



NRC-CNRC

Technical Review of Modeling and Simulation-Based Testing of CAV
Systems

Prepared for
Transport Canada
Motor Vehicle Safety Directorate

Prepared by
Taufiq Rahman, PhD
Andrew Liu, MASc
Automotive and Surface Transportation Research Centre
National Research Council Canada

Updated: August, 2021



National Research
Council Canada

Conseil national de
recherches Canada

Canada

© 2021 Her Majesty the Queen in Right of Canada, as represented by the National Research Council of Canada.

Cat. No. NR16-370/2021E-PDF

ISBN 978-0-660-40551-3

NRC.CANADA.CA



Table of contents

Table of contents.....	3
Executive Summary	12
1 Background	14
1.1 Objectives	15
1.2 Study Methodology & Scope	16
1.3 Limitations.....	16
2 Modeling & Simulation of Cyber-Physical Systems	17
2.1 Cyber-Physical Systems & CAV	17
2.2 CPS Modeling Paradigms	17
2.2.1 Physics Based Models	18
2.2.2 State Machine Based Models	18
2.2.3 Rule & Agent Based Models	18
2.2.4 Data-driven Models	18
2.3 Simulation-Based Testing for CPS Verification & Validation.....	19
2.3.1 Simulation Testing of Cyber Components	20
2.3.2 Validation of Autonomous Systems	22
3 Modeling & Simulation in Aerospace	24
3.1 Introduction.....	24
3.2 Rationale for M&S in Aerospace Systems	24
3.3 Regulatory Aspects of M&S in Aerospace	26
3.4 Aerospace M&S Use Cases	28
3.4.1 DO-254 Compliant Aerospace Sensor Development	29
3.4.2 Safety Assessment of Aircraft Landing Gear for Certification	30
3.4.3 Simulation Scenarios for Testing Autonomous Drones	30
3.5 Lessons Learned from Boeing 737 MAX Crashes	30
4 Modeling & Simulation in Rail Transportation	33
4.1 Introduction.....	33
4.2 Automation in Rail Transportation	33
4.3 Regulatory Standards.....	35

4.4 Railway Automation M&S Use Cases	36
4.4.1 Testing CBTC Systems.....	36
4.4.2 Validation of a Railway Signaling System.....	37
5 Modeling & Simulation in Marine Transportation	38
5.1 Introduction	38
5.2 Automation in Marine Transportation	39
5.3 Regulatory Aspects of M&S in Marine Transportation	41
5.4 Marine Automation M&S Use Cases	42
5.4.1 Vehicle-in-the-Loop Simulation of Marine Robot Swarm	42
5.4.2 Simulation of Energy Performance of Ships	42
5.4.3 Simulation-based Verification of Autonomous Navigation Systems	43
6 Modeling & Simulation in CAV	44
6.1 Introduction	44
6.2 CAV Definitions & Terminology	44
6.3 Rationale for M&S in CAV	46
6.4 CAV M&S Use Cases	49
6.4.1 Sensor Simulation	52
6.4.2 ADAS Testing.....	52
6.4.3 Virtual V2X Demonstrations/Testing	53
6.5 CAV Standards Relevant to M&S.....	54
6.6 Reported M&S Activities from Key Players	56
6.6.1 Waymo	56
6.6.2 Mercedes-Benz	57
6.6.3 Uber ATG	57
6.7 Overview of Software Tools.....	57
6.7.1 Description of Test Scenarios	57
6.7.2 Automated Driving Simulators	58
6.7.3 Traffic Simulators	60
7 Conclusion & Summary of Findings.....	62
7.1 Research Questions & Answers.....	62
7.2 Open Questions/Issues	65
7.3 Concluding Remarks	65

Acknowledgements	67
References	68
Appendix A: SAE J3016 Levels of Driving Automation.....	78

List of tables

Table 1: XIL simulation testing of the cyber layer of CPS.....	21
Table 2: DO-178C software levels.	27
Table 3: Essential aspects of model-based software simulation [34].	29
Table 4: MASS levels of automation (adapted from [72] and [73]).	40
Table 5: Definitions of the terms "scene", "situation" and "scenario" by Ulbrich <i>et. al.</i>	46
Table 6: Examples of text coverage required to demonstrate AV reliability [3].	47
Table 7: Results of the simulation study reported by Borg <i>et. al.</i> [102].	53
Table 8: BSI developed CAV standards.	55
Table 9: Comparison of automated driving simulators, adapted and updated from [113].	59

List of figures

Figure 1: V-model for system-level development of CPS [16].	19
Figure 2: Typical XIL simulation testing topology for software components.	20
Figure 3: Configuration of XIL simulation in various staged of the V-model of CPS development (adapted from [18]).	21
Figure 4: Validation technologies for autonomous systems (adapted from [19]).	23
Figure 5: Impact of CFD at Boeing (green: areas with strong CFD penetration, blue: areas with some penetration, red: future opportunities) [26]. © Royal Aeronautical Society 2016.	25
Figure 6: Relationship among aerospace certification standards [40].	27
Figure 7: MCAS system operation. © The Seattle Times	31
Figure 8: Typical dynamic analyses applied during railway engineering. ©Global Railway Review	34
Figure 9: Grades of automation (GoA) in rail transportation, adapted from [58] and [59].	35
Figure 10: Scope of major CENELEC railway application standards (adapted from [63]).	36
Figure 11: Wave Glider - an autonomous USV. © Liquid Robotics	40
Figure 12: Simulation test system proposed in [82].	43
Figure 13: Global vehicle target (GVT) specified by Euro NCAP. © Euro NCAP	48
Figure 14: Example of an edge case - a floating balloon can be challenging for AI-enabled perception. © AEye	49
Figure 15: Generic ADS testing process [85].	50
Figure 16: CAV system architecture.	51
Figure 17: Elements of a typical virtual driving simulator platform, as presented in [97].	52
Figure 18: ADS verification and validation tools, practices, and standards. ©Foretellix	55
Figure 19: Three roadway scenes generated from a single ~20 line Scenic file [111].	58
Figure 20: Categorization of the problem of CAV simulation testing.	65

List of Abbreviations

Abbreviation	Definition
ABS	Anti-lock Braking System
ACC	Adaptive Cruise Control
ADAS	Advanced Driver Assistance System
ADS	Automated Driving System
ADS-DV	Automated Driving System-Dedicated Vehicle
ADSM	Airworthiness Design Standards Manual
AEB	Automatic Emergency Braking
AHRS	Attitude Heading Reference System
AIAA	American Institute of Aeronautics and Astronautics
ALSE	Aviation Life Support Equipment
ANS	Autonomous Navigation Systems
AoA	Angle of Attack
ATO	Automatic Train Operation
ATP	Automatic Train Protection
AV	Automated Vehicle
BSI	UK British Standards Institution
CAD	Computer Aided Design
CAE	Computer Aided Engineering
CAV	Connected and Autonomous Vehicle
CBTC	Communication-based Train Control
CEN	European Committee for Standardization
CENELEC	European Committee for Electrotechnical Standardization
CFD	Computational Fluid Dynamics
CI/CD	Continuous integration and continuous development
CMVSS	Canada Motor Vehicle Safety Standard
COLREGs	Convention on the International Regulation for Preventing Collisions at Sea
CPS	Cyber-Physical Systems
DDT	Dynamic Driving Tasks

Abbreviation	Definition
DGA	Direction générale de l'armement (French Department of Defence)
DND	Canadian Department of National Defence
DNN	Deep Neural Network
DoD	US Department of Defense
DUT	Design Under Test
EASA	European Union Aviation Safety Agency
EN	European Standards
ESC	Electronic Stability Control
ESD	Energy Saving Devices
ETCS	European Train Control System
ETSI	European Telecommunications Standards Institute
Euro NCAP	European New Car Assessment Programme
FAA	US Federal Aviation Administration
FEA	Finite Element Analysis
FMEA	Failure Mode and Effects Analysis
FMVSS	US Federal Motor Vehicle Safety Standards
FPGA	Field Programmable Gate Arrays
FSS	Full Flight Simulator
FSTD	Flight Simulation Training Device
FTA	Fault Tree Analysis
FTD	Flight Training Device
GHG	Greenhouse Gas
GN&C	Guidance, Navigation and Control
GNSS	Global Navigation Satellite System
GoA	Grades of Automation
GVT	Global Vehicle Target
HAV	Highly Automated Vehicle
HCS	Heading Control System
HD Map	High Definition Map
HIL	Hardware-in-the-loop
IAPT	International Association of Public Transport

Abbreviation	Definition
IEC	International Electrotechnical Commission
IEEE	Institute of Electrical and Electronics Engineers
IMO	International Marine Organization
IMU	Inertial Measurement Unit
INCOSE	International Council on Systems Engineering
ISS	International Space Station
IV&V	Independent Verification and Validation
JAXA	Japan Aerospace Exploration Agency
JTAG	Joint Test Action Group
LKS	Lane Keeping System
M&S	Modelling and Simulation
MASS	Maritime Autonomous Surface Ships
MBSA	Model Based Safety Assessment
MBSE	Model Based Systems Engineering
MBT	Model Based Testing
MCAS	Maneuvering Characteristics Augmentation System
MIL	Model-in-the-loop
ML	Machine Learning
MSC	Maritime Safety Committee
M-SDL	Measurable Scenario Description Language
NASA	National Aeronautics and Space Administration
NHSTA	US National Highway Traffic Safety Administration
NLOS	Non-line-of-sight
NRC	National Research Council Canada
NSW	Nose Wheel Steering System
ODD	Operational Design Domain
OEDR	Object and Event Detection and Response
OEM	Original Equipment Manufacturer
OSI	Open Simulation Interface
PAS	Publicly Available Specifications
PIL	Processor-in-the-loop

Abbreviation	Definition
RAMS	Reliability, Availability, Maintainability and Safety
ROS	Robot Operating System
RTCA	Radio Technical Commission for Aeronautics
SAE	Society of Automotive Engineers International
SBST	Search-based Software Testing
SDLC	System Development Life Cycle
SE	System Engineering
SIL	Software-in-the-loop
SIL	Safety Integrity Level
SLAM	Simultaneous Localization and Mapping
SOTIF	Safety Of The Intended Functionality
STCW	Standard for Training, Certification and Watchkeeping
SUT	System Under Test
TC	Transport Canada
TCPS	Transportation Cyber-Physical Systems
TEVV	Test, Evaluation, Verification and Validation
TMR	Triple Modular Redundancy
UNECE	United Nations Economic Commission for Europe
US DOT	US Department of Transportation
USV	Unmanned Surface Vehicle
V&V	Verification & Validation
V2X	Vehicle to everything
VIL	Vehicle-in-the-loop
VSSA	Voluntary Safety Self-Assessments
VUT	Vehicle Under Test
XIL	X-in-the-loop

Executive Summary

Models are representations of physical systems and processes. Ranging from explicit internal structures derived from well understood system behavior to data driven descriptions, models can be developed in a number of ways. Regardless of the model development paradigm used, their fidelity (i.e., the degree to which they represent reality) may vary. Simulation models refer to a class of models that are executable in some way so that the time evolution of these models can be studied under different inputs and operating conditions. Since the advent of computer aided engineering (CAE) and computer aided design (CAD), modeling and simulation (M&S) have been used for a variety of applications. Most widespread use cases of M&S include design synthesis, design optimization, performance evaluation, etc. In every branch of engineering including the multi-modal transportation sector, M&S have been used as a scalable and resource efficient alternative to physical testing.

The evolution of traditional motor vehicles to connected and automated vehicles (CAV) has been driven by the continuous development and adoption of cyber elements. These cyber elements enable these systems to autonomously respond to the dynamic operating environment. As a result, humans have an increasingly diminishing role in operating these systems. By automating the operation of motor vehicles, the CAV community is trying to realize a vision of a safer, more efficient and more environmentally responsible road transportation. Although significant advancements have been made, there remains a number of challenges. One of these challenges, verification and validation of CAV is the central focus of this technical review.

There is a consensus in the CAV community that traditional testing practices used in automotive engineering cannot address the challenges involving validation of CAV systems. From a practical point of view, on road testing in test-tracks and on public roads cannot accomplish enough coverage of the diverse operating conditions CAV systems will encounter and be able to handle safely due to the resource intensive nature of the exercise. In addition, ensuring safety in physical tests of systems whose safety is yet to be validated can be challenging. As a feasible alternative, simulation based testing can provide sufficient test coverage so that these systems can be validated with acceptable statistical significance.

In order to understand how simulation testing can be implemented and adopted for CAV, this report investigated how M&S have been adopted and applied in other modes of transportation. It was found that the aerospace, rail and marine transportation sectors principally apply M&S for a wide spectrum of use cases. Themes of design synthesis and performance characterization problems mainly focusing on physical aspects of these systems were found to be prominent in the related M&S literature. Proliferation of cyber elements into these transportation systems is a relatively newer development. As such, development of standards for validating these cyber elements was found to be in early stages.

Arguably, CAV is leading the evolution of traditional systems into cyber-physical systems (CPS) amongst all transportation modes. The problem of validation of CAV systems can be characterized by the intersection of two highly evolving engineering fields: validation of cyber physical systems (CPS) and validation of AI-based autonomous systems. Despite the progress made due to significant research and development efforts in the form of public-private standardization initiatives, industrial consortiums, and public & privately funded academic research programs, etc., the validation frameworks composed of accepted methodologies and standards are still developing. Although the standards are developing, simulation based testing will

play a key role in not only developing CAV systems but validating the safety argument of these systems because of the obvious limitations of physical testing. It is widely accepted in the broader engineering community that M&S will never replace physical testing. However, M&S can offer detailed insight into the system behavior. Thus, the knowledge gained through a detailed M&S trial and reporting can be applied in physical testing, which can potentially contribute more value to the engineering design and validation process than what physical testing alone could accomplish.

1 Background

Connected & automated vehicles (CAV) utilize hardware (e.g., sensors, real-time embedded computers, actuators) and software (e.g., perception & path-planning algorithms) components to fully or partially automate the task of driving an automobile to achieve goals of improving safety, efficiency, convenience, and reducing greenhouse gas (GHG) emissions. Computer software and hardware (also colloquially known as the autonomy stack) in CAVs partially or fully perform the dynamic driving task (DDT). In CAV systems the autonomy stack is tasked to make and subsequently execute safety-critical decisions related to DDT in order to either assist human drivers (SAE¹ L1-L2²) or to render their input/role limited (SAE L3-L4) or even completely unnecessary (SAE L5). It can be argued that the immensity of operational scope of software in conjunction with sensors and actuators is the most significant distinction CAVs have in comparison to manually driven vehicles. In the Federal Automated Vehicles Policy published by NHSTA in 2016 [1], the term “highly automated vehicle” (HAV) was used to represent SAE L3-L5 vehicles with automated systems that are responsible for monitoring the driving environment. The term ADS-DV (automated driving systems dedicated vehicles) has also been used in a NSTHA report published as recently as 2020 [2] to refer to vehicles that are designed to be exclusively operated as SAE L4 and L5 vehicles for all trips and are not equipped with manual driving controls. Since autonomy hardware and software takes on the role of making and executing safety-critical decisions in HAV and ADS-DV, safety evaluation of these systems must evolve from traditional testing and validation methodologies that principally involve physical testing of mechanical components of an automobile. As a result, a robust safety validation framework must be established that seeks to ensure that CAV systems can safely operate in diverse roadway environments, and can safely navigate vehicles through all possible real-world situations. Some industry observers have suggested that an automated driving system (ADS) would have to drive billions of miles in the real-world to experience an adequate number of situations to statistically validate safety performance claims [3]. To reduce the time, cost, and risks of physical testing (i.e., track and real-world testing), industry and the international regulatory community are examining how simulations and simulation-based testing can help address some of these challenges.

Functional safety is defined as those aspects of overall safety of a system that rely on automatic protection to enable predictable and appropriate responses to anomalous operating conditions caused by factors such as human error, hardware/software failure and excessive operational/environmental stress. A technical paper presented in 2019 World Congress of Society of Automotive Engineers (SAE) by academic researchers from Ohio State University [4] argue that the current automotive functional safety standards, as defined by the automotive safety integrity levels (ASIL), focus on controllability of the system. Since controllability, which is greatly affected by human intervention, is a large consideration in ASIL definitions, current concepts of functional safety is not strongly relevant for emerging CAV systems. It is so because humans play increasingly diminishing roles as decision makers in operating these systems. Greater adoption of modeling & simulation (M&S), in conjunction with new physical testing methodologies³, can

¹ [SAE J3016 levels of automated driving](#)

² See Appendix A

³ Waymo accumulated more than 20 million real-world miles on public roads and more than 15 billion simulated miles of automated driving in their testing campaign (source: [Waymo Safety Report Feb 2021](#)).

address some of these evolving issues relating to CAV safety evaluation. Nonetheless, adherence to safety standards to inform design, manufacturing and life-cycle management of safety-critical systems has been a customary practice in many industries besides automotive. Industries that develop safety-critical systems have been observed to be driven by regulation and safety standards, and a large body of governing safety standards have been developed to address the uniqueness of each industry [5]. For example, ISO 26262, EN 50128 and DO-178C are respectively functional safety standard for automotive, railway and aerospace systems. In addition to functional safety (mitigating risk due to system failure), the concept of SOTIF (safety of intended functionality) in the ISO/PAS 21448:2019⁴ standard is concerned with assuring safety in the absence of a fault.

Simulations and simulation-based testing use a virtual environment, with simulated objects and other elements, to determine how the vehicle or its subsystems will respond to specific situations. While it is not intended to replace physical testing, simulation-based testing can augment physical testing by expanding test coverage. As it is more scalable, cost-effective, safe, and efficient, simulation-based testing provides a test administrator the ability to create a wide range of scenarios and permutations, including complex scenarios where a diverse range of elements are examined.

The use of simulation-based testing is not unique to the automotive sector. Simulation and simulation-based testing is used extensively in aviation for product design, development, verification and certification, as well as for the training and type testing of pilots. It is also used in other regulated sectors, such as the marine and rail transportation, which have incorporated simulation-based functions into their research and development, training, and safety validation and certification.

As Transport Canada and the international automotive regulatory community contemplates how simulations and simulation-based testing could be used as validation tools for emerging CAV technologies, it is helpful to understand the considerations, best practices and lessons learned from other sectors where simulation-based testing is currently utilized, as well as the state of development within the automotive sector.

1.1 Objectives

This report summarizes the findings of an extensive technical review of the literature related to automation of transportation systems to examine the use of modelling, simulations and simulation-based testing in aviation, rail and marine transportation, as well as the state of development within the automotive sector, in order to provide insight to the following key questions:

1. How modeling, simulation and simulation-based testing is defined within the sector examined.
2. How it is currently utilized within the sector examined.
3. Whether there are regulatory requirements or frameworks.
4. What are the best practices and lessons learned?

After reviewing the related literature answers to these questions are presented in Section 7.1.

⁴ [ISO/PAS 21448:2019 Road vehicles — Safety of the intended functionality](#)

1.2 Study Methodology & Scope

This technical review was conducted based on expert analysis of the related literature comprised of:

- Reports & guidelines published by regulatory, certification and industrial consortium bodies.
- Technology promotion materials provided by verification service providers and software vendors.
- White papers published by industry stakeholders.
- Peer-reviewed journals & conference papers published by research organizations including universities.

A large body of such literature was reviewed, and concepts, data and information from a subset of them (~150 references) were deemed relevant to this technical review. The subject matter was found to be very diverse. However, extensive effort was applied so that the relevant concepts can be discussed following a set of common themes.

1.3 Limitations

Although every effort was made to include the most current information available from literature found in the public domain, it cannot be guaranteed that all relevant information was reviewed. In order to minimize the likelihood of such unintended omissions, the report went through multiple rounds of internal and external reviews. In addition, content and data from reliable and, whenever possible, peer-reviewed sources were employed in this report. Integrity of the references was considered implied.

2 Modeling & Simulation of Cyber-Physical Systems

2.1 Cyber-Physical Systems & CAV

Connected and Autonomous Vehicle (CAV) systems are considered a subset of cyber-physical systems (CPS), which are defined as those systems that employ computer algorithms to control and/or monitor physical phenomena and processes [6]. Not surprisingly, the term “Transportation Cyber-Physical Systems” (TCPS) was coined in a review paper authored by US and Chinese university researchers and published in the highly respected Institute of Electrical and Electronics Engineers (IEEE) Transactions on Vehicular Technology journal in 2016 [7] to refer to the ever increasing proliferation of cyber elements in surface, marine and air transportation. In this research paper all systems related to vehicle connectivity, driving automation and smart infrastructure are categorized as TCPS. Correspondingly, the problem of testing CAV systems belongs in the broader field of verification and validation of CPS systems. Since CPS systems operate in a physical world, they are exposed to stochastic stimuli from the operating environment. This translates to an infinite number of permutations of inputs and conditions that the CPS system operates in. Exhaustive testing of these permutations is not combinatorically feasible. Correspondingly, generating a set of test cases in order to sufficiently evaluate safety and functionality of a CPS system is considered a hallmark problem in this field. Based on the testing categories for CPS systems provided in [8], simulation testing of CAV can be regarded as an intersection between model based testing (MBT) and search based testing. In MBT testing paradigm, a model of the CPS is tested to confirm that the exhibited behavior conforms to the specifications. In addition, search based test methods employ meta-heuristic search techniques such as genetic algorithm or simulated annealing algorithm to generate test cases to represent the real-world scenarios that the CPS is expected to be exposed to.

Simulation testing of CPS systems is a two part process. The first part involves model development or generation. An extensive body of literature has been published on the problem of modeling of CPS systems, and this field of model development is still very active, which indicates that our understanding of the problem is still evolving. Nonetheless, once a model is generated, it can be subjected to a *simulation* exercise to perform tests to verify and validate that the system indeed conforms to a pre-defined performance and/or safety target.

2.2 CPS Modeling Paradigms

Modeling refers to the process of developing a description or representation of a system, entity or phenomenon. The developed description or representation is called a “model” and typically the purpose of the modeling exercise is to gain better understanding and insights about the modeled object by studying it. Modeling paradigm refers to a set of rules or a philosophy which are followed during the model development phase. According to a survey paper published in IEEE Access in 2018 [9], there are four main categories of modeling paradigms for CPS systems: (a) physics based models, (b) state machine based models, (c) rule and agent based models, and (d) data-driven models. Since CAV is a subset of the broader concept of CPS, these modeling paradigms are also applied in CAV modeling and simulation (M&S) activities. These modeling paradigms are briefly discussed below.

2.2.1 Physics Based Models

Physics based models are developed in the form of equations from first principles. The underlying equations can be simple dependence equations or ordinary differential equations or partial differential equations. Depending on the goal, the model developer usually makes a subjective decision on how simple or complex the resulting model would be. Although assumptions are usually made to simplify the resulting model, depending on the requirement of model accuracy the degree to which it is done may vary. Physics-based models are generally used to understand how a system or process evolves over time under dynamic conditions. The turbocharger model for automotive diesel engine control applications described by university researchers from China in a 2019 SAGE journal paper [10] can be cited as a physics based model. This model can also be characterized as a multi-physics model because it incorporates more than one physics domain – multi-body mechanics and fluid dynamics. The mechanical architecture of the modeled turbocharger was described by multi-body mechanics (i.e., their inertia properties and the underlying kinematic constraints). In addition, the flow of air and its interactions with the turbocharger vanes were modeled with computational fluid dynamics equations. In both cases the two physical phenomena (fluid flow inside the turbocharger and its mechanical aspects) were modeled using well-established physics equations.

2.2.2 State Machine Based Models

State machine based models are used to represent systems showing discrete dynamics. While physics based models represent continuous dynamics (i.e., system behavior is known at all time instances), state machine based models represent system behavior at discrete time intervals. The algorithms that are part of a CPS system are evaluated in a discrete fashion, and state machine based modeling paradigm is the most appropriate for such cases. For example, a state machine based model was proposed for fault protection of the guidance, navigation and control subsystems of an aerospace vehicle in a 2019 Journal of Aerospace Information Systems article [11]. A functional state machine was developed to model system behavior in conjunction with a diagnostic state machine developed for on-board fault diagnosis. Since system behavior can be categorized as discrete states from a fault perspective (e.g., fault-free vs faulty operation) and failure events are also discrete events, this modeling paradigm was regarded as the most appropriate modeling paradigm.

2.2.3 Rule & Agent Based Models

Rule & agent based models attempt to capture the shared dependencies among the heterogeneous components a CPS system have. These models describe the modeled behavior using a set of agents that make decisions and execute actions based on a set of semantically defined rules. For example, in order to develop a model for a traffic intersection, this modeling paradigm can be considered most appropriate. Vehicular and pedestrian traffic can be modeled as agents and their behavior can be governed by a set of rules. The multi-agent model presented in an article published in the Journal of Sustainable Cities and Societies in 2018 [12] employed this modeling technique to describe the underlying components in a microgrid power system. Since the components of a microgrid power system can be highly diverse in terms of power storage and generation fluctuations, this modeling technique was deemed suitable in this case.

2.2.4 Data-driven Models

Data-driven models are developed from an existing set of input-output data acquired from the actual system or another simulation model. These models find the relationship between the inputs and outputs without

requiring a *priori* knowledge of the underlying dynamics. They are particularly useful in situations where the system dynamics is too complex to model explicitly (e.g., physics based model) or too computationally expensive for conducting a simulation study. There are many approaches described in the literature used for data-driven models. In the CAV ecosystem, DNN (deep neural network) based models that are used for image based classification and localization of roadway environment elements can be considered a type of data-driven model that represent visual human perception. Black-box, white-box and grey-box models are all variations of data-driven models. In a black-box model, the mathematical relationships between inputs and outputs are derived by some means such as statistical regression, and the physical significance of these relationships is poorly understood. On the other hand, the causal relationships between inputs and outputs are well understood in white-box models. Finally, grey-box models combine knowledge of explicit causal structure of the system with data to achieve greater accuracy in describing the modeled behavior. Grey-box models therefore combine aspects of white-box and black-box modeling together.

2.3 Simulation-Based Testing for CPS Verification & Validation

The lifecycle of a CPS system typically starts with a concept and ends in field deployment and operation is often summarized under the V-model (see Figure 1). Verification & validation (V&V) is an important part of that evolution. It should be mentioned that the V-model is rarely a linear one in real-life scenarios. There are usually many iterations involving reviews of the design and integration tasks directed by the outcomes of the V&V process. Although an idealized representation of the reality, the V-model still provides a good description of the different phases a development process goes through. Instead of the seemingly linear transitions from phase to phase, a real development process may jump from one phase to another as required by the evolving situation. A PhD thesis [13] from Carnegie Mellon University and published in 2019 categorizes V&V activities into two main types: formal methods and simulation-based methods. The major difference between formal and simulation-based verification methods is derived from how mathematically rigorous the testing protocol is [14]. For a given design property, formal methods prove that the design property holds for every point in the search space (i.e., any permutation of the inputs). On the other hand, simulation-based models only use a subset of the search space for testing purposes. As the complexity of the CPS under test rises, adopting formal methods becomes exponentially difficult, even sometimes an untenable proposition. Simulation-based testing methods are preferred for their scalability and tractability when formal methods are no longer a practical alternative. It should be noted that if required V&V practitioners may combine both methods to extend test coverage. For example, the problem of demonstrating computer network survivability was solved using a combination of formal and simulation based methods in a conference paper published in 2010 [15].

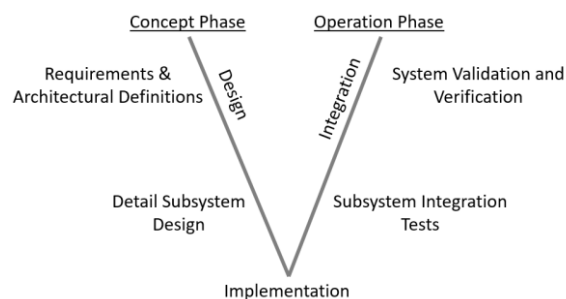


Figure 1: V-model for system-level development of CPS [16].

Simulation-based testing of CPS systems are based on system models that can predict or estimate the state of the system at a future time instance given past knowledge of current states, inputs and operating conditions. Depending on the characteristics of the system, any combinations of the four modeling paradigms described in Section 2.2 can be applied. For example, let's consider the CPS modeling and simulation testing exercise described in a conference paper presented at the 2nd International Workshop on Autonomous Systems Design (ASD 2020) [17]. In this paper, the automatic transmission of an automobile was examined. The developed model follows a hybrid approach where physics based equations are used to describe vehicle dynamics aspects of the systems, and the control logic parts are described as a discrete state machine.

2.3.1 Simulation Testing of Cyber Components

In a CPS the cyber layer represents the computing hardware and software that monitor and regulate the physical layer. A CAV system can be regarded as a traditional automotive system augmented with cyber elements (sensors, computers, algorithms, and actuators) to implement driving automation functions. In regards to simulation-based testing of cyber components of a CPS, a number of testing paradigms are used that vary in scope and the degree to which they represent the final design. The overarching term X-in-the-loop (XIL) can be found in the literature to refer to these simulation testing methodologies. Depending on the use case and testing objectives (e.g., design development vs obtaining certification), X can be replaced with model, software, processor, hardware etc. to represent the design under test (DUT). The DUT is usually tested in a simulated environment, which represents the closed-loop interaction between the DUT with the outside world through its interfaces. The DUT may impose one or several inputs on the simulated environment based on the information received from it (see Figure 2). Brief descriptions of the XIL testing methodologies for the cyber layer of the CPS systems are provided in Table 1.

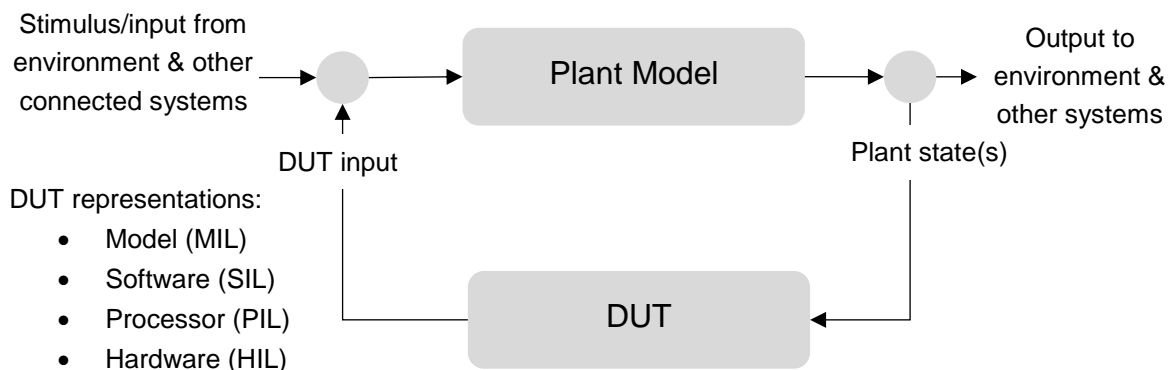


Figure 2: Typical XIL simulation testing topology for software components.

XIL Paradigm	DUT Implementation/representation	Typical Testing Objective
Model-in-the-loop (MIL)	Model of the DUT is derived from the specifications	Proof of concept verification

XIL Paradigm	DUT Implementation/representation	Typical Testing Objective
Software-in-the-loop (SIL)	DUT functionalities are implemented as software code	Verification and validation of software implementation
Processor-in-the-loop (PIL)	DUT software code deployed on the target microprocessor typically embedded in a <i>development board</i> ⁵	<ul style="list-style-type: none"> • Verification of real-time performance requirements • Identify and debug run-time errors
Hardware-in-the-loop (HIL)	DUT software code deployed on the target hardware (e.g., security certified computing hardware platform to be used in the final design)	<ul style="list-style-type: none"> • Verification of previous verification results (SIL & PIL) • Confirmation of performance in real-life conditions

Table 1: XIL simulation testing of the cyber layer of CPS.

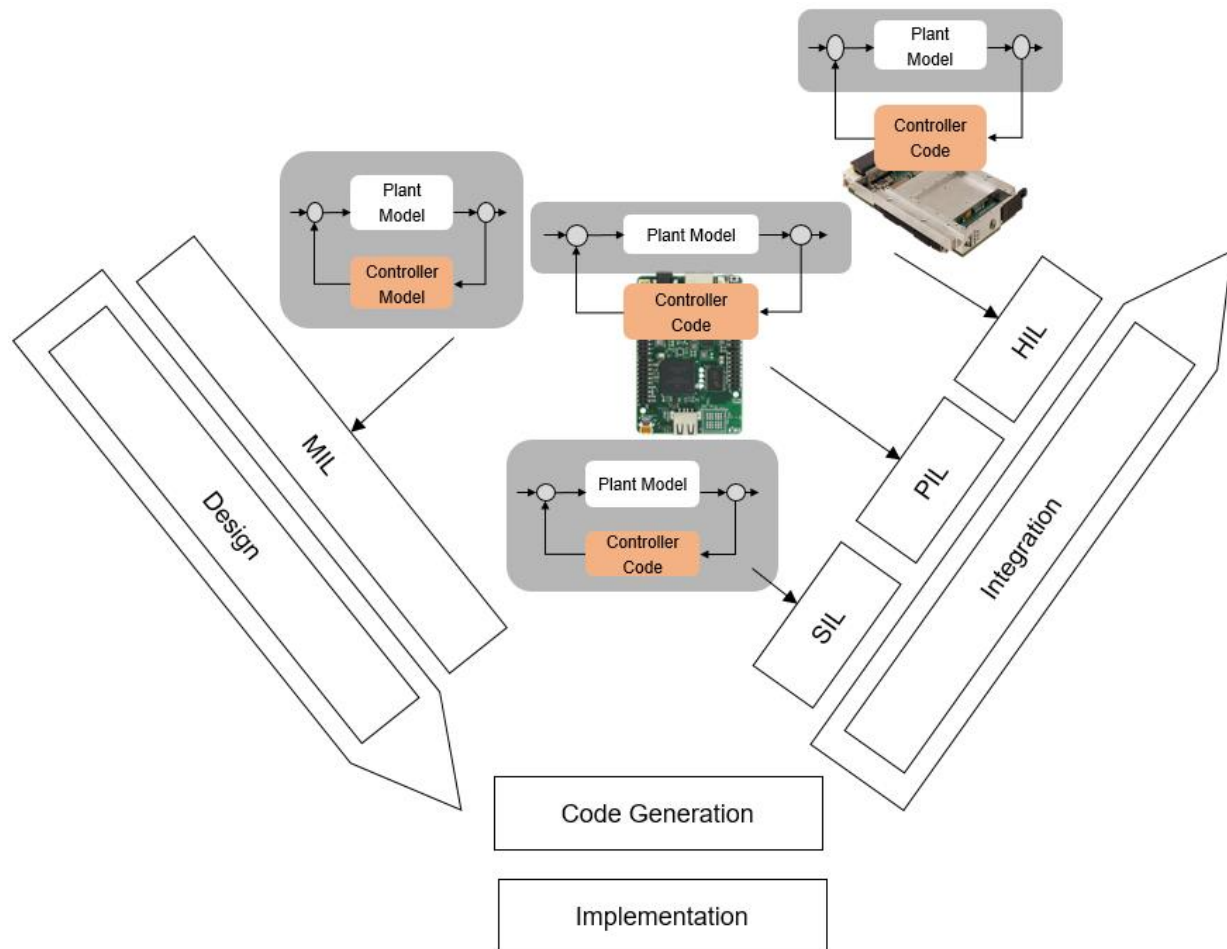


Figure 3: Configuration of XIL simulation in various staged of the V-model of CPS development (adapted from [18]).

⁵ A *development board* is a printed circuit board featuring a target microprocessor and support hardware (e.g., communication and debug interfaces) to facilitate development/testing activities for embedded applications.

A conference paper primarily authored by researchers from the German Aerospace Center (DLR) and presented in the 2019 American Institute of Aeronautics and Astronautics SciTech Forum [18] illustrated how XIL testing methodology relates to the V-model of CPS development. The various configurations of XIL simulation used in different stages of system development life cycle (SDLC) are shown in Figure 3.

2.3.2 Validation of Autonomous Systems

Validation of autonomous systems are relatively more challenging compared to other types of CPS systems because these systems are expected to operate in unstructured operating environments. The fact that the permutations and the combinations of inputs and operating conditions that an autonomous system can be exposed to approach infinity renders the proposition of evaluating these systems employing traditional CPS validation methods impractical. A number of articles in the reviewed literature have discussed these challenges. For example, an editorial article published in IEEE Software journal in 2019 [19] lists a number of open questions that the research community is trying to answer about the validation of autonomous systems:

- How can reliability be defined for these systems?
- How these systems can be supervised?
- How do we define liability in the event of failure?

In the same vein, a US Department of Defense (DoD) report published in 2017 [20] observes: “*for the most demanding adaptive and non-deterministic systems, a new approach to traditional TEVV (test, evaluation, verification and validation) will be needed.*”

Ebert & Weyrich in their editorial [19] used the pictorial shown in Figure 4 to provide an overview of the current validation technologies used for autonomous systems. Some of the terms such as MIL, SIL and HIL have already been discussed in previous sections. Some additional concepts, namely, FMEA (failure mode and effects analysis), FTA (fault tree analysis), fault injection and function test, that were not introduced before are described below:

- FMEA refers to the sequential approach applied to identify all possible failure modes (i.e., ways a design or system might fail resulting in it not functioning as intended) and their corresponding effects.
- FTA is a deductive failure analysis method that applies a top-down approach to characterize a system failure as a causal effect of lower-level events. Simply put, it aims to identify the ways a system may fail with a view to introduce countermeasures to reduce the associated risks.
- Fault injection is a testing technique that characterizes system behavior when it is under fault (e.g., anomalous input, defective components and unusual operating conditions).
- Function test is the process of evaluating a system’s performance against the stated design requirements.

Validation Handling	Automatic	<ul style="list-style-type: none"> Simulation environments with MIL, HIL and SIL 	<ul style="list-style-type: none"> Simulation environments with MIL/SIL Brute-Force Usage in the Real World, while running realistic scenarios Intelligent validation (e.g., cognitive testing and AI testing)
	Manual	<ul style="list-style-type: none"> Function test Fault injection Negative requirements with misuse, abuse and confuse cases FMEA and FTA for safety Simulation environments with MIL, HIL and SIL 	<ul style="list-style-type: none"> Experimental and empirical test strategies Simulation environments with MIL/SIL Brute-force usage in the real world, while running realistic scenarios Specific quantity requirements; e.g., penetration testing and usability
		White Box	Black Box
		Validation Strategy	

Figure 4: Validation technologies for autonomous systems (adapted from [19]).

3 Modeling & Simulation in Aerospace

3.1 Introduction

Although autopilots in aircrafts have existed for quite some time now, a team of university and industry researchers in a 2019 Journal of Transportation Research Record article [21] argue that these systems are considered to be parallel to Society of Automotive Engineers International (SAE) Level 2 driving automation systems because they need to be supervised by flight crew. Nonetheless, similar to road vehicles, automating the operations of an aircraft (i.e., self-flying) is an emerging trend in the aerospace industry. This trend is being driven by the anticipated shortage of trained pilots over the next couple of decades. A report published by the American Institute of Aeronautics and Astronautics (AIAA)⁶ estimates 600,000 trained pilots will be needed in the next 20 years. For reference, only 200,000 pilots have been trained since the start of commercial aviation.⁷ Although road vehicles and aircrafts have significantly different operating environments, they are anticipated to use similar enabling technologies to implement automation. Driven by the desire of automakers and technology companies to establish early market dominance in the automated vehicle sector, these enabling technologies for road vehicles have seen tremendous growth in recent years. Flying automation technologies is directly benefitting from the advances made in driving automation research.⁶ For example, a report from Boeing [22] reviews M&S tools developed for driving automation systems with a view to transpose this knowledge for developing safe automated flight systems. M&S as a tool have played a number of roles in the aerospace industry including system design, performance validation, certification, training etc., and it is expected to be a vital development and validation tool for flying automation systems. Some example use-cases are described in Section 3.4.

3.2 Rationale for M&S in Aerospace Systems

Aerospace systems intersect many engineering disciplines including fluid mechanics, electrical engineering, mechanical engineering, software engineering etc. Given the complexity of these systems, it is of paramount importance to verify that designs perform to set specifications and targets. Depending on the scope of simulation, M&S based activities can be performed at the component level, the system level or the mission level. Component level M&S activities have limited scope where a single component of the aircraft is considered (e.g., fuel pump, sensor, etc.). On the other hand, system level M&S tasks attempt to understand how components interact with each other (e.g., an engine with many components, avionics systems, etc.) under different conditions. Finally, mission level M&S focuses on the overall performance of the entire aircraft (e.g., flight simulation with or without pilot in the loop).

At the component level, physics based M&S tools are often used to synthesize and evaluate designs of physical components such as engines, airframes etc. Since physical prototyping and testing can be a cost intensive exercise, finite element analysis (FEA) is used to evaluate materials and to develop structural designs. For example, FEA was used (a) to determine the combined structural performance of an aerospace-grade aluminum alloy and the manufacturing process friction drilling in [23], (b) to accelerate

⁶ AIAA report: <https://aerospaceamerica.aiaa.org/features/achieving-autonomy/>

⁷ News report: <https://www.sdbj.com/news/2018/apr/08/solving-pilot-shortage-starts-small-airports/>

design cycle time of composite materials to build airframes in [24], (c) to optimize designs of gas turbine blades in [25] etc. Another widely adopted physics based simulation tool is computational fluid dynamics (CFD) which have been used for the design and evaluation of airframes (structural and thermal loads, aerodynamics performance), engine design, operational issues such as icing, etc. Figure 5 can be referenced to illustrate how far-reaching applications of CFD have been in the aerospace industry. Generally both CFD and FEA have been used in the aerospace industry as an alternative to physical prototyping and testing to facilitate virtual prototyping activities for design development, optimization, and performance verification etc. CFD and FEA are enabling aerospace engineers to design and implement better components to an extent that cannot be practically replicated with physical testing because of the time and the cost needed to implement it. Besides physics based models, software components of an aircraft system may employ state machine based models to aid development and to prove functionalities/performance in a simulated environment.

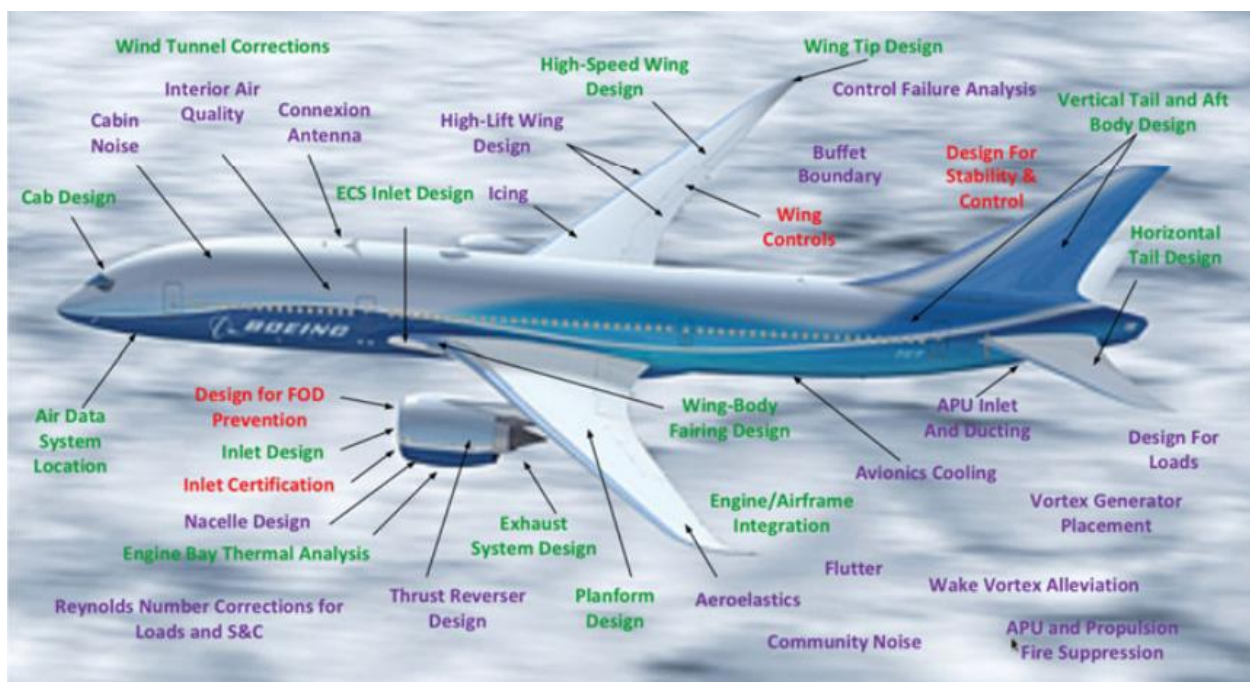


Figure 5: Impact of CFD at Boeing (green: areas with strong CFD penetration, blue: areas with some penetration, red: future opportunities) [26]. © Royal Aeronautical Society 2016.

At the system level, the model based systems engineering (MBSE) paradigm have gained traction in recent years to find a better methodology to handle the increasing complexity of aircraft systems⁸. MBSE has been defined by the International Council on Systems Engineering (INCOSE) as “*the formalized application of modeling to support system requirements, design, analysis, verification and validation activities beginning in the conceptual design phase and continuing throughout development and later life cycle phases*”⁹. Simulation is an important design and verification tool in the MBSE paradigm because it moves away from a document-centric approach to capture information about a system in the form of an executable description

⁸ SAE news: [MBSE is transforming aerospace engineering, systems integration](#)

⁹ Systems Engineering Vision 2020: http://www.cose.org/media/upload/SEVision2020_20071003_v2_03.pdf

(i.e., a model) [27]. Lind & Andersson of Saab Aeronautics in [27] described how MBSE is being adopted for aircraft systems design. The potential for programmatically automating design and audit tasks such as measuring change impact, integrity, quality, completeness and accuracy that MBSE offers is enabled by using the models for simulation¹⁰. In contrast to the traditional document based design approach, automation in MBSE can *potentially* achieve greater safety by design.

One of the most prominent mission-level M&S use case is flight simulators of various types such as Full Flight Simulator (FSS), Flight Simulation Training Device (FSTD) and Flight Training Device (FTD). Although mainly used for pilot training, these mission level simulators can also be used for cockpit design and evaluation (e.g., [28]), aircraft handling characteristics research (e.g., [29]), etc. All these training, design development and evaluation tasks are performed with relative ease in a cost-efficient and safe manner because mission-level M&S tools are available.

3.3 Regulatory Aspects of M&S in Aerospace

Aerospace is a highly regulated industry. Regional and national regulatory bodies such as United States Federal Aviation Administration (FAA), Transport Canada, and European Union Aviation Safety Agency (EASA) author policies, standards and guidelines, provide certification and oversight, and enforce compliance in relation to safety of security aspects of civil aviation. Ensuring airworthiness, which is a measure of suitability of an aircraft for safe flight, can be considered as one of the major objectives of these regulatory activities. While performing literature search for this technical review, it was found that the term “simulation” has been used scarcely in documents related to airworthiness published by aerospace regulatory bodies. For example, the Airworthiness Design Standards Manual (ADSM) [30] published by the Canadian Department of National Defence (DND) in 2020 recognises simulation as a testing means to show compliance for certification besides flight test, ground test, functional test, etc. However, in this document simulation test has been explicitly identified in the context of proving compliance of aviation life support equipment (ALSE) and pilot training. Canadian Aviation Regulation (current to June 28, 2021) [31] mentions simulation in the context of pilot training and emergency situations only. A report published by FAA in 2016 titled “Safety Issues and Shortcomings with Requirements Definition, Validation, and Verification Processes” [32] recognizes simulation as a prototyping tool, but it was suggested to apply this tool with caution because “*simulations that model a system may not be accurate in a specific condition.*” Therefore, this report suggests to carefully assess the fidelity of models/simulation being used.

In relation to M&S, autonomous operation enabled by software can be regarded as the common ground between the aerospace and the CAV sectors. The aerospace industry (civil, military and space systems) have published a number of standards and specifications for software V&V tasks. Some of these standards, specifications and guidance documents are:

- **RTCA DO-178C Software Considerations in Airborne Systems and Equipment Certification** [33] serves as a guide for the production of software for airborne systems that must comply with airworthiness requirements. Regulators such as FAA, EASA and Transport Canada use this as the primary document to approve all commercial software-based aerospace systems [34]. According

¹⁰ Presentation given to INCOSE chapter meeting by Laura E. Hart of Lockheed Martin - [Introduction to Model-Based System Engineering \(MBSE\) and SysML](#)

to the safety impact of the software component, the DO-178C standard categorizes software into five levels, and each level must show compliance with a number of objectives detailed in the document to gain certification. See Table 2 for details. According to the DO-178C standard, preferred use of a test environment “includes the software loaded into the target computer and tested in an environment that closely resembles the behavior of the target computer environment.”

- **NASA-STD-8739.8A Software Assurance and Software Safety Standard** [35] defines the requirements to implement a systematic approach to software safety in safety critical systems throughout the lifecycle of software components starting from the concept phase.
- **NASA-GB-8719.13 Software Safety Guidebook** [36] is intended to guide the creation and assurance activities involving safety critical software, firmware (computer programs deployed on embedded systems) and programmable logic; e.g., field programmable gate arrays (FPGA) etc.
- **ARP4761 Guidelines and Methods for Conducting the Safety Assessment Process on Civil Airborne Systems and Equipment** [37] is a “recommended practice” document that describes guidelines and methods of assessing safety for the certification of civil aircraft.
- **ARP4754A Guidelines for Development of Civil Aircraft and Systems** [38] is a guidelines document to address the development cycle for aircraft and systems with consideration to overall aircraft operating environment and functions. Although it does not specifically deal with software or electronic hardware development, NASA applied this guideline in the development of computer-based aircraft systems [39].

Failure Outcome	Software Level	# of Objectives to Meet
Catastrophic	A	71
Hazardous	B	69
Major	C	62
Minor	D	24
No Safety Impact	E	None

Table 2: DO-178C software levels.

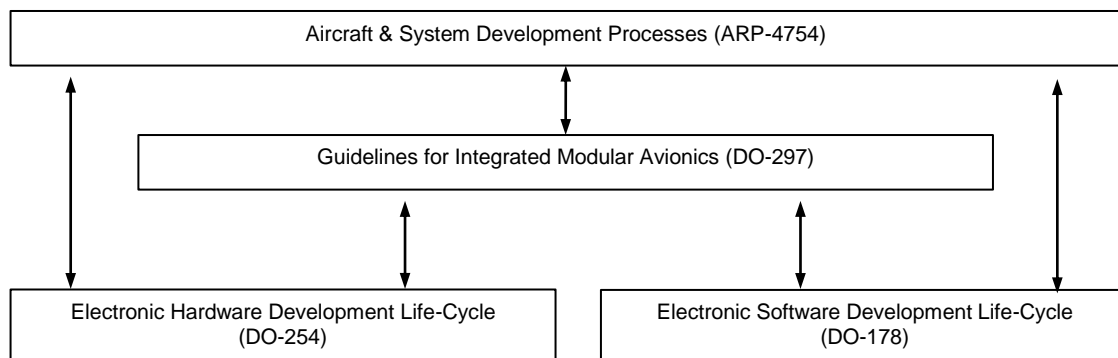


Figure 6: Relationship among aerospace certification standards [40].

Among the above mentioned standards, specifications and guidance, the DO-178C standard has been prominently mentioned in the context of certification of safety critical software. Since software must be deployed on computing hardware to be able to function in aerospace systems, these hardware elements must be safety certified as well. Correspondingly, **RTCA DO-254 Design Assurance Guidance for Airborne Electronic Hardware** [41] is the document followed by designers to develop aerospace compliant computing hardware. Another related certification document is the **RTCA DO-297 Integrated Modular Avionics (IMA) Development Guidance and Certification Considerations** [42], which is used by FAA and EASA to approve modular avionics devices. How these different standards are interconnected is pictorially described in a paper presented at the 10th Aerospace Technology congress in 2019 [40] (see Figure 6).

3.4 Aerospace M&S Use Cases

Typical of any complex engineering system development in this era of CAD (computer aided design) and CAE (computer aided engineering), M&S activities are heavily utilized in the development of aerospace systems. As automation of safety critical systems have gained prominence in the aerospace industry, M&S activities have been expanded to verification & validation of these components. In a technical paper published in the SciTech forum of American Institute of Aeronautics & Astronautics in 2019 [34], author Michael D. Rubin (Lead, Flight Software Independent V&V, Red Canyon Software¹¹) have discussed essential model-based software simulation for space vehicles in the context of independent verification and validation (IV&V). These aspects of model-based software simulation is tabulated in Table 3, which demonstrates how diversified the scope and the applications of simulation testing of software elements can be in the aerospace domain. At a microscopic level, individual components are tested against the design specifications. However, as these individual components are integrated into sub-systems, which in-turn are organized as systems to ultimately form the overall architecture of an entire aerospace vehicle, their interactions with each other through various interfaces can be characterized as a combinatorial explosion. Therefore, simulation testing of aerospace software components must be performed at the component level, system level and vehicle level scopes, as shown in Table 3. Some reported use cases of M&S activities in the development and verification of aerospace systems are discussed next.

Functional Area	Key Simulation Uses for IV&V
Guidance, navigation and control (GN&C)	<ul style="list-style-type: none"> • Kinematics & dynamic properties of the space vehicle • Functionality testing of requirements-based GN&C algorithm models • Fault injection for erroneous guidance and/or navigation data
Sensor data collection & fusion	<ul style="list-style-type: none"> • Functional models of the sensors, including any redundant data collection methods • Simulation of the fusion of the various sensor and data sources, including any special calculations • Fault injection for erroneous sensor data, by one or more sensors

¹¹ Red Canyon Software: <https://redcanyonsoftware.com/>

Functional Area	Key Simulation Uses for IV&V
Sub-system interfaces	<ul style="list-style-type: none"> • Simulation of software/software, hardware/software and hardware/hardware interfaces • Types, ranges and rates of different data elements
Vehicle commanding and command sequencing	<ul style="list-style-type: none"> • Modeling of command and data busses relating to vehicle commanding • Sending and receiving of vehicle commands, including relevant timing and performance requirements • Fault injection of one or more erroneous command sent or received • Need to look at usage and verification different sequence engines
Vehicle fault tolerance	Where possible, model fault tolerant aspects of the software itself, such as mode/bus/commanding switching in the event of a redundant computer failure, or in the event of a partial or complete sub-system failure.
Robotic or payload manipulation system	<ul style="list-style-type: none"> • Kinematic and dynamic modeling and simulation of the robotic system • Kinematic and/or actuator redundancy • Considerations and modeling of different types, sizes, and shapes of payloads
Individual sub-system simulation	<ul style="list-style-type: none"> • Verifying and validating individual sub-systems or components as a “black box” to ensure that its’ behavior and performance function per requirements and design for nominal and off-nominal conditions • Verifying and validating the sub-system or component in a “white box” testing scenario
Overall vehicle simulation	<ul style="list-style-type: none"> • Modeling and simulation of all of the sub-systems in a spacecraft integrated together is ideal, as this would provide information on overall vehicle behavior and performance in both nominal and off-nominal conditions.

Table 3: Essential aspects of model-based software simulation [34].

3.4.1 DO-254 Compliant Aerospace Sensor Development

Attitude Heading Reference System (AHRS) is a sensor that incorporates inertial measurement unit (IMU) and on-board processing to provide motion information of an aircraft including attitude, heading, roll, pitch and yaw. Since the DO-254 standard demands a top-down approach with traceability of individual requirements to the final design. Proving compliance can be difficult for a number of reasons. Since it is unlikely that all components of an AHRS system can be found in-house for a system-wide certification, it is challenging to demonstrate compliance for components sourced from elsewhere. Engineers from the defense contractor Northrop Grumman have addressed these challenges and detailed the design process in a conference paper published in 2018 in [43]. In addition using to the physics-based simulation tool ANSYS for designing the mechanical components of the AHRS system, they have used MATLAB/Simulink for developing and verifying the control algorithms needed for the on-board processing in a model-in-the-loop (MIL) testing methodology. In later design stages they have used FPGA (field programmable gate array) design and verification tools to implement the developed models and algorithms on the target hardware.

3.4.2 Safety Assessment of Aircraft Landing Gear for Certification

How the concept of Model Based Safety Assessment (MBSA) was applied to perform safety and risk analysis of an aircraft landing gear system according to SAE ARP 4761 standard was described in a conference paper [44], which was published in 2020 and was authored by a design practitioner from the aerospace industry. Closely related to the MBSE methodology, MBSA is the formalized application of modeling to perform safety analysis by simulating a system model with nominal (non-failure) functionalities in conjunction with an augmented fault model to study system responses under one or more combinations of faults and failures. In this case, MBSA was applied on a multi-physics model of the nose wheel steering system (NSW) characterized by a combination of electro-mechanical, mechanical, hydraulics and embedded controller sub-systems. This multi-physics model was simulated along with a fault injection model to understand and quantify the safety risks. Simulation of system behavior under fault conditions is the basic premise of MBSA, and in this case it was applied to identify modes of failures and hazards to eventually remove them to obtain a certifiable design.

3.4.3 Simulation Scenarios for Testing Autonomous Drones

German aerospace researchers in 2020 published a Robot Operating System (ROS) based simulation architecture for autonomous drones in [45]. Rather than focusing on certification, this effort used scenario-based testing in a ROS simulation environment to verify the different components of autonomous drones. It is interesting to note that major toolchains used in this work (ROS¹² and Gazebo¹³) have strong CAV connections, which is further indication that automation research in the aerospace industry is being directly benefitted from the advances made in CAV system. Another related example includes a conference paper authored by researchers from Japan Aerospace Exploration Agency (JAXA) in 2021 [46] where a ROS-based software architecture was proposed for space robotics. This architecture is being planned for space demonstration using a drone in the International Space Station (ISS).

3.5 Lessons Learned from Boeing 737 MAX Crashes

Despite being a highly regulated industry laser focused on safety, aerospace sector has observed a few failure events with catastrophic and fatal outcomes over the years. The catastrophic failures of space shuttles Challenger and Columbia had been due to hardware issues. However, the more recent and highly publicized Boeing 737 MAX crashes were attributed to failure to properly design and deploy a software component in the flight control system.

The Boeing 737 MAX aircraft was developed as a fourth generation variant of the widely successful Boeing 737 series, which has a long history dating back to its first flight in 1967. The new generation was announced in August 2011 and the maiden flight was conducted in January 2016. It obtained certification from FAA and EASA in March 2017. Global regulators grounded the MAX variant in March 2019 after two fatal crashes (Lion Air Flight 610 & Ethiopian Airlines Flight 302) that had claimed a total of 346 lives. In both cases the Maneuvering Characteristics Augmentation System (MCAS), a software component of the flight control system designed to trigger autonomously without pilot involvement, was cited as contributing

¹² <https://www.ros.org/>

¹³ <http://gazebo.org/>

causes involved in the two catastrophic failures.¹⁴ In order to understand how the MCAS system operates, the pictorial from the Seattle Times¹⁵, as shown in Figure 7, can be referenced.

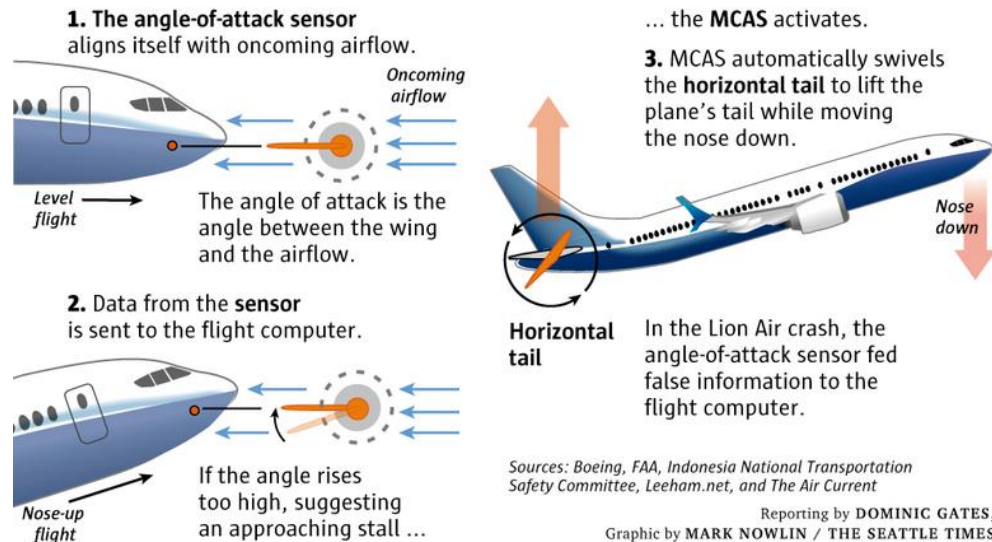


Figure 7: MCAS system operation. © The Seattle Times

The MCAS system was introduced to address the handling characteristics of the MAX variant's airframe outfitted with large engines which caused it to behave differently than the previous generation aircrafts. The choice of large engines were driven by the objective of achieving greater fuel economy. Because of their size, MAX's engines were positioned farther away from the fuselage and well in front of the wings than previous designs of 737 [47]. This caused the airframe to have a high propensity to "pitch-up" or raise its nose when the engines are delivering high power. This increases the angle of attack (see Figure 7) and if it is large enough, the aeroplane could enter in an aerodynamic stall (i.e., reduction in lift force that prevents the aeroplane to maintain its altitude). Because of the placements of the large engines, the aerodynamic behavior of MAX series was significantly different from its predecessor despite sharing the same airframe. Since Boeing wanted to leverage the existing certification of the previous generation 737, it had the motivation to project the impression that the new design "feels and flies like other 737s."¹⁴ In addition, engineering the new design to feel and fly like other 737s also resulted in not having to retrain pilots who have been flying previous generation 737 aircrafts. This provided significant cost savings for Boeing's customers (i.e., commercial airlines) [47]. Motivated by these factors, the MCAS system was introduced to operate in the background to avoid "pitch-up" conditions by engaging the horizontal tail of the aircraft. Initially Boeing did not include MCAS training in the pilots' manual.¹⁴ The MCAS system was designed to take input from a single angle of attack (AoA) sensor despite there being two of these sensors in the airframe.¹⁵ It was unclear from the reviewed literature what factors led Boeing engineers to this design decision. In [48], however, it was mentioned that Boeing offered a second AoA vane and a "disagree" light as an add-on option/feature, which provides visual indication to the pilot when one of the AoA sensor reading is different from the other one. It should be mentioned that neither of the two crashed aircrafts were

¹⁴ Business Insider news report: [The first Boeing 737 Max crash was 2 years ago today. Here's the complete history of the plane that's been grounded since 2 crashes killed 346 people 5 months apart.](#)

¹⁵ The Seattle Times news report: [Flawed analysis, failed oversight: How Boeing, FAA certified the suspect 737 MAX flight control system](#)

equipped with this add-on features. The authors Johnston and Harris concluded in [48] that Boeing did not sufficiently test their 737 MAX aircrafts to identify the lack of redundancy in the critical sensor supported MCAS system.

In the US House Committee on Infrastructure and Transportation report [49], it was determined that malfunctioning AoA sensors unnecessarily triggered the MCAS systems in both crashes. The pilots, in both cases, were not trained to deal with this malfunction, which ultimately led to the catastrophic outcomes. Although the MCAS software was wrongly classified as “hazardous failure” (see Table 2) during the safety assessment, the warranted input redundancy (i.e., having at least two AoA sensors instead of one) was not implemented¹⁵. Johnson and Harris in [48] characterize the MCAS as a “*quickly applied software patch*” driven by Boeing’s desire to use the already certified 737 airframe with large engines to achieve greater fuel efficiency. At the same time, Boeing did not want to impose the cost of retraining pilots to handle the MAX’s significantly different flying characteristics on their customers. Although this design decision of augmenting the flight control software with an autonomous MCAS system without providing sufficient training to the pilots should have been identified and rectified by industry standard practices such as design review, functional hazard analysis and safety certification. All the reviewed literature (i.e., [48], [49], [50]) determine that the safety certification/assessment processes were rushed. FAA managers encouraged delegation of safety assessments from FAA safety engineers to Boeing itself, for the sake of speedy approval of the resulting analysis.¹⁵ Correspondingly, Gipson, a law professor from the University of Memphis, remarks that the workplace culture and administrative processes at the Federal Aviation Administration (FAA) with respect to aircraft certification require an overhaul in order to insure public safety [50]. Celebrated retired airline pilot Chelsey B. “Sully” Sullenberger III¹⁶ also reflected a similar sentiment when he said in his testimony to the US congressional Subcommittee on Aviation in a June 2019 meeting [49]: “*these crashes are demonstrable evidence that our current system of aircraft design and certification has failed us.*”

Separate inquiries from the US Senate¹⁷ and the US Congress¹⁸ have found both FAA and Boeing at fault for the 737 MAX crashes. Specifically, the US Congress report [49] found that “*a culture of concealment*” at Boeing and the “*fundamentally flawed*” nature of current regulatory system contributed to the two crashes. As a rectification measure FAA is planning to reform its certification process of airplanes.¹⁹ However, two years before 737 MAX was certified, it was argued in a NASA commissioned report published in 2015 [51] that the current aerospace standards are not conducive to certifying adaptive and autonomous systems. This observation underscores the fact that the MCAS system was designed as an autonomous system to operate in the background without requiring the pilots being aware of its existence. The details of the safety analyses for the MAX certification process were found to be scarce in the publicly available literature. Despite this lack of information, one could argue that exhaustive M&S activities for MCAS validation with component, system and vehicle level scopes could potentially expose the design flaw and would ensure the appropriate designation of safety risks during the design, V&V and safety certification processes.

¹⁶ Captain Sullenberger III performed safe emergency landing of an Airbus 320 aircraft that had lost all engine power on the Hudson River off Midtown Manhattan, NY on January 15, 2009. Despite this catastrophic failure, all 155 onboard souls were saved, thanks to his skillful handling of the situation.

¹⁷ NPR news report: [Senate Report Faults FAA And Boeing For Failures In Review Of 737 Max](#)

¹⁸ CNBC news report: [Congressional report faults Boeing, FAA for 737 Max failures, just as regulators close in on recertification](#)

¹⁹ Reuters news report: [FAA to reform new airplane safety approvals after 737 MAX crashes](#)

4 Modeling & Simulation in Rail Transportation

4.1 Introduction

The reviewed literature showed modeling and simulation can be used in a number of ways in the railway industry. M&S has been traditionally applied as a design and performance characterization tool focusing mainly on the static and dynamic performance of tracks and vehicles in the context of multi-body physics, structural dynamics, terramechanics, passenger comfort, etc. Some examples include:

- Safety evaluation of high speed railway embankment under heavy rainfall and dynamic train loads using a finite element model in [52].
- Understanding a complex railway network's (i.e., the Dutch railway system) macroscopic behavior with a view to innovate the core processes to manage it using game simulation technology in [53].
- Study of fault conditions in the traction power supply system in an electrical high speed railway using a physics-based model developed in the MATLAB/Simulink environment [54].
- Evaluation of railway passenger comfort using a multi-domain energy based bond graph model in [55].
- Early detection of potential collisions and resolution techniques in a complex railway network applying dynamic programming techniques in [56].
- Seismic performance of a high speed railway bridge system using a finite element model in [57].

In a magazine article published in 2008, it was detailed how the train manufacturer Bombardier Transportation utilized simulation for performing a number of dynamic analysis required for developing and producing bogies for railway vehicles.²⁰ The scope of these dynamic analyses are described in Figure 8. Physics-based models and finite element models are typically employed for conducting the simulation studies in railway engineering. The test coverage simulation studies can provide cannot be replicated with physical testing, which shows how important the roles M&S play are in the railway industry. Since one of the major focus of this report is transportation automation, a brief account of the current state of development and adoption of automation in the rail industry is provided next.

4.2 Automation in Rail Transportation

The railway industry is incorporating cyber elements into its operation to implement increasingly greater degrees of automation. Much like other transportation sectors, this trend is being driven by goals of efficient infrastructure utilization, mitigating effects of labor shortage, reduction of operating costs, enhancing safety, greater sustainability etc. Unlike road vehicles, aerospace or marine sectors, operating environments are relatively less dynamic for rail transportation, which potentially renders development of reliable and performant automation systems a more achievable proposition. Among different types of railways, urban train systems operate in a closed and simple environment (e.g., tunnels, elevated tracks, fenced tracks, etc.), and unsurprisingly automation to the highest level is already a reality in this transportation modality.²¹

²⁰ Global Railway Review article, 2008: [Simulations of running dynamics in bogie design and development](#)

²¹ Management consultant firm Wavestone report: [World's Best Driverless Metro Lines 2017](#)

Proliferation of automation in the rail industry has been observed in freight train systems as well. For example, a multi-organization automatic train operation (ATO) project has been recently announced that will take place in Finland over the European Train Control System (ETCS).²² Another example is the heavy-haul long distance rail network operated by Rio Tinto in Australia that has been fully automated to provide safety and productivity benefits.²³

Task				Type of analysis	Calculation method				
					Eigenvalue analysis	Qual-static analysis	Simulation		
							Straight track	Full curve	Curve transition
Internal Need				Eigenbehaviour	X		X		
Risk assessment	Customer's specification	Vehicle acceptance	Safety	Carbody sway in curve		X		X	X
				Safety against derailment		X		X	
				Track shift force			X	X	
				Stability	X		X		
				Ride characteristics			X	X	
				Track loading		X		X	
		Support of the specialists	Ride comfort			X	X	X	
			Wear		X		X		
			Gauging			X	X	X	
			Influence of external loads			X	X	X	
			Load collectives			X	X	X	






Figure 8: Typical dynamic analyses applied during railway engineering. ©Global Railway Review

The International Association of Public Transport (UITP) defined a Grades of Automation (GoA) (see Figure 9) framework to define the degree of automation deferred to an automated train control system in a published report [58]. GoA categorizes the role of automation into two main tasks: (a) Automatic Train Protection (ATP) and (b) Automatic Train Operation (ATO). While the functions of ATP involve basic safety such as avoiding collisions, preventing red signal overrunning, complying with speed limits etc., ATO implements partial or complete automation of train piloting and driverless functionalities [59]. It is claimed that urban rail systems (metro) have been operating with GoA 2 automation for more than 40 years now. More than 70 metro lines in 40 cities around the globe operate GoA 4 systems.²⁴ The “Canada Line” rapid transit is a GoA 4 system operating in Vancouver, BC.²¹

²² News report: [Autonomous rail freight transport at GoA 4 to be tested in Finland](#)

²³ News report: [Successful rollout of AutoHaul™ is celebrated by Rio Tinto](#)

²⁴ International Railway Journal article: [Automatic Train Operation takes to the main line](#)

Grade of Automation (GoA)	Type of train operation	Setting train in motion	Stopping train	Operation in event of disruption	Door closure
GoA 0 	No ATP	Driver (no supervision)	Driver (no supervision)	Driver	Driver
GoA 1 	ATP with driver	Driver	Driver	Driver	Driver
GoA 2 	ATP & ATO with driver	Automatic	Automatic	Driver	Driver
GoA 3 	Driverless	Automatic	Automatic	Train attendant	Train attendant
GoA 4 	Unmanned Train Operation (UTO)	Automatic	Automatic	Automatic	Automatic

Automatic Train Protection (ATP)

Automatic Train Operation (ATO)

Figure 9: Grades of automation (GoA) in rail transportation, adapted from [58] and [59].

4.3 Regulatory Standards

European Standards (EN) drafted and maintained by CEN (European Committee for Standardization), CENELEC (European Committee for Electrotechnical Standardization) and ETSI (European Telecommunications Standards Institute) are the most prominently cited in the related literature. These standards are described below:

- **EN 50126: Railway applications - The specification and demonstration of reliability, availability, maintainability and safety (RAMS)** [60] is the sector specific application of IEC 61508²⁵ standard. This standard aims to establish a systematic and coherent approach towards managing RAMS in all railway application.
- **EN 50128: Railway applications - Communication, signalling and processing systems - Software for railway control and protection systems** [61] standard applies to programmable electronic systems used in control and monitoring applications for railway protection. It defines a number of methods and techniques for developing and evaluating safety critical software used in railway sector [62].

²⁵ International Electrotechnical Commission (IEC) published the IEC 61508 standard as a basic functional safety standard applicable to all industries.

- **EN 50129: Railway applications - Communication, signalling and processing systems – Safety related electronic systems for signalling** [63] defines requirements for the acceptance of safety-related electronic systems used in railway signaling.
- **EN 50159: Railway applications - Communication, signaling and processing systems - Safety-related communication in transmission systems** defines performance requirements that a data transmission system must meet in order to be considered safe in railway applications. These data transmission systems are considered integral parts of electronic safety-related systems and facilitate information sharing between two or more locations.

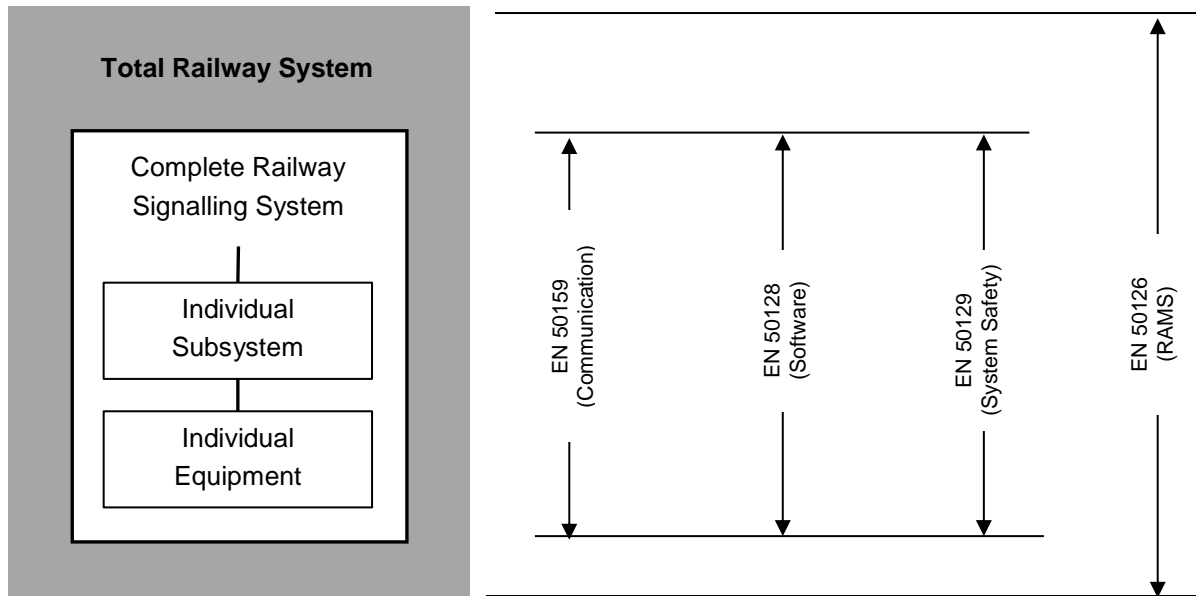


Figure 10: Scope of major CENELEC railway application standards (adapted from [63]).

The scopes of these aforementioned standards are pictorially shown in Figure 10. The overarching EN 50126 standard recognizes M&S as a means to validate system safety requirements. Although this standard accepts real or “*simulated*” conditions to collect “*objective evidence*” of requirements validation, additional details about M&S was found to be absent.

4.4 Railway Automation M&S Use Cases

A number of traditional use case of M&S have already been mentioned in Section 4.1. In addition, use cases that focus on cyber elements of railway systems are discussed in the following.

4.4.1 Testing CBTC Systems

Communication-based Train Control (CBTC) is a railway signaling system that utilizes telecommunication between train and track equipment for the purposes of traffic management and infrastructure control. Since it is characterized as a large scale (i.e., distributed over a large area with mobile and fixed equipment) safety-critical system, verification and validation activities can be complicated. Although CBTC is typically tested in a simulation environment, conducting the tests is manual which leads

to long test cycle, low test accuracy and low efficiency according to a conference paper presented at the 2020 IEEE International Conference on Intelligent Transportation Engineering [64]. This paper also proposes an automated test generation scheme for CBTC systems. They have utilized the open source test automation tool Robot Framework²⁶ to automate the testing tasks. This testing scheme belongs to the “experimental and empirical test strategies” methodology listed in the bottom right cell in Figure 4. The authors have claimed 10x efficiency improvement and reduction of bug identification failure rate to 0.08%. This test automation exercise is one example of how the rail industry is adopting M&S validation methodologies for a large scale safety critical system like the CBTC system.

Another example of applying simulation to validate the performance of wireless data communication pipeline used in a CBTC system is described in 2015 IEEE International Conference on Intelligent Transportation system [65]. The authors have adopted two simulators BRaVE and OMNeT++²⁷ for this purpose. BRaVE allowed the authors to simulate railway networks and accompanying CBTC system at a microscopic level, while OMNeT++ was used to simulate wireless networks with dynamic elements such as physical obstacles, fading, interference etc. By coupling these two simulators together the authors were able to study wireless communication performance for a simulated railway network. Finally, a failure mode triggered by loss of wireless connectivity between a train vehicle and a wayside access point was studied.

4.4.2 Validation of a Railway Signaling System

A team of industrial and academic researchers from Portugal and Hungary in 2017 authored a book chapter to detail a fault injection validation methodology for a safety critical railway signaling system [66]. The system under test features Triple Modular Redundancy (TMR) because it was rated as a Safety Integrity Level (SIL) 4 system, in accordance with EN 50126, 50128 and 50129 standards. The embedded hardware elements of the system were tested by simulating hardware faults through a JTAG (Joint Test Action Group)²⁸ interface. Equipped with an automated fault injection system, this study was able to trigger the system to enter a fail-safe mode. The authors demonstrated this fault injection validation scheme as an efficient method to perform V&V on safety critical systems.

²⁶ <https://robotframework.org/>

²⁷ <https://omnetpp.org/>

²⁸ JTAG is an electronics industry standard interface used for flashing firmware, debugging, boundary scan testing, digital simulation, etc.

5 Modeling & Simulation in Marine Transportation

5.1 Introduction

The reviewed literature showed that M&S activities in the marine industry has been used for many design and development tasks ranging from design of hull & propulsion systems design to development of the recent trend of adoption of autonomy. Application of CFD and FEA for design development and evaluation are two main ways the marine industry has adopted M&S. Some examples of related CFD-based M&S activities include:

- CFD simulation was employed to optimize the hull form of a trimaran²⁹ ship in order to reduce the hydrodynamic resistance in an article published in the Ships and Offshore Structures Journal in 2020 [67].
- In order to improve propeller efficiency, CFD simulation was used to optimize propeller design in a conference paper presented in 2018 [68].
- In an effort to conform with the regulation related to energy efficiency of ships mandated by international maritime organization (IMO),³⁰ viscous CFD simulation was used to evaluate performance of three energy saving devices (ESD)³¹ and hull combination in a paper published in the Ocean Engineering Journal in 2019 [69].

While CFD is applied to understand the hydrodynamics aspects of marine vessel design, FEA is used mainly on structural aspects of marine vessels and its components. Some examples include:

- Structural performance of large container ship hulls in severe waves was simulated by coupling CFD and FEA models in a research paper published in the Marine Structures journal in 2018 [70]. The CFD model evaluated the hydrodynamic and hydro-elastic phenomena in severe wave conditions, and the FEA model represented the deformable structural aspects of the hull. Tank testing of a scaled ship model was used to validate the coupled CFD and FEA models.
- In order to investigate the potential advantages marine propellers made with composite materials offer over traditional metal propellers (e.g., weight reduction, energy efficiency, superior hydrodynamic performance, easier maintenance), FEA was used in a research project funded by the French Department of Defence (DGA) to characterize the mechanical properties [71]. Specifically inter-laminar shear strength and fatigue behavior was studied. Later physical testing of a prototype was used to evaluate the accuracy of the FEA model.

The marine sector is unique because it has a long history of certification activities, dating back to 1760.³² These certification and regulations heavily rely on human input. For example, Rule 5 of internationally accepted marine regulation COLREGs (Convention on the International Regulation for Preventing

²⁹ A trimaran ship is a multi-hull marine vehicle consisting of one main hull and two smaller ones in each side. This hull form is typically used for recreational and racing vessels.

³⁰ [Energy efficiency measures mandated by IMO](#)

³¹ Some examples of ESD from industrial supplier Wartsila: [Energy Saving Devices Improve your vessel performance](#)

³² Lloyd's register: [A brief history](#)

Collisions at Sea, 1972) requires that “every vessel shall at all times maintain a proper look-out by sight and hearing as well as by all available means appropriate in the prevailing circumstances and conditions so as to make a full appraisal of the situation and of the risk of collision.”³³ This indicates more emphasis has been put on human perception with other means of environmental characterization (e.g., by sensors) playing an auxiliary role. This is just one example of how the long history of internationally accepted human-centric certification and regulatory practices may render it challenging to accommodate disruptive technologies like marine automation. In alignment with this report’s focus on transportation automation and the corresponding role M&S plays, a brief account of the current state of marine automation is provided next.

5.2 Automation in Marine Transportation

Supervised automation in the form of marine autopilots³⁴ or the regulatory term Heading Control System (HCS) already exists. The marine autopilots, just like aerospace autopilots, are intended to assist helmsmen to maintain a set heading. The CAV analogue would be the LKS (lane keeping system) which centers a cruising automobile inside the driving lane, while the driver is expected to continuously evaluate the safety of the vehicle and take back control as soon as it is necessary to do so. Marine autopilots are still subjected to COLREGs, which means navigational watchkeeping is still a human responsibility. However, the international marine regulatory body IMO addressed marine automation in the 100th session of its senior technical body, the Maritime Safety Committee (MSC). The term Maritime Autonomous Surface Ships (MASS) was used to refer to marine automation and 4 different levels of autonomy was defined to facilitate the regulatory scoping exercise in a 2018 press release by IMO [72] (see Table 4).

Although the regulatory, standardization and the legal frameworks addressing marine automation are still developing, there have been a few instances of automation demonstration of cargo and passenger ships. In December, 2018, Rolls-Royce Marine and Finnish state-owned ferry operator Finferries demonstrated autonomous and remote controlled operation of the Falco, a 53.8m long passenger ferry.³⁵ Kongsberg Maritime is building an autonomous and electric container ship Yara Birkeland with a planned launch in late 2021.³⁶ It was unclear from the reviewed literature how these projects have demonstrated compliance with existing IMO regulations. Marine automation for smaller vessels outside the scope of IMO regulations has already shown significant progress. For example, the California-based robotics company Liquid Robotics (later acquired by Boeing) developed the unmanned surface vehicle (USV) Wave Glider³⁷ which became the first autonomous robot to cross the Pacific Ocean in 2012.³⁸ Marine research organization ProMare³⁹ and technology giant IBM built a 15m long a fully autonomous ship Mayflower with a trimaran configuration that was scheduled to make a fully autonomous trans-Atlantic voyage in June, 2021 starting from Plymouth,

³³ IMO website article: [Convention on the International Regulations for Preventing Collisions at Sea, 1972 \(COLREGs\)](#)

³⁴ Cruising World magazine article: [Modern Sailboat Autopilots](#)

³⁵ Rolls-Royce news release: [Rolls-Royce and Finferries demonstrate world's first Fully Autonomous Ferry](#)

³⁶ Electrek news article: [Meet the world's first electric autonomous container ship](#)

³⁷ [Liquid Robotics Wave Glider](#)

³⁸ IEEE Spectrum article: [Liquid Robotics' Wave Glider Completes Pacific Crossing](#)

³⁹ <https://www.promare.org/>

UK with the final destination of Plymouth, Mass., USA. A mechanical issue forced ProMare to abort the mission.⁴⁰

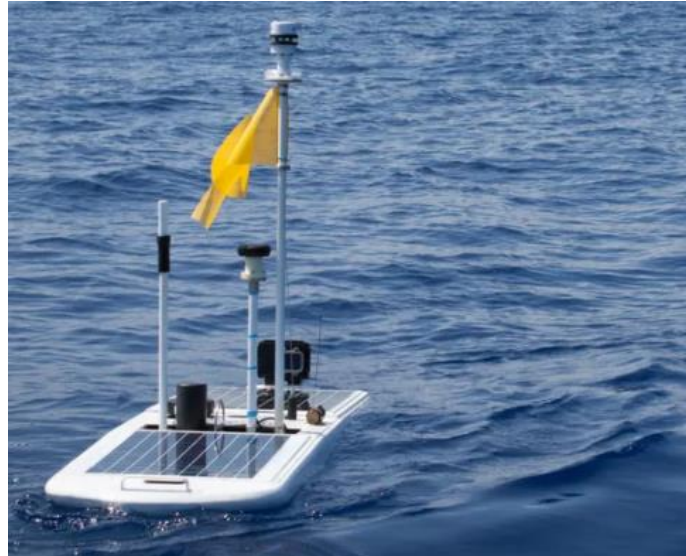


Figure 11: Wave Glider - an autonomous USV. © Liquid Robotics

Level of autonomy	Human presence	Operational Control	Human Role
Degree 1: Ship with automated processes and decision support	Yes	Seafarers are onboard to operate and control shipboard systems and functions. Some operations may be automated and at times be unsupervised but with seafarers onboard ready to take control.	Supervision and operation
Degree 2: Remotely controlled with seafarers onboard	Yes	Ship is controlled and operated from another location. Seafarers are available onboard as a fall-back measure.	Backup to maneuver and supervise the systems
Degree 3: Remotely controlled without seafarers onboard	No	Ship is controlled and operated from another location without any onboard seafarers.	Monitoring and remote control
Degree 4: Fully autonomous	No	The operating system of the ship is capable of making decisions and determines actions autonomously.	Monitoring and emergency management

Table 4: MASS levels of automation (adapted from [72] and [73]).

The abovementioned technology demonstrations and achievements suggest that the enabling technologies involving environmental perception, wireless communication and control algorithms have reached the

⁴⁰ Washington Post article: [An autonomous ship's first effort to cross the Atlantic shows the difficulty of the experiment](#)

maturity required to implement autonomous operation. The maritime classification society DNV GL rightly observes in a 2018 publication [74]: *“the main challenge for implementing fully automated systems controlled by remote operators or by algorithms is not to make them work, but to make them sufficiently safe. What is sufficiently safe, or has a tolerable risk level, will most likely be defined by a competent authority such as the International Maritime Organization (IMO) and flag states for any given operation.”*

5.3 Regulatory Aspects of M&S in Marine Transportation

The ecosystem of regulatory and standardization frameworks in marine transportation is composed of a number of different types of organizations. At the national level, the flag state (i.e., the jurisdiction in which a marine vessel is registered and licensed) have their own regulation and certifications practices, which are derived mainly from the international regulatory organization IMO standards and regulations. For example, Transport Canada (TC) is the authority that exercises flag state control for Canadian-flag vessels to ensure that they are inspected in accordance with Canadian and international regulations (if the vessel takes on international voyages).⁴¹ The inspection and certification process is usually delegated to a third party, known as *classification society*, by the flag state authority. Some TC recognized classifications societies are Lloyd’s Register, American Bureau of Shipping (ABS), Bureau Veritas, DNV GL, etc. Major functions of classification societies include establishing technical standards for construction and operation of ships based on experience and research, confirm conformance of designs and construction with these standards, survey ships during construction and commissioning, and periodically survey ships to ensure continued conformance with the applicable standards. Some of these classification societies have been in operation for hundreds of years (e.g., Lloyd’s Register and Bureau Veritas were founded respectively in 1764 and 1828), and they have developed deep technical and research expertise in ship design, building and operation. Correspondingly, national and international regulatory bodies such as IMO rely on their services to ensure safety and environmental sustainability in marine transportation. Some examples of how the classification societies have developed specific standards, recommended practice, guidelines for various M&S activities used in design synthesis, evaluation and training include:

- **ABS Guidance Notes on Safe Hull Finite Element Analysis of Hull Structures** [75] is a guideline on how FEA should be applied to evaluate ship hull structure strength. Modeling specifics such as mesh arrangements for a discretized representation of a hull, model verification criteria, etc. and simulation specifics such as definition of loading and boundary conditions to characterize model responses are detailed. Demonstration of conformance to this guideline in FEA analysis allows a ship designer to obtain ABS certification.
- **DNVGL-ST-0033 Maritime Simulator Systems** [76] is standard to specify requirements of performance of maritime simulator systems in order to demonstrate compliance with the Standard for Training, Certification and Watchkeeping – STCW.⁴² Marine simulators are used for training and certification of seafarers by creating certain condition by means of a model or to simulate situations related to maritime operation.
- **ABS Guidance Notes on Gas Dispersion Studies of GAS Fueled Vessels** [77] provides guidelines on how the CFD methodology should be applied for simulating dispersion of vented gas

⁴¹ Transport Canada publication: [Flag State Control](#)

⁴² Transport Canada publication: [Standard for Training, Certification and Watchkeeping - STCW](#)

from a pressure relief valve. Marine vessels that use combustion engines produce exhaust gases after fuel is burnt, and these harmful gases must be dispersed without posing any hazards to humans onboard. This document provides general guidelines with case studies for developing CFD-based gas dispersion models for gas fueled ships.

- **MED A.1/4.16 Heading Control System (HCS)** [78] is a type approval standard published by classification society DNV GL for certifying marine autopilots. This document details the certification procedure which involves submitting documentation describing the architecture of the system, built-in data security measures, functional diagrams for components, etc. This standard recognizes performance tests in the form of simulated noise and errors on the sensor input signals.
- **DNVGL-CG-0557 Data-driven Verification** [79] is a guideline by DNV GL that recognizes the need for introducing new validation & verification methods in lieu of “*methods that traditionally have relied on deploying personnel to the vessel*” in order to evaluate digital and connectivity technologies that are being deployed on marine vessels and offshore structures. This high level guideline identifies simulation as a means to verify functionality especially when it is being performed on a digital twin.⁴³

Besides these abovementioned use cases of M&S-based design verification and performance evaluation, it is expected that development of MASS will heavily rely on M&S. In regards to MASS, Henrik Ringbom, a Finnish maritime law professor, argues that the related regulatory standards are still in development, and “*the novelty of the subject represents an argument in favor of developing a new instrument to specifically address the various aspects of highly automated and autonomous ships*” [80].

5.4 Marine Automation M&S Use Cases

5.4.1 Vehicle-in-the-Loop Simulation of Marine Robot Swarm

A group of researchers from University of Zagreb, Croatia have developed a vehicle-in-the-loop (VIL) test environment for a swarm of marine robots that are being developed for long-term deployment to acquire and record environmental data from the lagoon of Venice, Italy [81]. The heterogeneous swarm consists of three types of autonomous robotic agents: floating, submerged and mobile and bottom dwelling. Swarm behavior of these robotic agents were studied in a ROS-based VIL simulator to demonstrate the long term deployment potential of the proof-of-concept system. This M&S use-case can be regarded as an example of mid-TRL technology development and demonstration exercise.

5.4.2 Simulation of Energy Performance of Ships

Reducing environmental impact of shipping has become a major focus of the industry, which is evidenced by the energy efficiency measures implemented by IMO.⁴⁴ Although CFD-based simulations can be used to demonstrate compliance to these IMO mandated energy efficiency, evaluation of a high fidelity CFD model under wide ranging sea-state conditions can be computationally expensive. Academic researchers from Egypt and the UK have developed a ship model equipped with appropriate sensors and data acquisition systems to acquire model-scale data from lake based and towing tank tests [82]. In order to support the test campaign, the ship model was deployed with autonomous self-propulsion features. The

⁴³ Digital twin is a virtual representation of a system or physical asset that makes system information available or evaluates performance through integrated models and data for the purpose of providing decision support [74].

⁴⁴ IMO article: [Energy Efficiency Measures](#)

data collected from the model tests were used to validate a physics-based ship model developed in a MATLAB/Simulink environment. Model-scale experiments and corresponding simulation tests were performed and the results were applied to identify energy efficient hull configurations for tanker ships.

5.4.3 Simulation-based Verification of Autonomous Navigation Systems

Researchers from the maritime classification society DNV GL have proposed a simulation-based verification system in a conference paper presented in the International Seminar on Safety and Security of Autonomous Vessels in 2019 [83]. The researchers observe that the real-life testing of autonomous navigation systems (ANS) cannot generate the test coverage required for performance assurance, which aligns with the RAND corporation study findings described in [3]. They had proposed a simulation-based testing environment as shown in Figure 12. The COLREGs regulations are used for test case generation. It should be noted that this work was intended as a roadmap for future verification activities for marine automation systems. Furthermore, the proposed system shares many philosophical and architectural similarities with simulation based testing approaches of CAVs.

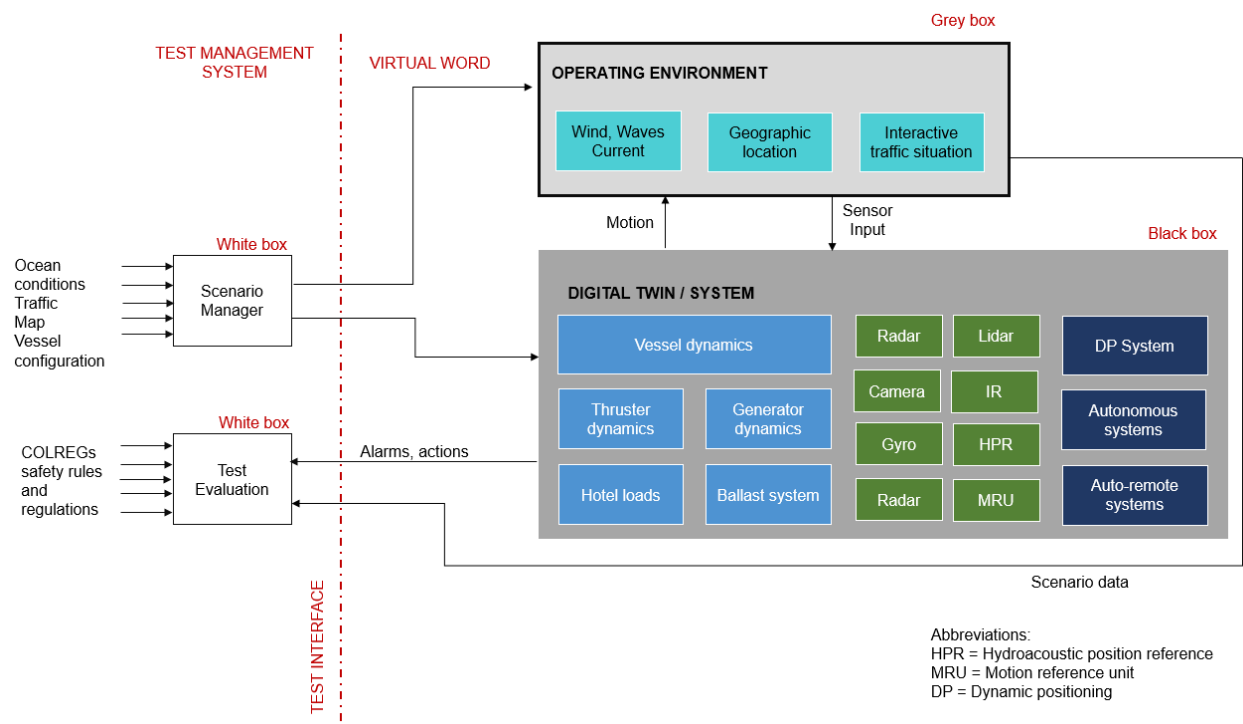


Figure 12: Simulation test system proposed in [82].

6 Modeling & Simulation in CAV

6.1 Introduction

This report excludes topics involving traditional application of M&S in the automotive industry such as CFD simulation to study and improve aerodynamic performance of the vehicle body or studying the thermo-fluidic phenomena inside an internal combustion engine, FEA modeling to characterize vehicle body's crash performance, or vehicle dynamics simulation for electronic stability control (ESC) or traction control (Anti-lock Braking System – ABS). Instead, this report examines how M&S is being applied to characterize emerging CAV technologies. Specifically, it reviews the related literature to summarize how M&S is being currently used or being developed for the future to study the safety case of these technologies. In order to frame the subsequent discussions with consistent nomenclature, a glossary of CAV definitions and terminology is provided next.

6.2 CAV Definitions & Terminology

A report published by the US DOT in 2018 [84] observes “*clear definitions and consistent use of terminology is critical to advancing the discussion around automation.*” Given the highly evolving nature of the CAV field, a plethora of terms can be found in the literature that have been used to communicate the idea of automating the driving function. Some of the more cited ones are defined below for the sake of consistency and clarity. Generally the definitions/concepts provided and described in the two US DOT reports published in 2018 and 2019 [84, 85], the SAE J3016 standard, and the 2021 United Nations Economic Commission for Europe publication on “New Assessment/Test Method for Automated Driving” [86] followed here.

Advanced Driver Assistance System (ADAS): Systems or features that assist human drivers in some aspects of performing the dynamic driving tasks (DDT) either by engaging in vehicle actuation (e.g., lane keeping system, adaptive cruise control, automated emergency braking) or by simply alerting the driver (e.g., audio-visual warning to driver when a potential collision is detected). Improving safety and reducing the cognitive load of DDT are two main value propositions of ADAS. Depending on the scope, ADAS features can be categorized as SAE L0, L1 or L2 automation.

A Priori Map: A HD map that is available to an ADS system during operation, which may have been constructed *a priori* to provide information such as detailed description of road topology, vertical signs, road markings or even a spatial representation of the road surroundings [87].

Automated Driving System (ADS): Technology stack composed of hardware and software that perform DDT on a sustained basis including those that are limited to specific operational design domain (ODD). Specifically, SAE L3, L4 and L5 systems are considered as ADS.

ADS-Dedicated Vehicle (ADS-DV): A vehicle designed to be operated in conformance with Level 4 or Level 5 SAE J3016 standard automation features for all trips. For brevity, henceforth SAE J3016 autonomy levels are referred to as SAE L1, L2, etc. to indicate the autonomy level of a vehicle.

Deterministic Model: A model whose temporal evolution can be predicted exactly. Example, a simple quarter car vehicle dynamics model.

Dynamic Driving Task (DDT): Real-time tasks involving environmental perception, decision making and motion control of a motor vehicle, which are required to operate a vehicle in on-road traffic. It excludes mission-level functions as trip scheduling, selection of destination and way points. Depending on the automation levels, DDT can be either a share or sole responsibility of human drivers or ADS systems. As the degree of automation increases in a vehicle, necessity of human input for DDT diminishes correspondingly.

Highly Automated Vehicle (HAV): Vehicles that are designed to be exclusively operated as SAE L4 and L5 vehicles for all trips and are not equipped with manual driving controls [2].

High Definition Map (HD Map): A spatial representation of a roadway environment up to centimeter-level precision and detailed roadway information including lane width, location of road markings, street signs, directions of travel, road junction information, speed limits, etc. HD maps can be built dynamically using simultaneous localization and mapping (SLAM) techniques or it can be supplied as an *a priori* map.

Object List: A map of all static and dynamic elements of a roadway environment detected by the sensors in an ADS, which are not part of an *a priori* map. Dynamically constructing and maintaining an object list in order to acquire comprehensive characterization of the operating environment is the role of sensor fusion and processing in many ADS architectures.

Object and Event Detection and Response (OEDR): Tasks that are considered subsets of DDT involving monitoring the driving environment in terms of detecting and characterizing static and dynamic objects and events, planning appropriate responses, and executing them as the evolving driving environment necessitates.

Operational Design Domain (ODD): ODD refers to the operating conditions under which an ADS or feature thereof is designed to function including, but not limited to, environmental, geographical, and time-of-day restrictions, and/or the requisite presence or absence of certain traffic or roadway characteristics (definition source: SAE J3016). ODD principally describes the limitations of an ADS so that it can be safely operated as it was designed to.

Sensor Simulation: It refers to the process of generating synthetic sensor signals from a model of a sensor operating in a virtual environment. For example, a LiDAR model can be used to generate point cloud data to represent a virtual roadway environment. The simulated sensor data then can be fed into a perception algorithm to evaluate its efficacy.

Stochastic Model: A model that incorporates uncertainty of the physical world that is typically used to represent real world phenomena with greater fidelity.

Validation of the Simulation Model: It refers to the process of determining how closely a simulation model represents the real world from the perspective of the intended uses of the tool.

Verification of the Simulation Model: It refers to the process of determining the degree of conformance of a simulation model or a virtual testing tool with a set of specifications and requirements that the model is expected to meet.

In addition to these abovementioned concepts and definitions, the terms “scene”, “situation” and “scenario” have been used in the CAV simulation testing literature interchangeably in variety of contexts and meanings. In an effort to disambiguate these terms German academic researchers Ulbrich *et. al.* defined them in a paper presented in the 2015 IEEE International Conference on Intelligent Transportation Systems [88]. These definitions are provided in Table 5.

Term	Definition by Ulbrich <i>et. al.</i>
Scene	A scene describes a snapshot of the environment including the scenery and dynamic elements, as well as all actors’ and observers’ self-representations, and the relationships among those entities. Only a scene representation in a simulated world can be all-encompassing (objective scene, ground truth). In the real world it is incomplete, incorrect, uncertain, and from one or several observers’ points of view (subjective scene).
Situation	A situation is the entirety of circumstances, which are to be considered for the selection of an appropriate behavior pattern at a particular point of time. It entails all relevant conditions, options and determinants for behavior. A situation is derived from the scene by an information selection and augmentation process based on transient (e.g. mission-specific) as well as permanent goals and values. Hence, a situation is always subjective by representing an element’s point of view.
Scenario	A scenario describes the temporal development between several scenes in a sequence of scenes. Every scenario starts with an initial scene. Actions & events as well as goals & values may be specified to characterize this temporal development in a scenario. Other than a scene, a scenario spans a certain amount of time.

Table 5: Definitions of the terms “scene”, “situation” and “scenario” by Ulbrich *et. al.*

6.3 Rationale for M&S in CAV

Researchers from the independent US proving ground Transportation Research Center and NHSTA argues in a conference paper published in 2017 [89] that the traditional testing methods for automotive safety systems cannot be directly adopted for high-level driving automation systems. Typical testing method of automotive safety systems involve identifying a finite set of scenarios in which the system under test (SUT) is designed to improve safety, and subsequently conducting a testing campaign using these scenarios. Test results thus obtained provide regulators physical evidence to quantify safety of the SUT. However, driving automation technologies cannot be tested following this traditional testing method because the scenarios that the CAV systems are expected to operate in can only be characterized as a combinatorial explosion of virtually infinite numbers of permutations and combinations of operating environment factors (e.g., dynamic traffic situation, weather, surrounding environment, road conditions) and the instantaneous state of the vehicle (e.g., vehicle speed, heading, sensor robustness, performance of perception and path-planning algorithms). Sufficient coverage of these virtually infinite number of scenarios cannot simply be achieved with traditional testing methodologies. Correspondingly, the authors observe that either large

number of physical testing or simulation testing of the algorithms in a validated simulation environment can generate sufficient test coverage to address this challenge.

A research paper from the San Jose State University, California published in 2017 [90] recognized “shadow driving” as the most common and well-reported automated vehicles (AV) testing methodology. Shadow driving involves a human driver supervising the safety of the automated driving performance, who must remain available at all times to take over the control of the vehicle whenever a potential unsafe event is imminent. A statistical treatise published by the research organization RAND corporation [3] estimated the mileage and the corresponding time requirement of AV testing in real traffic. A safety benchmark of 1.09 fatalities per 100 million miles was chosen [3], which reflects the number of fatalities caused by US drivers in 2013. Choice of this benchmark was not arbitrary, rather it attempted to set the minimum safety performance of AV platforms, which qualitatively refers to a performance standard that is at least as safe as human drivers. It was estimated in [3] that in order to achieve this benchmark with a statistical 95% confidence limit 275 million miles of test driving is required (see Table 6). A fleet of 100 AVs continuously driving for 12.5 years at an average speed of 25 mph can attain such a vast test coverage. In comparison, Waymo’s 55 vehicle AV fleet have been test driven approximately 1.3 million miles in autonomous mode, and was involved in 11 non-fatal crashes from 2009 to 2015, according to a report from Virginia Tech Transportation Institute [91]. In light of the immensity of the test coverage required to confirm the safety of an AV platform with some statistical significance, M&S have been identified as a feasible complementary testing method besides on-road testing by the AV development community.

Benchmark Failure Rate				
Statistical Question	How many miles (years*) would autonomous vehicles have to be driven ...	(A) 1.09 fatalities per 100 million miles?	(B) 77 reported injuries per 100 million miles?	(C) 190 reported crashes per 100 million miles?
	(1) without failure to demonstrate with 95% confidence that their failure rate is at most...	275 million miles (12.5 years)	3.9 million miles (2 months)	1.6 million miles (1 month)
	(2) to demonstrate with 95% confidence their failure rate to within 20% of the true rate of...	8.8 billion miles (400 years)	125 million miles (5.7 years)	51 million miles (2.3 years)
	(3) to demonstrate with 95% confidence and 80% power that their failure rate is 20% better than the human driver failure rate of...	11 billion miles (500 years)	161 million miles (7.3 years)	65 million miles (3 years)
*Assessment of the time it would take to complete the requisite miles with a fleet of 100 autonomous vehicles (larger than any known existing fleet) driving 24 hours a day, 365 days a year, at an average speed of 25 miles per hour.				

Table 6: Examples of text coverage required to demonstrate AV reliability [3].

Regulatory forums such as United Nations Economic Commission for Europe (UNECE) recognize the need for a multi-pillar approach for ADS assessment and testing which includes audit, simulation/virtual testing, test track, and real-world testing in a report published in 2021 [86]. This report identifies M&S as a powerful tool to assess the performance of an ADS under diverse and complex conditions, which are prohibitive in conventional physical testing. NHSTA has identified simulation as one of the five potential methods to verify compliance with the FMVSS safety standard in a report published in 2020 [2]. Solely software-based simulation activities and testing methodologies that interface with physical hardware components (i.e., hardware-in-the-loop/HIL) with a simulation framework have been mentioned as a potential verification method for CAV systems.



Figure 13: Global vehicle target (GVT) specified by Euro NCAP. © Euro NCAP

European New Car Assessment Programme (Euro NCAP) have published a test protocol for an important driving automation safety feature – automated emergency braking (AEB) in [92] where test conditions and tasks related to data acquisition and processing are specified. The physical test protocol make use of a Global Vehicle Target (GVT)⁴⁵ (see Figure 18) to provide a geometrically realistic target for the AEB sensors to simulate a stopped vehicle. Although perception sensors (LiDAR, camera, radar, etc.) and the corresponding software comprise a typical AEB stack, the protocol specifies a few test cases that the AEB must perform well in. The stochastic aspects of the AEB functionality such as robustness of AEB sensor performance against ambient noise/events (e.g., radio interference in radar, camera image sensor saturation due to solar exposure, false positive LiDAR pulses, etc.) or road conditions (i.e., the protocol specifies all testing must take place in dry conditions) are not taken into account. The large number of scenarios that an AEB system is expected to perform in are not sufficiently covered by this test protocol. Alternatively, if a model is available for the vehicle under test (VUT) that can simulate AEB behavior with some level of stochasticity integrated with an appropriate vehicle dynamics model, then the AEB feature can be tested virtually in a simulated environment in a large number of test cases to generate sufficient test coverage with a specified degree of statistical significance. Of course, the model being tested must be validated in some way so that its fidelity can be considered sufficiently high for the task at hand. Physical testing of AEB will still be required for tasks such as model development, calibration and validation.

⁴⁵ Euro NCAP Technical Bulletin: [Global Vehicle Target Specification](#)



Figure 14: Example of an edge case - a floating balloon can be challenging for AI-enabled perception. © AEye

Simulation testing is particularly useful for discovering edge cases in ADS operation. Edge cases are defined as those events that occur at extreme operating conditions. Typically they are ultra-low frequency events with high impacts. By definition, edge cases occupy only a fraction of the space defined by all combinations of operating environment factors and system states. Discovering edge cases, thus, through physical testing is difficult and can be achieved if a brute-force testing methodology is adopted. Physical testing is resource intensive, and thus brute-force testing is not a tenable proposition. Furthermore, physical testing of edge cases can be hazardous because of the nature of the situation. On the other hand, simulation testing is scalable and cost efficient. Once a simulation test is setup by developing, calibrating and validating the underlying models, scaling up the test coverage is relatively simple. However, it must be examined carefully what is the minimum fidelity of the models is required to achieve the test objective, because model complexity and the corresponding development effort typically increases exponentially as higher degree of fidelity is sought.

Finally, there is often a lack of legislation or it is too restrictive to road test autonomous vehicles, which further highlights the need for modeling and simulation. As highlight above, the reasons for this hesitancy range from technological risks, social risks, economic risks and adaption risks. Rosique *et. al.* in a 2019 literature review [93] found 12 countries permit partial public road access for testing and of these, only 6 countries and a select few US states allow unrestricted access. In relation to Canada, however, non-regulatory guidelines for public road testing have existed since 2018 (e.g., guidelines published for CAV testing published in 2018⁴⁶ and updated in 2021⁴⁷).

6.4 CAV M&S Use Cases

The seminal paper titled “How Many Miles of Driving Would It Take to Demonstrate Autonomous Vehicle Reliability?” from RAND Corporation is a highly cited work that brings attention to the stochastic nature of ADS validation. In addition, follow-up reports published in 2018 and 2020 continue to form the basis of key concepts for ADS safety and its validation. Fraade-Blanar *et. al.* in their 2018 report [94] observed “*competition among AV developers, varying approaches to simulation, and a lack of AV simulation*”

⁴⁶ TC document, 2018: [Testing Highly Automated Vehicles in Canada: Guidelines for Trial Organizations](#)

⁴⁷ TC document, 2021: [Guidelines for Testing Automated Driving Systems in Canada Version 2.0](#)

standards constrain comparisons of different AVs.” The report also warns against confirmation bias in simulation testing of ADS. Lack of understanding of the requirements for fidelity, operating ranges and insufficiently validated models can lead to drawing positive conclusions from bad simulation tests.

Blumental *et. al.* in their 2020 report [95] recommends to focus on ADS development processes so that the safety argument is considered a core element, especially because existing standards are not sufficiently developed to assess ADS safety. This report also identifies government at different levels as the authority responsible for data-driven communication of ADS safety to the public. Simulation testing is one of the more practical ways to generate this data-driven evidence with sufficient test coverage so that inferences can be made with some acceptable degree of statistical significance. Schnelle *et. al.* in the US DOT report titled “Review of Simulation Frameworks and Standards Related to Driving Scenarios” published in 2019 [85] provided an illustrative example of a generic simulation testing framework can complement test track and on-road testing of ADS (see Figure 15). The authors treated the models representing ADS system, features or subsystems as a black box, which contrasts with both the aerospace and the marine domains. As described in previous chapters, in both aerospace and marine transportation sectors certification of simulation based testing of cyber elements requires disclosure of information intrinsic to the system.

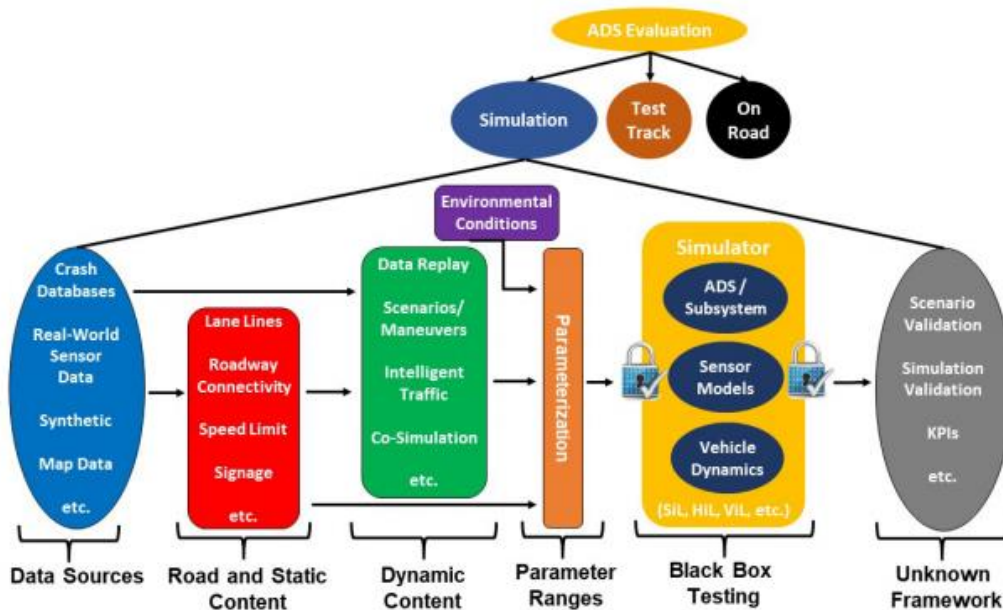


Figure 15: Generic ADS testing process [85].

In a review article published in the Journal of Advanced Transportation in 2019 [96], authors Do *et. al.* have provided an extensive survey on how simulation based studies have been applied for CAV development. Most of the reviewed use cases of simulation were focused on technology development and demonstration with a few focusing on GHG emissions and efficient utilization of roadways. Although the reviewed use cases can be categorized in a few themes (e.g., vehicle platooning, transportation system throughput, new ADAS algorithm), the applications represented a wide spectrum.

In order to understand how a simulation model for CAV can be built in a variety of ways, a representative subsystem level architecture of a generic CAV system is first reviewed. In this regard Figure 16 is referenced where information flow among various hardware and software elements of a CAV system is

described. In addition, how these elements dynamically interact with the physical world (i.e., the vehicle itself and the operating environment) is described. The various features of the operating environment (static & dynamic) including the motion state of the vehicle are captured by photonic (LiDAR and camera), radio (radar & GNSS) and mechanical (IMU & sonar) signals. These captured raw signals are then fed to perception algorithms so that (a) the operating environment can be both qualitatively and quantitatively characterized (i.e., what objects are around the vehicle and what are their locations relative to the vehicle), and (b) the position of the vehicle can be determined (i.e., localization). The situational awareness thus obtained then is fed to path planning algorithms that determine how the vehicle will negotiate the continuously evolving driving environment in terms of a prescribed path. Additional means of obtaining situational awareness may include application of *a priori* HD maps and NLOS (non-line-of-sight) situational awareness provided by V2X technologies. Finally, the low-level control algorithms follow the prescribed path by engaging the actuators while taking feedback from the instantaneous vehicle states. The vehicle states are also influenced by the aerodynamics and the road noise imposed by the operating environment.

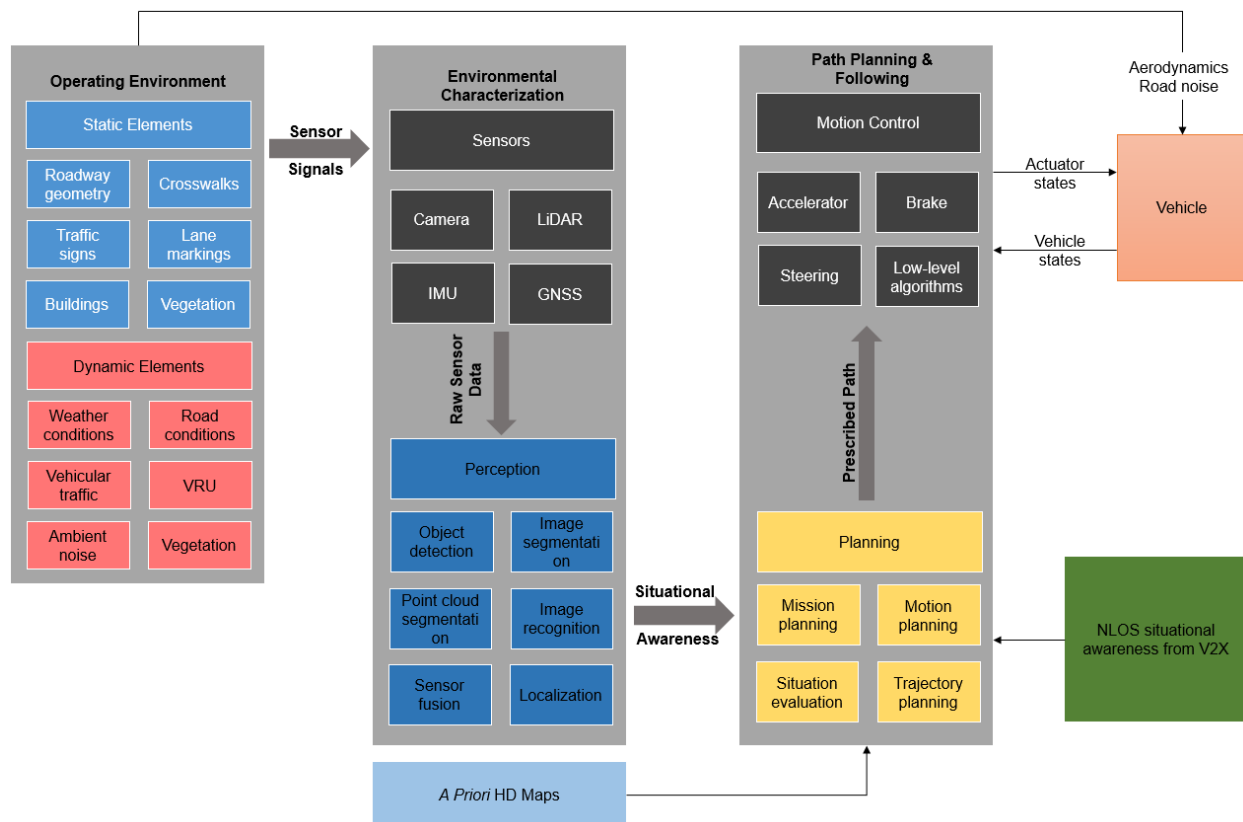


Figure 16: CAV system architecture.

Figure 16 describes how a CAV operates in the real world. Based on the objective and the requirements of a simulation exercise, some or all parts of this architecture are replaced by simulation models, as shown in Figure 17. In order to illustrate this point, a few simulation testing themes are described below. In each theme a few use cases are described to demonstrate how diversified simulation testing can be for CAV M&S activities.



Figure 17: Elements of a typical virtual driving simulator platform, as presented in [97].

6.4.1 Sensor Simulation

Typically sensor simulation is performed to understand sensor responses and quantify performance. A simulated operating environment represents the virtual test environment. Physics based equations are used to generate synthetic data of the environment (i.e., sensor simulation). The virtual environment provides the ground truth against which the synthetic data can be compared. A group of US academic researchers used such a scheme to study effects of rain on LiDAR in a paper published in the journal *Electronics* in 2019 [98]. The effects of rainfall was represented by an empirical model. The resultant simulation model can be used in a variety of ways such as understanding sensor performance degradation as a function of weather, improving robustness of a point cloud segmentation algorithm against rainfall, studying how rainfall can influence automation features that rely on LiDAR by augmenting additional models, etc. It should be noted the body of knowledge involving sensor simulation is still maturing. In a research paper published in the journal *IEEE Transactions on Intelligent Vehicles* in 2020 that examined methods and models for simulating all major ADS perception sensors [99], the authors remarked that “*significant work remains to be done in order to produce simulated data that is realistic and can be used to confidently test and evaluate autonomous agents in virtual proving grounds.*”

In 2021 academic and industry researchers from Germany published an interesting simulation study on radar sensors [100], which demonstrates how flexible simulation studies can be. In order to study radar-based ADAS performance, the authors employed a vehicle-in-the-loop (VIL) testing scheme where the vehicle under test (VUT) was deployed on a chassis dynamometer. A virtual test environment was integrated with the chassis dynamometer to represent realistic road load on the VUT. In addition, the radar sensor which is a part of the VUT’s ADAS stack was simulated in accordance with the virtual test environment. The authors present this testing scheme as a scalable alternative to physical testing.

6.4.2 ADAS Testing

Depending on the ADAS feature under test and the test objectives, a number of simulation model architectures can be devised. For example, academic researchers from Japan have used physics based vehicle dynamics models to develop a motion control algorithm for a common ADAS feature – Adaptive

Cruise Control (ACC), and published their findings in 2020 in a research paper [101]. Information related to cyber features (e.g., determining speed and distance from a lead vehicle) was assumed to be available to the ACC model, and an advanced motion controller was developed that optimized passenger comfort and GHG emissions. This study is a good example of how model fidelity should correspond to the test objectives. Since the test objective was to develop a motion controller that optimizes passenger comfort and GHG emissions, simulating cyber aspects of the ACC application with high fidelity models was deemed not necessary.

Some of the challenges involving validating autonomous systems have been demonstrated in a conference paper by academic researchers from Canada, Luxembourg and Sweden, which was presented at the 2021 IEEE conference on Software Testing, Verification and Validation [102]. The same model of a vision based pedestrian detection system as an ADAS feature was tested using a Search-based Software Testing (SBST) solution in two different commercial simulators; namely PreScan and Pro-SiVIC. This simulation study was based on 400 safety critical cases, and the results are presented in Table 7. Among all the test scenarios, in 229 and 236 cases respectively recorded from PreScan and Pro-SiVIC simulators the system under test did not detect the hazard, which led to a safety violation. It should be noted that the system under test was designed to issue only a warning if safety hazard was detected. Event triggered braking was part of its operational scope. Out of the total 800 safety-critical test scenarios, the two simulators could not produce the expected outcome (i.e., a collision) in total 59 cases. Because of the closed-source nature of the simulators, it was not possible to investigate this discrepancy. However, this observation underscores the challenges one may face in validating an autonomous system. After presenting qualitative arguments about the benefits of simulation testing of CAV, the authors recommended using more than one simulator for simulation based validation of ADAS features.

	PreScan	Pro-SiVIC
Safety-critical test cases	400	400
Safety violations (system failed)	229	236
Detections	171	164
Collisions	396	345

Table 7: Results of the simulation study reported by Borg *et. al.* [102].

Industry and academic researchers from Germany and Brazil in a joint effort employed simulated photorealistic roadway environments shown on a HD display to quantify performance of a proprietary multi-class object detection algorithm under various degrees of ambient photonic noise in a camera-in-the-loop simulation test, as described in a conference paper presented at the IEEE International Conference on Vehicular Electronics and Safety in 2018 [103]. Since this object detection algorithm was intended for ADAS applications, it must provide robust performance regardless of time-of-day or weather. This camera-in-the-loop simulation testing method enabled the authors to validate the underlying deep neural network (DNN) model.

6.4.3 Virtual V2X Demonstrations/Testing

V2X technologies enable NLOS (non-line-of-sight) situational awareness and maneuvers coordinated with neighboring vehicles to potentially further improve safety of CAV systems. Physical testing of V2X

applications is even more challenging because it requires multiple test vehicles and an appropriate wireless infrastructure to support connectivity. Unsurprisingly, the related literature prominently uses simulation based testing as the preferred method for V2X validation. In typical V2X simulation testing, more emphasis is put on the configuration of the roadway network and the composition of the vehicles and other autonomous agents (e.g., pedestrians, cyclists, etc.) that comprise the traffic environment. Wireless communication is also modeled to introduce physical events such as radio interference, latency, packet loss, etc. Simulation studies of a number of safety-critical and non-safety V2X applications have been described in an earlier NRC report titled “Technical Review of Safety Use Cases, Benefits and Safety Vulnerabilities Associated with Connected Vehicle Technologies.”⁴⁸

6.5 CAV Standards Relevant to M&S

US DOT recognizing the highly evolving nature of CAV technologies wishes to develop voluntary and consensus-based technical standards and approaches with input from CAV developers [84]. Correspondingly, they have invited CAV developers to disclose Voluntary Safety Self-Assessments (VSSA) to demonstrate that (a) developers are considering safety aspects and (b) they are collaborating and communicating with regulators. In addition, VSSA can potentially encourage self-establishment of industry safety norms and build public trust in CAV technologies. An index of the VSSA provided by leading CAV developers including Apple, Waymo, GM, Ford, Nvidia, and Mercedes can be found in the NHSTA website.⁴⁹ Most of these VSSAs mention simulation testing as a means for system verification, but mention of M&S specific standards was found to be scarce in these documents. These VSSA documents indicate that adoption and development of M&S specific standards remain in an early stage. Nonetheless, standards are expected to play a vital role in CAV development. The national standards body of the United Kingdom BSI⁵⁰ in a report published in 2020 identified three key purposes for CAV standards in [104]:

1. Encouraging collaboration among experts to help consolidate the state of the art.
2. Enabling interoperability of products from different manufacturers.
3. Discourage public deployment of immature or unsafe technology.

CAV standardization efforts from BSI have resulted in a number of publicly available specifications (PAS). Some of those standards are summarized in Table 8. Although these PASs do not strictly focus on M&S, some of the concepts inform how M&S activities for CAV should be developed and executed so that the corresponding results can achieve compliance to safety standards.

Publicly Available Specifications (PAS)	Description
PAS 1880:2020 Guidelines for Developing and Assessing Control Systems for Automated Vehicles	Initial guidelines for control system design for automated vehicles
PAS 1881 :2020 Assuring the Safety of Automated Vehicle Trials and Testing	Minimum requirements for assuring the safety case for automated vehicles trials and testing

⁴⁸ NRC technical report published in 2021: Technical Review of Safety Use Cases, Benefits and Safety Vulnerabilities Associated with Connected Vehicle Technologies

⁴⁹ NHSTA index: [Voluntary Safety Self-Assessment](#)

⁵⁰ <https://www.bsigroup.com/>

PAS 1882:2021 Data collection and management for automated vehicle trials for the purpose of incident investigation	Minimum requirements for data recording on a CAV
PAS 1883: Operational design domain (ODD) taxonomy for an automated driving system (ADS)	Minimum requirements for a hierarchical taxonomy of operational design domain (ODD)

Table 8: BSI developed CAV standards.

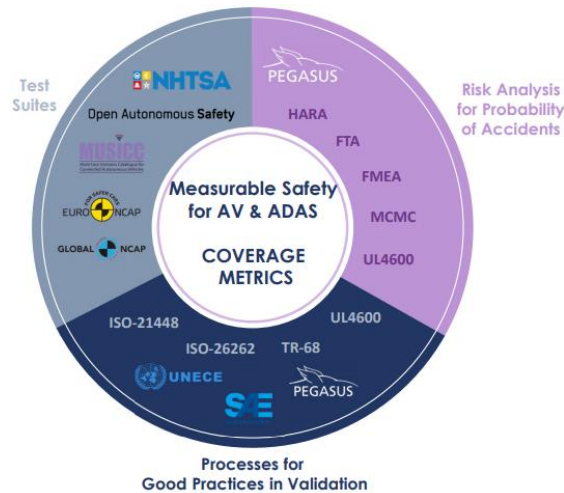


Figure 18: ADS verification and validation tools, practices, and standards.⁵¹ ©Foretellix

Another standardization body, Association for Standardization of Automation and Measuring Systems (www.asam.net), is in the process of developing a number of standards focusing solely on M&S activities for automated driving systems. The OpenSCENARIO, OpenLABEL and Open Simulation Interface (OSI) standards are specifically being developed to address the lack of standardized approach to realistic scenario based simulation. Since simulation testing of ADS systems will involve a number of heterogeneous industries such as automakers and their supply chain, simulation software vendors, standardization bodies, testing organizations, etc., the lack of a standardized approach can potentially impede progress. How the different developing standards and test protocols contribute to ADS verification is shown by automated driving verification firm Fortelix in a presentation provided to the UNECE forum (Figure 18).

SAE have developed a recommended practice document (SAE J3049) in 2015 [105] with the goal to establish a common framework for developing simulation models and virtual environments in order to address the need for simulation testing to be portable across a number of different types of organizations such as OEMs, automotive parts suppliers, government agencies, research institutes etc. This recommended practice guide provides an architectural structure of a vehicle system partitioned into subsystem models organized according to the real-world interactions of these subsystems. For example, a vehicle dynamics model interfaces with road topology models. This guide also defines standard interfaces so that subsystem models can be replaced in a “plug & play” manner to accelerate simulation model

⁵¹ UNECE Wiki article: [Coverage Driven Verification](https://www.unece.org/transport/automotive/actr/actr-projects/coverage-driven-verification/)

construction. Another relevant SAE standard is J2998⁵² which was first published in 2014 and was revised in 2020. This standard defines the recommended information content that should be included with dynamical model elements used for simulating ground vehicle systems. In addition, SAE is currently engaged in some precursory work that are expected to produce documents and standards in this area. Two such initiatives are On Road Automated Driving Simulation Task Force⁵³ and On Road Automated Driving Verification and Validation committee.⁵⁴

Although ISO 26262 (Functional safety) is an often cited standard in relation to automotive safety, some literature point out that the safety case of ADS cannot be proven by compliance to ISO 26262 alone. For example, the 2019 SAE paper [4] remarks that the ISO 26262 is not suitable for ADS verification. Schnelle *et. al.* in [85] observes “an ADS system that perfectly fulfills its design specifications according to ISO 26262 still could cause or be involved in road crash and harm.” Other relevant standardization documents/activities include:

- ISO/PAS 21448:2019 Road vehicles — Safety of the intended functionality⁵⁵
- ISO/TR 21934-1:2021 Road vehicles — Prospective safety performance assessment of pre-crash technology by virtual simulation — Part 1: State-of-the-art and general method overview⁵⁶
- ISO/TC 22/SC 33/WG 9 – Test scenarios of automated driving systems⁵⁷
- ISO/TC 22/SC 33/WG 11 – Simulation⁵⁸
- IEEE P2846 – Assumptions for Models in Safety-Related Automated Vehicle Behavior⁵⁹
- UNECE WP.29 Validation Method for Automated Driving (VMAD) sub-group on virtual testing⁶⁰

6.6 Reported M&S Activities from Key Players

It was mentioned before that VSSAs from notable CAV developers indicate simulation testing being an integral part of the development activities. However, publicly available information about these activities beyond what is available in the VSSAs was found to be scarce because disclosure of such information can potentially create disadvantageous situations from a business point of view. However, a summary of what little information is available in the public domain is given below.

6.6.1 Waymo

Waymo employs three basic system-level testing for their CAV development, which are simulation, closed-course and public roads [106]. They claim they have conducted simulation testing representing over 15

⁵² SAE standard: [Model Description Documentation Recommended Practice for Ground Vehicle System and Subsystem Simulation J2998](#)

⁵³ SAE task force: [On Road Automated Driving Simulation Task Force](#)

⁵⁴ SAE committee: [On-Road Automated Driving Verification and Validation](#)

⁵⁵ [ISO/PAS 21448:2019](#)

⁵⁶ [ISO/PAS 21934-1:2021](#)

⁵⁷ ISO working group on automated driving test scenarios

⁵⁸ ISO working group on simulation

⁵⁹ [IEEE Standards Association working group](#)

⁶⁰ [UNECE Validation method for automated driving \(VMAD\)](#)

billion miles of simulated driving and on-road testing of over 20 million miles [107]. Because Waymo's M&S activities may constitute industry secrets, publicly available information about these were found to be scarce. A report from Boeing [22] informs that Waymo used the "Carcraft" automated driving simulator for simulation testing. A 2020 magazine article describes Carcraft as computationally expensive as physics based sensor models to generate synthetic perception data.⁶¹ Waymo's ADS use this synthetic data to perform *simulation driving*. A 2020 conference article jointly authored by researchers from University of Texas, Austin, Waymo and Google Brain [108] suggest that Waymo is using a neural network based approach to create a synthetic driving environment representation.

6.6.2 Mercedes-Benz

Mercedes-Benz creates virtual driving scenarios consisting of maps, traffic participants (other vehicles, pedestrians, etc.) with behavioral modeling, static and dynamic objects [109]. Ambient conditions such as lighting and weather are also simulated. It is claimed that the simulation testing allows Mercedes to not only evaluate ADS performance within the defined ODD, but also extreme conditions to validate ADS fallback performance.

6.6.3 Uber ATG

Uber ATG (automated driving unit) employs simulation and HIL testing to evaluate ADS performance [110]. Similar to software releases, Uber ATG employs regression testing to validate ADS prior to deployment. Their regression testing involves a set of simulated on-road scenarios in which the ADS software must demonstrate acceptable performance. HIL testing validates that the execution of ADS software on deployment hardware can be performed correctly and without any anomalous incidents (e.g., processing entering a locked state while handling multiple threads). Uber ATG claims that they are focused on the reliability of simulation results, but how they achieve this is not clear in the available literature. It should be mentioned that Uber ATG was acquired by another automated driving company Aurora at the end of 2020.⁶²

6.7 Overview of Software Tools

6.7.1 Description of Test Scenarios

The reviewed literature generally agree that simulation-based validation of ADS will utilize a set of test scenarios under which virtual testing will be conducted. In order to make this approach efficient and effective a variety of organizations such as regulators, standardization bodies, academia and industry need to work together to build a common database of ADS test scenarios. The SAE J3049 recommended practice also recognizes the need for common interface and languages for simulation model development. Safety Pool (www.safetypool.ai) is one such initiative that aims to create a scenario database so that a common ecosystem wherein the international community of industry, academia and policymakers can collaborate together towards the goal of certifiable safety for ADS can be developed. In order to bring this vision to life, a common standardized method of describing scenarios must be developed. To this end, a number of scenario description languages have been found in the literature. For example, Scenic is one such scenario description language [111]. The efficacy of this new language was demonstrated by applying Scenic on a

⁶¹ Venturebeat magazine article: [Waymo is using AI to simulate autonomous vehicle camera data](#)

⁶² NASDAQ report: [Uber sells ATG self-driving business to Aurora at \\$4 billion](#)

car detection algorithm based on a convolutional neural network. Another example is OpenSCENARIO⁶³ which is a XML-based file format that is developed to describe dynamic content of driving and traffic simulators. GeoScenario is another XML-based scenario description language proposed by Canadian academic researchers from University of Waterloo [112]. The Measurable Scenario Description Language (M-SDL)⁶⁴ is from ADS validation firm Foretellix.



Figure 19: Three roadway scenes generated from a single ~20 line Scenic file [111].

6.7.2 Automated Driving Simulators

Automated driving simulators are software environments that typically provide a way to create a virtual environment with static and dynamic roadway elements. Previously discussed scenario description language may serve as an input to these simulators so that the described scenario can be rendered. In addition, a model or a physical implementation of the system under test (SUT) is interfaced with the virtual environment typically in a closed loop configuration (i.e., the SUT receives stimulus from the virtual environment to produce responses, which in turn changes the state of the virtual environment).

There are many automated driving simulators available. Some proprietary ones are CarCraft (Waymo), SurfelGAN (Waymo), Webviz (Cruise), The Matrix (Cruise), DataViz (Uber), etc. Widely adopted open source simulators include CARLA, LGSVL, Gazebo, etc. An overview can be found in a survey paper authored by a group of academic researchers from USA [113] (see Table 9).

Requirements	Description	MATLAB/Simulink	CARLA	Gazebo	LGSVL
Perception	Support for sensor models	Y	Y	Y	Y
Perception	support for different weather conditions	N	Y	N	Y
Camera Calibration		Y	Y	N	N
Vehicle Control	Support for physics-based vehicle dynamics	Y	Y	Y	Y
3D Virtual Environment		Y	Y	Y	Y

⁶³ ASAM OpenSCENARIO: <https://www.asam.net/standards/detail/openscenario/>

⁶⁴ Foretellix blog: [Expressing More Scenario Needs with the New M-SDL 20.7 Release](#)

Requirements	Description	MATLAB/Simulink	CARLA	Gazebo	LGSVL
Traffic Infrastructure	Traffic lights, signage etc.	Y	Y	Y	Y
Traffic Scenario Simulation	Support for different types of dynamic objects	Y	Y	N	Y
2D/3D Ground Truth		Y	Y	U	Y
Interfaces	With other software	CarSim, PreScan, ROS	ROS	ROS	Autoware, Apollo, ROS
Scalability	For example, via a server multi-client architecture)	U	Y	Y	Y
Open Source		N	Y	Y	Y
Well-maintained	Updated regularly	Y	Y	Y	Y
Portability	Multiple OS support	Y	Y	Y	Y
Flexible API	Application programming interface	Y	Y	Y	Y

Legends: Y = Yes, N = No, U = Unknown

Table 9: Comparison of automated driving simulators, adapted and updated from [113].

6.7.2.1 MATLAB/Simulink

Automated Driving Toolbox is part of the MATLAB/Simulink framework that supports the design, simulation and testing of ADAS and autonomous driving systems [114]. Core functionalities that the Automated Driving Toolbox help enable include vision and LiDAR perception systems, sensor fusion, path planning and vehicle control. The toolbox supports maps from HERE HD Live Map⁶⁵ and road networks from OpenDRIVE.⁶⁶ It also supports the automatic labeling of ground truth for training and evaluation of perception algorithms through the Ground Truth Labeler⁶⁷ app. Road networks can be designed and generated through RoadRunner,⁶⁸ which could then be used for hardware-in-the-loop testing for automated driving blocks such as perception, sensor fusion, path planning and control logic. This simulation environment also allows for the simulation of sensor output and detection rendered using the Unreal Engine.⁶⁹ Finally there are several examples already developed to simulate various ADAS features such as Adaptive Cruise Control (ACC), Automatic Emergency Braking (AEB), Forward Collision Warning (FCW), Lane Keeping Assist (LKA) and parking valet.

6.7.2.2 CARLA

CARLA is an open-source simulator for autonomous driving research [115]. The simulator was designed to be scalable and modular to address the wide range of tasks involved in the problem of autonomous driving. The API is flexible and is implemented in Python and C++. The simulation runs on the Unreal Engine and the standard used to define roads and maps is OpenDRIVE. Traffic scenarios are defined using the

⁶⁵ <https://www.here.com/platform/automotive-services/hd-maps>

⁶⁶ <https://www.opendrive.com/>

⁶⁷ MathWorks app: [Ground Truth Labeler](#)

⁶⁸ [MathWorks product RoadRunner](#)

⁶⁹ <https://www.unrealengine.com/en-US/unreal>

ScenarioRunner module. The sensor suite is configurable to include sensors such as LiDARs, cameras, depth sensors and GPS. CARLA is also fully compatible with ROS.

6.7.2.3 LGSVL

The LGSVL⁷⁰ is an open source simulator used for robotics and autonomous vehicle development. LGSVL Simulator consists of the simulation software along with the tools and ecosystem to enable the testing of tailored use cases. These use cases include multiple ego vehicles, configurable sensor layouts, traffic, dynamic external obstacles and pedestrians. The paid, premium version also features cloud simulation, allowing for simulation and scenario testing at large scales, as well as CI/CD⁷¹ integration.

Features of the LGSVL Simulator include the possibility to define and test scenarios involving complex traffic situations using real-world data, creation and debugging of localization modules in digital twin environments, test planning module in isolation with virtual ground truth detections, software-in-the-loop testing of entire autonomous vehicle stack, hardware-in-the-loop testing, and automatic execution of scenarios to ensure safety and functionality over interesting and edges cases. LGSVL Simulator runs on the Unity engine⁷² and is both Linux and Windows compatible. In regards to communication interfaces, LGSVL supports ROS, ROS2 and CyberRT⁷³ messages using default bridges, with the possibility to build the proper interface for custom or proprietary communication protocols.

6.7.2.4 Gazebo

Gazebo⁷⁴ is a popular open source simulator designed to rapidly test algorithms, perform regression tests and train AI robotic systems using realistic scenarios. Gazebo features a modular design, capable of using different physics engines including ODE, Bullet, Simbody, and DART. Gazebo is also built upon the Ogre3D, allowing for realistic renderings of the environment. The simulator also offers plugins, allowing for the integration of sensors and noise from LiDARs, stereo cameras, GPS, IMU and RADARs. Gazebo also supports cloud simulation on Amazon AWS, and can operate on both Linux and Windows platforms. Finally, what makes Gazebo one of the more popular simulators in the AV domain is the inherent integration with ROS and ROS2.

6.7.3 Traffic Simulators

In conjunction to automated driving simulators, traffic simulators represent traffic behavior within a roadway network. They are typically used to develop better strategies for more efficient system operation. For example, a traffic simulator can be used to study throughput improvement effected by a V2X application such as smart intersection management. These simulators can simulate traffic behavior at the macroscopic (e.g., traffic network of an entire city), mesoscopic (e.g., traffic network composed of a few neighboring intersections) and microscopic levels (e.g., traffic network involving a single intersection). Both proprietary

⁷⁰ [The SVL Simulator by LG Electronics America R&D Center](#)

⁷¹ CI/CD refers to a software engineering practice of continuous integration and continuous deployment that seeks to bridge the gaps between development and operation activities.

⁷² <https://unity.com/>

⁷³ CyberRT is an open source framework for autonomous driving scenarios.

⁷⁴ <http://gazebo-sim.org/>

and open source tools are available. Examples include SUMO (Simulation of Urban Mobility),⁷⁵ CORSIM,⁷⁶ MATSIM (Mutli-Agent Transport Simulation),⁷⁷ VISSIM,⁷⁸ etc. In some cooperative automation (i.e., automation enabled by V2X technologies through cooperative perception and coordination of maneuvers) applications traffic simulators can be integrated with automated driving simulators to cover both aspects of the application. For example, SUMO and CARLA were interfaced together in a co-simulation scheme to study cooperative driving automation by a group of US academic researchers in 2021 [116].

⁷⁵ <https://www.eclipse.org/sumo/>

⁷⁶ <https://ops.fhwa.dot.gov/trafficanalysistools/corsim.htm>

⁷⁷ <https://www.matsim.org/>

⁷⁸ <https://www.ptvgroup.com/en/solutions/products/ptv-vissim/>

7 Conclusion & Summary of Findings

7.1 Research Questions & Answers

This technology review set out to answer a few research questions. After reviewing the related literature, the following findings are presented as answers.

How modeling, simulation and simulation-based testing is defined within the examined transportation sectors?

Generally a model is a representation of a physical system that can be simulated to gain better understanding of system behavior. Typically M&S exercises are undertaken when physical testing is resource intensive to implement or it simply cannot provide the insight simulation testing can. The objectives of the task at hand (e.g., design development vs performance characterization) and the corresponding constraints (e.g., sufficient knowledge of system inner workings to enable white-box models vs data-driven black-box models that do not require any understanding of the inner structure of the system), models can take many forms. To illustrate this point, let's consider two simulation exercise scenarios for a marine ship: (a) designing a cargo ship hull for tropical waters, and (b) designing a polar ship hull that will encounter floating ice. In the first case, a typical FEA model may be sufficient to synthesize a design that will meet all application-specific requirements. In the latter case, the unique operational requirements involving ice interaction render the design exercise a lot more complicated. This design problem may need a detailed FEA model of the hull coupled with other models involving ice mechanics and CFD. Since applications of M&S in aerospace, railway, marine and CAV sectors can be represented by a wide spectrum, this question is answered with sector-specific examples below.

Aerospace

Aerospace systems can be regarded as an architecture of multi-domain subsystems that are interconnected with a complex topology. Traditionally CFD and FEA based M&S are used in the aerospace industry for developing mechanical, thermodynamic and aerodynamic aspects of aeroplane design. In addition, flight simulators are also prominently used not only for certification and training of pilots, but also to characterize aircraft performance in a simulated environment. Every aspect of aerospace engineering is represented by models under the MBSE (Model Based Systems Engineering) paradigm.

Rail Transportation

In the railway industry M&S is traditionally applied as a means to design development and performance characterization tool. The scopes of design problems that are solved using M&S methodologies may vary from microscopic (e.g., seismic performance of a high speed railway bridge) to macroscopic (e.g., optimization of entire railway networks). Dictated by the objective of the M&S exercise at hand, the models can take many forms ranging from physics based models to data driven models. With the advent of automation enabled by sophisticated computer algorithms and corresponding hardware elements, the railway sector is evolving to become cyber-physical systems. M&S in the context of railway transportation as a CPS system can be used to evaluate functionality of safety-critical systems such as communication-based train control (CBTC), railway signaling or automated operation.

Marine Transportation

Similar to aerospace industry, CFD and FEA are traditionally applied in the marine sector for efficient ship hull design and to prove the structural integrity is sufficient under a virtual prototyping philosophy. Design optimizations exercises such as new energy efficient propeller design, energy efficient ship hull configuration, etc. are also performed using M&S methodologies. Introduction of cyber elements into ships has recently paved the way for marine automation. M&S exercise are being performed for studying the performance and robustness of such new systems.

CAV

Adoption of automation for the operation of a motor vehicle delineates CAV from traditional motor vehicles. The models described/used in the related literature were diverse in nature covering all CPS modeling paradigms. For example, component level simulation involving sensor models to generate synthetic data represents one end of this spectrum. On the other end, simulation of a cooperative driving automation application involved two different categories of simulators (i.e., traffic simulator and automated driving simulator).

How is simulation testing currently utilized within the sectors examined? Are there regulatory requirements or frameworks?

Simulation testing is typically used as a resource efficient and more scalable alternative to physical testing and physical prototyping. Since automation is an emerging trend in all four sectors examined, this technical review was focused on those aspects of simulation testing that are related to automation.

Established simulation models are accepted as evidence of conformance to standards in the aerospace domain. However, for emerging technologies such as automation, the available body of literature suggests that it is a relatively new concept. Unsurprisingly, the ATTOL (autonomous taxi, take-off and landing) project⁷⁹ led by Airbus and completed in 2020 was claimed as the “world-first” demonstrating automation of routine aircraft operations tasks of taxi, take-off and landing, which indicates adoption of autonomy in the aerospace sector is still at its early stage. Correspondingly, application of M&S activities for V&V of autonomous systems was found to be limited in the aerospace domain. Although safe operation of traditional software components with deterministic behavior can be proven by demonstrating adherence to standards such as DO-178C, a number of research articles (e.g., [117], [118], [119]) argue that the current aerospace standards cannot accommodate the non-deterministic nature of operation of adaptive and autonomous systems, which are regarded as emerging technologies in the aerospace sector. These arguments, observations and recommendations are summarized below:

- Because current civil aviation processes are predicated on the idea the correct behavior of a system must be comprehensively specified and verified prior to operation, it is not clear how adaptive and autonomous systems can obtain the necessary certifications. As a result, these systems are more prominently found in the military aviation and space domains [117].
- Processes for certifying and verifying machine learning (ML) models for safety-critical applications are still evolving [118], and is an active area of research [117].
- Rapid growth of practical artificial intelligence (AI) technologies for autonomous systems has rendered modernization of system engineering (SE) methodologies an immediate necessity. It should be noted that application of SE methodologies are ubiquitous in the aerospace industry, and

⁷⁹ News report: [Airbus Concludes ATTOL Project That Featured ‘World-First’ Automated Takeoffs and Landings](#)

integration of AI technologies into V&V focused SE processes can be challenging. Therefore, V&V approaches must evolve to address autonomous system testing [119].

- Some researchers are developing aerospace software specific workflow so that existing standards framework can be used to obtain certification for emerging technologies. Once such work authored by German academic researchers was published in 2018 [120], which proposes a “lean” and “highly automated” software development process that can be potentially used for demonstrating compliance of software developed for unmanned systems, urban air mobility and general aviation. However, Bhattacharyya *et. al.* in [117] argue that the current aerospace certification processes without changes can only be adopted for adaptive and autonomous systems in limited scope.

In the railway sector, besides the typical use cases of design development, optimization and virtual prototyping, simulation testing can be accepted as demonstrable evidence for proving conformance of cyber elements with safety standards or to study robustness against failure to evaluate sufficiency of performance of safety critical systems. Details of these certification related aspects of simulation testing was found to be scarce in the publicly available literature. This contrasts with the aerospace industry, a conventionally highly regulated sector with abundance of information involving certification processes.

Classification societies in the marine sector play the role of surveyors and evaluators with authority delegated by regulators. Classification societies publish recommended practice and guideline documents for traditional modeling and simulation exercises so that the results can be used to validate a new design. However, certification of some marine automation systems such as marine autopilots is a well-reported process involving disclosure of proprietary system architecture and simulation testing of the electronic and software elements. For emerging marine automation technologies, the marine community recognizes the vital role simulation testing will play for validating these systems. However, the corresponding frameworks of standards and conventions are still developing.

In the CAV sector, most M&S exercises are being utilized for functionality demonstration, algorithm development, performance characterization etc. The CAV community have reached a consensus that simulation testing will be a one of the means of certification of driving automation systems. However, the standards are still developing. Regulatory activities regarding simulation testing of CAV from the EU are more prominently presented in the literature.

What are the best practices and lessons learned?

The following best practices and learned lessons are summarized from the reviewed literature:

- Understanding the fidelity of the simulation models and their limitations are prerequisites for drawing tangible conclusions from the results.
- If possible, a model must be validated and calibrated with data obtained from physical experiments so that it can be ascertained that the model is an adequate representation of the physical system.
- Diversity of test coverage is as important as volume of test coverage.
- Validating a CPS system that is expected to operate in a stochastic physical world with 100% reliability and 100% confidence level is not possible. Hence, some degree of residual risks must be accepted for deploying a CPS.
- When frameworks of standards and conventions are not sufficiently mature, stakeholders must collaborate together to develop them.

- Scenario based testing paradigm is being developed not only for CAV, but for any safety-critical applications that involve artificial intelligence in their operation.

7.2 Open Questions/Issues

Some standardization efforts such as traffic scenario description, road network description, virtual environment, labeling of ground truth data etc. will pave the way for exchange of information and knowledge among the stakeholders. This can lead to the development of a CAV simulation test scenario database that is widely adopted by regulators, standardization & certification bodies, technology developers, researchers etc. Similar practices already exist in the road transportation sector in the form of a national collision database. Simulating these scenarios will require vehicle level or component level models representing ADS and an appropriate automated driving simulator environment. While standardization and creation of test scenario databases is expected to be beneficial, employing multiple simulators to characterize reproducibility in a test campaign is advisable because it broadens the likelihood of discovery of anomalies and bugs. This vision of simulation testing of CAV to ensure safety leads to a few open questions and issues:

- Who develops and subsequently validates these models representing ADS vehicles and components?
- What standards these models must be validated to and who certifies them?
- How automated driving simulators themselves can be validated to execute the test scenarios with sufficient fidelity?

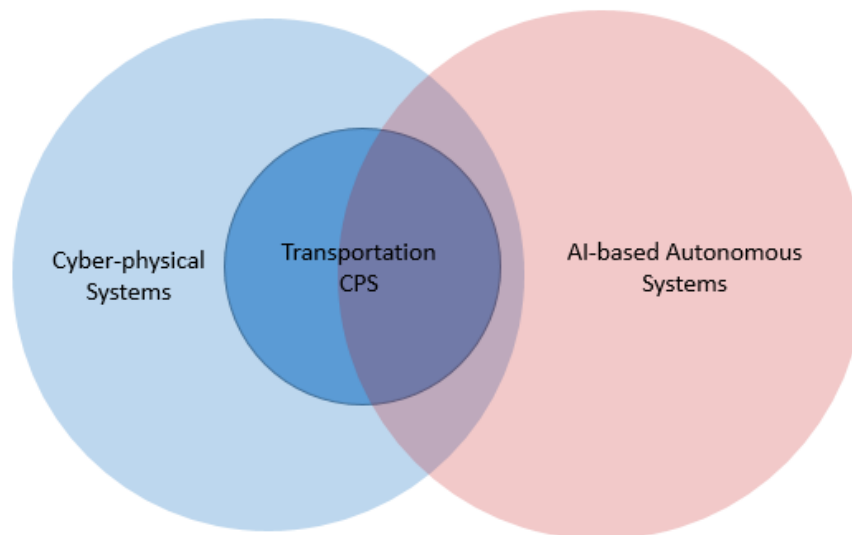


Figure 20: Categorization of the problem of CAV simulation testing.

7.3 Concluding Remarks

Simulation testing of CAV represents the intersection of two hallmark problems from two highly emerging fields: validation of cyber-physical systems and validation of AI-based autonomous systems (see Figure

20). After reviewing the M&S related literature from three other transportation sectors; namely, the aerospace, the railway and the marine sectors, it was found that modeling and simulation are traditionally applied for design synthesis, optimization and performance characterization. In some cases, simulation results were employed for safety certification with mandated demonstration of adherence to standards and best practices. For example, in the marine domain, FEM models are employed to prove structural integrity of the ship. Marine classification societies have published best practices documents so that these models can be developed accordingly. It should be noted that in cases where simulation results are accepted as certification evidence the underlying problem is well-defined; e.g., the ship hull is expected to experience this much load at the worst case scenario, the maximum allowable latency for a train signaling system can be determined from the maximum train speed, etc. In contrast, the CAV test and performance scenarios are not as well defined because these systems operate in a much more stochastic environment. Beyond the traditional applications of M&S, frameworks of standards were found to be developing for emerging use cases especially those involving cyber elements and particularly automation. Regulatory, standardization and policymaking documents involving CAV validation using simulation testing were found to focus on the philosophical aspects of the problem. Beyond these abstract ideas, specific regulations were found to be developing for CAV simulation-based validation. Nonetheless, scenario based testing and developing national, regional and international databases of CAV test scenarios are two related trends that are prominently discussed in the related literature.

Acknowledgements

Contributions from Stephen Sweeney (NRC), Stephanie Pavey (TC) and Benoit Ancil (TC) are gratefully acknowledged. Editorial feedback and suggestions/review of technical content from them improved the quality of this report.

References

- [1] National Highway Traffic Safety Administration, US Department of Transportation, "Federal Automated Vehicles Policy: Accelerating the Next Revolution In Roadway Safety," September 2016.
- [2] M. Chaka, M. Stow, L. Gabler, H. Weinstein, K. Gibbons and R. Fitchett, "FMVSS considerations for vehicles with automated driving systems: Volume 1 (Report No. DOT HS 812 796)," National Highway Traffic Safety Administration, 2020.
- [3] N. Kalra and S. 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*, vol. 94, pp. 182-193, 2016.
- [4] M. A. Appel and Q. Ahmed, "Intelligent Vehicle Monitoring for Safety and Security," in *SAE World Congress*, Detroit, Michigan, 2019.
- [5] A. V. Jayakumar, Systematic Model-based Design Assurance and Property-based Fault Injection for Safety Critical Digital Systems, PhD Thesis, Virginia Commonwealth University, 2020.
- [6] D. G. Broo, U. Boman and M. Törngren, "Cyber-physical systems research and education in 2030: Scenarios and strategies," *Journal of Industrial Information Integration*, vol. 21, p. 100192, 2021.
- [7] Y. Hou, Y. Zhao, A. Wagh, L. Zhang, C. Qiao, K. F. Hulme, C. Wu, A. W. Sadek and X. Liu, "Simulation-Based Testing and Evaluation Tools for Transportation Cyber-Physical Systems," *IEEE Transactions on Vehicular Technology*, vol. 65, no. 3, pp. 1098 - 1108, 2016.
- [8] X. Zhou, X. Gou, T. Huang and S. Yang, "Review on Testing of Cyber Physical Systems: Methods and Testbeds," *IEEE Access*, vol. 6, pp. 52179 - 52194, 2018.
- [9] R. Rai and C. K. Sahu, "Driven by Data or Derived Through Physics? A Review of Hybrid Physics Guided Machine Learning Techniques With Cyber-Physical System (CPS) Focus," *IEEE Access*, vol. 8, pp. 71050-71073, 2020.
- [10] K. Song, D. Upadhyay and H. Xie, "A physics-based turbocharger model for automotive diesel engine control applications," *Proceedings of the Institution of Mechanical Engineers, Part D: Journal of Automobile Engineering*, vol. 233, no. 7, pp. 1667--1686, 2019.
- [11] P. Z. Schulte and D. A. Spencer, "State Machine Fault Protection Architecture for Aerospace Vehicle Guidance, Navigation, and Control," *Journal of Aerospace Information Systems*, 2019.

- [12] S. Boudoudouh and M. Maâroufi, "Multi agent system solution to microgrid implementation," *Sustainable Cities and Society*, vol. 39, pp. 252-261, 2018.
- [13] R. Lee, *AdaStress: Adaptive Stress Testing and Interpretable Categorization of Safety Critical Systems*, Pittsburg, PA, USA: PhD Thesis, Carnegie Mellon University, 2019.
- [14] E. Goldberg, "On Bridging Simulation and Formal Verification," in *International Workshop on Verification, Model Checking, and Abstract Interpretation*, 2008.
- [15] P. Matousek, O. Rysavy, G. de Silva and M. Danko, "Combination of simulation and formal methods to analyse network survivability," in *Proceedings of the 3rd International ICST Conference on Simulation Tools and Techniques*, 2010.
- [16] M. Möstl, J. Schlatow, R. Ernst, N. Dutt, A. Nassar, A. Rahmani, F. J. Kurdahi, T. Wild, A. Sadighi and A. Herkersdorf, "Platform-Centric Self-Awareness as a Key Enabler for Controlling Changes in CPS," *Proceedings of the IEEE*, vol. 106, no. 9, pp. 1543 - 1567, 2018.
- [17] A. Donzé, "BreachFlows: Simulation-Based Design with Formal Requirements for Industrial CPS," in *2nd International Workshop on Autonomous Systems Design (ASD 2020)*, 2020.
- [18] P. Ulbig, U. Durak, D. Müller, T. Stripf and C. C. Insaurralde, "Flight simulator-based verification for model-based avionics applications on multi-core targets," in *AIAA Scitech 2019 Forum*, Sand Diego, USA, 2019.
- [19] C. Ebert and M. Weyrich, "Validation of Autonomous Systems," *IEEE Software*, vol. 36, no. 5, pp. 15-23, 2019.
- [20] US Department of Defense, "Unmanned Systems Integrated Roadmap 2017-2042," US DoD, 2017.
- [21] P. Junietz, U. Steininger and H. Winner, "Macroscopic safety requirements for highly automated driving," *Transportation research record*, vol. 2673, no. 3, pp. 1--10, 2019.
- [22] J. Fadaei, "The state of modeling, simulation, and data utilization within industry: An autonomous vehicles perspective," *arXiv preprint arXiv:1910.06075*, 2019.
- [23] M. Baraheni, A. Bagheri, B. A. Alaei and S. Amini, "Ultrasonic-assisted friction drilling process of aerospace aluminum alloy (AA7075): FEA and experimental study," *International Journal of Lightweight Materials and Manufacture*, vol. 4, no. 3, pp. 315-322, 2021.
- [24] B. C. Jin, "Design Optimization and Higher Order FEA of Hat-Stiffened Aerospace Composite Structures," in *Optimum Composite Structures*, IntechOpen, 2018.

- [25] A. SinghChauhana, B. Anirudhb, A. Satyanarayanaa and P. Rallapall, "FEA optimization of injection parameters in ceramic core development for investment casting of a gas turbine blade," in *10th International Conference of Materials Processing and Characterization*, 2020.
- [26] P. R. Spalart and V. Venkatakrishnan, "On the role and challenges of CFD in the aerospace industry," *The Aeronautical Journal* ,, vol. 120, no. 1223, pp. 209-232, 2016.
- [27] I. Lind and H. Andersson, "Model Based Systems Engineering for Aircraft Systems – How does Modelica Based Tools Fit?," in *Proceedings of 8th Modelica Conference*, Dresden, Germany, 2011.
- [28] A. Hebbar, A. Pashilkar and P. Biswas, "Using Eye Tracker To Evaluate Cockpit Design -- A Flight Simulation Study," *CoRR*, vol. abs/2106.07408, 2021.
- [29] D. H. Klyde, A. K. Lampton, D. G. Mitchell, C. Berka and M. Rhinehart, "A New Approach to Aircraft Handling Qualities Prediction," in *AIAA Scitech 2021 Forum*, 2021.
- [30] Department of National Defence Canada, "Airworthiness Design Standards Manual (ADSM)," Chief of the Defence Staff, 2020.
- [31] Ministry of Justice, Canada, "Canadian Aviation Regulation: SOR/96-433," 28 June 2021. [Online]. Available: <https://laws-lois.justice.gc.ca/PDF/SOR-96-433.pdf>.
- [32] P. D. a. D. Fogarty, "Safety Issues and Shortcomings with Requirements Definition, Validation and Verification Processes," US Department of Transportation, 2016.
- [33] Radio Technical Commission for Aeronautics, USA, "DO-178C - Software Considerations in Airborne Systems and Equipment Certification," RTCA, 2011.
- [34] M. Rubin, "The Role of Software Simulators in the Independent Verification and Validation of Commercial Space Vehicles," in *AIAA Scitech 2019 Forum*, 2019.
- [35] National Aeronautics and Space Administration (NASA), "NASA-STD-8739.8A: SOFTWARE ASSURANCE AND SOFTWARE SAFETY STANDARD," NASA, 2020.
- [36] National Aeronautics and Space Administration, USA, "NASA-GB-8719.13 Software Safety Guidebook," NASA, 2004.
- [37] SAE International, "ARP4761 Guidelines and Methods for Conducting the Safety Assessment Process on Civil Airborne Systems and Equipment," SAE, 1996.
- [38] SAE International, "Guidelines for Development of Civil Aircraft and Systems ARP4754A," SAE, 2010.

- [39] E. M. Peterson, "Application of SAE ARP4754A to Flight Critical Systems," NASA, 2015.
- [40] D. S. Loubach, J. C. Marques and A. M. da Cunha, "Considerations on Domain-Specific Architectures Applicability in Future Avionics Systems," in *FT2019. Proceedings of the 10th Aerospace Technology Congress*, Stockholm, Sweden, 2019.
- [41] Radio Technical Commission for Aeronautics, USA, "RTCA DO-254 Design Assurance Guidance for Airborne Electronic Hardware," RTCA Inc., 2000.
- [42] Radio Technical Commission for Aeronautics, USA, "DO-297 Integrated Modular Avionics (IMA) Development Guidance and Certification Considerations," RTCA, 2005.
- [43] U. Herberth, J. Rende and H. Lutz, "Development of Inertial Sensors for AHRS considering DO-254," in *2018 DGON Inertial Sensors and Systems (ISS)*, Braunschweig, Germany, 2018.
- [44] S. Haider, "Applying Model Based Safety Assessment for Aircraft Landing Gear System Certification," in *2020 Annual Reliability and Maintainability Symposium (RAMS)*, Palm Springs, CA, USA, 2020.
- [45] S. Gupta, U. Durak and S. Hartmann, "Simulation Scenarios for Testing Airborne Cyber-Physical Systems," in *Virtual ASIM Workshop 2021*, 2021.
- [46] H. Kato, D. Hirano, S. Mitani, T. Saito and S. Kawaguchi, "ROS and cFS System (RACS): Easing Space Robotic Development," in *2021 IEEE Aerospace Conference (50100)*, Big Sky, MT, USA, 2021.
- [47] B. S. Cruz and M. de Oliveira Dias, "Crashed Boeing 737-MAX: Fatalities or Malpractice?," *Global Scientific Journal*, vol. 8, no. 1, pp. 2615--2624, 2020.
- [48] P. Johnston and R. Harris, "The Boeing 737 MAX Saga: Lessons For Software Organizations," *Software Quality Professional*, vol. 21, no. 3, pp. 4-12, 2019.
- [49] Majority Staff of the Committee on Transportation and Infrastructure, "The Design, Development & Certification of the Boeing 737 Max," US House of Representatives Committee on Transportation and Infrastructure, September 2020.
- [50] R. R. G. Jr, "The FAA's Aircraft Design Approval Process Must be Overhauled," *Federal Bar Association Inland Empire Bar Review*, no. 20, p. 11, 2019.
- [51] S. Bhattacharyya, D. Cofer, D. Musliner, J. Mueller and E. Engstrom, "Certification Considerations for Adaptive Systems," National Aeronautics and Space Administration, 2015.

- [52] Y. Xie, S.-j. Feng, Y.-l. Xiong, L.-l. Zhang and G.-l. Ye, "Coupled hydraulic-mechanical-air simulation of unsaturated railway embankment under rainfall and dynamic train load," *Transportation Geotechnics*, 2020.
- [53] S. Meijer and X. Perpinya, "aming simulations for railways: Lessons learned from modeling six games for the Dutch infrastructure management," *Infrastructure design, signaling and security in railway*, pp. 275-294, 2012.
- [54] Z. Han, Y. Zhang, S. Liu and S. Gao, "Modeling and Simulation for Traction Power Supply System of High-Speed Railway," in *2011 Asia-Pacific Power and Energy Engineering Conference*, Wuhan, China, 2011.
- [55] S. Pradhan and A. K. Samantaray, "Integrated Modeling and Simulation of Vehicle and Human Multi-body Dynamics for Comfort Assessment in Railway Vehicles," *Journal of Mechanical Science and Technology*, vol. 32, no. 1, pp. 109-119, 2018.
- [56] P. Dalapati, A. Padhy, B. Mishra, A. Dutta and S. Bhattacharya, "Real-time collision handling in railway transport network: an agent-based modeling and simulation approach," *Transportation Letters*, vol. 11, pp. 458-468, 2018.
- [57] L. Chen, L. Jiang, Z. Zeng and W. Long, "Numerical modeling and simulation on seismic performance of high-speed railway bridge system," *Noise & Vibration Worldwide*, vol. 42, no. 10, pp. 15-21, 2011.
- [58] The International Association of Public Transport (UITP), "World Report on Metro Automation," 2018.
- [59] J. Athavale, A. Baldovin and M. Paulitsch, "Trends and Functional Safety Certification Strategies for Advanced Railway Automation Systems," in *2020 IEEE International Reliability Physics Symposium (IRPS)*, Dallas, TX, USA, 2020.
- [60] The British Standards Instituion, "Railway applications - The specification and demonstration of reliability, availability, maintainability and safety (RAMS)," BSI, 2017.
- [61] CENELEC, "CENELEC - EN 50128: Railway applications - Communication, signalling and processing systems - Software for railway control and protection systems," 2020.
- [62] HabibHadj-Mabrouk, "Application of Case-Based Reasoning to the safety assessment of critical software used in rail transport," *Safety Science*, vol. 131, 2020.
- [63] CENELEC, "EN 50129: Railway applications - Communication, signalling and processing systems – Safety related electronic systems for signalling," 2018.

- [64] S. Wang, Q. Shang, Q. Fang and F. Fang, "Research on Automated Testing Method of Railway Signaling System," in *IEEE 5th International Conference on Intelligent Transportation Engineering (ICITE)*, Beijing, China, 2020.
- [65] T. Wen, X. Lyu, D. Kirkwood, L. Chen, C. Constantinou and C. Roberts, "Co-simulation Testing of Data Communication System Supporting CBTC," in *IEEE 18th International Conference on Intelligent Transportation Systems*, Gran Canaria, Spain, 2015.
- [66] I. Irrera, A. Zentai, J. C. Cunha and H. Madeira, "Validating a Safety Critical Railway Application Using Fault Injection," in *Certifications of Critical Systems - The CECRIS Experience*, Gistrup, Denmark, River Publishers, 2017, pp. 227 - 246.
- [67] A. Nazemian and P. Ghadimi, "Shape optimisation of trimaran ship hull using CFD-based simulation and adjoint solve," *Ships and Offshore Structures*, pp. 1-15, 2020.
- [68] L.-C. Stan, "Optimization by CFD of the marine propulsion system," in *Advanced Topics in Optoelectronics, Microelectronics, and Nanotechnologies IX*, 2018.
- [69] N. Sakamotoa, K. Kume, Y. Kawanami, H. Kamiirisa, KenjiMokuo and M. Tamashima, "Evaluation of hydrodynamic performance of pre-swirl and post-swirl ESDs for merchant ships by numerical towing tank procedure," *Ocean Engineering*, vol. 178, no. 15, pp. 104-133, 2019.
- [70] T. Takami, S. Matsui, M. Oka and K. Iijima, "A numerical simulation method for predicting global and local hydroelastic response of a ship based on CFD and FEA coupling," *Marine Structures*, vol. 59, pp. 368-386, 2018.
- [71] D. Guillaume, D. Samuel, B. Franck, M. Pol and L. L. Frédérique, "Composite propeller in marine industry: first steps toward a technological breakthrough," in *Oceans 2019*, Marseille, France, 2019.
- [72] International Maritime Organization, "IMO takes first steps to address autonomous ships".
- [73] T.-e. Kim, A. Sharma, A. H. Gausdal and C.-j. Chae, "Impact of automation technology on gender parity in maritime industry," *WMU Journal of Maritime Affairs*, vol. 18, pp. 579-593, 2019.
- [74] DNV GL, "Remote-Controlled and Autonomous Ships," 2018.
- [75] American Bureau of Shipping, "ABS Guidance Notes on Safehull Finite Element Analysis of Hull Structures," Houston, TX, USA, 2014.
- [76] DNV GL, "Standard DNVGL-ST-0033: Maritime Simulator Systems," 2017.
- [77] American Bureau of Shipping, "ABS Guidance Notes on Gas Dispersion Studies of Gas Fueled Vessels," American Bureau of Shipping, 2019.

- [78] DNV, "Type Approval Programme No. 844.80 - MED A.1/4.16 Heading Control System (HCS)," 2012.
- [79] DNV GL, "Class Guideline DNV GL-CG-0557 Data Driven Verificaton," 2020.
- [80] H. Ringbom, "Regulating autonomous ships—concepts, challenges and precedents," *Ocean Development & International Law*, vol. 50, pp. 141-169, 2019.
- [81] A. Babić, G. Vasiljević and N. Mišković, "Vehicle-in-the-Loop Framework for Testing Long-Term Autonomy in a Heterogeneous Marine Robot Swarm," *IEEE Robotics and Automation Letters*, vol. 5, no. 3, pp. 4439-4446, 2020.
- [82] A. M. Bassam, A. B. Phillips, S. R. Turnock and P. A. Wilson, "Experimental testing and simulations of an autonomous, self-propulsion and self-measuring tanker ship model," *Elsevier Ocean Engineering*, vol. 186, p. 106065, 2019.
- [83] T. A. Pedersen, J. A. Glomsrud, E.-L. Ruud, A. Simonsen, J. Sandrib and B.-O. H. Eriksen, "Towards simulation-based verification of autonomous navigation systems," *Elsevier Safety Science*, vol. 129, p. 104799, 2020.
- [84] US DOT, "Preparing for the Future of Transportation: Automated Vehicles 3.0 (AV 3.0)," US Department of Transportation, 2018.
- [85] S. Schnelle, K. Salaani, S. J. Rao, F. S. Barickman and D. Elsasser, "Review of simulation frameworks and standards related to driving scenarios (Report No. DOT HS 812 815)," US DOT, 2019.
- [86] UN Economic Commission for Europe, "New Assessment/Test Method for Automated Driving (NATM) - Master Document," World Forum for Harmonization of Vehicle Regulations, 2021.
- [87] T. D. Dias, J. P. Ribeiro and L. T. Moura, "Cloud Based HD Maps in the 5G-MOBIX Project," in *2020 Virtual ITS European Congress*, 2020.
- [88] S. Ulbrich, T. Menzel, A. Reschka, F. Schuldt and M. Maurer, "Defining and Substantiating the Terms Scene, Situation, and Scenario for Automated Driving," in *IEEE 18th International Conference on Intelligent Transportation Systems*, Gran Canaria, Spain, 2015.
- [89] J. L. Every, F. Barickman, J. Martin, S. Rao, S. Schnelle and B. Weng, "A novel method to evaluate the safety of highly automated vehicles," in *25th International Technical Conference on the Enhanced Safety of Vehicles (ESV) National Highway Traffic Safety Administration, Detroit, Michigan*, Detroit, Michigan, 2017.

- [90] F. M. Favar`o, N. Nader, S. O. Eurich, M. Tripp and N. Varadaraju, "Examining accident reports involving autonomous vehicles in California," *PLoS one*, vol. 12, no. 9, p. e0184952, 2017.
- [91] M. Blanco, J. Atwood, S. Russell, T. Trimble, J. McClafferty and M. Perez, "Automated Vehicle Crash Rate Comparison Using Naturalistic Data," Virginia Tech Transport Institute, 2016.
- [92] European New Car Assessment Programme (Euro NCAP), "Test Protocol - AEB Systems," Euro NCAP, 2019.
- [93] F. Rosique, P. J. Navarro, C. Fernández and A. Padilla, "A Systematic Review of Perception System and Simulators for Autonomous Vehicles Research," *Sensors*, vol. 19, no. 3, 2019.
- [94] L. Fraade-Blanar, M. S. Blumenthal, J. M. Anderson and N. Kalra, Measuring automated vehicle safety: Forging a framework, RAND Corporation, 2018.
- [95] M. S. Blumenthal, L. Fraade-Blanar, R. Best and J. L. Irwin, Safe Enough: Approaches to Assessing Acceptable Safety for Automated Vehicles, RAND Corporation, 2020.
- [96] W. Do, O. M. Rouhani and L. Miranda-Moreno, "Simulation-based connected and automated vehicle models on highway sections: a literature review," *Journal of Advanced Transportation*, 2019.
- [97] H.-P. Schoener, "The Role of Simulation in Development and Testing of Autonomous Vehicles," in *Driving Simulation Conference – DSC 2017*, Stuttgart, Germany, 2017.
- [98] C. Goodin, D. Carruth, M. Doude and C. Hudson, "Predicting the Influence of Rain on LIDAR in ADAS," *Electronics*, vol. 8, no. 1, p. 89, 2019.
- [99] A. Elmquist and D. Negrut, "Methods and models for simulating autonomous vehicle sensors," *IEEE Transactions on Intelligent Vehicles*, vol. 5, no. 4, pp. 684-692, 2020.
- [100] A. Diewald, C. Kurz, P. V. Kannan, M. Giessler, M. Pauli, B. Gottel, T. Kayser, F. Gauterin and T. Zwick, "Radar Target Simulation for Vehicle-in-the-Loop Testing," *Vehicles*, vol. 3, no. 2, pp. 257-271, 2021.
- [101] Z. Nie and H. Farzaneh, "Adaptive Cruise Control for Eco-Driving Based on Model Predictive Control Algorithm," *Applied Sciences*, vol. 10, no. 15, p. 5271, 2020.
- [102] M. Borg, R. B. Abdessalem, S. Nejati, F.-X. Jegeden and D. Shin, "Digital Twins Are Not Monozygotic – Cross-Replicating ADAS Testing in Two Industry-Grade Automotive Simulators," in *2021 14th IEEE Conference on Software Testing, Verification and Validation (ICST)*, Porto de Galinhas, Brazil, 2021.

- [103] F. Reway, W. Huber and E. P. Ribeiro, "Test Methodology for Vision-Based ADAS Algorithms with an Automotive Camera-in-the-Loop," in *2018 IEEE International Conference on Vehicular Electronics and Safety (ICVES)*, Madrid, Spain, 2018.
- [104] S. Khastgir, "Enabling safe CAV deployment," BSI Group, 2020.
- [105] SAE International, "SAE J3049: Model Architecture and Interfaces Recommended Practice for Ground Vehicle System and Subsystem Dynamical Simulation," SAE, 2015.
- [106] Waymo LLC, "Waymo's Safety Methodologies and Safety Readiness Determinations," <https://storage.googleapis.com/sdc-prod/v1/safety-report/Waymo-Safety-Methodologies-and-Readiness-Determinations.pdf>, 2020.
- [107] Waymo LLC, "Waymo Safety Report," <https://storage.googleapis.com/sdc-prod/v1/safety-report/2020-09-waymo-safety-report.pdf>, 2020.
- [108] Z. Yang, Y. Chai, D. Anguelov, Y. Zhou, P. Sun, D. Erhan, S. Rafferty and H. Kretzschmar, "SurfelGAN: Synthesizing Realistic Sensor Data for Autonomous Driving," in *Conference on Computer Vision and Pattern Recognition*, 2020.
- [109] Mercedes-Benz & Bosch, "Reinventing Safety: A Joint Approach to Automated Driving Systems," [Online]. Available: <https://www.daimler.com/documents/innovation/other/vssa-mercedes-benz-and-bosch.pdf>.
- [110] Uber Advanced Technologies Group, "A Principled Approach to Safety," Uber ATG, 2018.
- [111] T. D. S. G. X. Y. A. L. S.-V. S. A. S. Daniel J. Fremont, "Scenic: A Language for Scenario Specification and Scene Generation," in *Proceedings of the 40th ACM SIGPLAN Conference on Programming Language Design and Implementation*, 2019.
- [112] R. Queiroz, T. Berger and K. Czarnecki, "GeoScenario: An Open DSL for Autonomous Driving Scenario Representation," in *IEEE Intelligent Vehicles Symposium (IV)*, 2019.
- [113] P. Kaur, S. Taghavi, Z. Tian and W. Shi, "A Survey on Simulators for Testing Self-Driving Cars," *arXiv preprint arXiv:2101.05337*, 2021.
- [114] MathWorks, "Automated Driving Toolbox," [Online]. Available: <https://www.mathworks.com/help/driving/>.
- [115] CARLA Team, "CARLA Documentation," [Online]. Available: <https://carla.readthedocs.io/en/0.9.12/>.

- [116] R. Xu, Y. Guo, X. Han, X. Xia, H. Xiang and J. Ma, "OpenCDA: An Open Cooperative Driving Automation Framework Integrated with Co-Simulation," in *2021 IEEE Intelligent Transportation Systems Conference (ITSC)*, 2021.
- [117] S. Bhattacharyya, D. Cofer, D. Musliner, J. Mueller and E. Engstrom, "Certification considerations for adaptive systems," in *2015 International Conference on Unmanned Aircraft Systems (ICUAS)*, Denver, CO, USA, 2015.
- [118] S. L. Brunton, J. N. Kutz, K. Manohar, A. Y. Aravkin, K. Morgansen, J. Klemisch, N. Goebel, J. Buttrick, J. Poskin, A. Blom-Schieber, T. Hogan and D. McDonald, "Data-Driven Aerospace Engineering: Reframing the Industry with Machine Learning," *AIAA Journal*, pp. 1-26, 2021.
- [119] A. K. Raz, E. P. Blasch, C. Guariniello and Z. T. Mian, "An Overview of Systems Engineering Challenges for Designing AI-Enabled Aerospace Systems," in *AIAA Scitech 2021 Forum*, 2021.
- [120] M. Hochstrasser, S. Myschik and F. Holzapfel, "Application of a Process-Oriented Build Tool for Flight Controller Development Along a DO-178C/DO-331 Process," in *Model-Driven Engineering and Software Development*, Springer, Cham, 2018.

Appendix A: SAE J3016 Levels of Driving Automation

SAE J3016 is a graphic document that defines six levels of driving automation to represent the wide spectrum of the scope and capabilities of these technologies. Starting from Level 0 (colloquially L0 – no automation) to Level 5 (colloquially L5 – full vehicle autonomy), these levels were created to clearly explain the features and the capabilities of these systems. The graphic below was recreated from the SAE J3016 document to pictorially explain these different levels of driving autonomy.

SAE J3016™ LEVELS OF DRIVING AUTOMATION

	LEVEL 0	LEVEL 1	LEVEL 2	LEVEL 3	LEVEL 4	LEVEL 5
What does the human drivers seat have to do?	You are driving whenever these driver support features are engaged – even if your feet are off the pedals and you are not steering			You are not driving when these automated driving features are engaged – even if you are seated in “the drivers seat”		
	You must constantly supervise these features; you must steer, brake or accelerate as needed to maintain safety			When the feature requests, you must drive	These automated driving features will not require you to take over driving	
	These are driver supported features			These are automated driving features		
What do these features do?	These features are limited to providing warnings and momentary assistance	These features provide steering / OR brake / acceleration support to the driver	These features provide steering AND brake / acceleration support to the driver	These features can drive the vehicle under limited conditions and will not operate unless all required conditions are met		This feature can drive the vehicle under all conditions
Example features	<ul style="list-style-type: none">Automatic emergency brakingBlind spot warningLane departure warning	<ul style="list-style-type: none">Lane centering OR Adaptive cruise control	<ul style="list-style-type: none">Lane centering AND Adaptive cruise control	<ul style="list-style-type: none">Traffic jam chauffer	<ul style="list-style-type: none">Local driverless taxiPedals/steering wheel may or may not be installed	<ul style="list-style-type: none">Same as level 4, but feature can drive everywhere in all conditions

