UNIVERSITÀ DEGLI STUDI DI NAPOLI “FEDERICO II”

DIPARTIMENTO DI INGEGNERIA CIVILE, EDILE E AMBIENTALE

DOTTORATO IN

INGEGNERIA DEI SISTEMI CIVILI

XXXIII CICLO

DEFINING PROCEDURES AND SIMULATION TOOLS TO TEST HIGH LEVELS OF AUTOMATION FOR CARS IN REALISTIC TRAFFIC, DRIVING AND BOUNDARY CONDITIONS

COORDINATORE:

Prof. Ing. Andrea Papola

CANDIDATO:

Dott. Ing. Claudio D’Aniello

TUTOR:

Prof. Ing. Gennaro Nicola Bifulco

«How many roads must a "car" run down before you can call it "tested"? »

Bob Dylan (more or less)

CONTENTS

Cap.1: Background 5
1.1 Introduction 5
1.2 Organismi e stakeholder 6
Society of automotive engineers (SAE) 6
Defense advanced research projects agency (DARPA) 8
Testing companies and standards 9
1.3 The simularion 12
1.4 Approaches 14
1.5 Making scenarios 20
1.6 Summary of literature known methods and unresolved problems 25
Cap.2: The proposed approach 27
2.1 An impersonal approach to evaluate all the unknown scenarios 27
2.2 A top-down approach for the analysis of road casistic, based on the road regulations and technical law 28
2.3 Identification of the activities compatibles with the road sections 34
2.4 Elaborating the Scene/Activity table and specifications within the activity of designing control logics 39
2.5 Low Resolution Scenario (LoReS): description of the new scenario tool, potential, problems, areas of application and technical needs. 43
Cap.3: Methodological and operative tools 46
3.1 Sensitivity analysis 46
Extraction methods 46
Sensitivity indices 48
Bootstrapping confidence intervals 49
Regional sensitivity analysis 50
Scatterplot analysis 52
3.2 Simulation tools 53

Vehicle dynamics simulation 53
MIL e HIL 54
Cap.4: Method developement 56
4.1 Design of experiments 56
4.2 Experiment 1: Experimenting the simulation tools dealing with scenario complexity 56
4.3 Experiment 2: Perfecting analysis tools and comparison with known testing methods 62
4.4 Experiment 3: Reducing the computational weight 68
Cap.5: Proof of concept: project MEDIATOR 76
5.1 Context and main objectives of the MEDIATOR project 77
5.2 Research contribution to the project and case study 79
5.3 Experiment description 81
5.4 Implementation: 82
5.5 Results: 86
Cap.6: Conclusions and discussions 92
6.1 Conclusions from the experiment results 92
6.2 Method perfectioning and results 93
6.3 Advantages/opportunity and criticity of the proposed tools and methods 94
6.4 Possible future developments of the research 95
Bibliografia: 98

Cap.1: Background

1.1 Introduction

Automated driving topic has featured prominently in research over the last decade, since autonomous and connected driving technology represents a major paradigm shift in the mobility and transport system. These technologies find their primary motivation in the increase of road safety, which still today represents a problem because of many road victims[1]. Europe decided to radically improve road safety[2] and look at drivers as the first cause of accidents[3]. Other areas of great improvement since the introduction of automated driving relate to the user experience and travel comfort, increased capacity of road infrastructure, fuel consumption and pollution. In urban areas, connected and automated driving can enable new mobility models. Many car manufacturers and global technology players are interested in full automation. The Google Car reached 100 miles of uninterrupted driving in 2009, and for some years now Waymo has been testing truck-based freight transport services[4]. At the World Artificial Intelligence Conference in 2020, Elon Musk announced that he is close to creating fully automated vehicles[5], and Uber has been experimenting robo-taxi for some time. Although the recent technological advances in the field of automation have been very significant, the main problem for bringing safe devices to market is the adequacy and comprehensiveness of the tests performed[6]. In Maurer et al. (2013)[7] the authors predicted that testing would be the bottleneck to get autonomous vehicles on the road. Many issues makes difficult to handle the problem. Firstly, there is insufficient behaviour data of self-driving vehicles (individually or as traffic flows) as they cannot simply be equated with human-driven vehicles. This implies unawareness of the possible risk conditions that may lead to accidents. Wachenfeld et al. (2016)[8] define the substantial problem as the absence of a driver. Although this statement may seem trivial, it includes the whole issue. The presence of a human driver does not only concern the driving activities he or she performs, but also the ethical and legal

---

[1] 'Indicatori Di Incidentalità Stradale' <http://dati.istat.it/Index.aspx?DataSetCode=DCIS_INDINCIDENT> [accessed 20 July 2021].

[2] 'Home - Toward Zero Deaths' <https://www.towardzerodeaths.org/> [accessed 20 July 2021].

[3] Area Professionale Statistica ACI ed Acinformatica, 'Veicoli e Incidenti Stradali Anno 2016'.

[4] 'Company' <https://waymo.com/company/#blog> [accessed 11 June 2021].

[5] 'Elon Musk Delivers Virtual Speech for WAIC - SHINE News' <https://www.shine.cn/biz/economy/2007091766/> [accessed 11 June 2021].

[6] Klaus Bengler and others, 'Three Decades of Driver Assistance Systems: Review and Future Perspectives', *IEEE Intelligent Transportation Systems Magazine*, 6.4 (2014), 6–22 <https://doi.org/10.1109/MITS.2014.2336271>.

[7] M Maurer, *Automotive Systems Engineering: A Personal Perspective*, *Automotive Systems Engineering*, 2013 <https://doi.org/10.1007/978-3-642-36455-6_2>.

[8] Walther Wachenfeld and others, 'Use Cases for Autonomous Driving', in *Autonomous Driving: Technical, Legal and Social Aspects*, 2016, pp. 9–37 <https://doi.org/10.1007/978-3-662-48847-8_2>.

responsibility of his or her choices. There are many issues that encompass these two areas. Nevertheless, although the scientific debate on the ethical issue is very broad and difficult to deal with, it is not the subject of this research. Similarly, the legal field, although crucial in determining a faster and more complete adoption of increasingly automated solutions, is also outside the limits of this work.

On the other hand, the importance of the responsibilities linked to the legal aspect, automotive companies are pushing research into the identification of appropriate and exhaustive testing methodologies and, in particular, aimed at safety and the analysis of risk factors. Adequate testing has a twofold function: i) to identify the weak points of the automation systems installed on vehicles, allowing a cyclical improvement in order to minimize the risks for users (passengers and non-passengers); ii) to obtain awareness of possible scenarios, in order to ensure with data the responsibilities of the manufacturer within a percentage of acceptable risk for all actors. In Reschka et al. (2016)[9] the authors elaborates solutions and interventions to manage scenarios in which the autonomous vehicle falls in a situation that it cannot autonomously deal with. Obviously, many approaches can be found, depending on the considered level of automation[10]. This type of issue is the focus of debates in the scientific community, leading to projects such as PEGASUS[11] and MEDIATOR[12]. The latter aims to create a Human Machine Interface (HMI) system that facilitates the transition from human driving to automatic driving, at various levels of automation, intervening properly on the human driver when he/she is needed. These projects of human-machine mediation and intervention aim, once again, to manage uncertainty in all driving situations that may occur. The immense number of variables, their wide variability and different nature, make the testing and risk assessment problem not easy to deal with[13].

1.2 Organismi e stakeholder

Society of automotive engineers (SAE)

Over the years, the problem of driver automation has been divided into different classes and levels, with separate problems and responsibilities. Numerous nomenclatures have been used to distinguish the types and levels of vehicle automation. First of all, the

---

[9] Andreas Reschka, 'Safety Concept for Autonomous Vehicles', in *Autonomous Driving: Technical, Legal and Social Aspects*, 2016, pp. 473–96 <https://doi.org/10.1007/978-3-662-48847-8_23>.

[10] SAE, 'Taxonomy and Definitions for Terms Related to Driving Automation Systems for On-Road Motor Vehicles', *Surface Vehicle Recommended Pratice*, 2021.

[11] Hermann Winner, Walther Wachenfeld, and Philipp Junietz, 'Safety Assurance for Highly Automated Driving – The PEGASUS Approach', in *Transportation Research Board*, 2017.

[12] 'MEDIATOR behind the Scenes: Use Cases | Mediator', 2019 <https://mediatorproject.eu/blog/mediator-behind-the-scenes-use-cases> [accessed 11 June 2021].

[13] Philip Koopman and Michael Wagner, 'Challenges in Autonomous Vehicle Testing and Validation', *SAE International Journal of Transportation Safety*, 4.1 (2016), 2016-01–0128 <https://doi.org/10.4271/2016-01-0128>.

distinction between Driver Assistance Systems (DAS) or Advanced DAS (ADAS) and automated or autonomous vehicles is significant. DAS indicates devices for driver assistance, which help the driver by signaling (passive devices) or correcting (active devices) specific driving maneuvers without replacing the driver. Autonomous and automated driver systems concern systems of devices that replace, partially or totally, the driver. Villagra et al. (2017)[14] distinguish autonomous vehicles from automated and connected vehicles, indicating that:

"Automated vehicles can be defined as those in which at least some safety-critical aspects occur without direct driver input."

Connected: "When these vehicles, (…) can communicate among them and with the infrastructure/cloud, a very relevant socioeconomic impact can be obtained, namely safety, congestion and pollution reduction, capacity increase, etc."

"Autonomous cars have theoretically the ability to operate independently and without the intervention of a driver in a dynamic traffic environment, relying on the vehicle's own systems and without communicating with other vehicles or the infrastructure"

According to this definition, autonomous vehicles are capable of totally replacing the human driver, while the term automated refers to the automation of certain specific tasks in certain contexts. An autonomous vehicle is also automated, while the opposite is not necessarily true. The term 'connected' refers to the ability of the vehicle to exchange information in real time with other vehicles or infrastructures and may or may not belong to autonomous or automated vehicles. This ability gives the possibility to manage driving choices through various information sources, giving the possibility to develop completely new control logics, such as the management of vehicles as flows rather than as single independent vehicles.

The most used distinction in the literature, now adopted as an unofficial standard, is the SAE designation for levels of driving automation. In the SAEJ3016 standard (in its first version in 2014[15]) a scale of 6 levels of automation (from 0 to 5) is first identified. It has been refined several times over the years, most recently in the 2020[16] update shown in Figure 1.

[14] Jorge Villagra and others, *Automated Driving, Intelligent Vehicles: Enabling Technologies and Future Developments* (Elsevier Inc., 2017) <https://doi.org/10.1016/B978-0-12-812800-8.00008-4>.

[15] SAE, 'SAE International Standard J3016: Taxonomy and Definitions for Terms Related to On-Road Motor Vehicle Automated Driving Systems', *Surface Vehicle Recommended Pratice*, 2014.

[16] SAE, 'Taxonomy and Definitions for Terms Related to Driving Automation Systems for On-Road Motor Vehicles'.

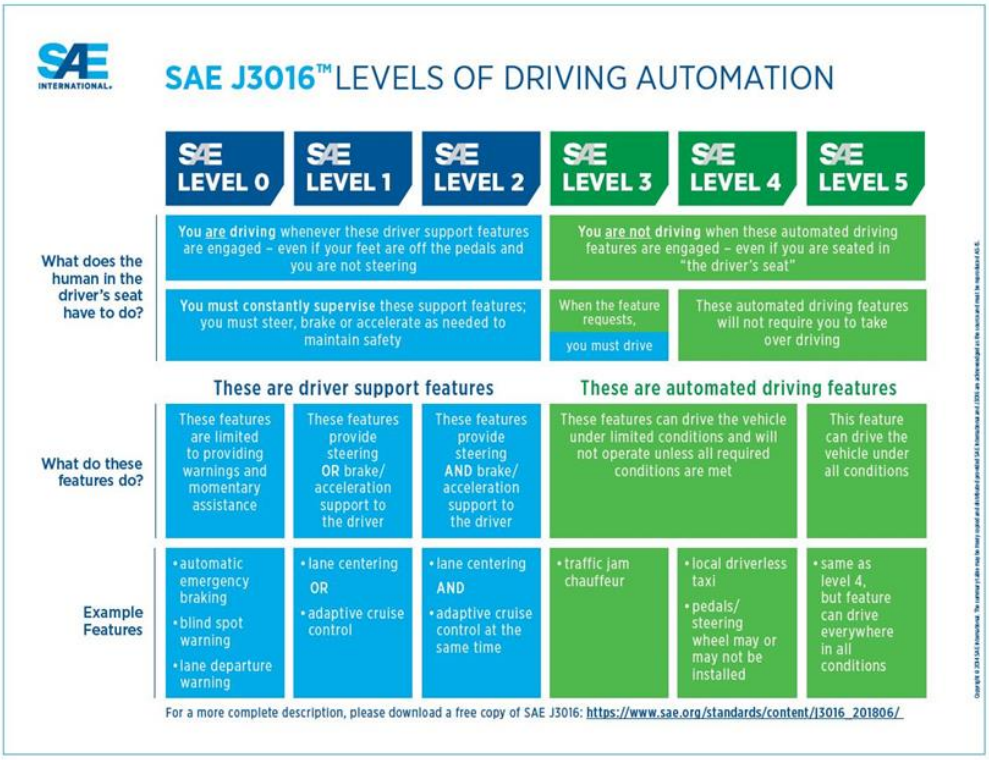

Figure 1: SAE levels of driving automation

This scheme is divided into two parts: levels 0 to 2 indicate more or less advanced DAS systems that never replace the driver, but assist the driver while driving with warnings (level 0, e.g. lane departure warning), with assistance in specific manoeuvres (level 1, e.g. adaptive cruise control) or with devices to assist the driver. adaptive cruise control) or with multiple devices working together to assist in more complex manoeuvres (liv 2, e.g. lane centering and adaptive cruise control together). The levels from 3 to 5 fall into the domain of automated vehicles, replacing the driver altogether in certain contexts that depend on operational design demand (ODD) and may still require a vigilant driver (liv 3, e.g. traffic jam chauffeur), stable but in controlled environments (liv. 4, e.g. local driveless taxi) or full automation that can drive autonomously in any context (liv.5).

Defense advanced research projects agency (DARPA)

Nowadays, the automation level on the market reach at most some level 3 functions. Fully automated driving logic has been under development for several decades. In order to fully replace the driver, an automation logic must make choices, interpret their effects and predict the reactions of the traffic jam to them. Mainly two things make the implementation of these logics complex: the correct interpretation of the environmental and traffic context in which the vehicle is placed, and the correct prediction of the 'intentions' of traffic participants. To stimulate and accelerate the development and demonstration of the effectiveness of driving automation systems, in 2004 DARPA (US Department of Defence) set up the 'Grand Challenge': an annual competition in which the various vehicle manufacturers can display their prototypes in a confrontation with other participants[17]. The first decision making system in the automotive sector dates back to 2007, demonstrated at the DARPA Urban Challenge and based on logical composite systems of finite state machines, decision trees and heuristic methods[18]. As described by Villagra et al. (2017)[19], Intelligent-Decision-Making Systems (IDMS) have evolved in multiple directions, passing through trajectory optimization, cognitive systems[20], agent systems[21], fuzzy systems[22], neural networks[23], evolutionary algorithms[24] and rule-based methods[25].

Testing companies and standards

As already stated, the technical obstacles to the realization of driving automation logics do not relate exclusively to the control logic itself, but to the methodological tools for their verification, control and subsequent validation; in other words, the testing methods.

The currently known testing methods do not provide sufficient knowledge to guarantee the safety of highly automated driving devices[26], nor are there any official

[17] DARPA, 'DARPA Grand Challenge', 2004 <https://archive.darpa.mil/grandchallenge04/overview.htm>.

[18] Chris Urmson and William Whittaker, 'Self-Driving Cars and the Urban Challenge', *IEEE Intelligent Systems*, 23.2 (2008), 66–68 <https://doi.org/10.1109/MIS.2008.34>.

[19] J Villagra and others, *Automated Driving, Intelligent Vehicles: Enabling Technologies and Future Developments*, 2017 <https://doi.org/10.1016/B978-0-12-812800-8.00008-4>.

[20] M Czubenko, Z Kowalczuk, and A Ordys, 'Autonomous Driver Based on an Intelligent System of Decision-Making', *Cognitive Computation*, 7.5 (2015), 569–81 <https://doi.org/10.1007/s12559-015-9320-5>.

[21] Bo Chen and Harry H. Cheng, 'A Review of the Applications of Agent Technology in Traffic and Transportation Systems', *IEEE Transactions on Intelligent Transportation Systems*, 2010, 485–97 <https://doi.org/10.1109/TITS.2010.2048313>.

[22] Rudwan Abdullah and others, 'Autonomous Intelligent Cruise Control Using a Novel Multiple-Controller Framework Incorporating Fuzzy-Logic-Based Switching and Tuning', in *Neurocomputing* (Elsevier, 2008), LXXI, 2727–41 <https://doi.org/10.1016/j.neucom.2007.05.016>.

[23] Long Chen and others, 'Deep Neural Network Based Vehicle and Pedestrian Detection for Autonomous Driving: A Survey', *IEEE Transactions on Intelligent Transportation Systems*, 22.6 (2021), 3234–46 <https://doi.org/10.1109/TITS.2020.2993926>.

[24] Debasri Chakraborty, Warren Vaz, and Arup Kr Nandi, 'Optimal Driving during Electric Vehicle Acceleration Using Evolutionary Algorithms', *Applied Soft Computing Journal*, 34 (2015), 217–35 <https://doi.org/10.1016/j.asoc.2015.04.024>.

[25] Alexander G. Cunningham and others, 'MPDM: Multipolicy Decision-Making in Dynamic, Uncertain Environments for Autonomous Driving', in *Proceedings - IEEE International Conference on Robotics and Automation* (Institute of Electrical and Electronics Engineers Inc., 2015), MMXV-JUNE, 1670–77 <https://doi.org/10.1109/ICRA.2015.7139412>.

[26] Winner, Wachenfeld, and Junietz, 'Safety Assurance for Highly Automated Driving – The PEGASUS Approach' TRB annual meeting 2017

regulations concerning the tests that such control devices must pass in order to be placed on the market. This regulatory gap corresponds to an opportunity for the testing market, and it is currently being filled by the New Car Assessment Programme (in Europe it is called EuroNCAP). The NCAP's are voluntary testing programs aimed at classifying automation devices for effectiveness in specific contexts considered to be more severe (worst cases).

European new car assessment programme (EuroNCAP)

EuroNCAP was established in 1997 with the support of the European Union to provide an objective assessment of the safety performance of vehicles placed on the market. This programme set new safety standards for passive passenger protection and its effect on the automotive market has been very positive[27]. Since 2013, EuroNCAP has also been dealing with active safety systems, introducing tests for AEB, entering the field of driver assistance systems and becoming de facto the only verification and validation company, external to the car manufacturers themselves, for DAS devices. It currently includes tests for the safety of vulnerable users (pedestrians and cyclists) and for some specific use cases in the field of partial automation (cut-in and cut-off on motorways for adaptive cruise control). The analysis method adopted is composed of a series of successive tests (test matrix) that reproduce very specific accident conditions, such as certain types of impact (frontal or small overlap) or particularly risky conditions for specific traffic participants (Vulnerable Road Users programme). These test series reconstruct accident scenarios considered to be very serious and extracted from the most recurrent ones found in the available accident databases. The accident analysis performed on the database is therefore the main tool used to identify the worst cases and carry out an assessment of the significant characteristics of these cases, aimed at their reconstruction in the laboratory, in order to define the specifications of the tests to be performed.

International Organization for standardization (ISO)

On the other hand, some mandatory standard protocols exist and are widely used in industry. These are not testing methods, but protocols act to guarantee some safety aspects. These are implemented by the International Organization for Standardization (ISO), which, in the area we are studying, defined standards and methodologies for the analysis and verification of automation systems and for the assessment and management of

[27] Johan Strandroth and others, 'Correlation Between Euro NCAP Pedestrian Test Results and Injury Severity in Injury Crashes with Pedestrians and Bicyclists in Sweden', *SAE Technical Paper Series*, 1.November 2014 (2018), 213–31 <https://doi.org/10.4271/2014-22-0009>.

accident risk. According to ISO 26262[28], risk is defined as the product of the severity of physical injuries due to an accident (harm) and its probability of occurrence. The standard was created to define terms and procedures for Hazard And Risk Analysis (HARA) and has been updated several times since its last version in 2018. ISO 26262, after a first part with the vocabulary of the terms used, identifies all the logical procedures to structure and perform a proper risk analysis of vehicle automation systems. Menzel et al. (2018)[29], in their analysis of ISO 26262, shows how the standard focuses on the risk strictly related to the components (hardware and software) with which the automation systems are composed (functional safety), neglecting, the elaborations related to the analysis of risks related to use and iteration with the context. i.e. the scenarios in which risks are related to the incorrect or misinterpreted use of automation logics. In scientific literature, the topic is known as Safety Of The Intended Functionality (SOTIF) and in 2019, ISO published the ISO-PAS-21448-2019[30] standard, which focuses precisely on this type of risk, defining the risk issues related to "some systems, which rely on sensing the external or internal environment, there can be potentially hazardous behaviour caused by intended functionality or performance limitation of system that is free from faults addressed in the ISO 26262 series". It also points out that "A proper understanding of the function by the user, its behaviour and its limitations (including the human/machine interface) is the key to ensuring safety". ISO/PAS 21448 covers, in fact, "hazards caused by a potentially hazardous system behaviour (...) both for use cases when the vehicle is correctly used and for use cases when it is incorrectly used in a reasonably foreseeable misuse". Like all ISO standards, ISO 21448 describes a procedure for the analysis through flowcharts, systematizing the approach to the analysis of the risk problems, but without going into detail on the type and method of execution of tests and checks. ISO 21448 distinguishes the functional safety phases from those of SOTIF and groups them in a V-shaped graph (Figure 2), defining their complementarity and linking them to the development and validation processes of the automation logics with a reference to the V-model.

[28] International Organization for Standardization, 'ISO/FDIS 26262 : Road Vehicles-Functional Safety', 2018.

[29] Till Menzel, Gerrit Bagschik, and Markus Maurer, 'Scenarios for Development, Test and Validation of Automated Vehicles', *ArXiv*, Iv, 2018, 1821–27; Johannes Bach, Stefan Otten, and Eric Sax, 'Model Based Scenario Specification for Development and Test of Automated Driving Functions', *IEEE Intelligent Vehicles Symposium, Proceedings*, 2016-Augus.Iv (2016), 1149–55 <https://doi.org/10.1109/IVS.2016.7535534>.

[30] International Organization for Standardization, 'ISO-PAS-21448-2019', 2019.

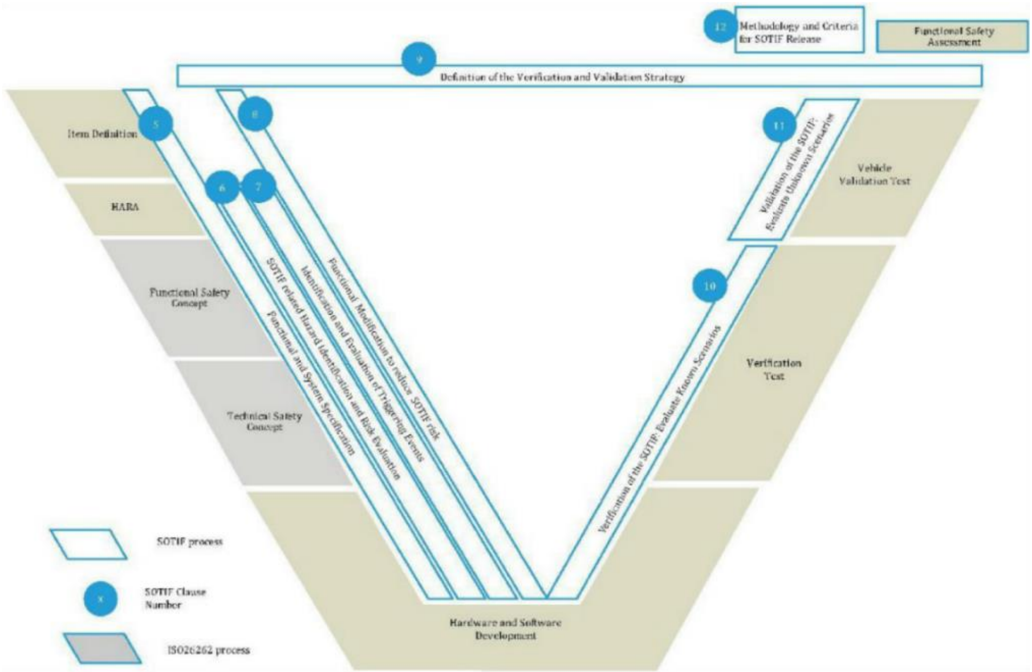

Figura 2: SOTIF, V-model

1.3 The simulation

Shifting from safety devices designed for a very specific use (e.g. seat belts) to driver assistance devices also needs a paradigm shift in testing. Indeed, whereas the former needs to verify the adequacy of the system to physical stresses (easy to reproduce in the laboratory), the latter needs to verify the system's adequacy to logical triggers (flag activation, recognition of scenarios, sensor-fusion efficiency, etc.). For this reason, the development and testing processes are longer and procedures are more complex. The "V" software development model has been applied to vehicles for many years[31], introduced in the MISRA[32] guidelines and definitively became standard through the ISO 26262. It is a procedural scheme that embraces all the phases from design to validation of driver assistance devices. It develops both vertically and horizontally, starting from the specification of the project requirements, arriving at the acceptance tests; crossing, in the two arms of the "V", all the design phases, placed on the left arm, and the validation phases, placed on the right arm. Testing is therefore not only a tool to verify the finished product, but a design tool active during all the device development phases. The use of simulation becomes essential in the realization of driver assistance devices. The V-model also

[31] P Koopman and M Wagner, ‘Challenges in Autonomous Vehicle Testing and Validation’, *SAE International Journal of Transportation Safety*, 4.1 (2016), 15–24 <https://doi.org/10.4271/2016-01-0128>.

[32] Motor Industry Software Reliability Association, ‘Development Guidelines for Vehicle Based Software’, 1994; Motor Industry Software Reliability Association, ‘Report 6: Verification and Validation’, 1995.

specifies the appropriate simulation phases for each stage of the development process. The kind of simulation is specified through a nomenclature that indicates the name of the test object, followed by 'in-the-loop'. This name indicates not only the test object, but also the simulation methods and tools. Typically, the test object represents the 'real' part insert in the virtual environment. As the test object varies, so does the simulated part, changing the entire simulation architecture. Each in-the-loop wording corresponds to different techniques and tools for carrying out the simulations.

Three types of simulation are most frequently used:

- model in-the-loop
- hardware in-the-loop
- driver in-the-loop.

The model in the loop (MIL): the test object is the automation logic in form of a model. In this case, the simulated part is the entire environment including the ego-vehicle, sensors, and all devices. Tests performed in MIL, as a complete simulation, are extremely fast, because they depend on the only software and computational power of the used computer.

The hardware in-the-loop (HIL): the test object is a physical control unit, a group of sensors or, in any case, a hardware on which logics are mounted. This hardware is put in direct communication with the simulation environment through a connection, which can be various, depending on the nature of the test object or the connection technology to be adopted/tested. Simulation software simulate the environment and the ego-vehicle through a *real-time-hardware* specifically designed to ensure that simulation time is equal to the real time. These simulations needs longer time than MIL simulations, because the processes must take place in real time to ensure a realistic simulation of the signals sent to the control unit under test, which has real, not simulated, processing times. The advantages of this type of simulation are the extremely low cost and time compared to road tests, and the immediate possibility of changing simulation scenarios or running long series of tests without interruption.

Driver-in-the-loop (DIL): the test involve human driver sitting in a simulator designed to be as similar as possible to the cabin of a vehicle. In this kind of simulations, the virtual environment represents the road and traffic context and the dynamics of the ego-vehicle. The vehicle dynamics may or may not be connected to a mechanical system that gives a realistic physical feedback on the vehicle/cabin in which the driver sits. These simulations requires to be performed in real time, as HIL simulations. The costs and simulation time of DIL are greater than HIL due to both hardware requirements and the need to test a sufficient population of drivers to obtain statistically significant results.

There are many other types of simulation, less used or more recent, such as:

Scenario in-the-loop (ScHL), in which a self-driving vehicle is driven on the road, receiving simulated signals through its sensor system. The simulated environment concerns only the context of traffic and signalling, while the vehicle travels in the real road. This kind of simulation requires very high costs and a purpose-built test track.

Vehicle in-the-loop (VIL), in which the real vehicle is mounted on a mechanical device that simulates the physical interactions between the vehicle and the simulated environment. This kind of simulator is very expensive and requires a prototype vehicle to be built.

All these kinds of simulators respond to various testing needs at different levels of detail or simulation. The need for the use of simulation stems from the impossibility of obtaining an exhaustive evaluation of the logic of driving automation through tests carried out exclusively on the road, in real traffic contexts, ensuring adequate execution times and guaranteeing minimum safety measures. Testing on public roads open to traffic is still the final stage of a development process, but it is reduced to the minimum.

1.4 Approaches

Simulations are based on the representation of a scenario.

The scenario approach is not the only one known in the literature. Li et al. (2016) [33] distinguishes testing into two approaches: functional and scenario-based

Functional testing serves the purpose of timely verification of the correct functioning of an elementary component of a logic. E.g., a system of lane departure warning, which has the purpose of signaling to the driver when the vehicle crosses the carriageway lines, can be verified by sending a vehicle on purpose beyond the lines and verifying that the device activates the signaling systems or sends the signal (flag) for their activation. This kind of tests can only take place at the end of the whole process of making the devices and for a limited number of cases. The increase in the complexity of automation logics and the increase in the number of use cases in which they must be correctly activated make it

[33] L Li and others, 'Intelligence Testing for Autonomous Vehicles: A New Approach', *IEEE Transactions on Intelligent Vehicles*, 1.2 (2016), 158–64 <https://doi.org/10.1109/TIV.2016.2608003>.

essential to verify the behavior of these logics in a very wide range of cases. This makes impossible to verify punctually with functional tests. It is here that a more complex logical process of scenario analysis is required in order to enlarge the control on the events that can actually occur on the road.

The scenario tool was created with the aim of reconstructing complex events, involving a multiplicity of factors, to study their possible consequences on the test objects. Scenario reconstruction as a tool for knowledge, planning and risk management is widely used in many fields of study (from business risk to seismic risk), but in the last few decades it has become a fundamental tool in the automotive field, with regard to the design of driving automation devices. Go and Carroll (2004) [34] are among the first to systematize the theory of scenarios (which was already widespread), and to explore its use in engineering and human-computer interface. According to the authors, a scenario is a description that contains (1) actors, (2) background information and assumptions, (3) goals and expectations and (4) a sequence of events and actions. In a table explaining the constituent factors of scenarios, some "uncertain factors" are highlighted (Table 1).

Table 1: Go and Carroll, 2004, uncertain factor

COMMUNITY	UNCERTAIN FACTOR	GOAL OF SCENARIO-BASED APPROACH	DEVELOPMENT PROCESS	VIEW POINT	BACKGROUND GOAL OF COMMUNITY
Strategic Planning	Environment	List "what-if" questions and their answers	Iterative	Organization Technological changes Economics Social, political regulations Consumer attitudes	Plan a course of actions
Human-Computer Interaction	Use of system	Envision user requirements of (future) system use	Iterative Prototyping	Human Usability Cognition Emotion	Describe use of (future) systems Design usable computer system
Requirements Engineering	System requirements Functionality	Acquire user requirements and specify them	Waterfall Spiral	System architecture Development process	Specify systems Provide a good transition to the next development phase
Object-Oriented Analysis/ Design	Objects Data structures Class hierarchy	Identify objects, data structures and model class hierarchy	Iterative Incremental	System Object	Design a model of world

This element is central to scenario building and is in fact addressed by every author who uses it. Rosson and Carroll (2002)[35] had already defined a scenario as "simply a story about people carrying out an activity", but it was only in 2014 that a unified ontology for scenarios in the context of driver automation was developed[36]. Through the theater

[34] K. Go and John M Carroll, 'The Blind Men and the Elephant: Views of Scenario-Based System Design', *Interactions*, 2004.

[35] Mary Beth Rosson and John M. Carroll, *Usability Engineering: Scenario-Based Development of Human-Computer Interaction*, 2002.

[36] Sebastian Geyer and others, 'Concept and Development of a Unified Ontology for Generating Test and Use-Case Catalogues for Assisted and Automated Vehicle Guidance', *IET Intelligent Transport Systems*, 8.3 (2014), 183–89 <https://doi.org/10.1049/iet-its.2012.0188>.

metaphor, all the elements that make up the scenario and incorporate it are distinguished and explored in depth (Figure 3).

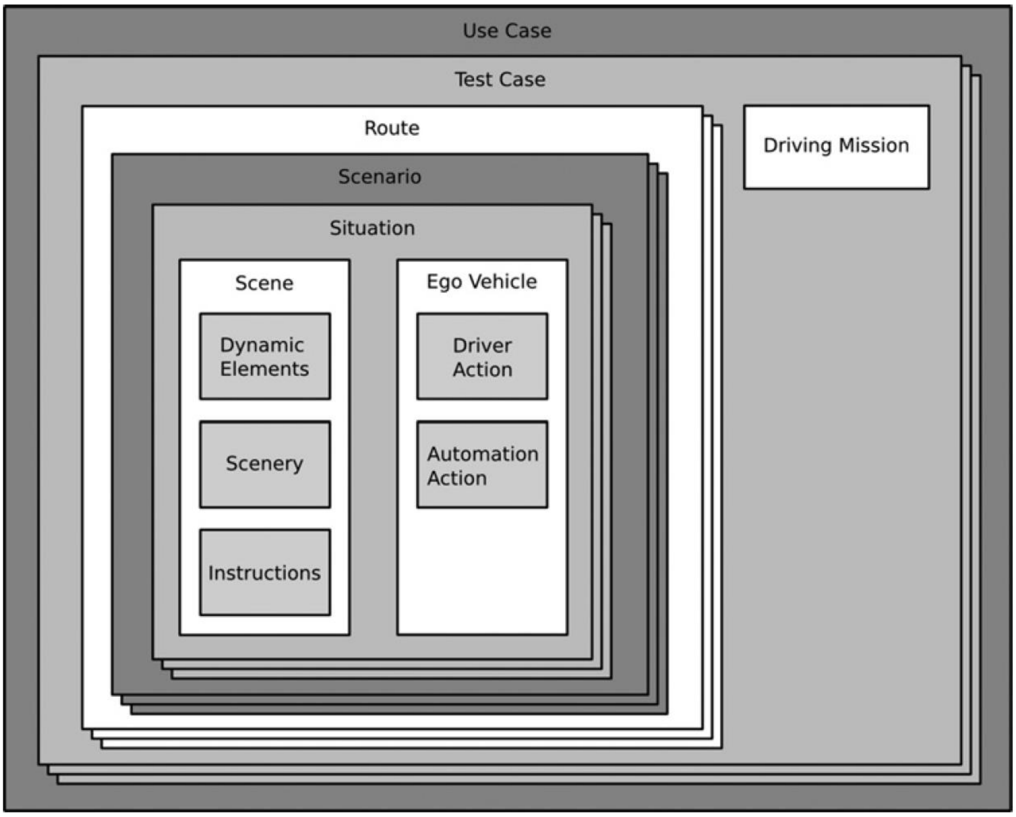

Figure 3: Geyer et al. hierarchy scheme from use-case to the elements composing a scenario

The Scenario is composed of a succession of Situations (the atomic element that characterizes an event: e.g. the Ego-Vehicle turns right, the blue car brakes, etc.), composed of Scene and Ego-Vehicle. The Scene is composed of Dynamic Elements, defined as elements whose state changes within the scene (with the exclusion of the Ego-Vehicle), Scenery, the structured set of the single static elements that constitute the context on which the scene takes place, and Instructions, optional characteristics of actors and the instructions they have to follow. Finally, the Ego-Vehicle is composed of its own actions, which can be originated by the driver or by the automation. The article also describes the elements of Driving mission and Route. The first constitutes the main objective or expectation addressed to the Ego-vehicle during the performance of the situations. The Route, on the other hand, describes the combination of all planned scenarios and the transitions between a scenario and another. This structure of interconnected elements describing the scenarios was later deepened in Ulbrich et al. (2015)[37], where the meanings of Situation, Scene and all their components are better specified. In the article the definitions of "scene" are studied in multiple authors of scientific literature, from Maurer (2000)[38] to Geyer (2014)[39] and Reschka (2015)[40]; concluding that: "The authors share the opinion, that a scene does not only cover environment aspects, but also the aspect of a self-representation.". The Scene is here described as a "Snapshot" of the environmental state (boundary conditions) and its self-representation. Moreover, the authors better detailed elements as the Dynamic elements, the Scenery and the Instruction, which the authors define as self-representations of actors and observers,.

Dynamic elements are no longer associated with the generic change of state, but with the act of moving in the space. Dynamic elements are, therefore, all the actors of the scenario that can move, whether or not they do so during the development of the scenario. Elements such as traffic lights are not dynamic elements because, although they can change state during the scenario, they are not able to move. This definition transforms the meaning of dynamic elements from that related to the change of information to all those elements that provide the scenario with the component of uncertainty related to the maneuver choices. This paradigm shift defines precisely which factor of uncertainty is being referred to Dynamic elements. The same type of clarification is carried out for static elements, identified by the scenery. Therefore, all elements whose information does not refer to the sphere of movement belong to the scenery. All infrastructural elements of the road, furniture, fixed and mobile, horizontal and vertical signs and traffic lights belong to the scenery. Furthermore, the meaning and necessity of the third component of the scenes is specified as Self-representations of actors and observers. It is fundamental and represents the relationships and interactions between all the elements of the scenario such as fields of view, occlusions, attributes and states of actors and observers (Figure 4). The authors also deepened the term Situation, highlighting a strong difference in the use of the term among the multiple authors in the literature analyzed. They suggest a broad and complex definition which is given in full here:

"A situation is the set of circumstances that must be considered in selecting an appropriate pattern of behaviour at a particular time. It includes all relevant conditions, options and determinants of behavior. A situation is derived from the scene by a process

[37] Simon Ulbrich and others, ‘Defining and Substantiating the Terms Scene, Situation, and Scenario for Automated Driving’, *IEEE Conference on Intelligent Transportation Systems, Proceedings, ITSC*, 2015-Octob (2015), 982–88 <https://doi.org/10.1109/ITSC.2015.164>.

[38] M. Maurer, ‘EMS-Vision: Knowledge Representation for Flexible Automation of Land Vehicles’, in *IEEE Intelligent Vehicles Symposium, Proceedings*, 2000, pp. 575–80 <https://doi.org/10.1109/ivs.2000.898409>.

[39] Geyer and others.

[40] Andreas Reschka and others, ‘Ability and Skill Graphs for System Modeling , Online Monitoring , and Decision Support for Vehicle Guidance Systems’, Iv, 2015.

of selection and augmentation of information based on transient (e.g. mission-specific) and permanent goals and values. Hence, a situation is always subjective, representing the point of view of the element" (Figure 6).

To summarize in a simplified way the image provided by the article which represents the relationship between scenario, use-case and scene (Figure 5).

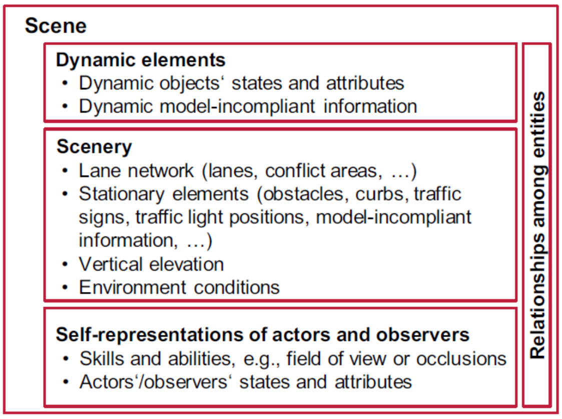

Figure 4: Ulbrich et al. Scene scheme

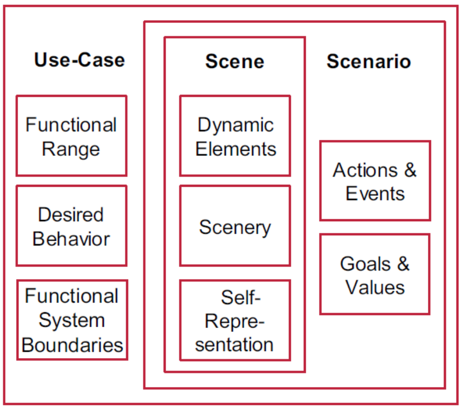

Figure 5: Ulbrich et al. elements hierarchy

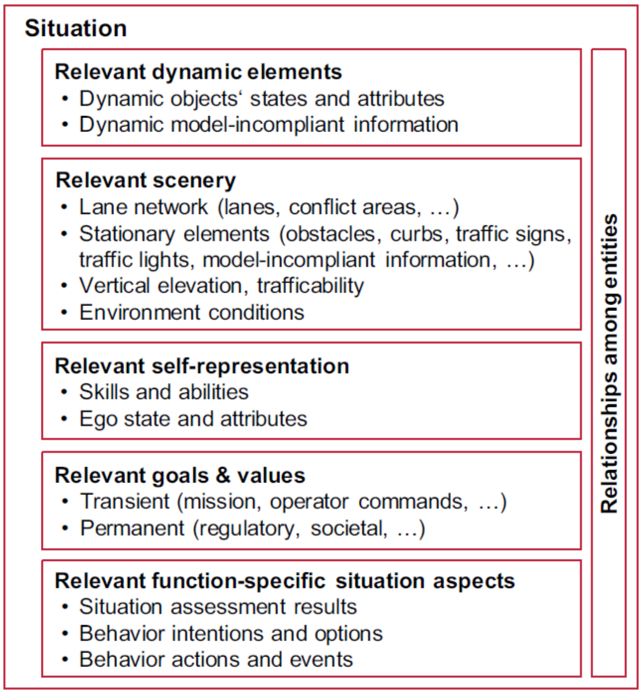

Figure 6: Ulbrich et al. Situation

1.5 Making scenarios

In order to build a scenario, it is necessary a detailed definition of all its components. In the literature, several approaches can be found to test automation logics. Statistical databases are frequently used as a basis to identify the characteristics of scenarios to be tested. In the Op den Camp et al. project (2017) [41], to perform the accident analysis, an extensive use of in-depth databases from France[42], Germany, Italy[43], Netherlands[44], Sweden[45] and UK[46] was done.

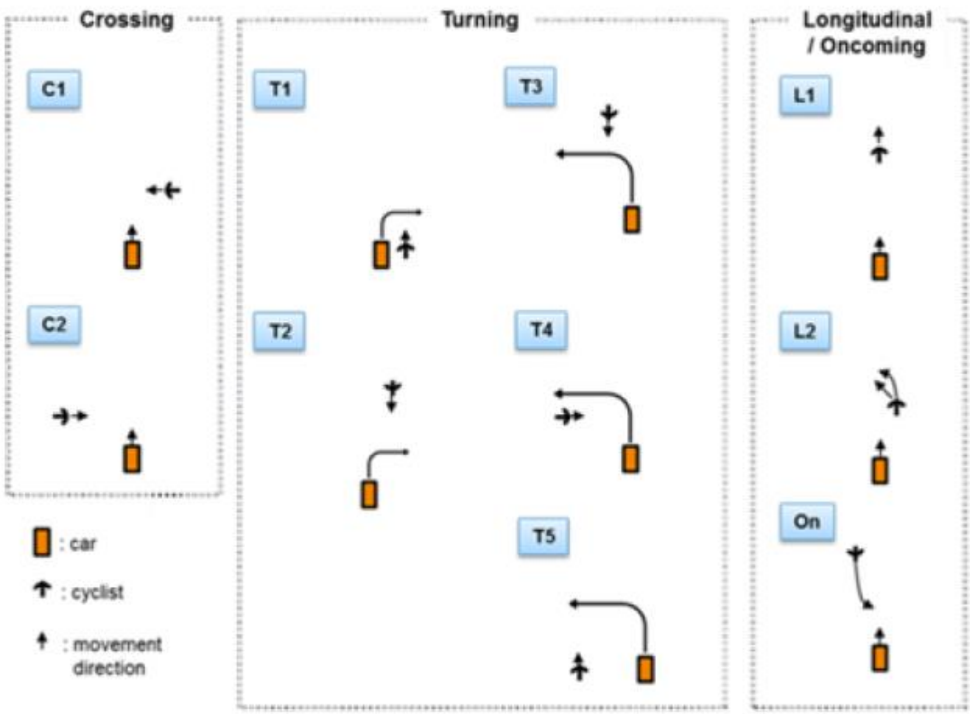

Figure 7: Op den Camp et al. maneuver and conflict scheme

The method mentioned by Op den Camp et al. (2017) [47] was also used for the project Assessment methodologies for forward looking Integrated Pedestrian and further extension to Cyclists Safety Systems Framework Programme (FP7 AsPeCSS) for AEB

[41] Olaf Op den Camp and others, 'Cyclist Target and Test Setup for Evaluation of Cyclist-Autonomous Emergency Braking', *International Journal of Automotive Technology*, 2017 <https://doi.org/10.1007/s12239-017-0106-5>.

[42] Yves Page and others, 'A Comprehensive Overview of the Frequency and the Severity of Injuries Sustained by Car Occupants and Subsequent Implications in Terms of Injury Prevention', *Annals of Advances in Automotive Medicine. Association for the Advancement of Automotive Medicine*, 2012.

[43] ISTAT, 'Year 2012 Road Accidents in Italy', 2012.

[44] BRON: Netherlands National Road and Register, 'Police Registered Numbers of Casualties, Drivers and Crashes', 2016 <http://www.swov.nl>.

[45] M Skyving, 'STRADA: Road Traffic Accident and Injury Data in Sweden', in *24th World Int. Traffic Medicine Association Congress*, 2015.

[46] Department for Transport, 'Reported Road Casualties Great Britain: 2015 Annual Report', 2016.

[47] Op den Camp and others, 'Cyclist target and test setup for evaluation of cyclist-autonomous emergency braking', *International Journal of Automotive Technology*, 2017.

effectiveness testing in pedestrian accident scenarios[48]. It firstly uses the abstract analysis of the maneuver's schemes involving a conflict between the actors involved in the accident (Figure 7). These are then compared with the databases to identify the most recurring ones (Table 2) and, basing on this data, the lowest number of scenarios representing the highest percentage of coverage of fatal or serious injury accidents are chosen.

Table 2: Op den Camp et al. mortal and heavy accident distribution

	Scenario	Weighted average K	Weighted average SI
Cyclist crossing	C1	25 %	28 %
	C2	29 %	28 %
Car turning	T1	0 %	3 %
	T2	0 %	1 %
	T3	2 %	5 %
	T4	1 %	2 %
	T5	2 %	3 %
Longitudinal	L1 + L2	24 %	7 %
On-coming	On	8 %	6 %
-	Remaining	10 %	16 %

The principle adopted is identifying the worst-cases. The system under test (SUT) will be tested on those cases that are considered worst-case and represent the highest percentage of serious accidents (with serious injuries or death). If the automation logic would be successful in these cases, it would reduce the largest risk according to the known accident statistics.

Once the most significant and recurrent cases have been identified, they are detailed by assuming the most recurrent values of the characteristic parameters. Op den Camp (2017)[49] dedicates a chapter to the parameters for car-to-cyclist scenarios extracted from databases. E.g., it is considered that 80% of accidents occur without rainfall, so it is assumed that rainfall is not an aggravating condition and therefore it is not necessary to

[48] Marcus Wisch and others, 'European Project AsPeCSS - Interim Result: Development of Test Scenarios Based on Identified Accident Scenarios', 2013.

[49] Op den Camp and others, 'Cyclist target and test setup for evaluation of cyclist-autonomous emergency braking', *International Journal of Automotive Technology*, 2017.

evaluate rainfall scenarios. The same applies to light conditions, because 75% ~ 90% of serious accidents occur in broad daylight.

This approach to worst-case bases only on recurrence, but there is no scientific basis for the association between recurrence and severity. For example, the drastic reduction in the number of bicycle accidents in the rain scenario compared to the non-rain scenario could be attributed to the fact that rain has a strong deterrent effect on the use of bicycles. In this case, the ratio between the number of cases would not give any significant information on the correlation between the rain factor and the accident risk, so we cannot be sure to find the real worst-cases using this approach.

Another approach widely used in scientific literature is the collection and analysis of data coming from Field Operational Tests (FOT)[50]. In the last decade, several European FOT[51] projects have been carried out, collecting data from vehicles equipped with on board units that harvest driving data or from roadside devices in the context of analysis of logics and communication systems between vehicles and vehicles or infrastructures (V2X). FOT projects have been carried out worldwide through various initiatives[52], defining numerous databases of driving behavior in real conditions. These databases can be used to reconstruct realistic driving situations and are often used as a basis for running simulation tests[53]. This kind of data is often used to perform MIL simulations with the montecarlo method. The random extraction of the boundary conditions of a scenario ensures that the scenario is explored without any possible accident analysis bias.

The FOT data, however, suffers the problem of rareness of accident data. It is estimated that on average a driver is involved in a serious accident every 38 years and a fatal accident every 6877 years[54]. The rarity of accident data is a well-known problem in the literature, and statistical analyses therefore suffer from a very large error due to a lack of confidence in the data. Zhao et al. (2016, 2017)[55] point out that simulations performed by Monte Carlo exploration are very time-consuming. Indeed, without an extraction rule focused on accident cases, the simulations performed have an extremely low probability to verify accident scenarios. In the paper, an accelerated evaluation is proposed based on

[50] M Aust, 'Evaluation Process for Active Safety Functions: Addressing Key Challenges in Functional, Formative Evaluation of Advanced Driver Assistance Systems', *Chalmers University of Technology*, 2012.

[51] Y Barnard and others, 'Methodology for Field Operational Tests of Automated Vehicles', in *Transportation Research Procedia*, 2016, XIV, 2188–96 <https://doi.org/10.1016/j.trpro.2016.05.234>.

[52] Devonshire LeBlanc, D., Sayer, J., Winkler, C., Ervin, R., Bogard, S. and T. J. Mefford, M., Hagan, M., Bareket, Z., Goodsell, R., and Gordon, *Road Departure Crash Warning System Field Operational Test: Methodology and Results. Volume 1: Technical Report*, 2006.

[53] Andreas Kendziorra, Peter Wagner, and Tomer Toledo, 'A Stochastic Car Following Model', *Transportation Research Procedia*, 15 (2016), 198–207 <https://doi.org/10.1016/j.trpro.2016.06.017>.

[54] NHTSA US Department of Transportation, *Facts, T. S. 2012*, 2012.

[55] Ding Zhao, 'Accelerated Evaluation of Automated Vehicles', 2016; Ding Zhao, Henry Lam, and others, 'Accelerated Evaluation of Automated Vehicles Safety in Lane-Change Scenarios Based on Importance Sampling Techniques', *IEEE Transactions on Intelligent Transportation Systems*, 18.3 (2017), 595–607 <https://doi.org/10.1109/TITS.2016.2582208>.

running random simulations of scenarios characterised by dangerous driving behaviour, extracted from the FOT databases through the application of importance sampling.

We underline that a scenario based on databases can only refer to known and recorded conditions. Winner et al. (2017)[56] show how referring to databases for accident risk scenarios involving vehicles with a high level of automation produces a dangerous false security. Databases refers to human driving conditions and do not contain within them novel conditions such as automated driving. This problem produces a dilemma known as the 'dark matter problem' [57], whereby the study of a problem implies knowledge of the problem itself. If a problem is unknown, it cannot be analysed except from an abstract and theoretical point of view.

The construction of testing scenarios, as we told, involves the use-case identification[58]. Once the use-case is known, it is possible to build scenarios even without referring to accident databases or FOT databases. In the case of very specific activities, the variables to be considered in the design of the scenario are limited. As the complexity of the use-case arise, so the contexts and variables that makes up the scenarios increases. Wachenfeld et al. (2016)[59] analyses use-cases where the autonomous vehicle stops working and requires human intervention to complete a specific mission (e.g. clear the roadway of the vehicle). The MEDIATOR[60] project analyses use-cases from an in-vehicle perspective, i.e. interaction between human and autonomous drivers. These use-cases represent novel driving condition, so they cannot even indirectly refer to databases and analyses collected on the road.

One of the most recent approaches in the literature involves an integrated simulation of traffic microsimulation and vehicle dynamics, reconstructing an entire road network in which the VUT runs for an indefinite time. The aim of this approach is to not predetermine the use-case, allowing all possible scenarios to occur during the simulation, leaving events to 'chance'. In this way, the principle of not basing testing on data collected in inhomogeneous contexts is respected, through the complete randomization of the events tested. Unfortunately, this method does not offer any certainty on the exhaustiveness of the tests, as it does not offer any reference on how many and which scenarios need to be addressed in order to consider the test completed.

---

[56] [56] Winner, Wachenfeld, and Junietz. 'Safety Assurance for Highly Automated Driving – The PEGASUS Approach' TRB annual meeting 2017

[58] Geyer and others, 'Concept and development of a unified ontology for generating test and use-case catalogues for assisted and automated vehicle guidance' *IET Intelligent Transport Systems*, 2014 ; Ulbrich and others, ' Defining and Substantiating the Terms Scene, Situation, and Scenario for Automated Driving', *IEEE Conference on Intelligent Transportation Systems, Proceedings, ITSC*, 2015.

[59] Wachenfeld and others, ' Use Cases for Autonomous Driving', *Autonomous Driving: Technical, Legal and Social Aspects*, 2016.

[60] 'MEDIATOR behind the Scenes: Use Cases | Mediator'.

On the other hand, defining the safety of a vehicle, a difficult question arises: what does 'safe' mean?

The safety of a device depends on the purposes for which it is made. Cars on the market today are a system of safety systems ranging from passive systems such as air bags and seat belts to active systems such as anti-lock braking systems (ABS) and emergency braking (AEB), but earlier vehicles were considered safe even without these devices. Defining "safe" the incorporation of a new device into the vehicle system means ensuring that this incorporation does not in itself constitute a new source of injury or damage to users and that it does not lead to a deterioration in comfort conditions. When the new device tends to replace the driver (partial or total automation), the disrupted system includes the entire driving activity and its consequences in traffic. In this case, the source of injury relates directly to the occurrence of road accidents. From this perspective, to prove that an automation system is safe means to prove that it does not produce more accidents than would occur without the new device. Kalra and Paddock (2016)[61] estimated that one fatal accident occurs every 100 million miles and estimate that to achieve 95% statistical confidence to claim the safety of autonomous vehicles, a fleet of 100 autonomous vehicles driven 24/7 for about 400 years would be needed. For that reason, simulation can make a great contribution in speeding up testing, ensuring that testing mileage targets are met in reasonable time. On the other hand, simulation does not guarantee the same degree of realism and randomness and variety of events as can be found in the real world. It would therefore be necessary to identify a *simulation equivalence ratio,* i.e. a ratio of equivalence between kilometres travelled in simulation and real-life kilometres. In that way, a simulated autonomous vehicle would be proved to have the same level of safety of a real vehicle, if it had at most the same number and severity of accidents as vehicles already on the market, after a number of test kilometres equal to that required in road tests to achieve 95% confidence multiplied by the simulation equivalence ratio.

Some studies, however, demonstrated the inadequacy of the statistical kilometre reference as a means of comparison between self-driving and human-driven vehicles[62]. Indeed, it seems immediately obvious that the amount of simulation kilometres required to demonstrate the efficiency of a vehicle is enormous. Moreover, ensuring that the set of simulations implemented to achieve the required mileage has a realistic correspondence

[61] Nidhi Kalra and Susan M. Paddock, 'Driving to Safety: How Many Miles of Driving Would It Take to Demonstrate Autonomous Vehicle Reliability?', *Transportation Research Part A: Policy and Practice*, 94 (2016), 182–93 <https://doi.org/10.1016/j.tra.2016.09.010>.

[62] Christian Amersbach and Hermann Winner, 'Defining Required and Feasible Test Coverage for Scenario-Based Validation of Highly Automated Vehicles', *2019 IEEE Intelligent Transportation Systems Conference, ITSC 2019*, 2019, 425–30 <https://doi.org/10.1109/ITSC.2019.8917534>.

with all possible driving situations that may occur on the road remains an extremely complex task.

1.6 Summary of literature known methods and unresolved problems

One of the most discussed problems in the literature regarding the exhaustiveness of tests is how quantify the use-cases to be tested. Testing an automation logic requires the reconstruction of scenarios representing vehicles interactions within a road context. How many and which scenarios should be tested is the first question that still remains unanswered in literature. Indeed, it seems evident that the number of scenarios is infinite because it is characterized by a high number of parameters that often vary continuously and can therefore take on infinite values. The question of what can make a series of tests exhaustive seems analytically unresolvable. For these reasons, the scientific community is divided into various schools of thought. Rather than exhaustiveness, the general used approach to the problem is looking at statistical significance within an acceptable confidence. As we saw, the approaches are many. Continuous testing in complex virtual environments or on the road, monitored with on-board instrumentation (OBU) [63] has the advantage of allowing unexpected conditions to occur and identify and solve logic bugs in a process of continuous refinement. However, this approach cannot demonstrate the safety of the logics in the timeframe required by the market. Naturalistic field operational tests[64] give real data in real conditions, resulting in an excellent source of information, but they do not constitute a validation tool for instrumented vehicles, because they suffer from the scarcity of accidental data, which determines unsustainable testing times. Test matrixes[65] only allow the improvement of systems in known conditions, which however are not at all generalizable and do not provide any information on unknown cases, making them the least effective tool to obtain statistically relevant results, because of the small number of scenarios that can be tested. The accelerated evaluations suggested by D. Zhao allow, by means of MIL simulations with random extraction[66], a very efficient exploration of the reconstructed use-cases, but it requires large FOT databases and applies to one use-case at a time, thus exploring only the use-cases known upstream. Each of these tools solves some problems without solving the whole problem. Maurer et al. (2016)[67] identifies and

[63] Hermann Winner and Walther Wachenfeld, 'Virtual Assessment of Automation in Field Operation A New Runtime Validation Method', *10. Workshop Fahrerassistenzsysteme*, 2015, 161–70 <http://tuprints.ulb.tu-darmstadt.de/5192/1/FAS_2015_Tagungsband.pdf#page=169>.

[64] LeBlanc, D., Sayer, J., Winkler, C., Ervin, R., Bogard, S. and J. Mefford, M., Hagan, M., Bareket, Z., Goodsell, R., and Gordon.

[65] Jeroen Uittenbogaard, Olaf Op den Camp, and Sjef van Montfort, 'CATS Deliverable 5.1: CATS Verification of Test Matrix and Protocol', 2016 <www.TNO.nl/CATS>.

[66] Ding Zhao, Xianan Huang, and others, 'Accelerated Evaluation of Automated Vehicles in Car-Following Maneuvers', 19.3 (2018), 733–44.

[67] Markus Maurer and others, *Autonomous Driving: Technical, Legal and Social Aspects*, *Autonomous Driving: Technical, Legal and Social Aspects*, 2016 <https://doi.org/10.1007/978-3-662-48847-8>.

discusses some of these problems and defines that, to be valid, a method must always start from the identification of actual worst cases. This concept, already addressed when we discussed the EuroNCAP method, is here explored in detail. The concept of worst case does not refer to the statistically most recurrent worst accident cases, but is a concept related to the SUT. A worst case is the scenario in which the device is likely to produce the greatest number of unexpected events with the worst effects. It is evident that these effects are linked to the specific logic and the contexts in which it will be applied, so that worst cases must be analyzed starting from the logic and context, instead of using the recurrence of data that are not dependent on them. Moreover, a strong standardization of test scenarios pushes the design of logics in the direction of overcoming the known test, instead of improving the system safety on unknown cases. The pesticide paradox[68] would thus arise, whereby the evolution of logics would tend to make them 'immune' to known tests, which would no longer provide any information on their efficiency.

[68] Boris Beizer, *Black-Box Testing: Techniques for Functional Testing of Software and Systems* (John Wiley & Sons, Inc., 1995); Philip Koopman and Michael Wagner, 'Toward a Framework for Highly Automated Vehicle Safety Validation', *SAE Technical Paper Series*, 1 (2018), 1–12 <https://doi.org/10.4271/2018-01-1071>.

Cap.2: The proposed approach

2.1 *An impersonal approach to evaluate all the unknown scenarios*

The state of the art shows how difficult to solve the problem of testing automation logics is. This difficulty due mainly to the large and complex variety of conditions to be tested, which results analytically in an infinite number of scenarios. The problem of testing requires a systematic approach to identify the real worst-cases in which the logic under test may occur[69], paying attention to be not affected by the false confidence provided by statistical databases collected in contexts other than those that should be tested[70]. It is necessary to reorganize the known information provided by the state of the art, in order to elaborate a methodological tool that takes into account these critical points, to use the known available tools in a new framework. The first problem to face is how to identify the scenarios to be tested. The number and complexity of these will indirectly define the possible testing tools (road, simulated, simulation method, etc.).

The effort we made is to approach the problem trying to break free from the consolidated categories built in the testing sector, which, although developed due to the evolution of technology, are sometimes a limit for a holistic approach that embraces the problem of testing such a disruptive technology as autonomous driving.

In this chapter, we will try to unravel the problem of the number of scenarios, providing a tool that performs the dual function of identifying useful scenarios for testing and limiting the testing space by identifying the limits of scenario variation.

In the literature, as told in the previous chapter, the scenario comes always from a use-case[71]. In this way, the use-case is a design choice made before the testing scenario design. For this reason, literature studies hardly focus on the analysis of the possible use-cases of a device, because use-cases are assumed to be the purpose of the device. This view is consistent with some ADAS designed for very specific purposes, such as, e.g., the autonomous emergency braking (AEB) designed to reduce the mortality of vulnerable users. For more complex devices, performing multiple functions in complex scenarios, this use-case approach may even be misleading. High levels of automation can lead to an

[69] Philip Koopman and Wagner, ‘Toward a Framework for Highly Automated Vehicle Safety Validation’.

[70] Winner, Wachenfeld, and Junietz‘Safety Assurance for Highly Automated Driving – The PEGASUS Approach’ TRB annual meeting 2017

[71] Geyer and others, 'Concept and development of a unified ontology for generating test and use-case catalogues for assisted and automated vehicle guidance' *IET Intelligent Transport Systems*, 2014; Ulbrich and others, ' Defining and Substantiating the Terms Scene, Situation, and Scenario for Automated Driving', *IEEE Conference on Intelligent Transportation Systems, Proceedings, ITSC*, 2015.

improper use or a reasonably foreseeable misuse[72] and, therefore, the use-cases to be tested should be the result of an analysis downstream of the device definition instead of upstream. Defining the use-case upstream of the process of elaborating test scenarios leads to an arbitrary choice of tests which, in the case of high levels of automation, compromises their exhaustiveness. It is therefore considered necessary to elaborate the series of tests starting from an in-depth analysis of the contexts that the vehicle/logic must face and to determine the use-cases descending from the logic.

This change of perspective may seem complex and difficult to manage, especially since it disrupts the typical scenario design process, right at the initial stage. However, we want to emphasize that the scenario / use-case dependency is univocal. In fact, even use-cases are limited by the context in which the tests take place. What a vehicle, a driver or an automation system can accomplish is strictly dependent on the context in which the vehicle is placed, the limits of the vehicle/device, the limits of the infrastructure, the participants of the traffic, and so on. All these elements on which the use-case depends are enclosed in two spheres: the scene and the ego-vehicle. Once these two elements are defined, the compatible use-case is determined.

In order to elaborate a method for a complete testing of high level of automation, in this research it is proposed not to define the use-cases a priori, but to identify all the possible contexts that the test object is likely to encounter, and then study the behaviour of the SUT varying the boundary conditions.

Considering that the objective of the proposed method is the construction of a solid system for testing, whether it is used in the design phase or in the validation phase, we will assume that the System Under Test is known. In the following paragraphs we will focus on the part related to the definition of the scene, its static and dynamic components, their use in defining test scenarios, the resulting scenarios and their applications.

2.2 A top-down approach for the analysis of road casistic, based on the road regulations and technical law

First of all, we should be remark that this study deals with vehicles used for on roads travelling. Despite the complex set of variables and boundary conditions is analytically infinite, the variability of the road environment is closely linked to and limited by the road regulations and the technical law. From that perspective, it is evident that the road context can be parameterized and distinguished into a finite number of categories.

[72] International Organization for Standardization, 'ISO-PAS-21448-2019'.

A first classification of roads by homogeneous categories is made by the "nuovo codice della strada"[73], which in Title I Article 2 gives the definition of a road and lists the types by distinguishing them in:

A. AUTOSTRADA: a suburban or urban road with independent carriageways or separated by an impassable dividing strip, each with at least two lanes, a paved left-hand lane if necessary and an emergency lane or paved right-hand lane. It has fencing and user assistance systems along the entire route and no intersections or private accesses. It is reserved for the circulation of certain categories of motor vehicles and marked by appropriate start and end signs. It must be equipped with service areas and parking areas, both equipped with accesses with deceleration and acceleration lanes.

B. STRADA EXTRAURBANA PRINCIPALE: a road with independent carriageways or separated by an impassable divider, each with at least two lanes and a paved right-hand side, marked by appropriate starting and ending signs, without level intersections. It can have access to the side properties. It is reserved for the circulation of certain categories of motor vehicles. Appropriate spaces must be provided for any other categories of users. It must be equipped with special service areas, including spaces for parking, equipped with accesses with deceleration and acceleration lanes.

C. STRADA EXTRAURBANA SECONDARIA: a single carriageway road with at least one lane in each direction and platforms.

D. STRADA URBANA DI SCORRIMENTO: road with independent carriageways or separated by a central reservation, each with at least two lanes of traffic, and a possible lane reserved for public transport, paved right and pavements. Any intersections is equipped with traffic lights. for parking are provided special areas or side strips outside the carriageway, both with inputs and outputs concentrated.

E. STRADA URBANA DI QUARTIERE: road with a single carriageway with at least two lanes, paved platforms and pavements. For parking there are areas equipped with a special maneuvering lane, outside the carriageway.

E. bis - STRADA URBANA CICLABILE: urban road with a single carriageway, paved banks, and pavements, with a speed limit not exceeding

[73] 'Nuovo Codice Della Strada' (decreto legisl. 30 aprile 1992 n. 285 e successive modificazioni).

30 km / h, defined by appropriate vertical and horizontal signs, with priority for bicycles.

F. STRADA LOCALE: an urban or suburban road suitably laid out for the purposes of paragraph 1 and not forming part of the other types of roads.

F. Bis - ITINERARIO CICLOPEDONALE: local road, urban, suburban or vicinal, intended mainly for pedestrian and cycle traffic and characterized by an intrinsic safety to protect the weak users of the road. Each of these categories is described in detail with a description of the general characteristics and minimum requirements, and speed limits, vehicles allowed and not allowed to circulate, and other characteristics are identified.

The DM 11-05-2001[74] details the technical and geometrical requirements, identifies the traffic categories, defines the constituent elements of the road space and specifies the technical and geometrical characteristics, sections, planimetric layouts, etc.

The regulatory instrument thus provides a fundamental indication of the limits between one road context and another, what objectively differentiates one road type from another, both through geometric indications and through indications of traffic types in terms of participants, speed, flows, conflicts, etc.

However, the regulation of traffic and infrastructures is not our aim. The indications of the standards must be used in order to identify a classification of road contexts from the point of view of its relationship with the vehicle/driver, i.e. a set of categories built focusing on the distinctive features that allow the realization of testing scenarios. Some purely bureaucratic distinctions (e.g. the distinction between urban and suburban) are used in order to identify the constituent parameters of scenarios and their range of variation, but do not necessarily distinguish one scenario category from another. The focus is therefore on the relationship between the ego-vehicle and the scenario context (scenes). This is expressed through the geometry of the section in which the scenario takes place and the traffic context that occupies it (static and dynamic parameters and self representations).

In order, therefore, to identify categories of scenarios that are homogeneous from the point of view of road geometry and traffic participants, it is possible to make an initial distinction of road types by the uncertainty they provide through the variability of participants, manoeuvres allowed, speed regimes, the presence or absence of conflict areas and its complexity.

[74] Ministero delle Infrastrutture e dei Trasporti, 'Norme Funzionali Geometriche per La Costruzione Delle Strade', *D.M. 5 Nov 2001*, 6792.

We therefore distinguish roads in:

- Low uncertainty: "fast-flowing" roads with a high control of traffic participants, the near absence of conflict areas and where all possible unexpected events are regulated.
    - At this category belong the roads referred to the new Highway Code type: A and B without private accesses and service roads.
- Medium uncertainty: "fast-flowing" roads characterized by low restrictions on traffic participants, some possible conflict areas, and lower control on unexpected events.
    - At this category belong roads referred to in the new Highway Code of type: B with private access and/or service roads and D
- High uncertainty: Roads characterized by a wide variability in terms of geometry, traffic participants, conflict areas, signals, and unexpected events.
    - At this category belong roads referred to the new Highway Code type: C, E, F

Each of these categories identifies a set of roads with similar geometric characteristics and known variables. From this first subdivision it is possible to distinguish elementary road sectors. The idea is to identify elementary sectors that constitute elements from the combination of which it is possible to reconstruct every possible road. Moreover, they will represent specific basic sceneries for the scenarios. The procedure implemented to identify elementary road sectors is based on the distinction between continuous variables and discrete variables. Each element which by its nature is discrete characterizes only one road sector, whereas all the continuously variable elements represent variations of the same elementary sector.

Undoubtedly there are many discrete elements such as, for example, the type of guardrail or the composition of the road surface, but many of these are not significant enough to represent a case itself or can be replaced by a continuously variable parameter (such as the coefficient of adherence between the tire and the asphalt).

Therefore, with reference to those discrete elements that significantly characterise a scene, by geometry or driving activity, we can distinguish:

Per 1. Low unchertainty:

Straight section with or without entrance or exit ramp or emergency or tunnel stop platform

Right or left curve section, with or without an entrance or exit ramp or in a tunnel

Motorway exit, service area, separate parking area

Per 2. Medium unchertainty:

Straight section with or without entrance or exit from service road or with private side accesses or with emergency stop or in tunnel

Right or left curve section, with or without entrance or exit to service road or private accesses or tunnel

Tollbooth, service area, parking area, traffic light flat intersection, roundabout

Per 3. High unchertainty:

Straight section with or without access to secondary roads, with or without side parking or dedicated side lanes or service areas, one-way, dual or alternating

Right or left curve section in the same cases as listed for straight sections

Intersection with or without traffic light, roundabout, pedestrian crossing, railroad crossing (in the various types of signaling), service stations and separate parking areas.

Some of these areas can be considered almost the same between one category of uncertainty and another (such as motorway toll booths and separate car parks). Others (especially in the urban sector), represent very specific cases that are entirely contained within other more complex scenes (e.g. pedestrian crossings are an integral part of intersections). If some of these scenes are included in the range of tests to be carried out, it is sufficient to consider the more complex ones or to extend the ranges of variation of the parameters in order to have a single but generic scene that includes several more specific ones.

If the specific case study requires it, the list of elementary scenes can be extended (or reduced) by distinguishing (or aggregating) cases relating to the variation of discrete parameters that are to be studied in detail (or for which specific analysis is not considered useful).

For example, it might be necessary to study the effects differentiated by type of guardrail, or, on the contrary, it might not be necessary to distinguish right-hand curves from left-hand curves, making all cases fall within the straight one and making the curvature of the road vary continuously.

A first subdivision made for the purposes of testing control logics is presented in the following:

Table 3: Road sectors distingued for uncertainty level (the groups divided with the use of letters, arabian numbers and roman numbers, so the letters A, B, C, etc. ***are not*** *to be considered as a reference to the "nuovo codice della strada")*

Low Unchertainty	Medium Unchertainty	High Unchertainty
		I. One way
A. Straight	**A. Straight**	**A. Straight**
1.Simple	1.Simple	1.Simple
2.With exit	2.With exit to a service road	2.Principal with input/exit
3.With input	3.With input from a service road	3.With linear parking
4.With emergency stop	4.With private exit/input	4.With angular parking
5.Gallery	5.With emergency stop	5.With side gas service
	6.Gallery	6.Not semaphored pedestrian crossing
B. Curve left	**B. Curve left**	**B. Curve left**
6.Simple	7.Simple	7.Simple
7.With exit	8.With exit to a service road	8.Principal with input/exit
8.With input	9.With input from a service road	9.With linear parking
9.Gallery	10.With private exit/input	10.With side gas service
	11.Gallery	11.Not semaphored pedestrian crossing
C. Curve right	**C. Curve right**	**C. Curve right**
10.Simple	12.Simple	12.Simple
11.With exit	13.With exit to a service road	13.Principal with input/exit
12.With input	14.With input from a service road	14.With linear parking
13.Gallery	15.With private exit/input	15.With angular parking
	16.Gallery	16.With side gas service
		17.Not semaphored pedestrian crossing
		II. Two directions
		from 18 to 34. Equal to one way
		III. Alternate direction
		from 35 to 51. Equal to one way
D. Special areas	**D. Special areas**	**D. Special areas**
14.Service area	17.Service area	52.Service area
15.Parking area	18.Parking area	53.Parking lot
16.Toll booth	19.Toll booth	
	20.Semaphored intersection	**E. Intersections**

	21.Roundabout	54.Semaphored
		55.Without traffic light
		56.Roundabout
		F. Rail crossing
		57.With light signal and barrier
		58.Without light signals or barriers

From this process we obtain 16 low uncertainty scenes, 21 medium uncertainty scenes and 58 high uncertainty scenes, for a total of 95 scenes. It is easy to notice that the road sectors more controlled with more restrictions such as motorways result in a very small number of scenes, while in urban areas the number of road sectors increases almost 4 times. As explained, this list of road sectors can be further extended or reduced. For our purposes, we will consider curvature as a continuously variable parameter in two directions from a straight road, and tunnel cases not different from non-tunnel cases using the variation of light as proxy for the differences between natural and artificial light. This approximation neglects the effects produced by the rapid change of brightness entering in and exiting from tunnels, but this has effect mainly on human drivers, while, for control logic analysis (unless specific analysis needs) the effects can be traced to the specifications of video optics (if used). These approximations reduce the number of scenes to a total of 42 which can also be reduced if they are redundant when performing tests that embraces multiple categories of uncertainty.

2.3 Identification of the activities compatibles with the road sections

After the definition of the road geometry and all features and variants related to the infrastructure, only a part of the static elements of the scenario is fixed. The method proposed by Schuldt et al. (2017)[75] and refined within the PEGASUS project[76] by Bock et al (2018)[77] identifies 6 layers for the definition of a scenario (Figure 8).

[75] Fabian Schuldt, 'Towards testing of automated driving functions in virtual driving environments', *PhD Thesis*, 2017.
[76] *Safeguarding Automated Driving Functions* <www.pegasusprojekt.de> [accessed 11 June 2021].
[77] Bock J. and others, 'Data Basis for Scenario-Based Validation of HAD on Highways', in *27th Aachen Colloquium Automobile and Engine Technology*, 2018, p. (pp. 8-10).

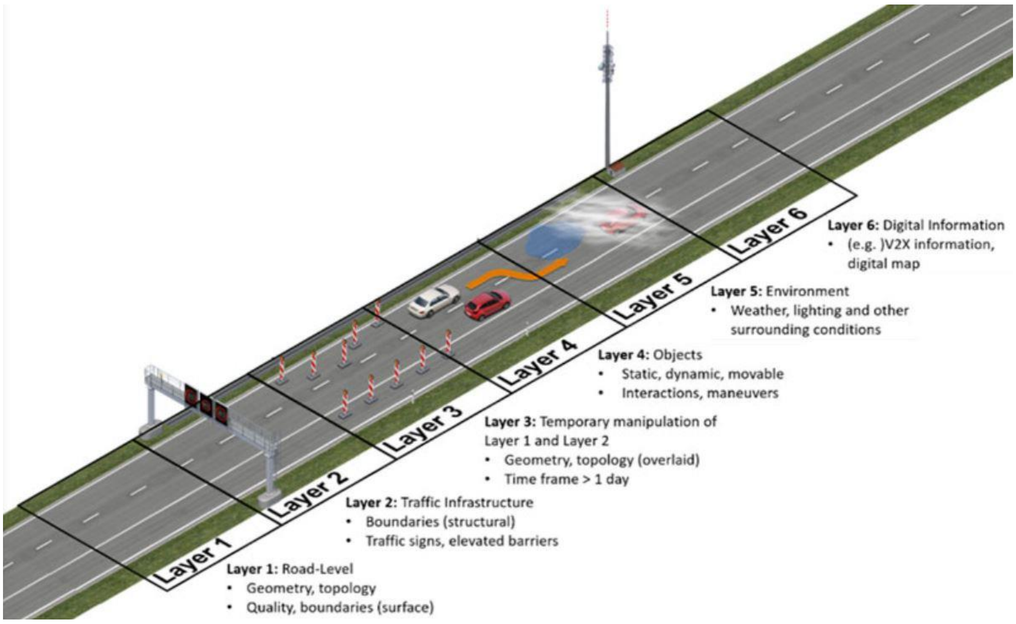

Figure 8: Six-layer scheme to the systematic analysis of the scenarios, perfected by Bock

The diagram in the previous paragraph defines the elements of the first two layers. There are other static components that can significantly define the scenario such as temporary signs (layer 3) or extraneous and unexpected elements (e.g., a boulder, a serious road damage or a crashed vehicle). These constitute the temporary static elements (temporary manipulation). In addition, static parameters must be associated with dynamic parameters and their representations[78] (layer 4), weather conditions (layer 5) and digital information such as V2X technologies (layer 6). Weather conditions, unless specifically required, can be treated as continuous variables (light, visibility, friction, wind direction and speed). The detail related to layer 6 is not useful for the purposes of this analysis on generic scenario categories, because it depends on the purposes of the experiments and can anyway be implemented from more general categories. Therefore, we will detail the elements of layers 3 and 4. These elements determine different situations which may occur in the same road sector and which, combined with it, configure very diverse scenarios.

To complete the scene, it is therefore necessary to specify the temporary static elements and the dynamic elements.

Temporary static elements are another discrete variable that can be applied to almost every road sector and therefore constitute separate groups of scenarios. For dynamic elements, the problem become complex, but we can make some generalizations to groups

[78] Ulbrich and others, ' Defining and Substantiating the Terms Scene, Situation, and Scenario for Automated Driving', *IEEE Conference on Intelligent Transportation Systems, Proceedings, ITSC*, 2015.

them by categories distinguishing discrete from continuous parameters due to our intent to obtain a discrete list of generic scenarios.

Reschka (2017)[79] identifies a list of elementary maneuvers, consisting of nine maneuvers: drive up, approach, follow, lane change, pass, turn left, turn right, turn back, stop, already used in Bagschik et al. (2016)[80] to identify the potential risks associated with an autonomous signaling highway vehicle.

Each maneuver sequence consists of a combination of some of these maneuvers and each controller must be able to perform these elementary maneuvers in order to drive a vehicle. However, their use simply in combination with scenes does not produce realistic scenarios, as they represent only uncontextualized maneuvering elements and not logically accomplished maneuvers. Due to the purpose of identifying testing scenarios, this maneuvers list cannot be used. To obtain a finite list of scenarios through the combination of contexts and driving objectives, we require a list of logically accomplished activities which can constitute stand-alone cases which, when combined with a scene, completes the scenario.

In order to identify categories that encompass all cases, a general to detailed approach has been adopted. From the user's point of view, it is possible to identify which driving objectives are possible in each road sector and which situations can occur.

For this analysis it is necessary to consider each elementary scenario as a piece of a generic vehicle journey. Each scenario has initial conditions that are to be considered as the final conditions of the previous piece of the journey and ends by producing the conditions that can be considered as the initial ones for the next piece of the journey. The analysis of each possible tile allows the knowledge of each possible combination of these and, therefore, the knowledge of each possible journey. In each scenario the ego-vehicle has a history that translates into the initial conditions of the scenario, has a path it is following and an initial speed. Analyzing what the possible activities are to be carried out for each scene, we can easily state that the ego-vehicle can have the objective of overcoming it (reaching the final section of the road in order to reach the next step of the journey) or remain inside it to carry out actions. In the first case, it may have the objective of staying in the path it is following (lane/road sector) or it may want to leave the path it is following and change road sector (locate and take the motorway exit, leave the road it is on and take another one). Thus, there are three possible objectives: to reach the final section

[79] A Reschka, 'Fertigkeiten-Und Fähigkeitengraphen Als Grundlage Für Den Sicheren Betrieb von Automatisierten Fahrzeugen in Städtischer Umgebung-English Title: Skill and Ability Graphs as Basis for Safe Operation of Automated Vehicles in Urban Environments', *Announced, PhD Thesis, Technische Universität Braunschweig*, 2017.

[80] Gerrit Bagschik and others, 'Identification of Potential Hazardous Events for an Unmanned Protective Vehicle BT - 2016 IEEE Intelligent Vehicles Symposium, IV 2016, June 19, 2016 - June 22, 2016', 2016-Augus.Iv (2016), 691–97 <https://doi.org/10.1109/IVS.2016.7535462>.

of the current road sector while keeping the current road, to change road or to stay within the current sector. These objectives can be declined in several ways, defining driving activities.

In the first case (reach the final section of the current road sector), the objective may be:

1. possible in a simple way: it involves only driving in more or less congested traffic (relatively simple maneuvers like: approach, car-following, pass)

2. with signaled difficulties: it requires the recognition of a temporary sign overlapping the pre-existing one and, possibly, the change of lane in safe conditions.

3. with non-signaled difficulties: an obstacle to be avoided (an accident, an animal on the road, a ditch, a bulky object)

4. impossible: reaching a traffic jam or an obstacle that takes up the entire carriageway

In the second case (change road), the target may:

5. require crossing explicit conflict areas such as (crossroads, roundabout, pedestrian crossing, etc.). It represents the activity of recognizing and giving correct precedence and performing maneuvers safely

6. does not explicitly require overcoming and resolving conflicts between vehicles. It includes the activities of identifying the new route, preparatory phases (signaling, slowing down, identifying the trajectories of other vehicles and possible conflicts) and execution of maneuvers.

The third case (execution of specific maneuvers to carry out an activity, remaining in the same road sector) concerns:

7. Parking: finding the parking space and act the parking maneuvers

8. Refueling, toll booth, drive-in, etc.: these are all activities that develop with the same dynamics. They involve locating and positioning the vehicle in a specific position, carrying out an activity (paying the toll, refueling), driving off and overtaking the road section to enter the next one (generally a conflict resolution sector).

The activity 1 can be distinguished into the set of driving activities associated with hypocritical traffic and those associated with hypercritical traffic. In the first case the vehicles can perform a greater number of maneuvers, lane changing is more possible as well as overtaking maneuvers, while in the second case the vehicles are limited to a simple car-following. However, the set of maneuvers that a single vehicle can performed in the hypercritical traffic case are entirely contained in the hypocritical case, so that, unless specific needs arise (e.g. analysis aimed at flow control logic rather than a single ego-vehicle), it may be not necessary to run the scenarios obtained from both conditions, assuming that control logics that adequately overcomes hypocritical scenarios are also capable of driving in stop-n-go conditions. It may be useful to split Activity 1 into the various driving models it contains. In this case Activity 1 would configure differentiated scenarios for drive-up, car-following, approach and overtaking activities, each tested within the range of variation of traffic conditions that allow them. It must be considered, however, that this kind of analysis may be useful during the design process to have a detailed analysis of the failure conditions of the single control models, but it produces redundant and sometimes not exhaustive analyses for safety validation purposes. Indeed in the case of high levels of automation, for the safety validation purpose we consider more exhaustive to test the generic driving activity in a wider variable traffic flow. In this way it is possible to verify not only the ability of the vehicle to execute the control models adequately, but also the verification of safety in the changes and in the choices between models (e.g. passage from car-following to overtaking). Activities 2 and 3 are the variants that add the temporary static elements to the road sectors. In these cases, traffic needs to be considered hypocritical. Otherwise, the scenario would configure a gridlock, that turn useless the presence of the characteristic temporary elements and configures a maneuver scheme equal to the activity 4. Activities 5 and 6 are distinguished by the clarification of a conflict to be resolved. While activity 5 concerns to across the resolution of an explicit conflict problem (e.g., an intersection or entering a motorway), activity 6 concerns situations in which the ego-vehicle should not normally come into conflict with other vehicles, either because no intersections of trajectories are configured or because it has an explicit priority (e.g., exiting the motorway). Therefore, the resolution of possible conflicts in activity 6 is part of safety or emergency maneuvers, while activity 5 explicitly concerns the activity of identifying priorities and respecting them.

These eight activities bring together multiple maneuvers and innumerable interaction dynamics between ego-vehicle and other traffic participants, but identify unique and separate categories of driving activities, distinguishing logical driving problems into activities that do not intersect each other and can be treated separately, embracing the totality of driving activities.

By combining activities and scenery, scenes and objectives of the ego-vehicle are completed, constituting almost all scenario elements (Figure 5). By specifying the value of all the characteristic parameters of the scenes and immersing the test vehicle in it, we obtain complete testing scenarios.

Obviously, not all activities are compatible with all sceneries and not all possible combinations are useful for testing purposes. A first analysis of the compatibility between sceneries and activities is therefore necessary. Afterwards, it will be possible to choose the set of scenery for specific tests.

2.4 Elaborating the Scene/Activity table and specifications within the activity of designing control logics

A first analysis of the compatibility between scenes and activities is obtained simply by comparing one by one the combinations, producing a table that lists in the rows the scenes and in the columns the various activities. The analysis carried out uses the 42 scenes obtained considering the curvature of the streets as a continuous variable parameter and considering that the tunnel cases provide the same information obtained from the exploration of the continuous variable of brightness, as explained in paragraph 2.2. Table 4 shows with an X the scenarios in which scenes and activities are compatible. It is possible to identify: 33 low uncertainty scenarios (motorways), 49 medium uncertainty scenarios (urban and suburban highways) and 132 high uncertainty scenarios (urban roads). Assuming the parameter grouping choices and the simplifications performed, these scenarios represent the whole possible scenario range of vehicles on the road. Of course, depending on the tests, it will never be necessary to run all scenarios because they are redundant with each other. For example, special areas, such as toll booths, service areas and dedicated car parks, are present in several uncertainty categories but do not differ significantly from each other, providing redundant informations. In such cases, it will therefore be useful to test only the scenarios with higher uncertainty (e.g. service areas with a greater variety of vehicles). This happens also between activities. It can be noted that activities 2 and 3 (man at work and obstacle) can be applied to all scenes, but sometimes activities performed in 1 (free flow) and/or 5 (intersection crossing) can be included in them. In these cases, the choice of which scenarios to test depends on the analysis objectives. Scenarios involving an obstacle on the roadway (marked or unmarked) include all activities that can be performed in free flow or intersection crossing, with the added difficulty of recognizing the obstacle and safely avoiding it. Due to the safety validation purpose, it can be considered that passing these scenarios implicitly includes passing the respective unobstructed scenarios. On the contrary, for design and verification of simpler functions, it is more useful to execute the unobstructed scenarios separately and then verify a limited number of scenarios with obstacles or narrowing of the roadway. For the same principle, in the high uncertainty sector some road sectors are repeated for the three directional cases, but execution of all cases is unlikely to be useful. The alternating one-way case configures the most difficult and complex case to solve, so it could, alone, be chosen for the set of scenarios for the final validation of the control logics. On the contrary, during the design phase, it may be useful to carry out tests in both one-way and two-way cases with separate lanes. As an example of an all-inclusive test set designed for the design phase Table 5 is given below.

Table 4: Compatibility scheme between Scene set and Activities

Context Unchertainty level	Scene Set – Generic set	Activities – Subset	Free Flow	Avoid accident *(obstacle)	Man at work **(bottleneck)	Intersection crossing	Approach to congestion	Change road without explicit conflict	Approach to a specific position (e.g. tool booth)	Parking
Low: *(Highways)*	Streight and curve	Simple	x	x	x		x			
		Exit road	x	x	x		x	x		
		Input road	x	x	x	x	x			
CdS Type:		Emergency stop	x	x	x	x	x		x	
A - B (part)	Special areas	Toll booth		x	x		x		x	
		Service area		x	x		x		x	
		Parking area		x	x		x		x	x
Medium: *(Extra-Urban)*	Streight and curve	Simple	x	x	x		x			
		Exit to a service road	x	x	x		x	x		
		Input from a service road	x	x	x	x	x			
CdS Type:		Private exit/input	x	x	x	x	x	x		
B (part) - D		Emergency stop	x	x	x	x	x		x	
	Special areas	Service area		x	x		x		x	
		Toll Booth		x	x		x		x	
		Parking area		x	x		x		x	x
		Roudabout		x	x	x	x	x		
		Semaphored intersection		x	x	x	x		x	
High: *(Urban)*	One way	Simple	x	x	x		x			
		Secondary input/exit	x	x	x	x	x	x		
		Linear parking	x	x	x	x	x		x	x
CdS Type:		Angular Parking	x	x	x	x	x		x	x
C - E - F		Side gas service	x	x	x	x	x		x	
		Pedestrian crossing	x	x	x		x			
	Two directions	Simple	x	x	x		x			
		Secondary input/exit	x	x	x	x	x	x		
		Linear parking	x	x	x	x	x		x	x
		Angular Parking	x	x	x	x	x		x	x
		Side gas service	x	x	x	x	x		x	
		Pedestrian crossing	x	x	x		x			
	Alternate directions	Simple	x	x	x		x			
		Secondary input/exit	x	x	x	x	x	x		
		Linear parking	x	x	x	x	x		x	x
		Angular Parking	x	x	x	x	x		x	x
		Side gas service	x	x	x	x	x		x	
		Pedestrian crossing	x	x	x		x			
	Intersections and conflict areas	Semaphored intersection		x	x	x	x			
		No traffic light		x	x	x	x			
		Roundabout		x	x	x	x	x		
	Rail crossing	Light signals and barriers		x	x	x	x			
		Only rail signal		x	x	x	x			
	Dedicated areas	Parking area		x	x		x		x	
		Service area		x	x		x		x	x

* Obstacle is a generic element to avoid, variable in dimensions and position. It can be present in any scene, so that it will replace with an additional difficulty the same scenario exploration done with another activity; therefore, it is useless to do both scenarios.

** Bottleneck is a narrowing of carriageway, variable in left, right or both.

Table 5: A LoReS example, finalized to a design purpose

Context Unchertainty level	Scene Set – Generic set	Activities – Subset	Free Flow	Avoid accident *(obstacle)	Man at work **(bottleneck)	Intersection crossing	Approach to congestion	Change road without explicit conflict	Approach to a specific position (e.g. tool booth)	Parking	
Low:	Streight and curve	Simple	x	x	x		x				
(Highways)		Exit road	x	*(simple)*	*(simple)*		*(simple)*	x			
		Input road	x	*(simple)*	*(simple)*	x	*(simple)*				
CdS Type:		Emergency stop	x	*(simple)*	*(simple)*	x	*(simple)*		x		
A - B (part)	Special areas	**Toll booth**		*(appr. to a s.p.)*	*(appr. to a s.p.)*		*(appr. to a s.p.)*		x		
		Service area		*(appr. to a s.p.)*	*(appr. to a s.p.)*		*(appr. to a s.p.)*		x		
		Parking area		*(parking)*	*(parking)*		*(parking)*		x	x	
Medium:	Straight and curve	Simple	x	x	x		x				
(Extra-Urban)		Exit to a service road	x	*(simple)*	*(simple)*		*(simple)*	x			
		Input from a service road	x	*(simple)*	*(simple)*	x	*(simple)*				
CdS Type:		Private exit/input	x	*(simple)*	*(simple)*	*(input)*	*(simple)*	*(exit to service)*			
B (part) - D		Emergency stop	x	*(simple)*	*(simple)*	x	*(simple)*		x		
	Special areas	**Service area**		*(appr. to a s.p.)*	*(appr. to a s.p.)*		*(appr. to a s.p.)*		x		
		Toll Booth		*(appr. to a s.p.)*	x		*(appr. to a s.p.)*		x		
		Parking area		*(appr. to a s.p.)*	*(parking)*		*(parking)*		x	x	
		Roudabout		*(curve left)*	*(curve left)*	x	*(curve left)*	*(curve left)*			
		Semaphored intersection		x	x	x	*(simple)*		x		
High:	One way	Simple	x	*(alternate directions)*	x		*(alternate directions)*				
(Urban)		Secondary input/exit	x	*(alternate directions)*	*(simple)*	x	*(alternate directions)*	*(alternate directions)*			
		Linear parking	x	*(alternate directions)*	*(simple)*	x	*(alternate directions)*		*(alternate directions)*	*(alternate directions)*	
CdS Type:		Angular Parking	x	*(alternate directions)*	*(simple)*	x	*(alternate directions)*		*(alternate directions)*	*(alternate directions)*	
C - E - F		Side gas service	x	*(alternate directions)*	*(simple)*	x	*(alternate directions)*		*(alternate directions)*		
		Pedestrian crossing	x	*(Inters. without traff. light)*	x		*(alternate directions)*				*** Only for Design purpose
	Two directions	Simple	x	x	x		x				
		Secondary input/exit	x	*(simple)*	*(simple)*	x	*(simple)*	x			
		Linear parking	x	*(simple)*	*(simple)*	x	*(simple)*		x	x	
		Angular Parking	x	*(simple)*	*(simple)*	x	*(simple)*		x	x	
		Side gas service	x	*(simple)*	*(simple)*	x	*(simple)*		x		
		Pedestrian crossing	x	*(Inters. without traff. light)*	x		*(simple)*				
	Alternate directions	Simple	x	x	x		x				
		Secondary input/exit	x	*(simple)*	*(simple)*	x	*(simple)*	x			
		Linear parking	x	*(simple)*	*(simple)*	x	*(simple)*		x	x	
		Angular Parking	x	*(simple)*	*(simple)*	x	*(simple)*		x	x	
		Side gas service	x	*(simple)*	*(simple)*	x	*(simple)*		x		
		Pedestrian crossing	x	*(Inters. without traff. light)*	x		*(simple)*				
	Intersections and conflict areas	**Semaphored intersection**		x	x	x	*(any simple case)*				
		No traffic light		x	x	x	*(any simple case)*				
		Roundabout		*(curve left)*	*(curve left)*	x	*(any simple case)*	*(curve left)*			
	Rail crossing	Light signals and barriers		*(Semaphored inters)*	x	x	*(any simple case)*				
		Only rail signal		*(Not sem. Int. and int. cros.)*	x	x	*(any simple case)*				
	Dedicated areas	**Parking area**		*(parking)*	*(parking)*		*(parking)*		x		
		Service area		*(appr. to a s.p.)*	*(appr. to a s.p.)*		*(appr. to a s.p.)*		x	x	

*** Only for Design purpose:
The "one way" scenario series can be considered a simplified version of the "two directions" scenario series. The "two directions" scenario series can be considered a simplified version of the "alternate directions" series.
Therefore, for validation purpose the "one way" and the "two directions" scenario series are useless. Their testing utility is limited to the design purpose.

* Obstacle is a generic element to avoid, variable in dimensions and position. It can be present in any scene, so that it will replace with an additional difficulty the same scenario exploration done with another activity; therefore, it is useless to do both scenarios.
The test series can be built in two way: 1. test all the "simple cases" and do some additional "obstacle scenarios". 2. execute all the "obstacle scenario" without execute the simplest versions.
The choice of the approach depends on the purpose of the test. Our suggest is to use the 1st approach to SUT design purpose and the 2nd approach to the SUT validation purpose.
The reason is that the results of the "simplest" tests gives more specific informations about the SUT behaviour in the generic activity (free flow or intersection crossing).
If the aim of the test is understanding the SUT behaviour in normal conditions, adding an obstacle to avoid in the scenario can give misleading informations.
Obiouvsly, if the SUT pass correctly the "obstacle test", it will pass the "siplest test". So, for validating purpose, the "obstacle test" can cover all the "free flow" and "intersection crossing" tests.
The table shows the 1st approach, because we think that the simplest tests (free flow or intersection crossing without obstacles) gives more specific informations about the SUT behaviour in the specific tasks.

** Bottleneck is a narrowing of carriageway, variable in left, right or both. In specific cases it gives no forter informations, because, once the SUT has correctly detected the temporary signals and behave consequentely, the scenario presents the same conditions of the free flow one with a narrow lane or numben of lanes (just explored)

(scene/activity) The activity/situation is possible, but it does not gives new informations because the case is just explored in another scene for the same activity or is a part of another activity for the same scene
In parenthesis there is the referral scene/activity for the case

Subset in red The subsets in red are scnenes to eliminate from a comprehensive test because they are repeated in lower control scene sets.

Subset in bold The subset in bold are the reference scene in a comprehensive test. They represent the worst case or the wider complexity case for the specific scene.

This table marks with an X the scenarios that need to be performed in order to obtain an exhaustive analysis for a vehicle with an SAE automation level of 4 or 5. In the table it is possible to distinguish in red some scenes. These are the scenarios that are redundant in a test series that embrace several uncertainty categories. The scenarios that has gray scripts in brackets do not need to be simulated because they represent conditions already present in other scenarios (indicated in the brackets).

2.5 Low Resolution Scenario (LoReS): description of the new scenario tool, potential, problems, areas of application and technical needs.

Basing on the choices made, the set of scenarios shown in Table 4 indicates the whole range of scenarios needed to be tested for automation logics design, and it includes 89 scenarios. So, assuming the choices made in the previous paragraph, the scenarios suitable to ensure a complete safety check of all driving activities in all possible road configurations would be 55. As we can see, the problem of identify the scenarios to be tested in order to obtain an exhaustive knowledge of the behavior of highly automated logics can be reduced to a finite number of scenarios. To achieve this, it was necessary to change not only the approach to scenario identification, but the concept of scenarios themselves. This new scenario paradigm can be described as a low-resolution scenario, as it does not represent a defined testing scenario, but it constitutes a category of scenarios that shares the same scenery, a given macroscopic traffic context and the same driving objectives. For convenience, from now on, low resolution scenarios will be called LoReS. In order for a LoReS to be transformed into a typical (high resolution) testing scenario, we have to define in detail the values of all characteristic parameters of the starting LoReS (the value of geometric quantities, vehicles speed, traffic and individual vehicle parameters, etc.).

The LoReS can be described as the space of scenarios characterised by the same discrete factors. In testing, the LoReS can be seen as a function that associates key parameters to the system under test in a given context. The context consists of several continuous variables defined in a range of values.

Each function (LoReS) becomes a test scenario (high resolution scenario) by giving a value to all variables (boundary conditions).

This implies that each LoReS specified in the table identifies a number of detailed scenarios equal to infinity raised by the number of characteristic variables. The problem may seem to be back to square one: the scenarios to be tested are infinite. The main advantage of the analysis conducted is that it places limits on infinity. On the analytical level, this allows first of all to know what the directions are and, therefore, the dimensions of each LoReS. This makes possible to know and measure the uncertainty of each automation logic. Knowing the dimensions of the space of the tests to be performed allows the measurement of what is known and what is not yet known about the behaviour of a given logic placed on the road, and therefore allows a confidence analysis. Each LoReS is therefore defined in a univocal space. This allows it to be explored using appropriate analysis tools. In this way it is possible to identify, in the systems tested, which parameters or values correspond to criticalities or determines them and allow to clearly identify the worst-cases, with a method free from any prejudice or false confidence obtained from data collected in non-homogeneous areas[81].

It is evident that the number of detailed scenarios arising from each LoReS implies that the analyses conducted on them cannot be carried out on the road, at least in the initial stages. Exploring a LoReS in detail implies the exploration of a continuous n-dimensional sample space. Random exploration and statistical analysis tools are therefore necessary. To this end, the experiments we will conduct will be carried out exclusively in MIL mode, which will guarantee the rapid execution of tens of thousands of detailed scenarios extracted from the same LoReS, allowing their analysis.

As already mentioned, each scenario is composed of static and dynamic elements. The latter constitute the traffic participants. Traffic is the most complex component to manage in the realization of detailed scenarios due to the discrete nature of the elements and the maneuvers they can perform. In fact, the typical implementation of testing scenarios involves defining the number and type of traffic participants, as well as the specific maneuvers they perform. To this end, in the analysis of LoReS, a macroscopic approach to the definition of traffic parameters is required. In this way, it becomes possible to manage the dynamic elements as traffic flows from the continuous variables that characterize them. This allows the variation of traffic by guaranteeing a complete exploration of the case studies. After analyzing the LoReS, it will be possible to identify the worst-cases, which can then be explored using the most appropriate methods.

A LoReS constitutes a space of scenarios which, in order to be properly explored, requires sample space exploration methods, sensitivity analysis methods and appropriate MIL simulation tools. The model of each LoReS constitutes a scenario that is variable in all its parameters. In order to realise one, a method of systematic analysis of the scenario parameters is needed, as well as programming tools for the simulation software used. This is helped by the method proposed by Schuldt (2017) [82], through which it is possible to probe, layer by layer, all the variable factors of the LoReS and identify those that can influence the SUT and those that do not need to be explored. Once the factors to be

[81] Winner, Wachenfeld, and Junietz'Safety Assurance for Highly Automated Driving – The PEGASUS Approach' TRB annual meeting 2017

[82] Schuldt, 'Towards testing of automated driving functions in virtual driving en- vironments', *PhD Thesis*, 2017.

explored have been identified, it is necessary to identify their range of variation. In the case of road geometric factors, limits and characteristics of traffic participants allowed, it will be necessary to refer to the technical standards.

Cap.3: Methodological and operative tools

3.1 Sensitivity analysis

LoReS scenario tool constitutes a space of scenarios which, to be properly explored, requires sample space exploration methods, sensitivity analysis methods and appropriate simulation tools. The dimensions defining LoReS are many, but their exploration depends on the SUT, which defines the analysis factors. To reach a goal, an ego-vehicle equipped with an automation device, an ADAS or a more complex logic, will rarely be susceptible to all the variable parameters of a LoReS. For this reason, Schuldt-Bock layer analysis can be the tool that discerns the parameters to be analyzed by mean of variation (factors) and the parameters that are not significant for the purposes of the tests, that will be constrained, reducing the size of the LoReS analyzed. After this operation we will obtain an n-dimensional scenario that will be explored by drawing the values of the analysed factors. The size of this scenario determines the tools that can be used to perform the tests. The greater the number of factors to be analyzed, the greater the number of tests to be performed, defining the analysis times. In the present research, MIL experiments were performed in order to run very large numbers of simulations in a short time. With the appropriate equipment, the experiments performed can also be carried out in HIL or VIL, exploiting multiple devices to reduce computational times. It is considered unfeasible to perform this kind of analysis on the road (or even in ScIL or DIL) because the number of cases would require test sessions that would be unsustainable both economically and in terms of time. However, it is possible to use the results obtained through simulation to identify the most relevant parameters and the most significant values in order to carry out the optimal road test series.

Extraction methods

The scenario space contained in a LoReS is characterised by continuous and finite dimensions. These can be explored by extracting the values of each factor, defining a sample that constitutes the scenario vector.

In the automotive field, test samples are often made by composing a regular grid, because it allows a sufficiently homogeneous exploration in all directions and the resulting analysis methods are widely used. This is the case of test matrixes which perform the variation of a parameter in a series of repeated values under different boundary conditions (variation of the other parameters)[83]. This approach is not very efficient, because the number of tests it requires increases exponentially as the number of factors analyzed increases. Its use is therefore feasible in cases where the number of factors explored is sufficiently limited (two or three).

A more effective method is pseudo-random exploration, for which the amount of samples analyzed is arbitrary. This means that the quantity of tests depends on the computational possibilities and the degree of confidence to be achieved (more tests imply better reliability of results). This approach is advantageous in terms of extraction and number of samples but has the disadvantage that the resulting analysis methods are fewer and less efficient than those based on a more structured construction of the test matrix. Pseudo-random extractions result in non-homogeneous explorations, creating denser regions of samples and less explored regions (holes) in the sample space. This problem is known as discrepancy and results in an increased need for test samples due to the inefficiency of the sample extraction method, and also in an error in the evaluation of the results obtained[84]. More densely explored regions have a higher statistical weight than the less explored regions, leading to errors in the evaluation of the results. These errors decrease as the number of extractions increases, leading to a need for very large experiments to obtain acceptable degrees of confidence.

Regular grid and pseudo-random are not the only methods for exploring and analyzing the sample space. There are, in fact, numerous quasi-random extraction methods Figure 9.

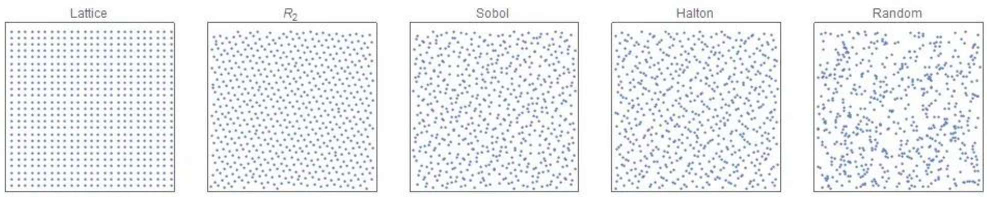

Figure 9: Some extraction methods: regular grid, quasi-random and pseudo-random

Quasi-random extraction is a hybrid method between Monte Carlo extraction and regular grid, in order to get the best of both[85]. Quasi-random (or quasi-Monte-Carlo) methods try to explore space homogeneously by reducing the number of samples, selecting them through an algorithm that ensures known homogeneity in all directions through an optimisation process. To this method belong the series of Halton, Kroneker, Niederreter, Sobol and others. The literature contains many studies aimed at identifying

[83] European New Car Assessment Programme, 'Test Protocol – AEB VRU Systems' <https://cdn.euroncap.com/media/26997/euro-ncap-aeb-vru-test-protocol-v20.pdf>.
[84] William Chen, Anand Srivastav, and Giancarlo Travaglini, *A Panorama of Discrepancy Theory*, 2014.
[85] Chen, Srivastav, and Travaglini, 'A panorama of discrepancy theory', 2014.

the best algorithms and quasi-random series[86]. In the experiments carried out within the framework of this research, both pseudo-random and quasi-random extraction methods were tested, in particular with reference to the Sobol sequence because, besides being among the best known in the literature for low discrepancy, it allows the integration of new series of experiments with the series already carried out, without the need to reconfigure the experiments from scratch. In addition, it is related to some of the sensitivity analysis methods used subsequently.

Sensitivity indices

The sensitivity analysis techniques analyzed in this research belong to the family of variance-based techniques, which were first employed by Cukier et al. (1973)[87], generalised by Sobol (1993)[88], through a quasi-random implementation based on the Monte Carlo method. The studies conducted by Saltelli et al. (2010)[89] optimize the computational efficiency by implementing Sobol-type extractions and identifying the most effective sensitivity indices. As shown in the literature, these methods have proven to overcome most of the limitations of other common approaches, such as One-At-Time Analysis (OAT), differential methods and regression/correlation analysis[90].

The idea of the method is based on the assumption that the variance is a valid proxy for output uncertainty. The method is based on the following variance decomposition formula[91]. Given a model $Y = f(X_1, X_2, \ldots, X_k)$, where $X_i \ \ \forall i \in [1, k]$ are the factors, and Y is the Output quantity of interest, output variances can be described as:

$$V(Y) = V_{X_i}\left(E_{\overline{X}_{\sim i}}(Y|X_i)\right) + E_{X_i}\left(V_{\overline{X}_{\sim i}}(Y|X_i)\right)$$

[86] Martin Roberts, 'The Unreasonable Effectiveness of Quasirandom Sequences', *Extreme Learning*, 2018 <http://extremelearning.com.au/unreasonable-effectiveness-of-quasirandom-sequences/> [accessed 19 June 2021]; Nadia A Mohammed, Quasi-monte Carlo Sobol, and Nadia A Mohammed, 'Comparing Halton and Sobol Sequences in Integral Evaluation', *Zanco Journal of Pure and Applied Sciences*, 31.1 (2019), 32–39 <https://doi.org/10.21271/zjpas.31.1.5>.

[87] R. I. Cukier and others, 'Study of the Sensitivity of Coupled Reaction Systems to Uncertainties in Rate Coefficients. I Theory', *Journal of Chemical Physics*, 1973 <https://doi.org/10.1063/1.1680571>.

[88] I.M. Sobol', 'Sensitivity Estimates for Nonlinear Mathematical Models', *Mathematical Modeling and Computational Experiment*, 1993; Ilya M. Sobol and Boris V. Shukhman, 'On Global Sensitivity Indices: Monte Carlo Estimates Affected by Random Errors', *Monte Carlo Methods and Applications*, 2007 <https://doi.org/10.1515/MCMA.2007.005>.

[89] Andrea Saltelli, Marco Ratto, and others, *Global Sensitivity Analysis. The Primer*, *Global Sensitivity Analysis. The Primer*, 2008 <https://doi.org/10.1002/9780470725184>; Andrea Saltelli, Paola Annoni, and others, 'Variance Based Sensitivity Analysis of Model Output. Design and Estimator for the Total Sensitivity Index', *Computer Physics Communications*, 181.2 (2010), 259–70 <https://doi.org/10.1016/j.cpc.2009.09.018>.

[90] Andrea Saltelli and Paola Annoni, 'How to Avoid a Perfunctory Sensitivity Analysis', *Environmental Modelling and Software*, 25.12 (2010), 1508–17 <https://doi.org/10.1016/j.envsoft.2010.04.012>.

[91] A Mood, F A Graybill, and D C Boes, 'Random Variables, Distribution Functions, Expectation', in *Ntroduction to the Theory of Statistics*, 3rd ed. (New York, NY, USA, 1974), pp. 51–72.

where X_i is the i-esim factor, and $\overline{X}_{\sim i}$ is the vector of all factors except X_i.

The first component, $V_{X_i}\left(E_{\overline{X}_{\sim i}}(Y|X_i)\right)$ is called "main effect (first order effect)" of X_i. The associated sensitivity measure, called "first-order sensitivity index" is equal to the first-order effect normalized to the total variance, i.e. unconditioned:

$$S_i = \frac{V_{X_i}\left(E_{\overline{X}_{\sim i}}(Y|X_i)\right)}{V(Y)}$$

It can be interpreted as the part of the variance of the output that is only due to the variation of the input factor X_i. Therefore, the first-order effect captures the stand-alone effect of the input factor on the output of the model. The first-order index of a factor coincides with the standardized coefficient of a least-squares linear regression of the input-output map of the factor.

However, for non-additive models, factor X_i also contributes to the variance of the output in its interaction with the other factors. This influence is called the interaction (or higher-order) effect relative to X_i. The sum of the first-order and higher-order effects for all factors "explains" all the variance of the output. Therefore, to quantify the total effect of a factor, the so-called "total sensitivity index" is introduced:

$$ST_i = \frac{E_{\overline{X}_{\sim i}}\left(V_{X_i}(Y|\overline{X}_{\sim i})\right)}{V(Y)} = 1 - \frac{V_{\overline{X}_{\sim i}}\left(E_{X_i}(Y|\overline{X}_{\sim i})\right)}{V(Y)}$$

which is the sum of the first-order effect of X_i and all higher-order effects involving X_i. Since the higher-order effects are included in the ST of each factor involved in the interaction (e.g. $S_{i,j} = S_{j,i}$ is included in both ST_i and ST_i) it results $\sum_{i=1}^{k} ST_i \geq 1$, where equality holds only for perfectly additive models (so $S_i = ST_i$, $\forall$ i $= 1, \dots,$ k).

Bootstrapping confidence intervals

The calculation of sensitivity indices improves as the number of tests performed increases and, depending on the model being tested, may stabilize more or less rapidly at a sufficiently small range of values. The Sobol extraction performed for a number equal to a power of two ensures a homogeneous exploration of the space. The composition of the sample matrix, given in Saltelli et al (2008)[92], is designed to follow this extraction rule and defines a number of tests equal to 2^N×(k+2), with k = number of factors. Thus, as N (and

[92] Saltelli, Ratto, and others,' Global Sensitivity Analysis. The Primer', 2008.

hence the number of samples extracted) increases, the model is gradually explored in all directions. This allows a calculation of the indices which, barring some initial jump, will gradually tend towards the constant function (Figure 10). Complex models can, however, exhibit unexpected behavior that can affect the calculation of the indices even after several tests. For this reason, it was decided to associate the calculation of the indices with a measure of the confidence of the results obtained.

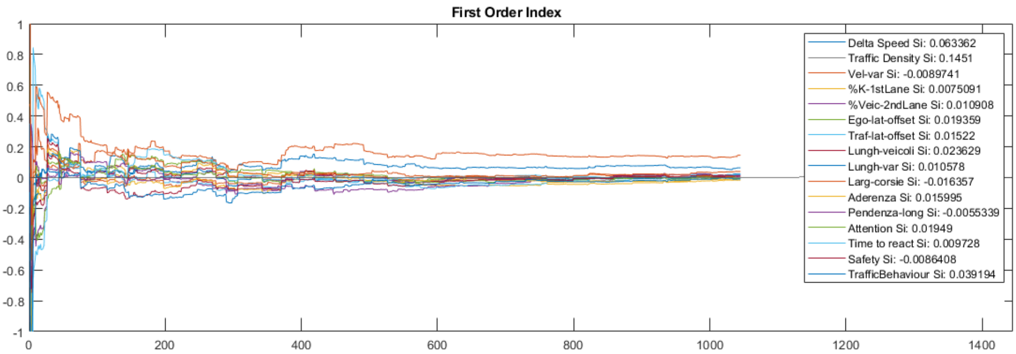

Figure 10: Sensitivity indices calculated at the increasing of the number of simulaztions

In scientific literature, the calculation of sensitivity indices is often associated with the calculation of the confidence interval, in particular by bootstrapping[93]. Therefore, the indices calculated in the experiments framework carried out in the course of this research will always be assisted by a confidence measure through bootstrapping confidence interval (BCI). There are various methods to perform this analysis. The basic principle is to extract a certain population of data in order to estimate its statistical distribution. In this case, from the set of all outputs, a population is extracted, and its indices are calculated. By repeating this operation a sufficient number of times it is possible to evaluate the distribution of the indices, making statements about its percentiles.

Regional sensitivity analysis

Regional analysis is another tool that can be used to assess the effects of input variance on output variance. It consists of dividing outputs into two categories based on whether a significant threshold of values is exceeded. This threshold divides desirable (behavioral)

[93] G. E.B. Archer, A. Saltelli, and I. M. Sobol, ‘Sensitivity Measures, Anova-like Techniques and the Use of Bootstrap’, *Journal of Statistical Computation and Simulation*, 1997 <https://doi.org/10.1080/00949659708811825>; Saltelli, Annoni, and others, ' Variance based sensitivity analysis of model output. Design and estimator for the total sensitivity index', *Computer Physics Communications*, 2010.

from undesirable (non-behavioral) outcomes, and consequently the inputs (factors) that generated them also fall into the behavioral or non-behavioral regions. Each factor can therefore be analyzed by comparing the behavioural and non-behavioral values. If the trends are similar, this implies that the variance of the factor analyzed does not affect whether the output belongs to a region or another. If this is not the case, the variance of the factor will be significant for the variance of the output, having an influence on the classification of the output in the behavioral and non-behavioral regions. Furthermore, in that way it is possible to identify for which values of the factor analyzed, the outputs will tend more towards one region than the other[94].

These analyses can be performed as pdf or histogram comparisons. In the first case, we evaluate the slope of the curves and the distance between them. The greater the distance between the curves, the greater the impact of the factor on the position of the output with respect to the chosen threshold. This comparison can be performed by means of standard statistical tests, such as the Kolmogorov-Smirnov test. In this test the "distance" between the two distributions, D_i, is defined by the cumulative distribution function of X_i^B and X_i^{NB}, as follow: $D_i = \max|F(X_i^B) - F(X_i^{NB})|$. Where the difference between the slopes of the curves changes sign, a threshold value for the input will correspond, resulting in a change of trend in the output (Figure 12). Using histograms, the comparison is again made on the trend of the behavioral and non-behavioral curves and the identification of differences focuses on the values where the recurrence of the data is significantly dissimilar (Figure 11).

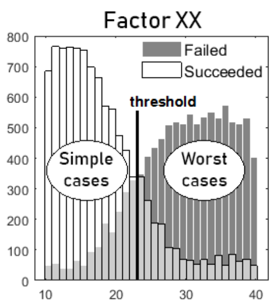

Figure 11: example1, behavioral histogram against non-behavioral histogram

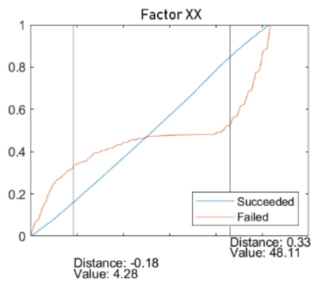

Figure 12: example 2, cumulative distributions comparison

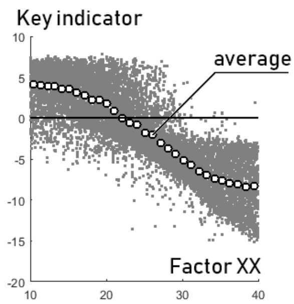

Figure 13: example 3, reading scatterplots

Among the advantages of this method is the fact that it is not linked to the extraction method, and it can be applied to any type of output as long as it is traced back to a Boolean

[94] Saltelli, Ratto, and others,' Global Sensitivity Analysis. The Primer', 2008.

data. Furthermore, this kind of analysis highlights for which range of input values there is the greatest tendency to obtain unwanted output. This makes it an extremely useful tool to identify worst-cases for the SUT based on statistical analysis customized on the SUT, without the need to rely on external and/or inhomogeneous databases.

Scatterplot analysis

Another method to explore the effects and the significance of a factor with respect to the variance of the outputs is scatterplot analysis. Quoting Saltelli et al. (2008) [95]

“*Input/output scatterplots are in general a very simple and informative way of running a sensitivity (…) they can provide an immediate visual depiction of the relative importance of the factors.*”

By arranging the outputs in a graph against a single factor, it is possible to observe the resulting point cloud and see whether it is equally distributed, following the distribution of the extraction function used, or whether it shows a shape that correlates it more or less strongly with the factor taken into analysis (Figure 14).

This kind of analysis makes it possible to identify with immediacy both the correlation between a factor and the output analyzed, but also whether there are unexpected dependencies, anomalies linked to a certain range of values of the factor. For example, in some cases it is possible to display any false positive activations of a logic and to which factor they are related.

In addition, for each scatterplot it is possible to analyze the trend of the averages of the output values as the factor changes. This tool allows a further visualization of the correlations, which is manifested through a trend of the averages that deviates from the constant. This trend can show simple relations (linear, quadratic, etc.) or articulated relations (a peak in correspondence of a range of factor values) (Figure 13).

[95] Saltelli, Ratto, and others,' Global Sensitivity Analysis. The Primer', 2008.

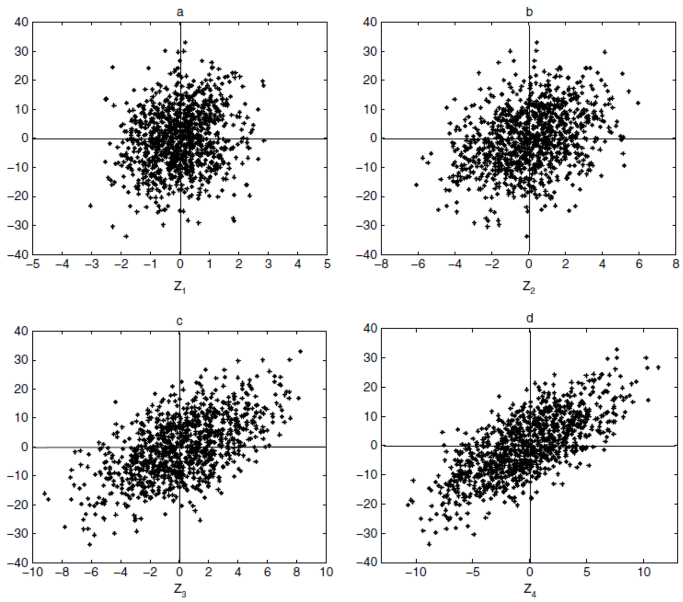

Figure 14: Example of scatterplot of a simulation series output against four different factors, which shows various influences on the output variance

3.2 Simulation tools

Vehicle dynamics simulation

Simulation software are a fundamental tool to carrying out a large quantity of tests quickly and at a very low cost. As we told in the previous chapters, simulation tools are many and consist of different systems depending on requirements. In the context of MIL and HIL ADAS and AV's testing, the simulation of vehicle dynamics is a crucial component. MIL and HIL simulations, in fact, test logic and hardware in virtual environments composed of the model of the vehicle, the model of the surrounding environment and interactions with the vehicle, the model of the sensors and the simulation of the data supplied to them.

Typically, simulation software performs a series of calculations, solving the equations that determine the interactions that occur in an extremely short period of time (simulation-step) between the simulated elements (models). In vehicle dynamics software, the focus is on a vehicle model (that can be more or less complex) and its interaction with the road (in the case of cars). Beyond this focus, the software can provide other features or can arrange communications with other tools dedicated to simulate some other elements such as, e.g., the signals perceived by the virtual sensors of the vehicle model or a graphic engine that makes photorealistic render images for sensor camera testing.

There are numerous software tools that perform these functions, both commercial and free or open source. Matlab provides its vehicle dynamics blockset, consisting of various vehicle models, which can be linked to a rendering engine (specifically Unreal engine) [96]. Simulink is a simulation tool that allows the creation of logic models, and it is often used to create simulation models of control logics, vehicles or components such as engines or tires[97]. GAZEBO is an open-source software that allows the modeling of robots in a virtual environment and has often been used to simulate vehicles with partial or total control logic[98]. Commercial software from companies like D-Space, IPG Automotive and VI-Grade are also widely used. The latter provide highly specific software for the simulation of vehicle dynamics and solutions for the testing of ADAS and AV systems. Within the framework of this research, extensive use was made of the simulation software provided by IPG under student license, which supplied a large part of the simulated environments needed for the experiments performed. In many cases, it was necessary to integrate further models developed specifically for this purpose into the software provided or provided by other interconnected software. There are many other products for the simulation of vehicle dynamics, such as RF-pro which also provides very detailed models for road surface simulation and rendering, but for the purposes of the research project, it was not necessary to go into the details of hyper-realistic tools which would have increased the computational needs for the simulated environments without providing significant contributions to the experiments performed. As Philip Koopman and Michael Wagner write:

"Realism for its own sake is an inefficient, and ultimately unaffordable, use of test resources. The key to simulation validity is having just the right amount of realism (simulation fidelity) to get the job done." [99]

MIL e HIL

The degree of realism depends on the needs for simulation. This gives rise to different approaches to simulation. As mentioned, model-in-the-loop tests are carried out in full

---

96 MathWorks, 'Vehicle Dynamics Blockset' <https://it.mathworks.com/products/vehicle-dynamics.html>.

97 Dennis Assanis and others, 'Validation and Use of SIMULINK Integrated, High Fidelity, Engine-Ln-Vehicle Simulation of the International Class VI Truck', *SAE Transactions*, 2000, 384–99.

98 H Singh and S Jha, 'Simulated Environment for Autonomous Driving Using ROS Based on Mahindra E2O Electric Car', in *ACM International Conference Proceeding Series*, 2019 <https://doi.org/10.1145/3352593.3352667>; B Vieira and others, 'Towards a Realistic Simulation Framework for Vehicular Platooning Applications', in *Proceedings - 2019 IEEE 22nd International Symposium on Real-Time Distributed Computing, ISORC 2019*, 2019, pp. 93–94 <https://doi.org/10.1109/ISORC.2019.00028>; M Nithya and M R Rashmi, 'Gazebo - ROS - Simulink Framework for Hover Control and Trajectory Tracking of Crazyflie 2.0', in *IEEE Region 10 Annual International Conference, Proceedings/TENCON*, 2019, MMXIX-OCTOB, 649–53 <https://doi.org/10.1109/TENCON.2019.8929730>.

99 Philip Koopman and Wagner, 'Toward a Framework for Highly Automated Vehicle Safety Validation'.

simulation. This guarantees extremely fast simulation times, but the simulated models are simplified allowing only the control logic testing. Structurally, MIL simulations do not differ much from HIL simulations, because only a few elements (hardware) are not simulated, while everything else is simulated as in MIL. This small difference, however, leads to a big change in the simulation hardware structure and in the test execution times. In order to test the real effectiveness of the hardware under test and whether it responds correctly and on time, HIL simulations must be run in real-time. The simulated environment is processed through hardware that ensures real-time execution and translates the patterns perceived by the simulated sensors into realistic signals that are sent to the hardware under test. This makes possible to test real sensors and/or the actual response of the logic mounted as controller. The difference in execution time between the two simulation approaches results in a significant difference in the number of simulations that can be performed.

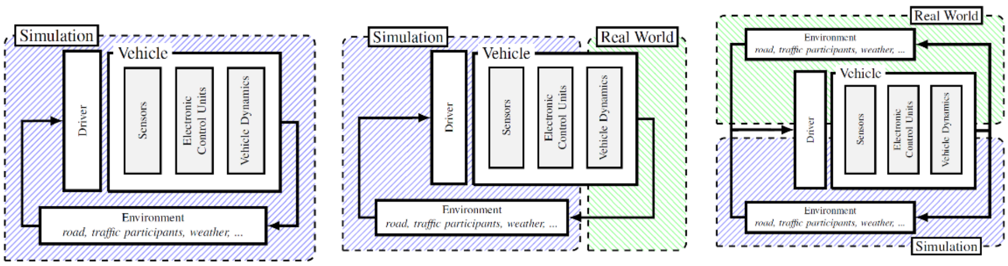

Figure 15: Stellet et al., simulation approaches

In the research carried out, only model-in-the-loop experiments were performed, since the research is aimed at exploring very complex scenarios that require the execution of a large amount of tests. The experimented method is therefore in the design and validation stages of the V-model linked to MIL experiments. However, we underline that the proposed method is not limited to the MIL phases, but embraces the entire approach to testing and validation. The analyses carried out in the MIL phases allow the identification of worst-cases and produces the results that determine the elaboration of the subsequent HIL phases. The latter can be carried out with the same approach as proposed because, downstream of the MIL simulations, the test space to be explored can be extremely reduced and, with it, the number of tests to be performed.

Cap.4: Method developement

4.1 Design of experiments

The proposed scenario tool and the methodologies identified requires appropriate simulation tools and the verification of their applicability, effectiveness and time required. To this end, a design of experiments has been drawn up to identify the most effective simulation tools and methodologies and verify their compliance with research needs.

The first experiment aims to test and verify the potential of the vehicle dynamics simulation tools provided by IPG Automotive (partner in the research project). The methods studied for the realization of testing scenarios[100], their applicability to the simulation software, the realization of random series of scenarios, the extraction of outputs and their post-processing are tested.

The second experiment investigates sensitivity analysis tools, identifying the most effective methodologies to analyze spaces of scenarios, the application of low discrepancy extraction methods, the simulation time of long series of complex scenarios and it will validate the method in comparison with known ADAS testing methodologies.

The third experiment applies the developed simulation and analysis method in an industrial environment. A use-case is analyzed, its uncertainty factors are identified and the main factors are analyzed, allowing the preparation of high efficiency testing experiments while minimizing the computational load.

The last experiment inserts the traffic component into the factor system under analysis. It constitutes a proof of concept that brings together all the skills and methodologies developed to perform the analysis of a case study with LoReS. The case study was identified through collaboration with the industrial partner FCA (now Stellantis). The LoReS analysis was done, and the needed analyses were done to completely experiment the proposed method.

4.2 Experiment 1: Experimenting the simulation tools dealing with scenario complexity

The first experiment performed the representation of a scenario characterized by multiple variable factors. To design of the scenario, we followed schemes indicated by

[100] Geyer and others, 'Concept and development of a unified ontology for generating test and use-case catalogues for assisted and automated vehicle guidance' *IET Intelligent Transport Systems*, 2014; Ulbrich and others, ' Defining and Substantiating the Terms Scene, Situation, and Scenario for Automated Driving', *IEEE Conference on Intelligent Transportation Systems, Proceedings, ITSC,* 2015; Schuldt, 'Towards testing of automated driving functions in virtual driving en- vironments', *PhD Thesis*, 2017.

Geyer, Ulbrich et al. (2014 and 2015)[101], whereby the use-case was defined first. Then, we identified all the scenario parameters using the Schuldt-Bock layer analysis[102], which identified the parameters which does not influence the SUT, which were fixed, and those that may influence the SUT, whose value was varied between tests. For each parameter the variation range was defined, and a random sequence of vehicle dynamics simulations was programmed.

Aim of the experiment due to research purposes: Verify the effectiveness of the method to define the scenario, test the potentialities of the simulation tools, observe the execution time for long series of tests and some of the possible methods to analyze results.

Aim of the test: Verify the effectiveness of an AEB system during a rapid change of target in the radar, with a partial obstruction of the radar field of view.

Use-Case: Ego-vehicle cut-in during an overtaking maneuver of a vehicle that reduces the field of view.

Scenario description: The scene takes place on a motorway. The ego-vehicle drives on the right lane, follows a truck. The truck follows a third vehicle (a small car). The ego-vehicle overtakes the truck and, before it completes its overtaking manoeuvre, the small car in front of the truck changes lanes, taking position in front of the ego-vehicle.

Key Parameter: collision count. The test is considered passed if there are no collisions, failed otherwise.

Scenario details: The scenario takes place on a straight part of a motorway. The road consists of two lanes for each direction and a side bank. Additional lanes would have no effect on the experiment because the logic implemented only apply brakes and does not provide any additional maneuver. The stretch is characterized by a variable slope both longitudinally and laterally (both varying from -6% to 6%). Although the technical standards do not admit the variability of the lateral slope in straight motorway sections, we wanted to test the influence of this variable, because it could influence the test results, especially in combination with other factors such as the longitudinal slope and the adherence. It is a hypothetical condition aimed at experiment. About the road geometry

[101] Geyer and others, 'Concept and development of a unified ontology for generating test and use-case catalogues for assisted and automated vehicle guidance' *IET Intelligent Transport Systems*, 2014; Ulbrich and others, ' Defining and Substantiating the Terms Scene, Situation, and Scenario for Automated Driving', *IEEE Conference on Intelligent Transportation Systems, Proceedings, ITSC*, 2015.

[102] Schuldt, 'Towards testing of automated driving functions in virtual driving en- vironments', *PhD Thesis*, 2017; Bock, J. and others, 'Data basis for scenario-based validation of HAD on highways', *27th Aachen colloquium automobile and engine technology*, 2018.

(layer 1), variations in the following factors were also analyzed: lane width (between 3.5 m and 4.5 m), bank width (between 1.5 m and 2.5 m) and tread-asphalt grip coefficient (between 0.3 and 1, considered as a sum of asphalt and tyres conditions and the weather). Due to the test purpose, infrastructural details such as signs and guardrails (layer 2) do not affect the simulation, so they have been made invariant. The chosen scenario does not fall within carriageway reductions or other effects due to works in progress (layer 3). About the traffic participants and their manoeuvres (layer 4), the types of participants and the parameters strictly necessary to achieve the phenomenon under analysis are fixed, while the other arbitrary factors have been made variable. In particular, the hierarchy of positions and speeds is fixed, but the distance between the truck and the compact car is varied (from 5 m to 20 m), the speed of the truck and the compact car (equal to each other and varying between 50 km/h and 90 km/h), the speed of the ego-vehicle (between 91 km/h and 130 km/h), the length and the width of the truck (between 7 m and 19 m and between 2 m and 3 m respectively). In the case of the width of the truck, the range of values was also widened in order to better allow the reading of any effects of the factor on the experiment. To ensure that overtaking takes place, the minimum speed that the ego-vehicle can have corresponds to slightly more than the maximum speed that the other two participants in the traffic can have. The sequence of maneuvers was programmed in such a way that the change of lane by the compact car occurs after a variable time between 0 and 2 seconds from the start of the overtaking manoeuvre. Finally, it was chosen to vary the radar mounted on the ego-vehicle, assuming that the manufacturer of the ADAS logic could choose between three types of observation depth (30 m, 50 m and 70 m) and three possible viewing widths (10°, 15° and 20°). Considering a radar technology, it was considered that varying brightness and atmospheric visibility conditions would not affect the test (layer 5). Finally, the scenario in question does not have any devices or logic that make use of V2X communications (layer 6). All the factors and their ranges of variation are summarised in Table 6.

Table 6: Experiment 1, Uncertainty Factors

Factors	Lower bound	Upper bound	To be found
Lateral Slope	-6	6	meters per meter (m/m)
Longitudinal Slope	-6	6	meters per meter (m/m)
Lane width	3.5	4.5	meters (m)
Shoulder width	1.5	2.5	meters (m)
Friction	0.3	1	non-dimensional (-)

Spacing truck/car	5	20	meters (m)
Truck and car speed	50	90	kilometers per hour (km/h)
Ego-vehicle speed	91	130	kilometers per hour (km/h)
Truck length	7	20	meters (m)
Truck width	2	3	meters (m)
Time between the beginning of the overtaking menuver and the lane changing start	0	2	seconds (s)
Radar view length	30	70	meters (m)
Radar view angle	10	20	degrees

Implementation: IPG Carmaker was used as vehicle dynamics simulation software for this experiment. The scenery and traffic participants were programmed. The vehicle models provided by the program were used and the variable factors such as truck size were defined. The vehicle maneuvers were programmed. The lane change maneuver of the compact-car starts with a variable delay from the beginning of the overtaking maneuver of the ego-vehicle. Through the graphic user interface (GUI) it is possible to give a value to the parameters provided by the programme or initialise them with a name that can be used through scripting to manage the parameter values. The programme provides a script management tool and its own command library compatible with the tcl/tk language, so the scripts were developed in tcl/tk. A script has been writed. For a specified number of simulations, a for-cycle initializes the variables, assigns a random value and launches the simulations. During the simulation, a while cycle monitors collisions, interrupting the simulation if they occur. If no collisions occur, the simulation stops when the vehicles reach the scenery limit, which is always located downstream of the manoeuvres. At the end of the simulation, the for loop starts again with a new step.

```
For (da n=1 a numero di simulazioni)
{
        Inizializza fattore 1 : random tra min e max
        …
        Inizializza fattore n : random tra min e max

        Calcola parametri dipendenti dai valori estratti
        Inizia la simulazione
```

```
    While (simulazione attiva)
        if (collisioni > 0)
            Stop simulazione

    Controlla che la simulazione sia finita
    Incrementa n
}
```

The software ran one simulation every 3 seconds on average, producing 30,000 simulations in approximately one day.

The results were analyzed using a regional method, distinguishing between passed and failed tests with the collision discriminator 0 or >0. The results were graphed with histograms (Figure 16).

By comparing the histograms it is possible to see that the most significant factors are: 1) the longitudinal slope, which determines a linearly decreasing trand of accidents as the value of the factor increases, 2) the speeds of both the ego-vehicle and the other traffic participants, which are also linearly correlated to the number of accidents but in direct proportion, 3) the friction, which has a more than linear incidence and finds a threshold of trend change between the values of 0.5 and 0.6. The radar parameters used also seem to have an influence on the occurrence of impacts, but the scarcity of values tested does not allow a clear identification of the effects through this type of analysis. The other factors do not show correlations with the simulation output, showing no differences between the trends of the successful and failed simulations, which are equally distributed, almost following the constant distribution of the random extraction function used.

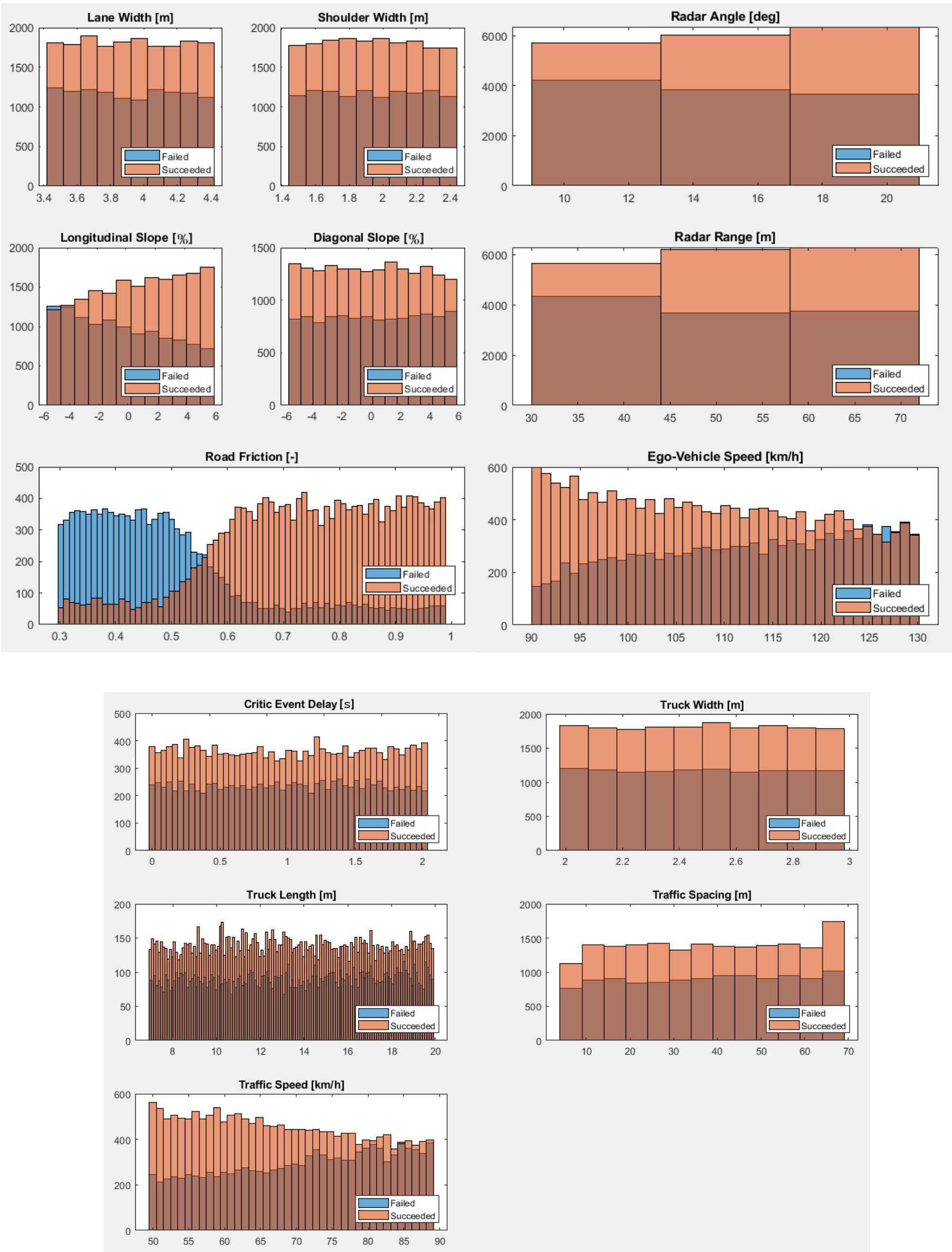

Figure 16: Experiment 1, comparison between histograms of accident and accident-free simulation.

After the experiments, we can say that the software used is suitable and versatile for carrying out long series of simulations, managing the parameters and series of simulations by scripting. Simulation time is short enough to obtain sufficient results to perform the analyses. A greater variety of extracted values is preferable, because when only a few values are tested, this kind of analysis does not show clear results.

4.3 *Experiment 2: Perfecting analysis tools and comparison with known testing methods*

The first experiment carried out showed some limitations of the regional analysis, so we decided to adopt furthermore in-depth analysis tools. To better validate the tools used and the results obtained, we decided to carry out an experiment based on a known use-case belonging to a widely used testing protocol: the EuroNCAP CBNA50 test for the validation of AEBs developed to reduce accidents involving vulnerable users. The experiment foresees the re-analysis of the use-case in order to identify the factors that may have significant effects on the test results and compare them with those actually tested in the official protocol. In addition, it was analyzed which values of the factors configures the most difficult conditions to pass the test (worst cases) and will be compared with the values tested in the official protocol.

In order to improve the exploration and analysis, it was decided to adopt the calculation of the sensitivity indices indicated by Saltelli et al. (2010)[103], which uses the Sobol sequence to optimize the exploration of the test space.

Aim of the experiment due to research purposes: Refine the analysis and extraction methods and validate them by comparison with a known test

Aim of the test: Validation of the AEB as a tool to reduce vulnerable users accidents

Use-Case: Ego-vehicle activates AEB to avoid an impact with a bicycle

Scenario description: EuroNCAP - VRU - CBNA50. The ego-vehicle reaches an intersection while a bicycle is approaching from the right side (in two scenarios: not covered and covered by a visual obstacle). The ego-vehicle must brake to avoid impact.

Key Parameter: Distance between the front of the ego-vehicle and the point where the impact would have occurred without AEB intervention, recorded at the end of the emergency braking.

Scenario details: The EuroNCAP scenario takes place with the ego-vehicle driving in a straight line along a dry road, in full light and visibility, for a sufficient distance to reach the test speed. A bicycle (Figure 17) is driven with a constant speed, in straight line, orthogonal to the vehicle. The trajectories and the speed of the vehicles ensures that if the ego-vehicle doesn't activate the brakes, an impact with the bicycle will occur at a

[103] Saltelli, Annoni, and others, ' Variance based sensitivity analysis of model output. Design and estimator for the total sensitivity index', *Computer Physics Communications*, 2010.

predetermined point (in the center of the front of the car and in the center of the left side of the torso of the dummy on the bicycle). The experiment is performed at various speeds of the vehicle under test (VUT), and in two cases: with a visual obstacle placed at 3.55 m from the central axis of the VUT (X position in Figure 18) or at a distance of 17 m to consider it as not influential.

To analyze the potentially impact factors, we release all the constraints that did not distort the aim of the chosen scenario. In the following, the Shuldt-Bock analysis procedure is performed, the scenario structure is retraced, and the potentially impacting factors are released.

Figure 17: Dummy on bicycle used in the EuroNCAP test

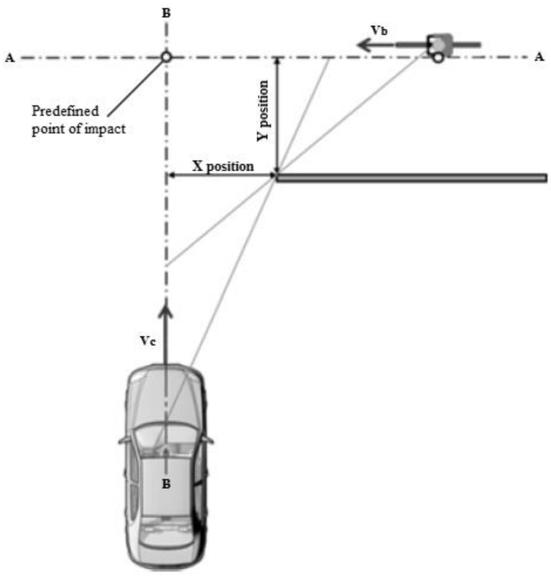

Figure 18: Op den Camp et al., trajectory scheme for the CBNA50 tests, with obstacle positioning to obstruct the radar field of view

Layer 1 (road geometry): The ego-vehicle travels in a straight line along a road and performs only the straight braking manoeuvre, so geometric parameters related to the width or number of lanes, the presence of side pavements or other are not significant. The longitudinal slope of the road, its regularity and the condition of the road surface could influence the experiment. The irregularity of the asphalt is simulated by means of a sequence of bumps varying in frequency and height.

Layer 2 (signs and static elements): the signs do not influence the experiment, but the positioning of the visual obstacle could be significant, so it will be varied by translating it in the plane for all the values that can influence the field of view of the AEB radar.

Layer 3 (temporary signs and works): the scenario does not include the presence of temporary signs or works in progress.

Layer 4 (dynamic elements): the ego vehicle travels on a road where bicycles also travel, so its speed may vary within the urban and suburban limits. The speed and size of the bicycle are variable factors, within a reasonable range compared to commercial vehicles.

Layer 5 (weather): Since the AEB logic is based on radar, light conditions are not significant. Wet or icy road conditions can affect braking, so they are considered in the grip factor.

Layer 6 (V2X): This experiment doesn't involve V2X devices.

Ten uncertainty factors, varying in the ranges indicated in the table, are therefore identified (Table 7).

Table 7: Experiment 2, Uncertainty Factors

Factors	Lower bound	Upper bound	To be found
Longitudinal Slope	-6	+6	meters per meter (m/m)
Friction	0.15	1.00	non-dimensional (-)
Road irregularity Intensity	0.00	0.05	meters (m)
Road irregularity Frequency	2	20	meters (m)
Ego-vehicle Speed	20	60	kilometers per hour (km/h)
Bike Speed	10	40	kilometers per hour (km/h)
Bike Length	1.40	2.00	meters (m)
Bike Width	0.50	0.65	meters (m)
X Position of the obstacle (Figure 18)	2.0	36.6	meters (m)
Y Position of the obstacle (Figure 18)	0.0	75.7	meters (m)

Implementation: The implementation of the chosen sensitivity analysis methods required a more complex and articulated intervention with respect to the previous experiment. It was necessary to integrate MatLab scripting, an emergency braking model built in SimuLink and co-simulated with CarMaker through the tool "CarMaker for SimuLink", supplied by IPG.

Three scripts were written in matlab, dividing the process into three phases:

1. Extraction of factor values and calculation of dependent parameters
2. Acquisition of simulation outputs, conversion into matlab format and calculation of key parameters
3. Performing sensitivity analyses

The first step makes use of the Matlab function LPTAU51 [104] for the quasi-random extraction of variables from 0 to 1. For each simulation, each factor takes on the value

$$f(x) = x(Ub - Lb) + Lb$$

In which: x extracted value corresponding to the factor for the specific simulation, Ub maximum value of the factor and Lb minimum value of the factor. After extracting the factor values and calculating the dependent parameters for each simulation, the script creates a text file containing the input values for the simulations in CarMaker.

Between phase one and phase two, there is an intermediate step: the simulation running. A tcl/tk script translates the text file containing the inputs into a CarMaker-readable file containing the sequence of simulations to be run. CarMaker runs the simulations and saves the results in time-series files in the software native format. During the phase two, a script in Matlab collects the simulation output files and converts them into a format that can be read by Matlab. From each simulation, the time series of ego-vehicle position values over time is extracted. The last position corresponds to the end of the emergency braking. The difference between the predicted impact position and the last position of the vehicle (vehicle front) is the key parameter extracted from the simulation. Each simulation is therefore examined through a single output value. Negative values correspond to the point of impact being exceeded, indicating that the AEB system activated late and was unable to avoid the hazard. Positive values indicate that the impact point was not reached and, therefore, that the AEB system was efficient. The continuous nature of the simulation output allows for a more in-depth analysis of AEB efficiency, including scatterplot analysis (Figure 19).

A total of 24576 simulations were carried out, in less than 40% of which the AEB intervened effectively, stopping the vehicle before the predicted point of impact.

[104] I. M. Sobol and others, 'LPTAU51', 1991.

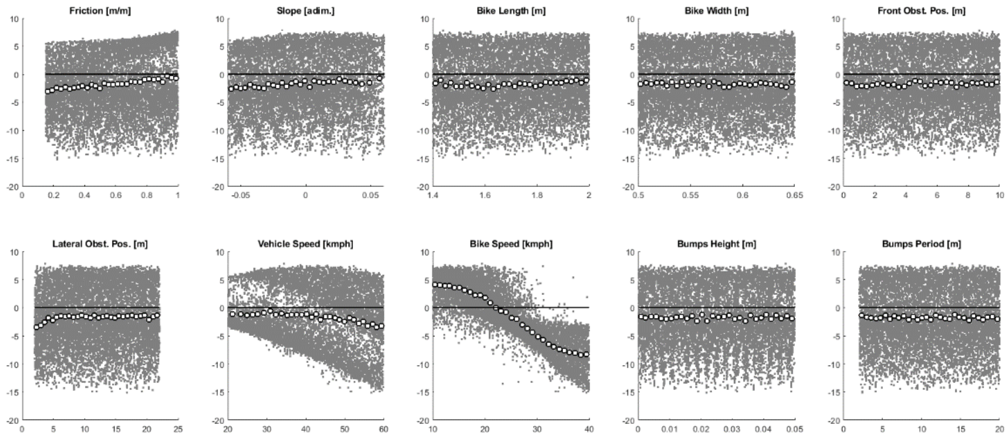

Figure 19: Experiment 2, scatterplots of the distance between the ego-vehicle and the pre-defined point of impact against the values of each factors for each simulation

From the scatterplots it is evident that the factor bicycle speed has a significant influence on the outcome of the tests, which are mostly passed for speeds below 20 km/h and almost exclusively failed for speeds above 30 km/h. Another factor that shows a significant influence on the variance of the results is the speed of the ego-vehicle, although the average of the results is little affected. The friction factor also shows a correlation with both the variance and the mean of the output, which is directly proportional to it. Finally, the lateral position of the visual obstacle seems to negatively influence the results only for a very narrow range of values around the minimum. The other factors do not show clear correlations in the scatterplots.

The calculation of the sensitivity indices confirms the significance of the speed of the bicycle above all other factors in the variance of the outputs, with a first-order value between 0.7 and 0.8 of the variance (Figure 20). Ego-vehicle speed and grip have similar first-order values, but below 0.1, but the significance of the ego-vehicle speed factor

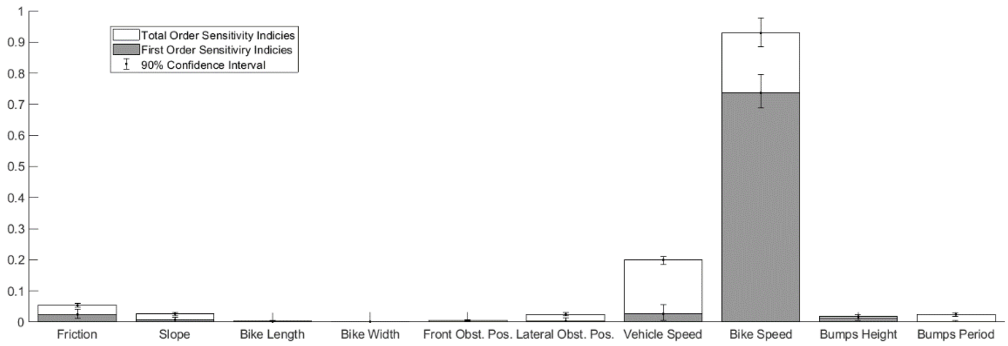

Figure 20: Experiment 2, summary of the sensitivity indices (first and total order) with the indication of the confidence interval calculated with the use of bootstrapping.

is way more significant in interaction with the other factors, showing a total order index that explains between 0.2 and 0.3 of the variance of the outputs.

To further investigate the analysis, an additional Boolean key-parameter was chosen to perform a regional analysis, specifically, the impact or non-impact between car and bicycle. The regional analysis further confirms the statements on the significance of the factors also through a different key-parameter and, in addition, allows the identification of threshold values between the simplest and the worst cases. Overlapping the behavioral histograms on the non-behavioral ones (Figure 21) it is possible to notice that the worst cases are those with: bike speed values higher than 22 km/h, lateral position of the visual obstacle lower than 4 m and grip lower than 0.4. We do not consider the slope as it showed negligible influence on the outputs, while the ego-vehicle speed does not identify clear thresholds in the regional analysis (presumably because its significance is mainly interactional).

The experiment performed shows the potential of the analysis tools used. The standard EuroNCAP test explores very few points in the analysed space. Despite the EuroNCAP test aims at testing worst cases, the values used in the tests are all within the behavioral threshold of the analysis outputs. In particular, in the EuroNCAP tests, the speed of the bicycle is fixed at 15 km/h, well within the threshold of the simple cases. Indeed, despite the tested AEB failed in more than 60% of the cases tested (Figure 22), the simulations performed with the EuroNCAP test values shows only successfully passed samples.

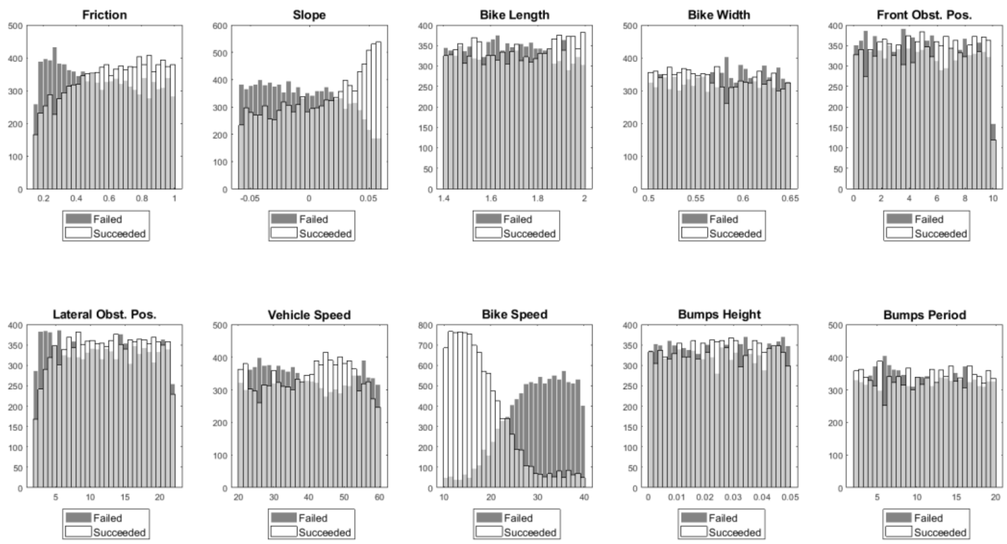

Figure 21: Experiment 2, regional analysis executed comparing histograms of the simulations in which accidents occurred against the accident-free simulations

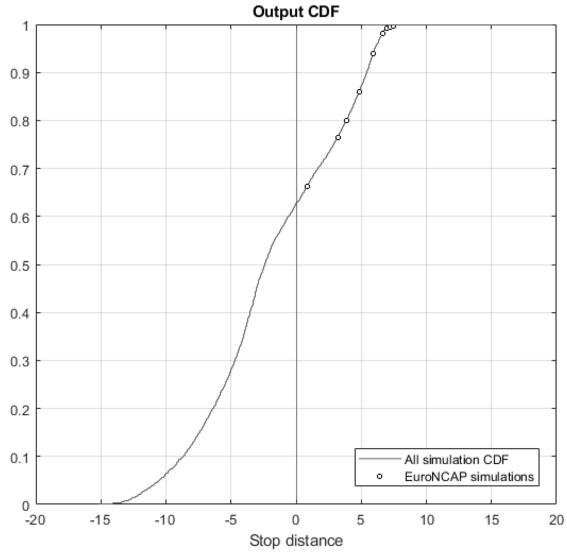

Figure 22: Experiment 2cumulative distribution function of the simulation outputs, with the indication of the outputs of the simulations done following the EuroNCAP standards

This experiment showed the effectiveness of the extraction method and the analysis performed. The choice of a continuous key-parameter allowed a clear identification of the conditional variance of the outputs with respect to the factors analysed. The regional analysis allows to identify the threshold values for the worst cases, given an identification of weak points in the EuroNCAP test. In terms of implementation, the scripts took some time to be written. The use of the Simulink-Carmaker co-simulation caused a significant slowdown in the execution of the simulations, which took several days to be completed.

4.4 Experiment 3: Reducing the computational weight

The following experiment uses the developed method on a scenario suggested by the company IPG Automotive (research partner).

Aim of the experiment due to research purposes: Verify the applicability of the developed tools methods in a business context, on a use-case explicitly requested by the car manufacturers.

Aim of the test: Identify the significant factors for the characterization of a testing scenario optimized for the specified use-case.

Use-Case: AEB in motorway cut-in on curve scenario

Scenario Description: The ego-vehicle is traveling in the passing lane, while a vehicle to its right changes lanes, facing the ego-vehicle, which is forced to brake to avoid impact

Key Parameter: time to collision (TTC), minimum distance between vehicles recorded during the simulation, impact occurred or not

Scenario details: the scenario requested by the company takes place on a motorway with two traffic participants, including the ego-vehicle and a generic car.

Layer 1: the road consists of two lanes with variable width and only two vehicles acting very specific manoeuvres. Since it is a motorway, there are no appreciable irregularities in the road surface. The radius of curvature is not specified, so it can vary from the minimum allowed by law to a maximum beyond which it is no longer considered a curve, due to the purposes of the regulations. The longitudinal slope is variable within the legal limits and the lateral slope varies according to the technical standards that relate it to the radius of curvature. An error in the lateral slope of 0.05 m/m is set as reasonable variation.

Layer 2: signs and other infrastructures do not influence the scenario

Layer 3: the scenario does not consider the presence of works in progress

Layer 4: the two vehicles represent all dynamic objects and can vary in speed, size and maneuver details. In particular, considering the two vehicles keeping constant speed until the maneuver change, they can vary from scenario to scenario keeping the speed of the ego-vehicle higher than the speed of the other traffic participant so that the cut-in maneuver and the consequent activation of the AEB takes place during the simulation. So, two speed factors are considered: the speed of the ego-vehicle and the speed difference with the other vehicle. The cut-in maneuver depends on two parameters: when it occurs and how long it lasts. So, two maneuver factors are considered: the distance between the vehicles at which the cut-in maneuver starts and the time it takes for the vehicle to change lane. In addition, the lateral offset of each vehicle with respect to the axis of the lane to which it belongs, and the three main dimensions of the traffic vehicle are considered as factors. The shape of the traffic car can influence both the maneuver and the perception of the radar used by the AEB, in combination with the radius of curvature and the slope of the road. The vehicle size was varied with reference to lower and upper limits of commercial cars.

Table 8: Experiment 3, Uncertainty factors

Factors	Lower bound	Upper bound	To be found
Friction	0.5	1	non-dimensional (-)
Curve Radius	964	4820	meters (m)
Lane Width	3.4	3.85	meters (m)
Longitudinal Slope	-12	12	meters per meter (m/m)
Lateral Slope Error	-0.05	0.05	meters per meter (m/m)
Traffic Vehicle Width	1.3	2	meters (m)
Traffic Vehicle Length	3	5.7	meters (m)
Traffic Vehicle Heigth	1.3	1.9	meters (m)
Traffic Vehicle Lateral Offset	-1	1	meters (m)
Vehicle Speed Difference	3	30	chilometri orari (km/h)
Distance between vehicles at which the Lane-change starts	0	50	meters (m)
Manovre Time	1	8	seconds (s)
Ego-vehicle Speed	50	130	kilometers per hour (km/h)
Ego-vehicle Lateral Offset	0.0	75.7	meters (m)

Layer 5: as for the EuroNCAP CBNA50 scenario, weather and light conditions are not considered relevant because they do not influence the perception of the radar. Wet or icy asphalt conditions are considered within the grip factor.

Layer 6: no devices using V2X technologies are tested, so the information related to this layer does not influence the test results.

The complete list of the identified factors and their relative variation values is indicated in Table 8.

For this experiment there was no important implementation differences with reference to the previous experiment. The MatLab scripts carried out the three phases of data extraction, data conversion and sensitivity analysis. A logic embedded in IPG CarMaker was used for the AEB, which allowed the simulation process to speed up.

Results: In order to stabilize the sensitivity indices, it was necessary to carry out a number of extractions equal to 2^{11} which, considering 14 parameters, required the simulation of 32768 scenarios. The results obtained from the key-parameters minimum distance and TTC offer a very similar picture, which can be seen both from the scatterplot and from the comparison of the sensitivity indices (Figura 23, Figura 24). The factor that most describes the variance of the outputs, both in the first-order effects and in the interaction effects, is the speed difference between the vehicles. As the speed difference increases, both the TTC and the minimum distance tend strongly to decrease on average and pushing the variance towards the lowest values, configuring the most dangerous conditions. The total sensitivity index of this factor explain about 80% of the variance of both outputs, while the first-order sensitivity index is the highest. The second most significant factor is absolute vehicle speed, with a total sensitivity index of about 0.5. In the scatterplots, the outputs graphed against this factor show a trend that may seem counterintuitive. The values of TTC and minimum distance, in fact, increase as the speed of the vehicles increases. In order to understand this phenomenon it is necessary to keep in mind how the simulated scenario has been built. The speed of the two vehicles does not vary independently, so that higher speeds of one vehicle correspond on average to higher speeds of the other vehicle (within the variations determined by the speed difference factor). It is therefore necessary to read the graphs without considering the speed difference factor and aware of the link between the speeds of the two vehicles. It is therefore possible to state that as the absolute speeds of the vehicles increase, the variance of the outputs tends to increase in the direction of the maximum, dragging with it the trend of the average values.

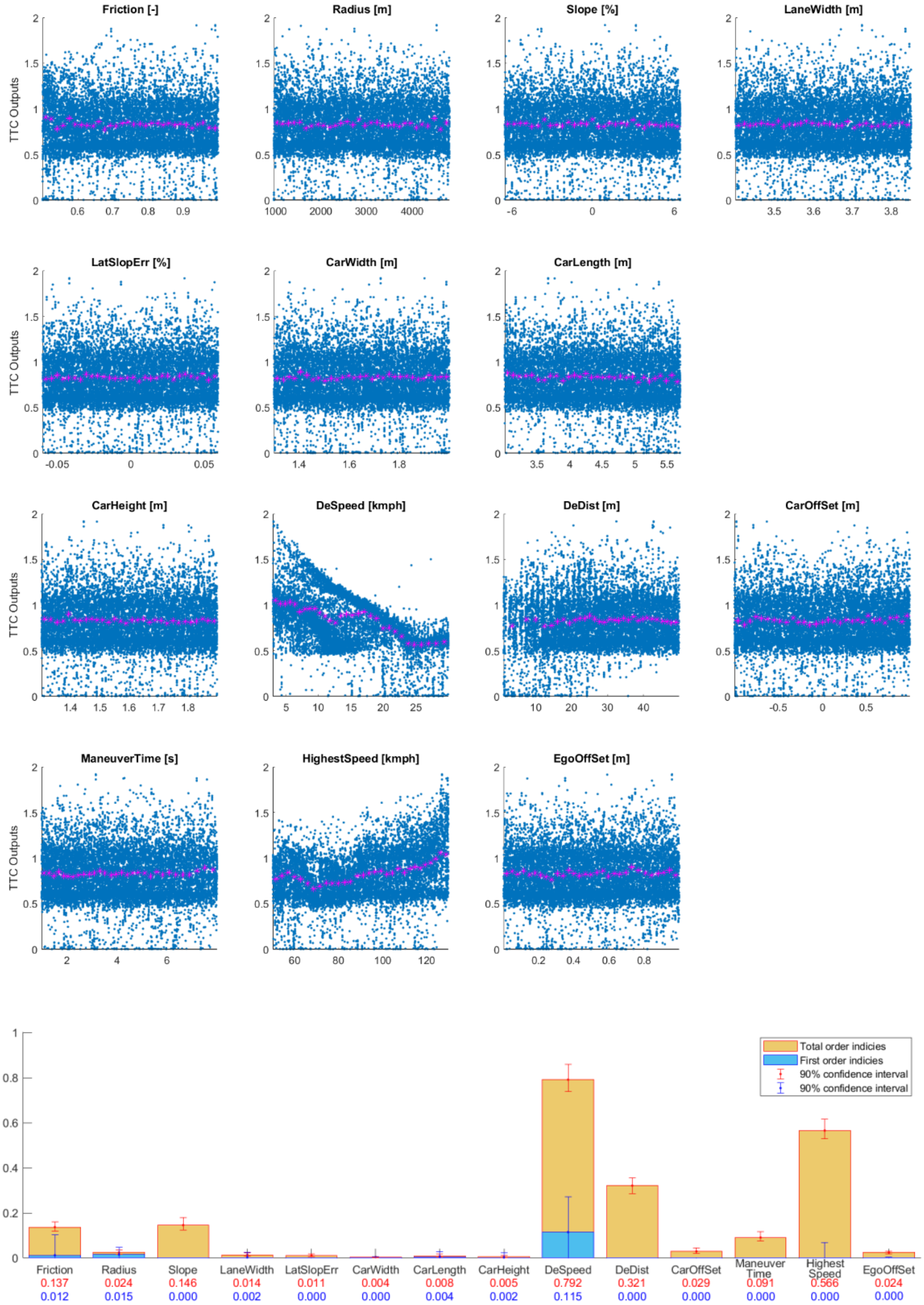

Figura 23: Experiment 3, scatterplot and summary of the sensitivity indices calculated with reference to the time to collision

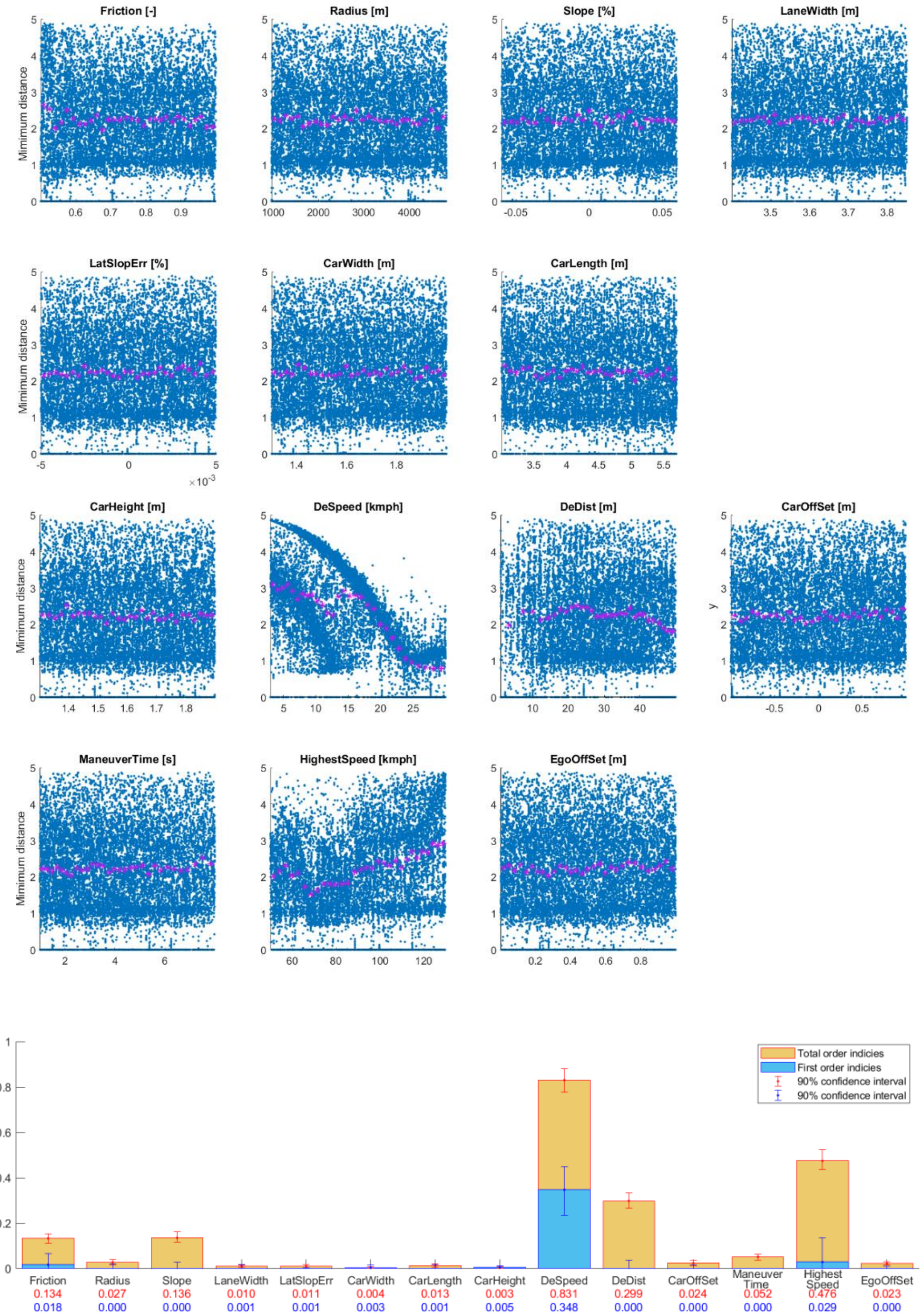

Figura 24: Experiment 3, scatterplot and summary of the sensitivity indices calculated with reference to the minimum distance between vehicles measured during each simulation

The distance between the vehicles at which the cut-in maneuver starts (DeDist) is the third factor in order of influence on the simulation output. This factor is much more relevant than the time of execution of the maneuver. Other significant factors are grip and longitudinal slope. From this analysis it is possible to state that, for the category of scenarios tested, the geometry of the vehicle has no influence on the effectiveness of the AEB for this use-case. Similarly, we can state that within the category of the simulated scenarios, for the ranges of values tested, the effects related to the factors vehicle offset, lane width, radius of curvature and side slope can be neglected.

The analysis carried out make it possible to define that, in order to adequately test this AEB logic, the experiments for this category of scenarios have to consider the variability of the conditions related to speed difference, absolute speed and a factor identifying the instant when the lane change occurs. Further analysis can explore the effects of longitudinal slope and adhesion, but it is not necessary to explore other scenarios produced by the variation of other factors. The size of the vehicle performing the lane change, the curvature of the road or the time of the lane change maneuver are not relevant.

Conclusions: The experiment showed how the space of tests can be reduced downstream of the analyses performed. Although the uncertainty was originally identified in 14 parameters, the analyses showed that the exploration of only 5 of them allows adequate confidence in the problem.

The experiment did, however, show some limitations of the simulation software. Observing the scatterplots relative to DeDist it is possible to notice a slight rarefaction of the graphed points for the lowest values of the factor. Given that the extractions of the values are performed uniformly, the only reason for the manifestation of gaps are simulation errors, which are discarded from the analysis. This means that the lowest values of this factor are mainly responsible for the occurrence of errors in the simulation software. This constitutes a problem for conducting this kind of analysis because, when weighing the influence of the factors on the variance of the outputs, some areas of the explored problem space are not considered due to simulation errors. This does not have a significant effect when simulation errors are also homogeneously distributed in the explored space. In this case, on the other hand, it is possible to identify areas of error depending on one single factor or some values of it, leading to a false estimation. By identifying the type of error, it is possible to operate in the simulation to avoid it or try to limit its occurrence. In this particular case, it was noted that for low DeDist values, especially in combination with low grip values, the ego-vehicle ended up off the road. IPG Carmaker, in fact, considers the cases where the ego-vehicle ends up off the road among the simulation errors. This may have had the effect of underestimating the significance of the DeDist and grip parameters. Considering that the phenomenon was limited and that the factors in question were already considered significant, it was considered unnecessary to reconfigure the experiments and repeat the analysis.

Cap.5: Proof of concept: project MEDIATOR

The European MEDIATOR project, carried out by a research consortium including FCA, is an excellent testbed to verify the scenario analysis methodology produced and, above all, to test its applicative versatility. In short, the project focuses on identifying the requirements for a human-machine interaction/interface system during the transition between human driving and partial or full automated driving, optimizing safety conditions. The focus is on the transition between automation and non-automation conditions (and vice versa), referring automation not to specific driving logics but rather to generic SAE scale levels and conditions of different driver involvement. For this reason, the testing scenarios required for this experiment are on a high level of abstraction, therefore, the design objectives are particularly appropriate to challenge the methodology.

The tools developed in the research followed are verified and applied within the partnership with FCA Automobiles. The company's interest was particularly focused on the methods of scenario creation and identification of significant factors, suggesting the collaboration on the MEDIATOR project.

The experiments conducted in the previous chapter explored and fine-tuned vehicle dynamics simulation tools, various sensitivity analysis tools and extraction methods. In order to carry out simulations that explore and analyze LoReS, a traffic and maneuver management by continuous parameters is as necessary as all the other factors analyzed. When the space of the scenarios descends from a very specific use-case, the number of vehicles and their maneuvers are also very specific, but in the cases that this research intends to explore, the problem has much more released constraint. To face it, we need to release it from the number of vehicles and the specification of maneuvers. As LoReS are characterized by simple road sections, dedicated to the characterization of a single arch or edge, it is possible to obtain an adequate traffic simulation even through the only vehicle dynamics simulation tool if it is assisted by an appropriate control of macroscopic and microscopic traffic factors. This can be achieved through a calculation software that adequately configures the initial conditions of the simulation (number of vehicles, type, positions, speeds and routes) and a realistic driver model that manages the behavior of the vehicles during the simulation. This ensures a consistent behavior of the simulation environment. To test this method, the driver logics offered by the CarMaker simulation software are used for this experiment. We highlight that these models are not the subject of this research and can be replaced if they are deemed inadequate.

5.1 Context and main objectives of the MEDIATOR project

The MEDIATOR project is part of the Horizon 2020 European research and innovation program, run by the consortium formed by the Dutch research institutes SWOV and NLR and the Swedish VTI, the technical universities of Delft, Chemnitz and Ben-Gurion and various industrial partners, including FCA Automobiles. The aim of the project is to realize interaction and transition systems between human driving and automation systems at various levels, to ensure a safe and comfortable transition, while respecting the SOTIF principles of reducing the risk of incorrect or misunderstood use of assistance and automation devices. The acronym MEDIATOR is the contraction of "MEdiating between Driver and Intelligent Automated Transport systems on Our Roads", using the description provided by the official project website:

“MEDIATOR will develop a mediating system for drivers in semi-automated and highly automated vehicles, resulting in safe, real-time switching between the human driver and automated system based on who is fittest to drive.

MEDIATOR pursues a paradigm shift away from a view that prioritises either the driver or the automation, instead integrating the best of both.”[105]

As can be seen from the above description, the mediation system approach is based on a principle of prioritize driving between human and automation systems, assessing the cases in terms of both external context and driver performance and comfort, making use of cognitive systems to monitor the driver's state. The deliverables show a focus on the systems and devices available for driver interaction, control, and assistance, focusing mainly on the human-machine interface, regardless of the interactions between vehicle and external context[106].

These studies distinguish three use-cases of the mediation system: Continuous Mediation (CM), where the driver has to be vigilant all the time even if assisted by automation systems (SAE levels 1 and 2), Driver Standby (SB), where the driver is allowed to be distracted for short periods of time, provided he/she is always ready to regain control of the vehicle (SAE levels 3 and 4), and the long out of the loop (LOotL) case, where the driver may leave the driving of the vehicle to the automation for long periods of time and be called upon only when necessary (other activities included in SAE level 4).

---

[105] ‘Home | Mediator’ <https://mediatorproject.eu/> [accessed 15 July 2021].

[106] A. Christoph, M., Cleij, D., Ahlström, C., Bakker, B., Beggiato, M., Borowsky, *Mediating between Human Driver and Automation : State-of-the Art and Knowledge Gaps*, 2019; B. Borowsky, A., Oron-Gilad, T., Chasidim, H., Ahlström, C., Karlsson J.G., Bakker, *Behavioural Markers for Degraded Human Performance Behavioural Markers for Degraded Human Performance*, 2019.

Based on this hierarchy of cases, distinguishing situations by hazard level, ten use-cases are identified, subsequently placed in different generic driving contexts (urban, suburban and motorway).

The cases, as described on the official project website[107], are the following:

UC 1: MEDIATOR initiated take over Human > Automation:
MEDIATOR detects degraded human fitness caused by A). drowsiness B). distraction, and initiates a forced take over to automation.

UC 2: Automation -> Human: Driver takes back control:
the human driver inidcates a desire to take back control via the HMI.

UC 3: Comfort take over Human -> Automation:
A). Driver initiated: driver is not motivated to drive and indicated a preference for automation to drive via the HMI.
B). Mediator initiated: Mediator detects an event (such as a text message or an upcoming traffic jam and used historical date to conclude that the driver would likely want to hand over control. Mediator proposes the Human --> Automation take over.

UC 4: Corrective Mediator action during standby:
the human gets drowsy while expected to be on standby.Mediator tries to improve the driver fitness and monitor the effect.

UC 5: Mediator initiated take over Automation -> Human:
A). Planned: the automation communicates that the current route will leave the ODD within the next seconds.
B). Unplanned: the automation communicates that its reliability is degrading rapidly and the human should take over within seconds. Mediator informs the human and guides an urgent take over.

UC 6: Comfort CM switch on:
A). Driver initiated: Human is not motivated to drive fully manually and indicated this via the HMI.
B). Mediator initiated: Mediator detects reliable automation and uses historical data to conclude that the human likely preferes to activate partial automation.

[107] 'MEDIATOR behind the Scenes: Use Cases | Mediator'.

UC 7: Prevention – CM Keep the driver in the loop:
While driving with L2 automation, Mediator tries to prevent underload of the human drive and keep him/her in the loop by providing an active task. What this task will entail is one of the research questions.

UC 8: Corrective – CM Get the driver back into the loop:
While driving with L2 automation drowsiness or distraction is detected. Mediator initiates a correction action such as a voice message to get the driver back in the loop.

UC 9: CM shuts off immediately:
While driving with L2 automation the road markings degrade and Mediator indicates L2 will shut off immediately.

UC 10: Smooth transition from Long Out of the Loop to Stand By:
The driver is fully out of the loop while driving on the highway with L4 automation when the route is approaching a highway exit. Mediator informs the drive that the standby mode (L3) will be switched on and monitors the driver fitness for this standby task.

5.2 Research contribution to the project and case study

The MEDIATOR project represents a challenging study case due to the methodology proposed. In fact, the MEDIATOR project does not identify a specific automation logic to be tested, since the device to be tested does not represent a control logic per se, but a human-machine interface that, potentially, can be applied to any scenario consistent with the context to which the use-case belongs. The method developed in this research has been used as a basis for answering two questions: i) which factors are the most significant for the success of a safe transition between automatic and non-automatic driving; ii) in which way the scenarios to be investigated by the MEDIATOR project should be more detailed, once the interfaces have been developed (and coupled with specific control logics). The use-cases addressed in the MEDIATOR project can be distinguished into two kinds: driver change and corrective actions.

In turn, driver change can be distinguished into cases where the human driver gives up control to the automation (voluntarily, suggested or forced) and cases where the automation system gives up control to the driver (requested from the human driver or the automation logic).

The analysis method presented can be applied to the specific case by reconstructing, in the appropriate scenarios, the conditions of driver change or variation of vehicle control levels, through the identification of the factors that characterize this type of event.

Since the LoReS matrix identified for testing in the design phase (Table 5) is based on the identification of driving activities, the control logic must be characterized in some way. To this end, we started from the identification of the level of automation on which the company requested a focus, as the only data regarding the logic provided by the use-cases of the project.

Therefore, we identified the activities related to level 2 and 3 of automation (Figure 1). These levels of automation include limited activities that take place in specific contexts, so it was decided to consider the case of automation in a motorway context.

As a case study, it was chosen to analyze the use-cases considered to be the most risky, so the focus of the study was on the transition from automation to the human driver due to a certain need; this is the case of the MEDIATOR 5, 9 and 10 use cases.

In the matrix of LoReS identified, the motorway context belongs to the first category of uncertainty. The driving activities considered are all those that can be carried out from the moment after entering the motorway until the moment before the exit. The seven LoReS falling within the identified activities are shown in Figure 25.

Context Unchertainty level	Scene Set / Activities		Free Flow	Avoid accident *(obstacle)	Man at work **(bottleneck)	Intersection crossing	Approach to congestion
	Generic set	Subset					
Low:	Streight and curve	Simple	x	x	x		x
(Highways)		Exit road	x	*(simple)*	*(simple)*		*(simple)*
		Input road	x	*(simple)*	*(simple)*	x	*(simple)*
CdS Type:		Emergency stop	x	*(simple)*	*(simple)*	x	*(simple)*
A - B (part)	Special areas	Toll booth		*(appr. to a s.p.)*	*(appr. to a s.p.)*		*(appr. to a s.p.)*
		Service area		*(appr. to a s.p.)*	*(appr. to a s.p.)*		*(appr. to a s.p.)*
		Parking area		*(parking)*	*(parking)*		*(parking)*

Figure 25: LoReS identification starting from the testing design LoReS matrix (Table 5)

As an example, the first case, from the straight simple scene, was developed for the free flow driving activity.

The low-resolution scenario represents the case in which the ego-vehicle is in a straight stretch of a motorway, dropped in a hypocritical traffic, and it is driven by a control logic that, for a reason due to reaching the limits of its functionality, can no longer guarantee a safe driving of the vehicle. Therefore, the MEDIATOR asks for the intervention of the human driver. The human driver may be ready to take control of the vehicle or take a few seconds to do so.

Given the generality of the MEDIATOR use cases and not having any information about the type of logic tested, we cannot make detailed assumptions about the reasons for reaching the functionality limits, so the implemented LoReS represents the most generic condition possible. For these reasons, the kind, the number and the range of variation of the factors span all related MEDIATOR use-cases, investigating the major risk factors in the context analyzed.

5.3 Experiment description

Aim of the experiment due to research purposes: Demonstrate the applicability of the proposed method by carrying out an experiment involving at least one LoReS in its entirety, exploring it appropriately, drawing useful conclusions for testing purposes, identifying the problems and possible solutions.

Aim of the test: the experiment concerns the analysis of the phenomena linked to the MEDIATOR use case that occur in the LoReS identified. It includes to identify the significant factors, the behavior of the model as they vary, the worst cases and the addresses for the in-depth study of the case study for subsequent levels of testing, as well as immediately highlighting suggestions on the critical aspects that a man/machine interface should address with priority attention.

Use-Case: transition from an automatic driver to a human driver. Use-case MEDIATOR 5 (SB to human) and 9 (CM to human).

Scenario description: LoReS 001, motorway context, simple section, free flow. The ego-vehicle is driven by an automatic driver that reaches its operating limits and therefore requires the intervention of the human driver. The human driver responds to the request according to his own levels of readiness and accuracy.

Key Parameter: minimum recorded distance between the ego-vehicle and other vehicles, number of collisions.

Scenario details:

Layer 1: the motorway context allows little variability on geometric factors. The factors varied are the friction, the longitudinal slope, and the width of the lanes. For sake of simplicity, we decided to consider the curved section case as a separate LoReS, so factors such as curvature and side slope will not be varied.

Layer 2: The nature of the experiment does not specify which ADAS logics are mounted on the vehicle nor which sensors or logics interact with the infrastructure, so vertical and horizontal signage is not considered in this experiment.

Layer 3: The LoReS being analysed does not include any works in progress or temporary signage.

Layer 4: the vehicles involved must be able to vary under all the realistic traffic conditions that the LoReS envisages. The motorway context provides numerous restrictions on the type of vehicles involved in the scenario. Four types of vehicles were considered for the experiment, to cover most of the differences in terms of size and influence on traffic. These are: motorbikes, cars, vans, and road tractors. In order to vary independent and continuous factors that characterize the traffic, some values, such as vehicle speed, are stochastically extracted from a normal probability function, characterized by two factors: mean and variance. Therefore, the factors analysed to vary, relative to traffic, are: mean and variance of traffic speeds, mean and variance of vehicle lengths (also used as a proxy for vehicle type), mean density of lanes, percentage of vehicles afferent to the ego-vehicle lane, percentage of remaining vehicles afferent to the lane closest to the ego-vehicle, variance of vehicle lateral offset, driver aggressiveness. For the ego-vehicle, the factors of speed, lateral offset and three driver-related factors are analyzed as variance: attention (reaction times expressed through pedal change, gear change and maneuver execution times), performance (longitudinal and lateral accelerations) and safety (safety distances and overtaking probability).

Layer 5: considering that the driver models used by the CarMaker software do not change their behaviour in reaction to different weather and lighting conditions, these elements are not varied in the current experiment.

Layer 6: For the same reasons as listed for layer 2, V2X devices and digital information are not considered for this scenario.

5.4 Implementation:

In order to reconstruct the use-case, it was necessary to program a sequence of events that resulted in a driver change. The ego-vehicle starts the simulation in a set of initial conditions representing the instant in which the automatic driver gives the vehicle to the human driver. For each simulation, a few simulation seconds have been passed maintaining the initial conditions in order to allow the simulated traffic drivers to recognize the context in which they are placed, so that they can assume realistic behavior. After this initialization time, the events were observed. To face the lack of driving logics to be tested, the embedded drivers of the CarMaker software were used, both for the ego-vehicle and to the other traffic vehicles involved.

For the experiment it was necessary to calculate all the characteristics of the traffic objects involved in the scenario, for each extracted scenario sample. A script in MatLab was therefore created to generate the traffic participants.

In order to speed up the execution of simulations in CarMaker, it was created a single test file that can vary its parameters by configuring all the scenarios to be simulated. To do this, it was necessary that this generic scenario model contains all the vehicles to be simulated in all possible scenarios. To associate the values with the CarMaker vehicles, it was therefore necessary to create a scenario model that included the maximum number of vehicles that could be generated by the MatLab script for each vehicle type. Five values are associated with each vehicle: longitudinal position with respect to the lane, lateral offset, afferent lane, speed and vehicle length. Vehicles not used in the simulation assume a standard size, zero speed and a position unrelated to the paths on which the scenario takes place. Regardless of the number of vehicles in the specific simulation, the vector of simulation parameter values always has the same length, defined by all road geometric parameter values, ego vehicle parameter values, driver parameter values and traffic parameter values. The latter are five per maximum number of vehicles in the densest simulation multiplied for the number of vehicle types simulated. There are four vehicle types simulated, so for each vehicle added to the simulations, the vector of parameters defining the scenario is lengthened by 20 elements that must be computed for each simulation.

To put vehicles in the scenario and run the simulation, a simple solution would be to generate vehicles from the beginning to the end of the simulation, inserting the ego vehicle into the simulation from the instant in which the entire simulated road section has reached the desired traffic conditions. This approach, however, would produce vehicles that do not interact with the ego-vehicle, some because they will exit the simulated road section before the ego-vehicle can reach them, and others, conversely, will not reach the ego-vehicle before the simulation end. This approach would make the computational problem unnecessarily heavy as well as impossible to handle with the available tools.

The only vehicles that are strictly necessary for the experiment are those that reach or are reached by the ego-vehicle (or its sensors) during the simulated time.

The simulated road section in the scenario can be divided into two stretches (distances): a densification distance within which vehicles are generated and an escape distance within which generated vehicles travel without further generation (Figure 26). With reference to the average speed of all vehicles (except the ego-vehicle), two cases can be distinguished: the case in which the ego-vehicle speed is higher than that of the traffic vehicles, and the case in which it is lower. The escape distance can be defined, therefore, as the space covered by the slowest vehicle in the time it takes the fastest vehicle to cover the entire simulated section. In the first case, the length of the simulated distance will be just the distance covered by the ego-vehicle in the simulation time.

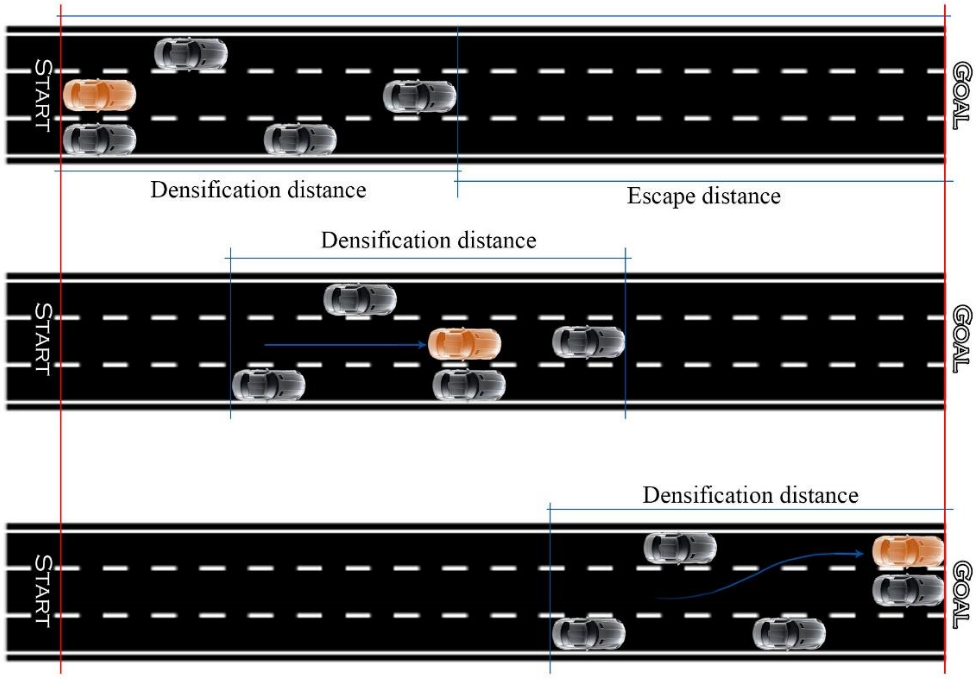

Figure 26: Densification Distance and Escape Distance

In this case, the ego-vehicle will only interact with the vehicles in front of it, and the length of the escape stretch will be the distance that the traffic vehicles can travel in the simulation time (Figure 26). Therefore, the length of the densification stretch will be equal to:

$$D_{dens} = D_{rif} - D_{fuga}$$

$$D_{dens} = D_{rif} - V_{tra}T$$

with D_{rif} reference distance (total lenght of the entire simulated road section), V_{tra} average traffic speed and T the simulation time.

Given that, in this case: $D_{rif} = V_{ego}T$

We can say that:

$$D_{dens} = V_{ego}T - V_{tra}T$$

$$D_{dens} = T\Delta V$$

In the second case, the ego-vehicle will only interact with vehicles coming from behind it, so the distance travelled by the ego-vehicle in the simulation time will define D_{fuga}, whereas the distance travelled by the generic traffic vehicle in simulation time will define the reference length D_{rif}

$$D_{dens} = V_{tra}T - V_{ego}T$$

$$D_{dens} = T\Delta V$$

It means that, regardless of the case, the length of the minimum section to be filled with vehicles for the ego-vehicle to interact with the desired traffic conditions for the duration of the simulation is equal to the simulation time for the difference in speed between the ego-vehicle and the traffic.

Knowing the length of the stretch to be densified and defining the traffic density per lane, the number of vehicles to be generated for each lane is also known. Once the number of vehicles is defined, it is possible to extract for each of them all the characteristic values from the probability distribution functions, built according to the mean and variance factors extracted for the simulation. The values of length, speed and lateral offset of each vehicle are thus defined. As mentioned before, the Sobol sequence is used to extract values for the simulations (other methods can be used). Considering that the dimensions of the carriageway vary according to the width of the lanes, the lateral offset of the ego-vehicle was varied from 0 to 1, as a non-dimensional value that identifies the position of the ego-vehicle in the carriageway, also defining the lane to which it belongs. The other vehicles are distributed following the density of each lane, so the lateral offset is varied as a non-dimensional value indicating the position within the reference lane.

Their longitudinal position follows as a reference the gap calculated through the average values of the simulated traffic.

$$Gap = Headway - Occupancy$$

$$G = \frac{D_{dens}}{V_{tra}N_i} - \frac{L_{tra}}{V_{tra}}$$

With N_i number of vehicles in the i-esim lane, and L_{tra} average of the vehicle length.

Placed the leader vehicle, the next one is placed behind it, at a distance equal to its speed for the gap considered for the simulation, and so on until the last vehicle. In order to ensure that the ego-vehicle interacts with the desired traffic conditions, an additional vehicle is generated on the "clearing" side from other traffic vehicles (behind the ego-vehicle if other vehicles are ahead, in front of the ego-vehicle if other vehicles are behind).

In this way, all the values to be associated with all the vehicles are obtained. Once the matrix of scenario vectors had been compiled, it was possible to produce the txt file that can be read by the script in tcl/tk, created previously in order to produce the test series file in the native format of CarMaker. Given the number of vehicles to be programmed, the vehicle defining process within the CarMaker scenario GUI tool would be unreasonably time consuming. So, a further script was created in MatLab to compile the part relating to the traffic vehicles in the creation of the generic scenario CarMaker file.

5.5 Results:

Performing 2^{11} extractions, 38912 simulations were carried out, after which the outputs were extracted and the key-parameters were calculated, allowing the calculation of the indices. The results show a strong relevance of vehicle speed, both with regard to ego-vehicle speed and traffic speed. In comparison, the indices calculated for the minimum distance recorded between the ego-vehicle and other vehicles during the simulation and the indices calculated for the number of collisions are very similar, especially for the total order indices (Figure 27). For the two speed factors, it is possible to note the significant weight of the total-order indices in comparison to the first-order indices. This may indicate that the high significance of these indices in the experiment is due to a direct interaction between the two factors. Analysing the scatterplots for the minimum distance output, we notice a non-monotonic trend of the averages of these two factors, which increase with the increase of the speed value up to a peak around 80 km/h, and then decrease (Figure 28). This scatterplots, together with the sensitivity indices recorded, suggests that the significant factors are not the absolute speeds, but the difference in speed between the ego-vehicle and the average traffic. A scatterplot comparing the speed difference to the minimum distances recorded during the simulations shows a trend in average distances that decreases sharply as the speed difference between the ego-vehicle and average traffic speed increases (Figure 29). The scatterplot (Figure 28) shows some vague link between the trend of the averages of the output values and the traffic densities afferent to the specific lanes, but not clearly distinguishable. Although the indices confirm a significant effect of vehicle densities on the variance of the outputs, the weight that the two speeds (traffic and ego-vehicle) have in interaction with each other could mask the effects of the other factors. Given that the speed of traffic flows is related to density by the traffic fundamental diagram, it was deemed necessary to repeat the experiments reconfiguring the factors to take due account of this kind of interaction.

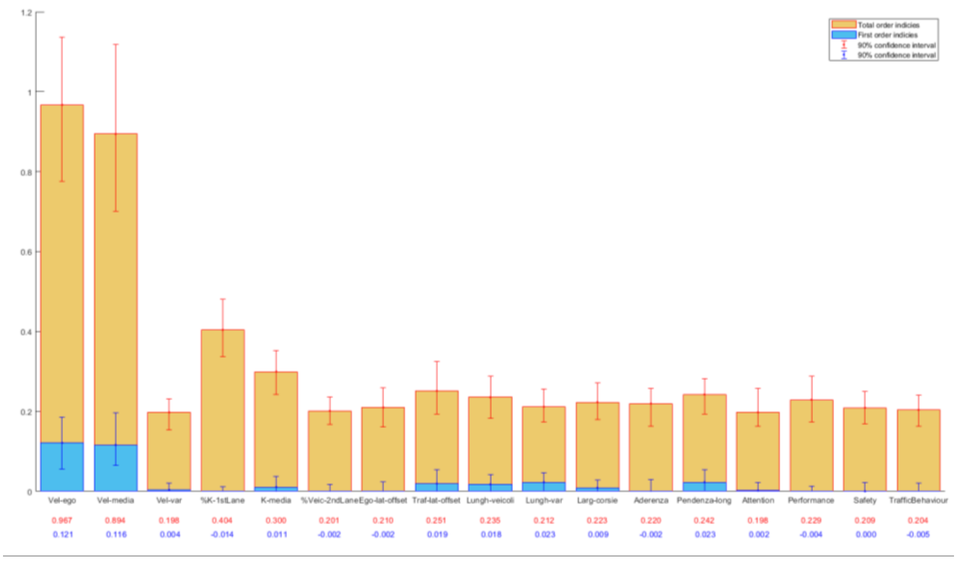

Figure 27: Sensitivity indices for the first MEDIATOR experiment, with reference to the minimum distance measured between the ego-vehicle and the other vehicles during each simulation

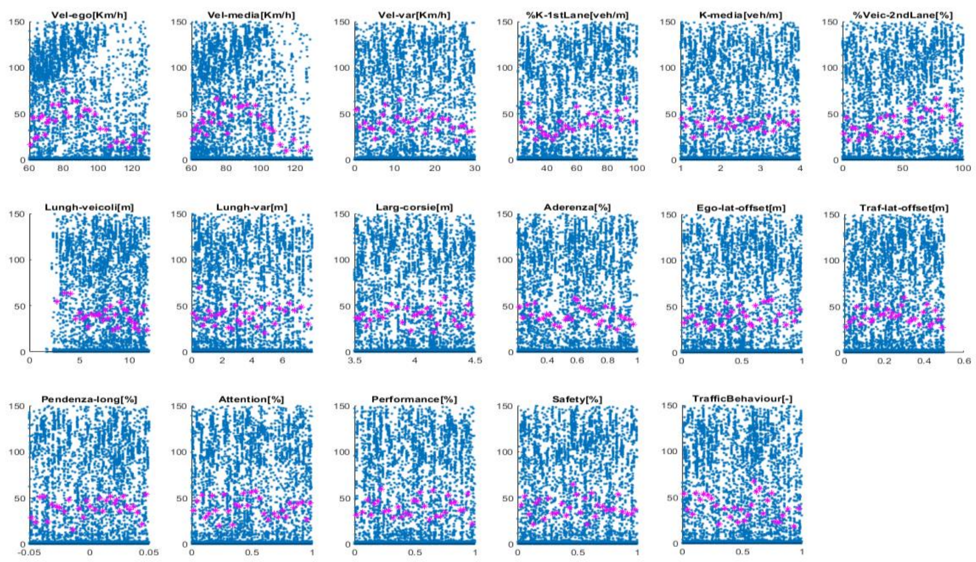

Figure 28: Scatterplot of the outputs of the first MEDIATOR experiment

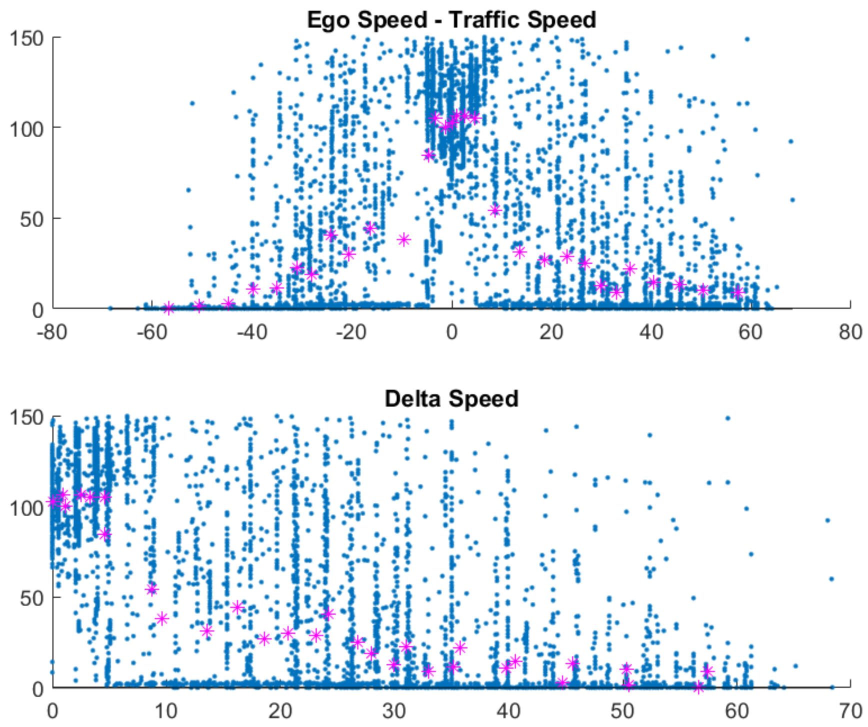

Figure 29: Scatterplot of outputs with respect to the speed difference between ego-vehicle and the traffic speed average.

A second experiment was carried out by achieving greater internal consistency of the traffic model, in particular by taking account, in the representation of the initial and boundary conditions of the experiment, of the flow relationship linking density and speed of traffic flows. The experiment was carried out by varying the average vehicle density in relation to the average traffic speed. Furthermore, traffic speed and ego-vehicle speed were not made to vary independently but were analyzed by varying the only speed difference factor. The factors thus constituted are summarized in Table 9. The LoReS analyzed represents a three-lane motorway, for which the Van Aerde Model was used, calculated according to values recorded on Italian motorways with similar characteristics[108]. Therefore, 16 factors were analyzed, and the results of the analysis performed on this new experiment show clearer dependencies. The scatterplots constructed on the minimum distance outputs recorded between ego-vehicles and other vehicles during the simulations show much clearer trends than the average values (Figure 30). The significance of the speed difference is still strong, producing greater distances for values around zero. The

[108] Andrea Pompigna, 'La Calibrazione Del Diagramma Fondamentale e La Valutazione Del Livello Di Servizio Operativo Sulle Autostrade Italiane' (Università di Bologna, 2017).

percentage of vehicles afferent to the ego-vehicle lane also shows a significant trend, showing a strong decrease in the values of the average distances as the value of the factor increases. The indices (Figura 31) do not show strong differences in the significance of the total order factors, varying from the maximum value of 0.59 (percentage of vehicles in the ego-vehicle lane) to the minimum 0.25 (longitudinal slope), while the first order values vary more (from 0 to 0.15).

Table 9: MEDIATOR Experiment 2, Uncertainty Factors

Factors	Lower bound	Upper bound	To be found
Speed difference between Ego-vehicle and Traffic flow average	-40	40	kilometers per hour (km/h)
Traffic flow Average Speed (Traffic Condition)	90	110	kilometers per hour (km/h)
Traffic flow Speed Variance	0	30	kilometers per hour (km/h)
Vehicle percentage on the Ego-vehicle lane	20	100	percentage (%)
Residual percentage of vehicles on the Ego-vehicle closest lane	0	100	percentage (%)
Ego-vehicle Lateral Offset	0	1	non-dimensional (-)
Traffic flow Lateral Offset Variance	0	0.5	non-dimensional (-)
Traffic vehicles Length Average	2	12	meters (m)
Traffic vehicles Length Variance	0	8	meters (m)
Lane Width	3.5	4.5	meters (m)
Friction	0.2	1	non-dimensional (-)
Longitudinal Slope	-0.05	0.05	meters per meter (m/m)
Ego-vehicle driver Attention	0	1	non-dimensional (-)
Ego-vehicle driver Reaction Time	0	10	seconds (s)
Ego-vehicle driver Safety attitude	0	1	non-dimensional (-)
Traffic flow drivers Safety attitude	0	1	non-dimensional (-)

The speed difference and the percentage of vehicles relative to the ego-vehicle lane are the most significant factors both in the first order (with respective values of 0.14 and 0.15) and in the total order (0.58 and 0.59). The traffic condition (average speed and traffic density) is not as significant and is similar to the least significant factors with a value of 0.3 total order index. Next, in order of significance, the percentage of vehicles relative to the lane closest to the ego-vehicle and the lateral offset of the ego-vehicle affect similarly both in the first order (0.09 and 0.07) and in the total order (0.43 and 0.41) presumably in interaction with each other. Adherence also shows similar values for both indices. Talking about the driver-related aspects, the safety distances seems to be the most significant, although only in interaction with other factors, with a total index of about 0.4 and a first order index of zero. The regional analysis carried out by distinguishing the simulations with collisions from those without collisions shows the significance of some factor values with respect to others in relation to the occurrence of collisions (Figura 32). The speed difference shows that the riskiest conditions occur for values far from zero. Specifically, the critical values are: speed differences greater than about 20 km/h when the ego-vehicle is speeding above the average traffic speed, and speed differences greater than about 10 km/h when the ego-vehicle is speeding below the average traffic speed. The percentage of vehicles in the ego-vehicle's lane configures a greater number of simulations with accidents, for values higher of 70%. An intuitive result concerns the lateral offset factor of the ego-vehicle, which shows the riskiest conditions when the vehicle across the lanes, leading to risky behavior by other vehicles that are forced to bypass the ego vehicle. Equally intuitive is the accident trend related to grip, which shows the most severe conditions for lower values and, in particular, below about 0.5. Less intuitive are the vehicle length indicators, both in absolute value and in variance. The trend of accidents increases considerably as vehicle length average decreases and variance increases, showing more dangerous conditions for motorbike and mixed traffic. The other factors do not show obvious trends, keeping the differences between the trends of the behavioral and non-behavioral simulations almost with a uniform distribution.

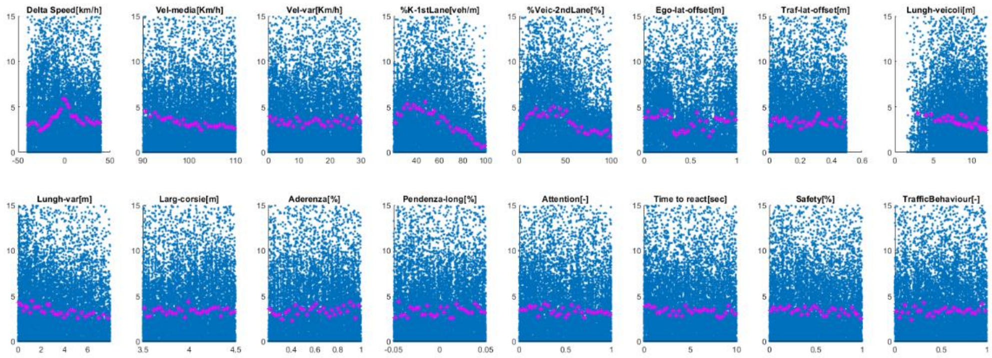

Figure 30: Scatterplot for the second MEDIATOR experiment.

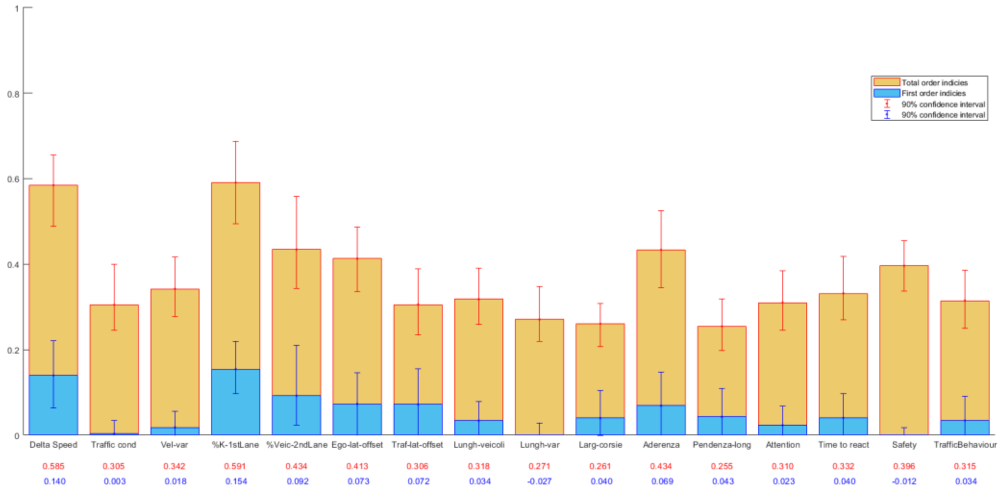

Figura 31: Sensitivity indices for the second MEDIATOR experiment.

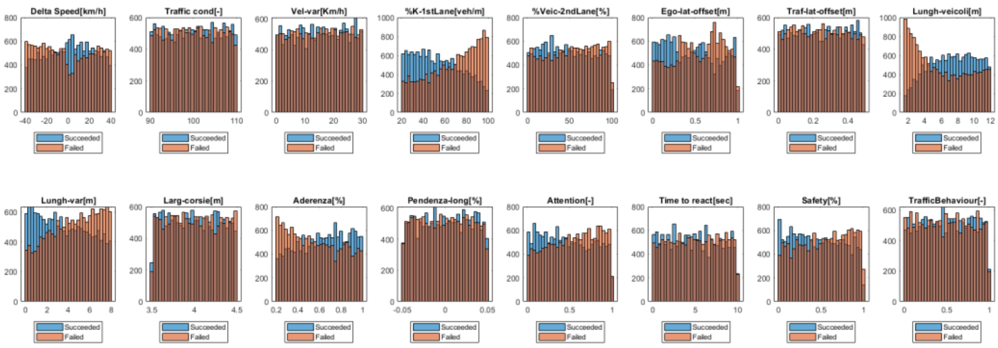

Figura 32: Regional analysis for the second MEDIATOR experiment. Simulations are divided in accident-free and accident occurred simulations.

Cap.6: Conclusions and discussions

6.1 Conclusions from the experiment results

The results obtained in the two experiments carried out within the MEDIATOR project shows how important is to take into account the traffic component and, moreover, that the correct identification of the factors relating to the macroscopic traffic conditions has a significant effect both on the simulations and on the analysis of the outputs, making the results clearer and more comprehensible. From the first analyses it became evident that, for the analyzed case, the phenomena are better described through the difference in speed between ego-vehicle and average traffic, compared to the absolute speeds analyzed separately. The first configuration of the factors under analysis, together with the varying behavior of the drivers of the traffic vehicles, did not ensure the correctness of the calculation and reading of the indices. This was due to the fact that some of the combinations of factors extracted determined initial conditions of speed and traffic density which were incompatible with each other and which, in the early simulation seconds, were varied by the traffic actors managed by the driver models, resulting in scenario conditions different from those extracted. The second set of experiments showed more balanced indices, which allowed a more conscious analysis of the phenomena. In particular, the speed difference determines more consistent simulation outputs and its significance is comparable to the quantities related to the vehicle density. Regarding the factors that manage vehicle density, the percentage of vehicles afferent to the ego-vehicle lane is more significant than the general traffic conditions on the whole carriageway. These results allow an understanding of the phenomena that determine the risk conditions for the ego-vehicle.

The scenarios analyzed represent the cases in which the control logic, for whatever reason, needs to abandon the control of the vehicle, requiring the intervention of the driver. The aims of the MEDIATOR project concern the realization of interaction models between the human driver and the automation, in order to facilitate that switch in safety and comfort conditions. Through the analyses carried out it is possible to define which are the least risky conditions in which to implement the driver change. For example, the analyses show that a decrease in vehicle speed does not trivially correspond to a reduction in risky conditions; on the contrary, the regional analysis (Figure 32) shows that the most onerous conditions are reached at lower speed differences when the ego-vehicle assumes lower than average traffic speeds. According to the results of the analysis, therefore, the least risky conditions correspond to those in which the ego-vehicle moves to the average speed of the traffic. This finding is an example of a suggestion that the analyses conducted can give to the human/machine interface, already at the stage of low-resolution scenarios: monitor the difference between vehicle speed and average speed of the traffic flow; try to keep the difference between the two speeds as low as possible and implement what is possible to release the automatic control of the vehicle to the human driver at the lowest possible speed difference. Other behaviors that reduce the risk of accidents include positioning the vehicle correctly in the center of the lane, choosing the lane with the lowest density of vehicles. In cooperative traffic conditions and with centralized control logics, it also seems preferable that separation conditions between different kinds of vehicles are maintained on the lanes as high as possible and that, in particular, the interaction of cars and motorbikes is reduced (although the significance of the factors of homogeneity and type of vehicles is relatively low compared to the difference in speed and density).

Regarding driver behavior, the most significant (although only in interaction with other factors) is the maintenance of safety distances. These have not been analyzed in detail, but the factor under analysis corresponds to an index linked both to distance and to the gap with the leading vehicle (between 2 m and 10 m and between 0.5 s and 3.0 s). The hierarchy of indices shows that the general conditions of the traffic context have a much greater impact than the characteristic parameters of the driver. This means that, regardless of the type of driver (more cautious, careful or ready), it is important to identify the safest available local traffic condition in order to decrease the risk. This could be achieved, for example, by moving the vehicle into a lane or applying a certain distance from the leading vehicle.

Having local traffic information in terms of average speed and localized vehicle density allows better management of vehicle risk conditions. This kind of information can be obtained, for example, through a V2I data exchange with the infrastructure.

In order to implement risk reduction logics based on the addition of maneuvers such as ad hoc calibrated lane changing, it would be appropriate to repeat the experiments to verify that there is not a greater risk per se in the act of performing this type of maneuver and also to verify which maneuvers, net of their execution, configure safer conditions.

All these latter considerations trace the direction in which the low-resolution scenarios should be made. They should be more detailed and explored when deepening the design or validation path within the typical automotive V-model.

6.2 Method perfectioning and results

The experiments carried out throughout the whole research, as well as the identification and development of the analysis tools used, demonstrated the great potential of the methods developed. The experiment carried out on the EuroNCAP scenario showed how weak the commonly used and internationally recognized approach to test matrix processing can be. The problem stems from an inadequate identification of the factors potentially impacting on the phenomena and, therefore, from a poor definition of the test scenarios. The proposed method largely succeeds in filling these gaps (as well as identifying them). The analysis methods identified have shown great versatility, demonstrating that they can also be applied to subjects other than the testing of control logics, as the last experiment showed.

In the experiments related to automation testing, the analysis method showed numerous advantages. First of all, it allowed a correct analysis of all the uncertainty factors, making it possible to obtain maximum awareness of the behaviour of the logics tested for the use-cases analysed. In addition, it allowed a significant reduction in the size of the space of significant scenarios to be explored. This allows, in the testing process foreseen in the V-model, the realisation of tests with greater significance and, at the same time, a considerable reduction in computational requirements.

6.3 Advantages/opportunity and criticity of the proposed tools and methods

Ultimately, the developed LoReS tool, assisted by the identified analysis tools that allow a proper exploration, allows, for the first time, the confinement of the problem of driving automation testing to a limited set of scenarios. That allows for the first time an exhaustive knowledge of the behavior of the logics under test. Furthermore, it allows the identification of failures, risk assessment, ensuring knowledge and management of all uncertainty factors.

The analysis methods identified are not the only ones possible and the extraction methods tested are only a few among all those available. However, they have proven to be highly effective, in various applications, with a rapid and efficient exploration of the test space, allowing multiple analyses that enable the realization of tests with a low computational impact together with a high significance of the examined samples, adequately identifying worst-cases.

This approach requires a large number of tests to obtain adequate confidence intervals. This determines that to conduct this type of analysis, a very abstract x-in-the-loop simulation is required, preferring MIL. The analyses carried out in the MIL, however, not only have applications in the design ambit (individuating the failures of the logics) and in the HARA ambit (individuating the sources of risk and the most onerous conditions), but they can be used as a design tool to identify the analysis factors for higher levels of simulations. E.g., it can be used to define the direction in which to deepen HIL experiments and of road testing (individuating the most relevant factors and values in the realization of the successive experiments).

One of the most relevant limitations of the analysis tools used is the need to analyse continuously changing factors. The first experiment showed how the readability of the analysis results and their interpretation strongly depend on the variability of the values that the input factors can assume. The analysis of factors that can take on few values can be managed through separate analyses carried out on different LoReS, which can be analysed by comparison in a qualitative manner. However, as we saw several times in the described experiments, it is possible to relate many discrete factors to continuous factors (such as the lane to which the ego vehicle belongs depends on its position with respect to the roadway, or the number of vehicles involved can be extracted from a distribution function characterised by continuous variable factors).

Another important limitation to be considered is the fact that the strong dependence of the proposed method on the execution of simulations means that the results obtained are closely linked to the models used by the simulation software. When these are unreliable or, for certain configurations of the input values, assume unrealistic behaviors, the resulting analyses are also affected by significant errors.

It should be emphasized that the LoReS tool, the methods of systematic scenario construction, uncertainty factor identification and sensitivity analysis are independent of the simulation tools used. Therefore, they can be performed with every simulation software and with every driver model, guaranteeing that results are obtained respecting correct principles of uncertainty analysis. They also exclude any evaluation error dependent on the use of incident databases that may be inhomogeneous with respect to the problem of highly automated driving.

Although it may seem trivial, it is necessary to underline the importance of verifying the reliability of the simulation tools to ensure the correctness of the analyses.

6.4 Possible future developments of the research

The strength of the method presented in this research for the identification of low-resolution scenarios lies in giving the limits on the uncertainty of the automation testing problem. This means that a finite number of cases can be defined to be analyzed. In addition, it has been shown that sensitivity analysis methods can be useful and appropriate for exploring the cases identified by the low-resolution scenario method. The LoReS matrix presented in this paper (Table 4) represents only one of the possible combinations of scenes/activities. As shown in Chapter 2, the distinction between scenes is arbitrary and various discrete factors may or may not be considered significant depending on the type of research to be performed, resulting in different sets of scenes and, consequently, a different scene/activities matrix.

The time for the study and representation of the LoReS analysed within the MEDIATOR project was not short, so it is considered that the elaboration of a complete set of LoReS may represent a non-trivial challenge. In addition, the experiments carried out analyzed only scenarios belonging to the low uncertainty category (motorway scene sets). A possible research development could be the configuration of a high uncertainty LoReS, such as, for example, the crossing of an urban intersection. In this particular case, one of the most relevant challenges that the research would have to face would be the management of discrete factors such as the type and number of arms of the intersection, the travel directions, the presence of lanes differentiated by traffic participants. It would be opportune to identify continuous factors that are able to describe and analyse some of these discrete factors (in a similar way to the factors of the last experiment performed), in order to express evaluations on their significance for defining a reduced number of cases to be tested separately.

The research carried out necessitated the development of lots of scripts to make compatible the simulation tools used with the scenario exploration methods implemented. Although the simulation software used are characterized by a wide versatility, it was often necessary to perform workarounds to compensate for the unpreparedness of the software for the methodological approaches carried out in this research. This eventuality demonstrates the innovativeness of the methodological approaches proposed, which have not yet been translated into adequate software functionalities, even in the very dynamic world of automotive development and testing. It would be appropriate to develop software tools specifically designed to work on low-resolution scenarios, guaranteeing control not only of detailed variables (such as maneuvers and ego-vehicle signals), but also the management of macroscopic tools for managing scenarios from the point of view of traffic conditions (e.g., traffic density, driver models, overtaking, lane change, etc.). Greater attention should also be paid to methods of varying inputs (types of extraction and methods of constructing test matrices).

Furthermore, we believe that the approach adopted to the analysis of the problem can be repeated in other areas. The generalization of the methodological approach referring to the delimitation of scenarios according to the amount of uncertainty and the classification of factors distinguished between discrete and continuous, is a method that combines the scenario approach with uncertainty analysis in a global vision that, I hope, can and will find other equally fertile fields of application.

Acknowledgement

This research was done in collaboration with FCA Automobiles and IPG Automotive. All the IPG software was given in Student license by IPG Automotive.

This research was co-financed with the resource of the "Programma Operativo Nazionale Ricerca e Innovazione 2014-2020"

La borsa di dottorato è stata cofinanziata con risorse del
Programma Operativo Nazionale Ricerca e Innovazione 2014-2020 (CCI 2014IT16M2OP005),
Fondo Sociale Europeo, Azione I.1 "Dottorati Innovativi con caratterizzazione Industriale"

UNIONE EUROPEA
Fondo Sociale Europeo

Ministero dell'Università e della Ricerca

PON
RICERCA E INNOVAZIONE
2014 - 2020

Bibliografia:

Abdullah, Rudwan, Amir Hussain, Kevin Warwick, and Ali Zayed, ‘Autonomous Intelligent Cruise Control Using a Novel Multiple-Controller Framework Incorporating Fuzzy-Logic-Based Switching and Tuning’, in *Neurocomputing* (Elsevier, 2008), LXXI, 2727–41 <https://doi.org/10.1016/j.neucom.2007.05.016>

Amersbach, Christian, and Hermann Winner, ‘Defining Required and Feasible Test Coverage for Scenario-Based Validation of Highly Automated Vehicles’, *2019 IEEE Intelligent Transportation Systems Conference, ITSC 2019*, 2019, 425–30 <https://doi.org/10.1109/ITSC.2019.8917534>

Archer, G. E.B., A. Saltelli, and I. M. Sobol, ‘Sensitivity Measures, Anova-like Techniques and the Use of Bootstrap’, *Journal of Statistical Computation and Simulation*, 1997 <https://doi.org/10.1080/00949659708811825>

Area Professionale Statistica ACI ed Acinformatica, ‘Veicoli e Incidenti Stradali Anno 2016’

Assanis, Dennis, Filipi Zoran, Steve Gravante, Dan Grohnke, Xinquin Gui, Lucas Louca, and others, ‘Validation and Use of SIMULINK Integrated, High Fidelity, Engine-Ln-Vehicle Simulation of the International Class VI Truck’, *SAE Transactions*, 2000, 384–99

Aust, M, ‘Evaluation Process for Active Safety Functions: Addressing Key Challenges in Functional, Formative Evaluation of Advanced Driver Assistance Systems’, *Chalmers University of Technology*, 2012

Bach, Johannes, Stefan Otten, and Eric Sax, ‘Model Based Scenario Specification for Development and Test of Automated Driving Functions’, *IEEE Intelligent Vehicles Symposium, Proceedings*, 2016-Augus.Iv (2016), 1149–55 <https://doi.org/10.1109/IVS.2016.7535534>

Bagschik, Gerrit, Andreas Reschka, Torben Stolte, and Markus Maurer, ‘Identification of Potential Hazardous Events for an Unmanned Protective Vehicle BT - 2016 IEEE Intelligent Vehicles Symposium, IV 2016, June 19, 2016 - June 22, 2016’, 2016-Augus.Iv (2016), 691–97 <https://doi.org/10.1109/IVS.2016.7535462>

Barnard, Y, S Innamaa, S Koskinen, H Gellerman, E Svanberg, and H Chen, ‘Methodology for Field Operational Tests of Automated Vehicles’, in *Transportation Research Procedia*, 2016, XIV, 2188–96

<https://doi.org/10.1016/j.trpro.2016.05.234>

Beizer, Boris, *Black-Box Testing: Techniques for Functional Testing of Software and Systems* (John Wiley & Sons, Inc., 1995)

Bengler, Klaus, Klaus Dietmayer, Berthold Farber, Markus Maurer, Christoph Stiller, and Hermann Winner, ‘Three Decades of Driver Assistance Systems: Review and Future Perspectives’, *IEEE Intelligent Transportation Systems Magazine*, 6.4 (2014), 6–22 <https://doi.org/10.1109/MITS.2014.2336271>

Bock, J., Krajewski, R., Eckstein, L., Klimke, J., Sauerbier, J., & Zlocki, A., ‘Data Basis for Scenario-Based Validation of HAD on Highways’, in *27th Aachen Colloquium Automobile and Engine Technology*, 2018, p. (pp. 8-10)

Borowsky, A., Oron-Gilad, T., Chasidim, H., Ahlström, C., Karlsson J.G., Bakker, B., *Behavioural Markers for Degraded Human Performance Behavioural Markers for Degraded Human Performance*, 2019

BRON: Netherlands National Road, and Register, ‘Police Registered Numbers of Casualties, Drivers and Crashes’, 2016 <http://www.swov.nl>

Chakraborty, Debasri, Warren Vaz, and Arup Kr Nandi, ‘Optimal Driving during Electric Vehicle Acceleration Using Evolutionary Algorithms’, *Applied Soft Computing Journal*, 34 (2015), 217–35 <https://doi.org/10.1016/j.asoc.2015.04.024>

Chen, Bo, and Harry H. Cheng, ‘A Review of the Applications of Agent Technology in Traffic and Transportation Systems’, *IEEE Transactions on Intelligent Transportation Systems*, 2010, 485–97 <https://doi.org/10.1109/TITS.2010.2048313>

Chen, Long, Shaobo Lin, Xiankai Lu, Dongpu Cao, Hangbin Wu, Chi Guo, and others, ‘Deep Neural Network Based Vehicle and Pedestrian Detection for Autonomous Driving: A Survey’, *IEEE Transactions on Intelligent Transportation Systems*, 22.6 (2021), 3234–46 <https://doi.org/10.1109/TITS.2020.2993926>

Chen, William, Anand Srivastav, and Giancarlo Travaglini, *A Panorama of Discrepancy Theory*, 2014

Christoph, M., Cleij, D., Ahlström, C., Bakker, B., Beggiato, M., Borowsky, A., *Mediating between Human Driver and Automation : State-of-the Art and Knowledge Gaps*, 2019

‘Company’ <https://waymo.com/company/#blog> [accessed 11 June 2021]

Cukier, R. I., C. M. Fortuin, K. E. Shuler, A. G. Petschek, and J. H. Schaibly, ‘Study of the Sensitivity of Coupled Reaction Systems to Uncertainties in Rate Coefficients. I Theory’, *Journal of Chemical Physics*, 1973 <https://doi.org/10.1063/1.1680571>

Cunningham, Alexander G., Enric Galceran, Ryan M. Eustice, and Edwin Olson, ‘MPDM: Multipolicy Decision-Making in Dynamic, Uncertain Environments for Autonomous Driving’, in *Proceedings - IEEE International Conference on Robotics and Automation* (Institute of Electrical and Electronics Engineers Inc., 2015), MMXV-JUNE, 1670–77 <https://doi.org/10.1109/ICRA.2015.7139412>

Czubenko, M, Z Kowalczuk, and A Ordys, ‘Autonomous Driver Based on an Intelligent System of Decision-Making’, *Cognitive Computation*, 7.5 (2015), 569–81 <https://doi.org/10.1007/s12559-015-9320-5>

DARPA, ‘DARPA Grand Challenge’, 2004 <https://archive.darpa.mil/grandchallenge04/overview.htm>

Department for Transport, ‘Reported Road Casualties Great Britain: 2015 Annual Report’, 2016

‘Elon Musk Delivers Virtual Speech for WAIC - SHINE News’ <https://www.shine.cn/biz/economy/2007091766/> [accessed 11 June 2021]

European New Car Assessment Programme, ‘Test Protocol – AEB VRU Systems’ <https://cdn.euroncap.com/media/26997/euro-ncap-aeb-vru-test-protocol-v20.pdf>

Geyer, Sebastian, Martin Kienle, Benjamin Franz, Hermann Winner, Klaus Bengler, Marcel Baltzer, and others, ‘Concept and Development of a Unified Ontology for Generating Test and Use-Case Catalogues for Assisted and Automated Vehicle Guidance’, *IET Intelligent Transport Systems*, 8.3 (2014), 183–89 <https://doi.org/10.1049/iet-its.2012.0188>

Go, K., and John M Carroll, ‘The Blind Men and the Elephant: Views of Scenario-Based System Design’, *Interactions*, 11.6 (2004), 44–53 <https://doi.org/10.1145/1029036.1029037>

‘Home - Toward Zero Deaths’ <https://www.towardzerodeaths.org/> [accessed 20 July 2021]

‘Home | Mediator’ <https://mediatorproject.eu/> [accessed 15 July 2021]

‘Indicatori Di Incidentalità Stradale’ <http://dati.istat.it/Index.aspx?DataSetCode=DCIS_INDINCIDENT> [accessed 20 July 2021]

International Organization for Standardization, ‘ISO-PAS-21448-2019’, 2019

———, ‘ISO/FDIS 26262 : Road Vehicles-Functional Safety’, 2018

ISTAT, ‘Year 2012 Road Accidents in Italy’, 2012

Kalra, Nidhi, and Susan M. Paddock, ‘Driving to Safety: How Many Miles of Driving Would It Take to Demonstrate Autonomous Vehicle Reliability?’, *Transportation Research Part A: Policy and Practice*, 94 (2016), 182–93 <https://doi.org/10.1016/j.tra.2016.09.010>

Kendziorra, Andreas, Peter Wagner, and Tomer Toledo, ‘A Stochastic Car Following Model’, *Transportation Research Procedia*, 15 (2016), 198–207 <https://doi.org/10.1016/j.trpro.2016.06.017>

Koopman, P, and M Wagner, ‘Challenges in Autonomous Vehicle Testing and Validation’, *SAE International Journal of Transportation Safety*, 4.1 (2016), 15–24 <https://doi.org/10.4271/2016-01-0128>

Koopman, Philip, and Michael Wagner, ‘Challenges in Autonomous Vehicle Testing and Validation’, *SAE International Journal of Transportation Safety*, 4.1 (2016), 2016-01–0128 <https://doi.org/10.4271/2016-01-0128>

———, ‘Toward a Framework for Highly Automated Vehicle Safety Validation’, *SAE Technical Paper Series*, 1 (2018), 1–12 <https://doi.org/10.4271/2018-01-1071>

LeBlanc, D., Sayer, J., Winkler, C., Ervin, R., Bogard, S., Devonshire, and T. J. Mefford, M., Hagan, M., Bareket, Z., Goodsell, R., and Gordon, *Road Departure Crash Warning System Field Operational Test: Methodology and Results. Volume 1: Technical Report*, 2006

Li, L, W.-L. Huang, Y Liu, N.-N. Zheng, and F.-Y. Wang, ‘Intelligence Testing for Autonomous Vehicles: A New Approach’, *IEEE Transactions on Intelligent Vehicles*, 1.2 (2016), 158–64 <https://doi.org/10.1109/TIV.2016.2608003>

MathWorks, ‘Vehicle Dynamics Blockset’ <https://it.mathworks.com/products/vehicle-dynamics.html>

Maurer, M., ‘EMS-Vision: Knowledge Representation for Flexible Automation of Land Vehicles’, in *IEEE Intelligent Vehicles Symposium, Proceedings*, 2000, pp. 575–80 <https://doi.org/10.1109/ivs.2000.898409>

Maurer, M, *Automotive Systems Engineering: A Personal Perspective*, *Automotive Systems Engineering*, 2013 <https://doi.org/10.1007/978-3-642-36455-6_2>

Maurer, Markus, J. Christian Gerdes, Barbara Lenz, and Hermann Winner, *Autonomous Driving: Technical, Legal and Social Aspects*, *Autonomous Driving: Technical, Legal and Social Aspects*, 2016 <https://doi.org/10.1007/978-3-662-48847-8>

'MEDIATOR behind the Scenes: Use Cases | Mediator', 2019 <https://mediatorproject.eu/blog/mediator-behind-the-scenes-use-cases> [accessed 11 June 2021]

Menzel, Till, Gerrit Bagschik, and Markus Maurer, 'Scenarios for Development, Test and Validation of Automated Vehicles', *ArXiv*, Iv, 2018, 1821–27

Ministero delle Infrastrutture e dei Trasporti, 'Norme Funzionali Geometriche per La Costruzione Delle Strade', *D.M. 5 Nov 2001*, 6792

Mohammed, Nadia A, Quasi-monte Carlo Sobol, and Nadia A Mohammed, 'Comparing Halton and Sobol Sequences in Integral Evaluation', *Zanco Journal of Pure and Applied Sciences*, 31.1 (2019), 32–39 <https://doi.org/10.21271/zjpas.31.1.5>

Mood, A, F A Graybill, and D C Boes, 'Random Variables, Distribution Functions, Expectation', in *Ntroduction to the Theory of Statistics*, 3rd ed. (New York, NY, USA, 1974), pp. 51–72

Motor Industry Software Reliability Association, 'Development Guidelines for Vehicle Based Software', 1994

———, 'Report 6: Verification and Validation', 1995

NHTSA US Department of Transportation, *Facts, T. S. 2012*, 2012

Nithya, M, and M R Rashmi, 'Gazebo - ROS - Simulink Framework for Hover Control and Trajectory Tracking of Crazyflie 2.0', in *IEEE Region 10 Annual International Conference, Proceedings/TENCON*, 2019, MMXIX-OCTOB, 649–53 <https://doi.org/10.1109/TENCON.2019.8929730>

'Nuovo Codice Della Strada' (decreto legisl. 30 aprile 1992 n. 285 e successive modificazioni)

Op den Camp, Olaf, Sjef van Montfort, Jeroen Uittenbogaard, and Joke Welten, ‘Cyclist Target and Test Setup for Evaluation of Cyclist-Autonomous Emergency Braking’, *International Journal of Automotive Technology*, 2017 <https://doi.org/10.1007/s12239-017-0106-5>

Page, Yves, Sophie Cuny, Thierry Hermitte, and Maxime Labrousse, ‘A Comprehensive Overview of the Frequency and the Severity of Injuries Sustained by Car Occupants and Subsequent Implications in Terms of Injury Prevention’, *Annals of Advances in Automotive Medicine. Association for the Advancement of Automotive Medicine*, 2012

Pompigna, Andrea, ‘La Calibrazione Del Diagramma Fondamentale e La Valutazione Del Livello Di Servizio Operativo Sulle Autostrade Italiane’ (Università di Bologna, 2017)

Reschka, A, ‘Fertigkeiten-Und Fähigkeitengraphen Als Grundlage Für Den Sicheren Betrieb von Automatisierten Fahrzeugen in Städtischer Umgebung-English Title: Skill and Ability Graphs as Basis for Safe Operation of Automated Vehicles in Urban Environments’, *Announced, PhD Thesis, Technische Universität Braunschweig*, 2017

Reschka, Andreas, ‘Safety Concept for Autonomous Vehicles’, in *Autonomous Driving: Technical, Legal and Social Aspects*, 2016, pp. 473–96 <https://doi.org/10.1007/978-3-662-48847-8_23>

Reschka, Andreas, Gerrit Bagschik, Simon Ulbrich, Marcus Nolte, and Markus Maurer, ‘Ability and Skill Graphs for System Modeling , Online Monitoring , and Decision Support for Vehicle Guidance Systems’, Iv, 2015

Roberts, Martin, ‘The Unreasonable Effectiveness of Quasirandom Sequences’, *Extreme Learning*, 2018 <http://extremelearning.com.au/unreasonable-effectiveness-of-quasirandom-sequences/> [accessed 19 June 2021]

Rosson, Mary Beth, and John M. Carroll, *Usability Engineering: Scenario-Based Development of Human-Computer Interaction*, 2002

SAE, ‘SAE International Standard J3016: Taxonomy and Definitions for Terms Related to On-Road Motor Vehicle Automated Driving Systems’, *Surface Vehicle Recommended Pratice*, 2014

———, ‘Taxonomy and Definitions for Terms Related to Driving Automation Systems for On-Road Motor Vehicles’, *Surface Vehicle Recommended Pratice*, 2021

Safeguarding Automated Driving Functions <www.pegasusprojekt.de> [accessed 11 June 2021]

Saltelli, Andrea, and Paola Annoni, ‘How to Avoid a Perfunctory Sensitivity Analysis’, *Environmental Modelling and Software*, 25.12 (2010), 1508–17 <https://doi.org/10.1016/j.envsoft.2010.04.012>

Saltelli, Andrea, Paola Annoni, Ivano Azzini, Francesca Campolongo, Marco Ratto, and Stefano Tarantola, ‘Variance Based Sensitivity Analysis of Model Output. Design and Estimator for the Total Sensitivity Index’, *Computer Physics Communications*, 181.2 (2010), 259–70 <https://doi.org/10.1016/j.cpc.2009.09.018>

Saltelli, Andrea, Marco Ratto, Terry Andres, Francesca Campolongo, Jessica Cariboni, Debora Gatelli, and others, *Global Sensitivity Analysis. The Primer*, *Global Sensitivity Analysis. The Primer*, 2008 <https://doi.org/10.1002/9780470725184>

Schuldt, Fabian, ‘Ein Beitrag Für Den Methodischen Test von Automatisierten Fahrfunktionen Mit Hilfe von Virtuellen Umgebungen - English Title: Towards Testing of Automated Driving Functions in Virtual Driving En-Vironments’, 2017

Singh, H, and S Jha, ‘Simulated Environment for Autonomous Driving Using ROS Based on Mahindra E2O Electric Car’, in *ACM International Conference Proceeding Series*, 2019 <https://doi.org/10.1145/3352593.3352667>

Skyving, M, ‘STRADA: Road Traffic Accident and Injury Data in Sweden’, in *24th World Int. Traffic Medicine Association Congress*, 2015

Sobol’, I.M., ‘Sensitivity Estimates for Nonlinear Mathematical Models’, *Mathematical Modeling and Computational Experiment*, 1993

Sobol, I. M., V. I. TURCHANINOV, Yu L. LEVITAN, and B. V. SHUKHMAN, ‘LPTAU51’, 1991

Sobol, Ilya M., and Boris V. Shukhman, ‘On Global Sensitivity Indices: Monte Carlo Estimates Affected by Random Errors’, *Monte Carlo Methods and Applications*, 2007 <https://doi.org/10.1515/MCMA.2007.005>

Strandroth, Johan, Simon Sternlund, Anders Lie, Claes Tingvall, Matteo Rizzi, Anders Kullgren, and others, ‘Correlation Between Euro NCAP Pedestrian Test Results and Injury Severity in Injury Crashes with Pedestrians and Bicyclists in Sweden’, *SAE Technical Paper Series*, 1.November 2014

(2018), 213–31 <https://doi.org/10.4271/2014-22-0009>

Uittenbogaard, Jeroen, Olaf Op den Camp, and Sjef van Montfort, ‘CATS Deliverable 5.1: CATS Verification of Test Matrix and Protocol’, 2016 <www.TNO.nl/CATS>

Ulbrich, Simon, Till Menzel, Andreas Reschka, Fabian Schuldt, and Markus Maurer, ‘Defining and Substantiating the Terms Scene, Situation, and Scenario for Automated Driving’, *IEEE Conference on Intelligent Transportation Systems, Proceedings, ITSC*, 2015-Octob (2015), 982–88 <https://doi.org/10.1109/ITSC.2015.164>

Urmson, Chris, and William Whittaker, ‘Self-Driving Cars and the Urban Challenge’, *IEEE Intelligent Systems*, 23.2 (2008), 66–68 <https://doi.org/10.1109/MIS.2008.34>

Vieira, B, R Severino, A Koubaa, and E Tovar, ‘Towards a Realistic Simulation Framework for Vehicular Platooning Applications’, in *Proceedings - 2019 IEEE 22nd International Symposium on Real-Time Distributed Computing, ISORC 2019*, 2019, pp. 93–94 <https://doi.org/10.1109/ISORC.2019.00028>

Villagra, J, L Acosta, A Artuñedo, R Blanco, M Clavijo, C Fernández, and others, *Automated Driving, Intelligent Vehicles: Enabling Technologies and Future Developments*, 2017 <https://doi.org/10.1016/B978-0-12-812800-8.00008-4>

Villagra, Jorge, Leopoldo Acosta, Antonio Artuñedo, Rosa Blanco, Miguel Clavijo, Carlos Fernández, and others, *Automated Driving*, *Intelligent Vehicles: Enabling Technologies and Future Developments* (Elsevier Inc., 2017) <https://doi.org/10.1016/B978-0-12-812800-8.00008-4>

Wachenfeld, Walther, Hermann Winner, J. Chris Gerdes, Barbara Lenz, Markus Maurer, Sven Beiker, and others, ‘Use Cases for Autonomous Driving’, in *Autonomous Driving: Technical, Legal and Social Aspects*, 2016, pp. 9–37 <https://doi.org/10.1007/978-3-662-48847-8_2>

Winner, Hermann, and Walther Wachenfeld, ‘Virtual Assessment of Automation in Field Operation A New Runtime Validation Method’, *10. Workshop Fahrerassistenzsysteme*, 2015, 161–70 <http://tuprints.ulb.tu-darmstadt.de/5192/1/FAS_2015_Tagungsband.pdf#page=169>

Winner, Hermann, Walther Wachenfeld, and Philipp Junietz, ‘Safety Assurance for Highly Automated Driving – The PEGASUS Approach’, in *Transportation Research Board*, 2017

Wisch, Marcus, Patrick Seiniger, Mervyn Edwards, Thomas Schaller, Mònica Pla, Andrés Aparicio, and others, ‘European Project AsPeCSS - Interim Result: Development of Test Scenarios Based on Identified Accident Scenarios’, 2013

Zhao, Ding, ‘Accelerated Evaluation of Automated Vehicles’, 2016

Zhao, Ding, Xianan Huang, Huei Peng, Henry Lam, and David J Leblanc, ‘Accelerated Evaluation of Automated Vehicles in Car-Following Maneuvers’, 19.3 (2018), 733–44

Zhao, Ding, Henry Lam, Huei Peng, Shan Bao, David J. LeBlanc, Kazutoshi Nobukawa, and others, ‘Accelerated Evaluation of Automated Vehicles Safety in Lane-Change Scenarios Based on Importance Sampling Techniques’, *IEEE Transactions on Intelligent Transportation Systems*, 18.3 (2017), 595–607 <https://doi.org/10.1109/TITS.2016.2582208>