. In *Automatisiertes Fahren 2019*, pages 155–164. Springer Fachmedien Wiesbaden, Wiesbaden, Germany, 2020. DOI: 10.1007/978-3-658-27990-5_14.

[Tru18] Mercedes-Benz Trucks. Operating instructions. In *Actros, Arocs, Antos*. Article - (Retrieval date: 01.06.2022), 2018.

[Tun19] C.E. Tuncali. Search-based Test Generation for Automated Driving Systems: From Perception to Control Logic. In *Arizona State University*. Dissertation - (Retrieval date: 01.06.2022), 2019.

[TuS19] TuSimple. TuSimple Self Driving Safety Report. In *Version 2.0*. Report - (Retrieval date: 01.06.2022), 2019.

[TZ15] J. Trost and Z. Zoltan. Method for Operating a Brake Assist Device and Brake Assist Device for a Vehicle. In *United States Patent and Trademark Office*. Daimler AG, 2015. AN: 201113822362.

[U+15] S. Ulbrich et al. Defining and Substantiating the Terms Scene, Situation, and Scenario for Automated Driving. In *Intelligent Transportation Systems*. IEEE, Gran Canaria, Spain, 2015. DOI: 10.1109/ITSC.2015.164.

[uEB22] Bundesverband Güterkraftverkehr Logistik und Entsorgung (BGL). In Deutschland bei Lkw-Unfällen Getötete und Schwerverletzte im Vergleich zur Lkw-Transportleistung 1992–2021. In *Statistisches Bundesamt*. DESTATIS - (Retrieval date: 01.06.2022), 2022.

[UL22] Underwriters Laboratories. Standard for Safety for the Evaluation of Autonomous Vehicles and Other Products. In *ANSI/UL 4600*. Standard - (Retrieval date: 01.06.2022), 2022.

[UNGO21] I. Urbieta, M. Nieto, M. Garcia, and O. Otaegui. Design and Implementation of an Ontology for Semantic Labeling and Testing: Automotive Global Ontology (AGO). In *Applied Sciences*, volume 11. MDPI, 2021. DOI: 10.3390/app11177782.

[UNMH14] S. Ulbrich, T. Nothdurft, M. Maurer, and P. Hecker. Graph-based context representation, environment modeling and information aggregation for automated driving. In *Intelligent Vehicles Symposium*, pages 541–547. IEEE, Dearborn, Michigan, USA, 2014. DOI: 10.1109/IVS.2014.6856556.

[VDA17] VDA. Automative SPICE - Process Assessment / Reference Model - Version 3.1. In *Quality Management in the Automotive Industry*. Standard - (Retrieval date: 01.06.2022), 2017.

[VGL+12] B. Vanholme, D. Gruyer, B. Lusetti, et al. Highly Automated Driving on Highways Based on Legal Safety. In *IEEE Transactions on Intelligent Transportation Systems*, volume 14, pages 333–347. IEEE, 2012. DOI: 10.1109/TITS.2012.2225104.

[vNC14] K. von Neumann-Cosel. Virtual Test Drive: Simulation umfeldbasierter Fahrzeugfunktionen. In *Technische Universität München*. Dissertation - (Retrieval date: 01.06.2022), 2014. Lehrstuhl für Echtzeitsysteme und Robotik.

[vNCDW09] K. von Neumann-Cosel, M. Dupuis, and C. Weiss. Virtual Test Drive provision of a consistent tool-set for [D, H, S, V]-in-the-Loop. In *Proceedings of the Driving Simulation Conference 2009 Europe VR*. Driving Simulation Association, Monte Carlo, Monaco, 2009.

[VR15] I.T. Voyles and C.J. Roy. Evaluation of Model Validation Techniques in the Presence of Aleatory and Epistemic Input Uncertainties. In *17th AIAA Non-Deterministic Approaches Conference*. Aerospace Research Central, Kissimmee, Florida, 2015. DOI: 10.2514/6.2015-1374.

[W+14] A. Welzel et al. Accurate camera-based traffic sign localization. In *17th International IEEE Conference on Intelligent Transportation Systems (ITSC)*. IEEE, Qingdao, China, 2014. DOI: 10.1109/ITSC.2014.6957730.

[W+15] M. Weiskopf et al. Absicherung eines Radarsensors im Systemverbund mit der Hardware-in-the-Loop Testtechnologie. In *Automotive-Safety & Security*, pages 29–40. Gesellschaft für Informatik e.V., Bonn, Germany, 2015. ISBN: 978-3-88579-634-3.

[W+16] W. Wachenfeld et al. The worst-time-to-collision metric for situation identification. In *2016 IEEE Intelligent Vehicles Symposium (IV)*, pages 729–734. IEEE, Gothenburg, Sweden, 2016. DOI: 10.1109/IVS.2016.7535468.

[W+18] H. Winner et al. Validation and Introduction of Automated Driving. In *Automotive Systems Engineering II*, pages 177–196. Springer International Publishing, Cham, Switzerland, 2018. DOI: 10.1007/978-3-319-61607-0_8.

[W+19] H. Watanabe et al. Methodology of Scenario Clustering for Predictive Safety Functions. In *9. Tagung Automatisiertes Fahren*. Article - (Retrieval date: 01.06.2022), 2019. Technische Universität München - Lehrstuhl für Fahrzeugtechnik.

[Wac17] W.H.K. Wachenfeld. How Stochastic can Help to Introduce Automated Driving. In *Technische Universität Darmstadt*. Dissertation - (Retrieval date: 01.06.2022), 2017. Fachgebiet Fahrzeugtechnik.

[Wag21] S.K. Wagner. Efficient Scenario-based Assessment of Automated Driving Systems through Virtual Testing. In *Technische Universität München*. Dissertation - (Retrieval date: 01.06.2022), 2021. Fakultät für Informatik.

[Wan21] C. Wang. Silent Testing for Safety Validation of Automated Driving in Field Operation. In *Technische Universität Darmstadt*. Dissertation - (Retrieval date: 01.06.2022), 2021. Fachgebiet Fahrzeugtechnik.

[Web19] H.and others Weber. A framework for definition of logical scenarios for safety assurance of automated driving. In *Traffic Injury Prevention*, volume 20, pages 65–70. Taylor & Francis, 2019. DOI: 10.1080/15389588.2019.1630827.

[Wei13] A. Weitzel. Objective Controllability Assessment for Unintended ADAS Reactions. In *Automotive Systems Engineering*, pages 135–145. Springer Verlag, Berlin, Heidelberg, 2013. DOI: 10.1007/978-3-642-36455-6_7.

[WHO20] WHO. Road Traffic Injuries. In *World Health Organization*. Article - (Retrieval date: 01.06.2022), 2020.

[Wie17] A. Wieczorek. The influence of self-driving transport vehicles on the field of logistics. In *Transport Economics and Logistics*, volume 66, pages 107–114. Research Journal of the University of Gdansk, 2017. DOI: 10.5604/01.3001.0010.5602.

[Wil15] B. Wilmes. Hybrides Testverfahren für Simulink/Targetlink Modelle. In *Technische Universität Berlin*. Dissertation - (Retrieval date: 01.06.2022), 2015. Fakultät IV - Elektrotechnik und Informatik.

[Wil16] B. Wilmes. TASMO: Automated Test Data Generation for Simulink Model Coverage. In *Simulation and testing for vehicle technology*, pages 123–133. Springer International Publishing, Cham, Switzerland, 2016. DOI: 10.1007/978-3-319-32345-9_-10.

[Win13] H. Winner. Challenges of Automotive Systems Engineering for Industry and Academia. In *Automotive Systems Engineering*, pages 3–15. Springer Verlag, Berlin, Heidelberg, 2013. DOI: 10.1007/978-3-642-36455-6_1.

[Win16a] T. Winkle. Development and Approval of Automated Vehicles: Considerations of Technical, Legal, and Economic Risks. In *Autonomous Driving: Technical, Legal and Social Aspects*. Springer Verlag, Berlin, Heidelberg, 2016. DOI: 10.1007/978-3-662-48847-8_28.

[Win16b] H. Winner. ADAS, Quo Vadis? In *Handbook of Driver Assistance Systems: Basic Information, Components and Systems for Active Safety and Comfort*, pages 1557–1584. Springer International Publishing, Cham, Switzerland, 2016. DOI: 10.1007/978-3-319-12352-3_62.

[WK15] M. Wagner and P. Koopman. A Philosophy for Developing Trust in Self-driving Cars. In *Road Vehicle Automation 2*, pages 163–171. Springer International Publishing, 2015. Article - (Retrieval date: 01.06.2022).

[WLFM19] H. Winner, K. Lemmer, T. Form, and J. Mazzega. PEGASUS—First Steps for the Safe Introduction of Automated Driving. In *Road Vehicle Automation 5*, pages 185–195. Springer International Publishing, Cham, Switzerland, 2019. DOI: 10.1007/978-3-319-94896-6_16.

[WPZ19] X. Wang, H. Peng, and D. Zhao. Combining Reachability Analysis and Importance Sampling for Accelerated Evaluation of Highly Automated Vehicles at Pedestrian Crossing. In *Dynamic Systems and Control Conference*, volume 3. ASME, 2019. DOI: 10.1115/DSCC2019-9179.

[WRM+19] M. Wood, P. Robbel, M. Maass, et al. Verification and Validation. In *Safety First For Automated Driving*. Report - (Retrieval date: 01.06.2022), 2019.

[WW16] W. Wachenfeld and H. Winner. The Release of Autonomous Vehicles. In *Autonomous Driving: Technical, Legal and Social Aspects*, pages 425–449. Springer Verlag, Berlin, Heidelberg, 2016. DOI: 10.1007/978-3-662-48847-8_21.

[WW19] C. Wang and H. Winner. Overcoming Challenges of Validation Automated Driving and Identification of Critical Scenarios. In *2019 IEEE Intelligent Transportation Systems Conference (ITSC)*, pages 2639–2644. IEEE, Auckland, New Zealand, 2019. DOI: 10.1109/ITSC.2019.8917045.

[Xi08] C. Xi. Requirements and concepts for future automotive electronic architectures from the view of integrated safety. In *Universität Karlsruhe*. Dissertation - (Retrieval date: 01.06.2022), 2008. Fakultät für Elektrotechnik und Informationstechnik.

[Z+18a] H. Zhang et al. An ontology-guided semantic data integration framework to support integrative data analysis of cancer survival. In *Medical Informatics and Decision Making*. BMC, 2018. ISSN: 1472-6947.

[Z+18b] D. Zhao et al. Accelerated Evaluation of Automated Vehicles in Car-Following Maneuvers. In *IEEE Transactions on Intelligent Transportation Systems*, volume 19, pages 733–744. IEEE, 2018. DOI: 10.1109/TITS.2017.2701846.

[ZBIG12] C. Zinoune, P. Bonnifait, and J. Ibañez-Guzmán. Detection of missing roundabouts in maps for Driving Assistance Systems. In *2012 IEEE Intelligent Vehicles Symposium*, pages 123–128. IEEE, Madrid, Spain, 2012. DOI: 10.1109/IVS.2012.6232245.

[Zha16] D. Zhao. Accelerated Evaluation of Automated Vehicles. In *University of Michigan*. Dissertation - (Retrieval date: 01.06.2022), 2016. Mechanical Engineering.

[Zio13] E. Zio. System Reliability and Risk Analysis. In *The Monte Carlo Simulation Method for System Reliability and Risk Analysis*, pages 7–17. Springer, London, UK, 2013. DOI: 10.1007/978-1-4471-4588-2_2.

[ZKK+16] M.R. Zofka, S. Klemm, F. Kuhnt, T. Schamm, and J.M. Zöllner. Testing and validating high level components for automated driving: Simulation framework for traffic scenarios. In *Intelligent Vehicles Symposium*, pages 144–150. IEEE, Gothenburg, Sweden, 2016. DOI: 10.1109/IVS.2016.7535378.

[ZMMFm13] J. Zürn, U. Mierisch, R. Müller-Finkeldei, and A. Köllermeyer. Der Neue Arocs von Mercedes-Benz. In *ATZoffhighway 6*, pages 24–36. Springer, 2013. DOI: 10.1365/s35746-013-0058-2.

[ZRBE20] A. Zlocki, D. Raudszus, J. Bock, and L. Eckstein. Methoden zur Absicherung von Komponenten und Funktionen. In *ATZextra 25*, pages 28–33. Springer, 2020. DOI: 10.1007/s35778-020-0111-6.

[ZWZ18] J. Zhu, W. Wang, and D. Zhao. A Tempt to Unify Heterogeneous Driving Databases using Traffic Primitives. In *Computer Vision and Pattern Recognition*. arXiv: 1805.04925 - (Retrieval date: 01.06.2022), 2018.

Project-Related Publications

Conferences and Journals

[BES+21] A. Birlet, M. Elgharbawy, A. Schwarzhaupt, U. Wiesel, C. Weber, and M. Frey. Kundenorientierte Absicherung mit Hilfe methodischer Testfallerstellung aus Felddaten. In *Proceedings of the 6th Commercial Vehicle Technology Symposium*, pages 431–458. Springer Vieweg, Wiesbaden, Germany, 2021. DOI: 10.1007/978-3-658-29717-6_29.

[E+16] M. Elgharbawy et al. An Agile Verification Framework for Traffic Sign Classification Algorithms in Heavy Vehicles. In *2016 IEEE/ACS 13th International Conference of Computer Systems and Applications (AICCSA)*, pages 1–8. IEEE, Agadir, Morocco, 2016. DOI: 10.1109/AICCSA.2016.7945719.

[E+18] M. Elgharbawy et al. A Data-driven Verification Framework for Active Safety Functions. In *Proceedings of the Driving Simulation Conference 2018 Europe VR*, pages 131–135. Driving Simulation Association, Antibes, France, 2018. ISBN: 978-2-85782-734-4.

[ESA+17] M. Elgharbawy, A. Schwarzhaupt, R. Arenskrieger, M. Frey, and F. Gauterin. A Testing Framework for Predictive Driving Features with an Electronic-Horizon. In *Proceedings of the Driving Simulation Conference 2017 Europe VR*, pages 17–18. Driving Simulation Association, Stuttgart, Germany, 2017. ISSN: 0769-0266.

[ESFG18] M. Elgharbawy, I. Scherhaufer, M. Frey, and F. Gauterin. A Scenario-based Verification Framework for Truck Platooning Functions. In *Proceedings of the Driving Simulation Conference 2018 Europe VR*. Driving Simulation Association, Antibes, France, 2018. ISBN: 978-2-85782-734-4.

[ESFG19a] M. Elgharbawy, A. Schwarzhaupt, M. Frey, and F. Gauterin. Ontology-based Adaptive Testing for Automated Driving Functions using Data Mining Techniques. In *Special TRF issue: Driving simulation*, volume 66, pages 234–251. Transportation Research Part F: Traffic Psychology and Behaviour, 2019. DOI: 10.1016/j.trf.2019.07.021.

[ESFG19b] M. Elgharbawy, A. Schwarzhaupt, M. Frey, and F. Gauterin. A Real-time Multisensor Fusion Verification Framework for Advanced Driver Assistance Systems. In *Special TRF issue: Driving simulation*, volume 61, pages 259–267. Transportation Research Part F: Traffic Psychology and Behaviour, 2019. DOI: 10.1016/j.trf.2016.12.002.

[ESFG20] M. Elgharbawy, A. Schwarzhaupt, M. Frey, and F. Gauterin. Measurable Safety of Automated Driving Functions using Stochastic Analysis Methods. In *Proceedings of the Driving Simulation Conference 2020 Europe VR*, pages 27–34. Driving Simulation Association, Antibes, France, 2020. ISSN: 0769-0266.

[ESm+19] M. Elgharbawy, A. Schwarzhaupt, R. Arenskrieger, H. Elsayed, M. Frey, and F. Gauterin. A Testing Framework for Predictive Driving Features with an electronic Horizon. In *Special TRF issue: Driving simulation*, volume 61, pages 291–304. Transportation Research Part F: Traffic Psychology and Behaviour, 2019. DOI: 10.1016/j.trf.2017.08.002.

[ESO+19] M. Elgharbawy, I. Scherhaufer, K. Oberhollenzer, M. Frey, and F. Gauterin. Adaptive Functional Testing for Autonomous Trucks. In *International Journal of Transportation Science and Technology*, volume 8, pages 202–218. Tongji University Press, 2019. DOI: 10.1016/j.ijtst.2018.11.003.

[ESS+16] M. Elgharbawy, A. Schwarzhaupt, G. Scheike, M. Frey, and F. Gauterin. A Generic Architecture of ADAS Sensor Fault Injection for Virtual Tests. In *2016 IEEE/ACS 13th International Conference of Computer Systems and Applications (AICCSA)*, pages 1–7. IEEE, Agadir, Morocco, 2016. DOI: 10.1109/AICCSA.2016.7945680.

[ESW+16] M. Elgharbawy, A. Schwarzhaupt, G. Weiskopf, M. Frey, and F. Gauterin. A Real-time Multi-Sensor Fusion Verification Framework for Advanced Driver Assistance Systems. In *Proceedings of the Driving Simulation Conference 2016 Europe VR*, pages 145–149. Driving Simulation Association, Paris, France, 2016. ISSN: 0769-0266.

[RESS21] F. Reisgys, M. Elgharbawy, A. Schwarzhaupt, and E. Sax. Argumentation on ADAS Simulation Validity using Aleatory and Epistemic Uncertainty Estimation. In *Proceedings of the Driving Simulation Conference 2021 Europe VR*, pages 25–32. Driving Simulation Association, Munich, Germany, 2021. ISSN: 0769-0266.

Supervised Theses

[Amm15] M.B. Ammar. Entwicklung und Integration von Umfeldsensormodellen für Fahrerassistenzsysteme in einer echtzeitfähigen Hil-gestützten Simulationsumgebung. In *Technische Universität Dresden*. Diplomarbeit, 2015. Lehrstuhl Fahrzeugmechatronik.

[Are16] R. Arenskrieger. HiL-Absicherungsumgebung für Fahrerassistenzsysteme mit elektronischem Horizont. In *Karlsruher Institut für Technologie*. Masterarbeit, 2016. Institut für Fahrzeugsystemtechnik.

[Ber16] B. Bernier. Conception and implementation of an algorithm for detection and classification of traffic signs in heavy vehicles. In *Karlsruhe Institute of Technology*. Master's Thesis, 2016. Institute for Vehicle System Technology.

[Els18] H. Elsayed. Data-driven analysis of naturalistic driving for automated driving functions. In *University of Stuttgart*. Master's Thesis, 2018. Institute for Visualization and Interactive Systems.

[Kan16] M. Kanberger. Evaluation of testability of camera-based driver assistance systems based on Hardware-in-the-Loop simulations by verification of CAN messages. In *University of Applied Sciences Esslingen*. Master's Thesis, 2016. Graduate School.

[Sch16a] G. Scheike. Entwicklung eines graphischen Tools zur Evaluierung der Testaktivität bei E/E-Systemen. In *Fachhochschule Dortmund*. Studienarbeit, 2016. Fachbereich Informations- und Elektrotechnik.

[Sch16b] G. Scheike. Konzeptionierung und Umsetzung einer gesamtheitlichen Datenverwaltung für einen FAS HiL Prüfstand. In *Fachhochschule Dortmund*. Studienarbeit, 2016. Fachbereich Informations- und Elektrotechnik.

[Sch17] G. Scheike. Entwurf und Umsetzung eines agilen Testverfahrens für Fahrerassistenzsysteme. In *Fachhochschule Dortmund*. Bachelorarbeit, 2017. Fachbereich Informations- und Elektrotechnik.

Invention Disclosures

[EDA20] M. Elgharbawy, J. Dickmann, and N. Appenrodt. Verfahren zur Evaluierung von systematischen und statistischen Fehlern von Multi-Radar basierten Erkennungssystemen. In *German Patent and Trade Mark*. Daimler AG, 2020. AN: 102020006644.

[ESD20] M. Elgharbawy, A. Schwarzhaupt, and J. Dickmann. Verfahren zum Testen einer automatisierten Fahrfunktion. In *German Patent and Trade Mark*. Daimler AG, 2020. AN: 102020005507.

[ESF19] M. Elgharbawy, A. Schwarzhaupt, and M. Frey. Verfahren zum Testen eines Spurhalteassistenzsystems für ein Fahrzeug. In *German Patent and Trade Mark*. Daimler AG, 2019. AN: 102017009971.

[ESS+19a] M. Elgharbawy, A. Schwarzhaupt, I. Scherhaufer, M. Gut, and M. Frey. Verfahren zum Testen eines Assistenzsystems für ein Fahrzeug. In *German Patent and Trade Mark*. Daimler AG, 2019. AN: 102018005865.

[ESS+19b] M. Elgharbawy, A. Schwarzhaupt, I. Scherhaufer, M. Gut, and M. Frey. Verfahren zum Testen eines Totwinkelassistenzsystems für ein Fahrzeug. In *German Patent and Trade Mark*. Daimler AG, 2019. AN: 102018005864.

[ESS+19c] H. Elsayed, A. Schwarzhaupt, I. Scherhaufer, M. Gut, M. Elgharbawy, and M. Frey. Verfahren zum Testen eines Bremsassistenzsystems für ein Fahrzeug. In *German Patent and Trade Mark*. Daimler AG, Karlsruher Institut für Technologie (KIT), 2019. AN: 102018004429.