Risk Quantification for Automated Driving Systems in Real-World Driving Scenarios

De Gelder, Erwin; Elrofai, Hala; Saberi, Arash Khabbaz; Paardekooper, Jan Pieter; Den Camp, Olaf Op; De Schutter, Bart

DOI
10.1109/ACCESS.2021.3136585

Publication date
2021

Document Version
Final published version

Published in
IEEE Access

Citation (APA)
De Gelder, E., Elrofai, H., Saberi, A. K., Paardekooper, J. P., Den Camp, O. O., & De Schutter, B. (2021). Risk Quantification for Automated Driving Systems in Real-World Driving Scenarios. *IEEE Access*, *9*, 168953-168970. https://doi.org/10.1109/ACCESS.2021.3136585

Received December 4, 2021, accepted December 16, 2021, date of publication December 20, 2021,
date of current version December 30, 2021.

Digital Object Identifier 10.1109/ACCESS.2021.3136585

Risk Quantification for Automated Driving Systems in Real-World Driving Scenarios

ERWIN DE GELDER[1,2], HALA ELROFAI[1], ARASH KHABBAZ SABERI[3], JAN-PIETER PAARDEKOOPER[1,4], OLAF OP DEN CAMP[1], AND BART DE SCHUTTER[2], (Fellow, IEEE)

[1]Department of Integrated Vehicle Safety, TNO, 5708 JZ Helmond, The Netherlands
[2]Delft Center for Systems and Control, Delft University of Technology, 2628 CD Delft, The Netherlands
[3]Automated Driving Product Unit, TomTom, 5656 AE Eindhoven, The Netherlands
[4]Donders Institute for Brain, Cognition and Behaviour, Radboud University, 6525 XZ Nijmegen, The Netherlands

Corresponding author: Erwin de Gelder (erwin.degelder@tno.nl)

ABSTRACT The development of safety validation methods is essential for the safe deployment and operation of Automated Driving Systems (ADSs). One of the goals of safety validation is to prospectively evaluate the risk of an ADS dealing with real-world traffic. ISO 26262 and ISO/DIS 21448, the leading standards in automotive safety, provide an approach to estimate the risk where the former focuses on risks due to potential malfunctioning of components and the latter focuses on risks due to possible functional insufficiencies. The main shortcomings of the approach provided in ISO 26262 are that it depends on subjective judgments of safety experts and that only a qualitative risk estimation is performed. ISO/DIS 21448 addresses these shortcomings partially by providing statistical methods to guide the safety validation, but no complete method is provided to quantify the risk. The first objective of this article is to propose a method to estimate the risk of an ADS in a more quantitative and objective manner. A data-driven approach is used to rely less on subjective judgments of safety experts. The output of the method is the expected number of injuries in a potential collision. Thus, the method is quantitative, the result is easily interpretable, and the result can be compared with road crash statistics. The second objective is to provide a method that supports the risk assessment as stipulated by the ISO 26262 and ISO/DIS 21448 standards by decomposing the quantified risk into the 3 aspects of risk as mentioned in these standards: exposure, severity, and controllability. The proposed methods are illustrated by means of a case study in which the risk is quantified for a longitudinal controller in 3 different types of scenarios. The code of the case study is publicly available.

INDEX TERMS Intelligent vehicles, intelligent transportation systems, ISO 21448, ISO 26262, performance evaluation, risk analysis.

I. INTRODUCTION

It is expected that ADSs will make traffic safer by eliminating human errors, enable more comfortable rides, and reduce traffic congestion [1]. Lower levels of automation systems, such as Adaptive Cruise Control (ACC) [2] and lane keeping assist systems [3], are already widely deployed in modern cars and trucks. Since the development of ADSs has made significant progress, it is expected that ADSs addressing higher levels of automation and covering the full dynamic driving task, i.e., SAE level 3 or higher [4], are soon to be introduced on public roads [5]–[7]. Before deploying an ADS on public roads, it is of paramount importance to ensure that there is no negative impact on the traffic safety. It has even been suggested [8] that vehicles controlled by an ADS should be at least 4 to 5 times as safe as human-driven vehicles in order to be accepted by the general public.

Safety validation of an ADS is essential to guarantee that the ADS is safe enough to be allowed on public roads. Retrospective safety validation alone, i.e., through test drives with prototypes, requires millions of kilometers of driving [9], which makes this practically infeasible. Therefore, prospective safety validation, i.e., before performing (test) drives in public traffic, is required. Scenario-based safety validation is an approach for prospective safety validation that is broadly supported by the automotive field [10]–[22].

The associate editor coordinating the review of this manuscript and approving it for publication was Zhenzhou Tang.

As a consequence of the broad support for scenario-based safety validation, significant research progress has been made. Recent research focuses, among others, on scenario terminology [23], [24], scenario-based requirement verification [25], virtual simulation of scenarios [26]–[30], generation of scenarios [31]–[34], and scenario databases [11]–[13]. All these components are vital parts for estimating the risk of the deployment of an ADS, where risk is the combination of the probability of occurrence of harm and the extent of that harm. This risk estimation itself, however, has received less attention in the literature despite its essential role in the well-known ISO 26262 [35] and ISO/DIS 21448 [36] standards, which capture the state of the art in automotive safety.

The contribution of this work is twofold. First, we propose a novel data-driven method for assessing and quantifying the risk of an ADS considering real-world driving scenarios. To provide a more objective method for risk quantification compared to existing approaches, the proposed method uses real-world data and relies less on the judgments of experts. Second, we show how the risk quantification can be decomposed into the terms exposure, severity, and controllability, such that the method can be applied to assess risks based on the ISO 26262 and ISO/DIS 21448 standards. Therefore, our method supports the risk assessment activities of these standards. In the remainder of this section, we elaborate more on the need for our work and we present the research questions that are addressed by the current article.

A. RISK QUANTIFICATION OF AUTOMATED DRIVING SYSTEMS

In [37]–[41], risk quantification is proposed for real-time use, e.g., to support the path planning of a self-driving vehicle, so this is not intended to be used for a prospective risk assessment. In [42], a method for quantifying the risk of scenarios is proposed based on the probability of occurrence of so-called environmental conditions and the probability that an error propagates in the fault tree given a specified environment condition. However, they [42] do not provide methods to estimate or specify these probabilities. Furthermore, the role of a back-up operator, e.g., a human driver that supervises the ADS, is not considered, whereas the influence of the back-up operator is an important aspect that is considered in the ISO 26262 and ISO/DIS 21448 standards. Another method to quantify the risk of a driving scenario is proposed in [43], but this method does not consider the likelihood of encountering the scenario and the role of the back-up operator is not explicitly considered. A quantitative assurance framework is proposed in [44], [45], but this framework assumes that the frequency of accidents is known, whereas this is unknown in a prospective assessment. Furthermore, similar to [42], [44] and [45] do not consider the role of a back-up operator.

To address the aforementioned shortcomings, the current work aims to answer the following question:

Research question 1: How to quantify the risk of an ADS in real-world driving scenarios?

To answer Research question 1, this article proposes a novel method for quantifying the risk of an ADS. The first step of the presented method is to identify the scenarios that the ADS encounters in real life. Next, the exposure, i.e., the likelihood of encountering these scenarios, is estimated. Using simulations, the probability that a scenario leads to a harmful event is calculated. Combining this probability with the exposure and the probability that a harmful event leads to an injury results in the estimated risk.

B. RISK QUANTIFICATION IN RELATION WITH ISO 26262 AND ISO/DIS 21448

ISO 26262 is the state-of-the-art standard in automotive functional safety that offers a framework for measuring risk in a qualitative manner. This standard concerns hazards that are the result of malfunctioning behavior of components. In the Hazard Analysis and Risk Assessment (HARA), an Automotive Safety Integrity Level (ASIL) is determined for each hazardous event based on the classification of three aspects: exposure, severity, and controllability (the definitions of these terms will be provided in Section II). The classification has two limitations. First, the classification relies on the judgments of experts of the three risk aspects, which renders the classification subjective [46], [47]. Second, the classification is qualitative and offers only 5 different levels of risk.

As an addition to the ISO 26262 standard, the ISO/DIS 21448 standard, also known as ''Safety Of The Intended Functionality (SOTIF)'', concerns hazardous behavior due to a functional insufficiency of the intended functionality at the vehicle level as opposed to malfunctioning behavior of components. A functional insufficiency may refer to a failure due to technological limitations of a sensor or an actuator. The hazard in the context of SOTIF is initiated by a so-called triggering condition. For the risk assessment, the ISO/DIS 21448 standard does not determine an ASIL, but the exposure, severity, and controllability can be used to determine the required validation effort. In [36, Annex C.6.4], the exposure, severity, and controllability are quantitatively expressed as probabilities, but no method is provided for determining these probabilities other than that the probabilities can be checked for consistency with the functional safety HARA of the ISO 26262 standard. Therefore, the risk estimation of the ISO/DIS 21448 standard inherits the limitations of the ISO 26262 standard, which means that there is no explicit risk estimation method that considers hazardous behavior caused by a triggering condition.

To address the lack of a quantitative and objective approach to assess risks related to the ISO 26262 and ISO/DIS 21448 standards, the current work also aims to answer the following question:

Research question 2: How to quantify the risk of an ADS using the 3 aspects of risk as defined by the ISO 26262 standard: exposure, severity, and controllability?

To answer Research question 2, we decompose the quantified risk into the terms exposure, severity, and controllability. Note that it is not our objective to provide a method that

replaces parts of the ISO 26262 and ISO/DIS 21448 standards, but rather to provide a method that supports the risk assessment activities of these standards.

C. ORGANIZATION OF THIS ARTICLE

This article is organized as follows. We first elaborate on the risk assessment in the ISO 26262 and ISO/DIS 21448 standards in Section II. Section III provides a method for risk quantification to answer Research question 1. In Section IV, Research question 2 is answered by describing how the proposed risk quantification method relates to the risk assessment according to the ISO 26262 and ISO/DIS 21448 standards. To illustrate the proposed method, a case study[1] involving the risk quantification of an ACC is presented in Section V and the results are reported in Section VI. This article ends with a discussion in Section VII and conclusions in Section VIII.

II. ISO 26262 AND ISO/DIS 21448

In Section II-A, the risk assessment approach provided in the ISO 26262 standard [35] and its shortcomings are described. Section II-B elaborates on the risk assessment approach described in the ISO/DIS 21448 standard [36].

A. ISO 26262

The ISO 26262 standard [35] captures the state of the art in automotive functional safety. It defines the safety lifecycle and the related safety activities such as the HARA. Other methodologies, such as Systems-Theoretic Processes Analysis (STPA) [48] and Failure Mode and Effect Analysis (FMEA) [49], give guidelines on safety engineering based on systems theory. Unlike the ISO 26262 standard, STPA and FMEA do not offer a framework for measuring risk.

The ISO 26262 standard gives guidelines to assess risk based on hazardous events. A hazardous event is the combination of an operational situation (or a scenario) with a potential source of harm caused by malfunctioning behavior of system components. The standard requires analyzing the risk of each hazardous event based on three aspects: exposure, severity, and controllability; see Table 1 for the definitions according to the ISO 26262 standard. In this framework, each aspect is classified in 4 or 5 levels:

- Exposure is classified as 0 (''incredible''), 1, 2, 3, or 4 (''high probability'');
- Severity is classified as 0 (''no injuries''), 1, 2, or 3 (''life-threatening injuries, fatal injuries''); and
- Controllability is classified as 0 (''controllable in general''), 1, 2, or 3 (''difficult to control or uncontrollable'').

The combination of these aspects contributes to constructing the ASIL ranking A, B, C, D. With an ASIL D, the most stringent requirements on system design, verification, and testing apply while an ASIL A requires the least additional safety measures. If the sum of the aspects is 10, an ASIL D is assigned, representing the most critical level. ASIL B or C is assigned when the sum equals 8 or 9, respectively. If the sum is 7 and no aspect has scored 0, then ASIL A is assigned. In all other cases, there are no requirements to comply with the ISO 26262 standard and the classification ''Quality Management (QM)'' applies because it is assumed that the QM system of the manufacturer suffices for reducing the risk.

TABLE 1. Definitions of exposure, severity, and controllability according to the ISO 26262 standard [35].

Term	Definitions
Exposure	State of being in an operational situation that can be hazardous if coincident with the failure mode under analysis
Severity	Estimate of the extent of harm to one or more individuals that can occur in a potentially hazardous event
Controllability	Ability to avoid a specified harm or damage through the timely reactions of the persons involved, possibly with support from external measures

A shortcoming of the ASIL ranking is that the results are subjective. Teuchert [46] mentions that the classification of the ASIL depends very much on the engineers that perform the classification. Also Khastgir *et al.* [47] mention the subjectivity of the ranking: ''The two distinct shortcomings of the current ISO 26262-2011 standard are guided by the subjective nature of the experts' mental models leading to unreliable ratings and the ability to identify a hazard (including the black swan events).'' Through an experiment, Khastgir *et al.* [47] demonstrated the low (intra-rater) reliability due to the subjectivity of the Hazard Analysis and Risk Assessment (HARA) process. Another disadvantage is the qualitative nature of the ASIL ranking. Because only 5 different levels of risk are considered (ASIL A to D and QM), only large changes in the overall risk are captured.

Since our proposed method for estimating the risk, presented in Section III, is based on data, it is more objective by nature. In addition, the results are quantitative. Therefore, our proposed method can be used to determine an ASIL-like indicator in a *quantitative* and *objective* manner.

B. ISO/DIS 21448

Whereas the ISO 26262 standard focuses on possible hazards caused by the malfunctioning behavior of system components, the ISO/DIS 21448 standard [36] addresses hazardous situations caused by the intended functionality, despite the systems being free from the faults addressed in the ISO 26262 standard. The absence of unreasonable risk due to these hazardous situations is defined as the Safety Of The Intended Functionality (SOTIF). Although no ASIL is determined for SOTIF-related hazards, the aspects exposure, severity, and controllability are still used to adjust the required evidence of the safe operation, including the number validation scenarios used for testing. For determining these aspects, the ISO/DIS 21448 standard refers to the ISO 26262 standard,

[1]The code is publicly available at https://github.com/ErwindeGelder/ScenarioRiskQuantification.

which — as reasoned earlier — provides a subjective and qualitative approach.

To address the qualitative nature of the ISO 26262 standard, Annex C of the ISO/DIS 21448 standard provides statistical methods to guide the SOTIF verification and validation. A method to quantify the validation targets is provided in [36, Annex C.2], while [36, Annex C.3] proposes to use on-road data or simulations to validate whether systems meet these targets. The use of importance sampling to lessen the amount of simulation testing is discussed in [36, Annex C.5], but no quantification of, e.g., the risk, is provided. In [36, Annex C.6], a statistical approach is presented for arguing that a safety criterion is met while considering the performance of the constituent components. Also in [36, Annex C.6], the risk aspects exposure, severity, and controllability are quantitatively expressed as probabilities. No method is provided, however, to estimate these probabilities. Thus, where Annex C of the ISO/DIS 21448 standard provides a step towards making the risk evaluation more quantitative, the current work aims to be more explicit in quantifying the risk. More specifically, in Section IV, we show how our proposed method for risk quantification can be used to quantify the aspects exposure, severity, and controllability. Furthermore, Section IV describes how to combine these aspects to quantify the risk.

III. METHOD FOR RISK QUANTIFICATION

To answer Research question 1, this section proposes a novel method for quantifying the risk of an ADS in real-world driving scenarios. To structure the risk quantification, the method consists of six steps:

1) Identify the scenarios that are part of the so-called Operational Design Domain (ODD) of the ADS.
2) Determine the exposure of these scenarios, i.e., the expected number of occurrences per hour of driving.
3) Simulate the response of the ADS in these scenarios.
4) Calculate the expected number of harmful events.
5) Calculate the expected injury rate.
6) Calculate the risk by combining the exposure and the injury rate.

These steps are schematically shown in Figure 1. Figure 1 also shows how the risk aspects exposure, severity, and controllability relate to these steps. This relation will be further explained in Section II.

Table 2 presents the definitions of the terms that are used in our proposed method. In this work, the probability of u is denoted by $\mathbb{P}(u)$, while $\mathbb{E}[u]$ denotes the expectation of u. In the following subsections, the 6 steps for quantifying the risk are described.

TABLE 2. The terms and definitions.

Term	Definition
Risk	Combination of the probability of occurrence of harm and the severity of that harm [35]
ODD	Operating conditions under which a given driving automation system or feature thereof is specifically 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 [4]
Scenario	Quantitative description of the relevant characteristics and activities and/or goals of the ego vehicle(s), the static environment, the dynamic environment, and all events that are relevant to the ego vehicle(s) within the time interval between the first and the last relevant event [24]
Scenario category	Qualitative description of the relevant characteristics and activities and/or goals of the ego vehicle(s), the static environment, and the dynamic environment [24]
Triggering condition	Specific conditions of a scenario that [may] serve as an initiator for a subsequent system reaction leading to hazardous behavior [36]

A. IDENTIFICATION OF SCENARIOS

An ADS is designed to operate within its ODD, which is defined by the ADS developer and typically consists of a geofence and some known operational conditions, e.g., see [50]–[52]. The ODD is used to confine the risk analysis [53]. To quantify the risk of an ADS in driving scenarios, the ODD of the ADS must be known. Once deployed, the ADS needs to deal with many scenarios and the ODD in which the ADS is operating determines the variety of these scenarios. It is our goal to provide a method for determining the risk for the ADS in these scenarios.

Considering the wide variety of scenarios, we propose to distinguish between quantitative scenarios and qualitative scenarios, where scenario categories refer to the latter, see Table 2. It is assumed that all possible scenarios within a given ODD can be categorized into one or more scenario categories. This assumption does not limit the applicability of the methodology proposed in this work, though it might require many scenario categories to describe all these scenarios. In the remainder of this section, we propose a method to calculate the risk for an ADS in all scenarios that are categorized by the same scenario category, i.e., for all $S \in \mathcal{C}$ [24], where S and $\mathcal{C}$ denote a scenario and a scenario category, respectively. For example, the scenario category "cut-in" comprises all possible cut-in scenarios in the ODD of the ADS. See [54] for more examples of scenario categories.

Remark 1: As part of the scenarios, some factors may cause hazardous behavior. These factors are called *triggering conditions*, because they may trigger some specific behavior [36]. Typically, triggering conditions may happen rarely, so the impact on safety may not be known. In our case, one or more triggering conditions could be part of a scenario category. Examples of triggering conditions are heavy rain, low road friction, poor lighting, or dirty sensor(s). For more examples, see [36, Annex B.2]. ◊

B. PROBABILITY OF EXPOSURE

To determine the exposure, we estimate the expected number of encounters of a scenario $S \in \mathcal{C}$. Let $n_{\mathcal{C}}$ denote the number of encounters per unit of time of a scenario belonging to scenario category $\mathcal{C}$. We express the exposure as $\mathbb{E}[n_{\mathcal{C}}]$, i.e.,

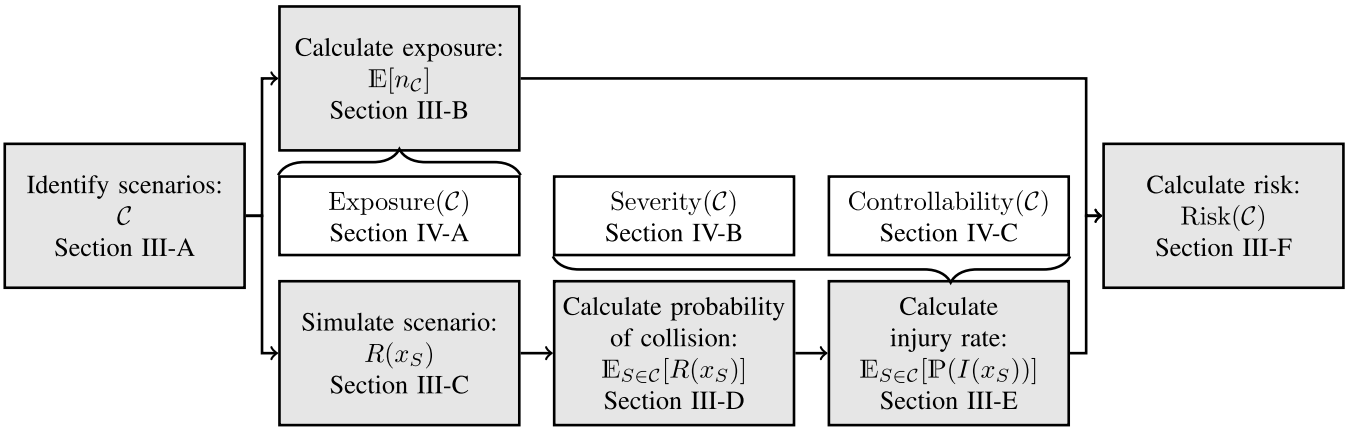


FIGURE 1. Schematic overview of the risk quantification method presented in this article. The gray blocks represent the six steps of the risk quantification method explained in Section III. Each of these steps is further explained in Sections III-A to III-F, respectively. The white blocks show the relation between our proposed method and the three aspects of risk as proposed in the ISO 26262 standard. Each of these aspects is further explained in Sections IV-A to IV-C, respectively.

the expected number of encounters per unit of time of a scenario belonging to scenario category $\mathcal{C}$:

$$\mathbb{E}[n_{\mathcal{C}}] = \sum_{n=1}^{\infty} n\mathbb{P}(n_{\mathcal{C}} = n), \tag{1}$$

where $\mathbb{P}(n_{\mathcal{C}} = n)$ denotes the probability of encountering n scenarios belonging to scenario category $\mathcal{C}$.

We propose to determine $\mathbb{P}(n_{\mathcal{C}} = n)$, $n = 0, 1, 2, \ldots$, based on data, because the data provide a quantitative way to estimate $\mathbb{P}(n_{\mathcal{C}} = n)$, $n = 0, 1, 2, \ldots$. Furthermore, assuming that the data are collected with the same conditions as specified by the ODD of the ADS, the data provides an objective way to estimate $\mathbb{P}(n_{\mathcal{C}} = n)$. The probability can be estimated by counting the number of occurrences of the scenarios in the data. The method to find the scenarios belonging to $\mathcal{C}$ is explained in [55]: First, tags are used to describe activities, such as lane changing and braking, and statuses, such as "leading vehicle" and "driving slower". Note that a tag is typically associated with an object and has a start time and an end time. Second, by searching for a particular combination of these tags that describes the scenario category $\mathcal{C}$, the start and end time of the scenarios are found.

Remark 2: It is not uncommon to assume that

- the occurrence of a scenario $S_1 \in \mathcal{C}$ does not affect the probability that a second scenario $S_2 \in \mathcal{C}$ occurs,
- the expected rate at which a scenario belonging to $\mathcal{C}$ occurs is constant, and
- two scenarios belonging to $\mathcal{C}$ cannot occur at exactly the same time.

In that case, the probability $\mathbb{P}(n_{\mathcal{C}} = n)$ simplifies to a Poisson distribution:

$$\mathbb{P}(n_{\mathcal{C}} = n) = \frac{\lambda^n}{n!}\exp\{-\lambda\}. \tag{2}$$

Here, $\lambda > 0$ is a parameter that determines the rate at which a scenario belonging to $\mathcal{C}$ occurs. Assuming (2), (1) simplifies to $\mathbb{E}[n_{\mathcal{C}}] = \lambda$. ◊

C. SIMULATION OF SCENARIOS

The next step of the risk quantification is to simulate how the ADS behaves in the scenarios belonging to scenario category $\mathcal{C}$. Let $x_S \in \mathbb{R}^d$ denote the d-dimensional parameter vector that describes scenario S. The (stochastic) outcome of a simulation of scenario S is denoted by $R(x_S)$, where $R(x_S) = 1$ means that the simulation of the scenario with parameters x_S ends with an unsuccessful outcome and otherwise $R(x_S) = 0$. An unsuccessful outcome may be a collision or a situation where the ego vehicle ends off the road. In addition to $R(x_S)$, the output of a simulation run provides information to rank multiple scenarios from "most critical" to "least critical" (see Section III-D2) and information to estimate the extent of harm in case of a collision (see Section III-E).

To enable the simulation of the scenarios, a simulation framework is set up. Figure 2 shows the scheme of the simulation framework, which is represented by the following five blocks:

- World: the relevant information about the environment of the system under test. This includes other vehicles.
- Sensors: mapping of the global information to sensor data that can be used by the ADS.
- ADS: the logic and control laws to perform an automated function. The ADS uses the information of the sensors to determine the input to the vehicle. The ADS provides input to the actuators of the vehicle and information to the operator, e.g., a beep in case of a collision warning.
- Operator: the actual driver that is behind the steering wheel or a remote driver.
- Vehicle: the system consisting of actuators for translating the inputs generated by the ADS and the operator into vehicle motions and subsystems for enhancing conspicuity via, e.g., lighting, signaling, and sounding the horn.

Note that the simulation framework is easily extended to consider multiple ADSs, operators, and automated vehicles.

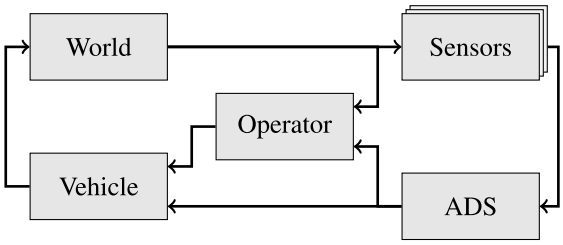


FIGURE 2. Scheme of the simulation framework.

This could be of interest for testing, e.g., (cooperative) ACC systems because one might be interested in the performance of a platoon of vehicles rather than the performance of a single vehicle.

D. PROBABILITY OF COLLISION

Instead of estimating the risk of an ADS in a specific scenario with parameters x, the goal is to calculate the risk for all scenarios from scenario category $\mathcal{C}$. The first step toward the calculation of the risk is to compute the expected outcome while averaging over all scenarios in $\mathcal{C}$: $\mathbb{E}_{S\in\mathcal{C}}[R(x_S)]$. Here, the subscript $S \in \mathcal{C}$ indicates that the expectation is computed while averaging over all scenarios belonging to scenario category $\mathcal{C}$.

To provide inputs to the simulation, scenarios are identified and characterized from real-world driving data, e.g., collected in field operational tests or naturalistic driving studies. In this way, the scenarios are most likely to represent real-world driving conditions [11], [13], [56]. One approach would be to simply simulate the scenarios that are recorded from data, but this gives two problems. First, because not all possible variations of the scenarios might be found in the data, the failure modes of the ADS might not be reflected in the simulations [57]. Second, because the set of extracted scenarios is largely composed of non-safety critical scenarios, many scenarios may be required to obtain the required statistical accuracy of rare events such as collisions [57], [58]. To overcome these problems, so-called importance sampling can be applied in order to put more emphasis on scenarios that are likely to trigger safety-critical situations [57]–[60].

In this section, we propose a nonparametric importance sampling method without requiring a-priori knowledge of what might be scenarios that are likely to trigger safety-critical situations given the ADS under test. First, crude Monte Carlo sampling is used, see Section III-D1. The nonparametric importance sampling method based on the simulation results of the crude Monte Carlo sampling is proposed in Section III-D2.

1) CRUDE MONTE CARLO SAMPLING

With crude Monte Carlo sampling, parameters are sampled from a probability density function (pdf). Let $f_{\mathcal{C}}(\cdot)$ denote the pdf of the parameters of the scenarios from scenario category $\mathcal{C}$. Typically, the pdf $f_{\mathcal{C}}(\cdot)$ is unknown, so it needs to be estimated. To estimate the pdf, we use the observed scenarios that have also been used to estimate the exposure (Section III-B). Since the shape of the pdf is also unknown beforehand, assuming a predefined functional form of the pdf for which certain parameters are fitted to the data may lead to inaccurate fits unless a lot of hand-tuning is applied. As an alternative, we propose to estimate $f_{\mathcal{C}}(\cdot)$ using Kernel Density Estimation (KDE) [61], [62]. Let $x_{S_1}, \ldots, x_{S_N}$ denote the parameters of the observed scenarios $S_i \in \mathcal{C}$, $i \in \{1, \ldots, N\}$, then the density $f_{\mathcal{C}}(\cdot)$ is estimated by

$$\hat{f}_{\mathcal{C}}(x) = \frac{1}{Nh^d} \sum_{i=1}^{N} K\left(\frac{1}{h}\left(x - x_{S_i}\right)\right). \tag{3}$$

Here, $K(\cdot)$ is the so-called kernel and h is the bandwidth. The choice of the kernel function is not as important as the choice of the bandwidth [63], [64]. We use the often-used Gaussian kernel,[2] but our method does not depend on the choice of kernel. The Gaussian kernel is given by

$$K(u) = \frac{1}{(2\pi)^{d/2}} \exp\left\{-\frac{1}{2}\|u\|_2^2\right\}, \tag{4}$$

where $\|u\|_2^2 = u^{\mathsf{T}}u$ denotes the squared 2-norm of u.

The bandwidth $h > 0$ is a free parameter that controls the width of the kernel. A larger bandwidth results in a smoother pdf, but choosing h too large may result in loss of details in the pdf. Methods for estimating the bandwidth range from simple reference rules like Silverman's rule of thumb [67] to more elaborate methods (for reviews, see [63], [68], [69]). We use leave-one-out cross validation to determine the optimal bandwidth because this minimizes the Kullback-Leibler divergence between the estimated pdf, $\hat{f}_{\mathcal{C}}(\cdot)$, and the unknown pdf that underlies the data, $f_{\mathcal{C}}(\cdot)$ [63]. With leave-one-out cross validation, the bandwidth equals:

$$\arg\max_h \prod_{i=1}^{N} \left(\frac{1}{(N-1)h^d} \sum_{j=1, j\neq i}^{N} K\left(\frac{1}{h}\left(x_{S_j} - x_{S_i}\right)\right)\right). \tag{5}$$

Note that with the one-dimensional bandwidth h, the same amount of smoothing is applied in every direction. Therefore, the parameters are first scaled, such that they have the same standard deviation in each dimension. Our method can easily be extended to a multi-dimensional bandwidth [64], [70].

Sampling from $\hat{f}_{\mathcal{C}}$ is straightforward. First, an integer $j \in \{1, \ldots, N\}$ is chosen randomly with each integer having equal likelihood. Next, a random sample is drawn from a Gaussian with covariance $h^2 I_d$ and mean x_{S_j}, where I_d denotes the d-by-d identity matrix.

With crude Monte Carlo, the probability of a collision is calculated by taking the mean of $R(x_S)$ over a large number, N_{MC}, of different realizations of x_S:

$$\mu_{\mathrm{MC}} = \frac{1}{N_{\mathrm{MC}}} \sum_{j=1}^{N_{\mathrm{MC}}} R(x_j), \quad x_j \sim \hat{f}_{\mathcal{C}}. \tag{6}$$

[2]The advantage of the Gaussian kernel is that it gives the possibility to calculate a metric that quantifies the completeness of the data [65] and to apply conditional sampling when generating scenario parameters [66]. Both these topics are out of scope of this article.

It is easy to see that the crude Monte Carlo result is unbiased, i.e., $\mathbb{E}[\mu_{\text{MC}}] = \mathbb{E}_{S\in\mathcal{C}}[R(x_S)]$. To estimate the potential approximation error, $\mu_{\text{MC}} - \mathbb{E}_{S\in\mathcal{C}}[R(x_S)]$, the estimated standard deviation of (6) is commonly used:

$$\sigma_{\text{MC}} = \frac{1}{N_{\text{MC}}}\sqrt{\sum_{j=1}^{N_{\text{MC}}}\big(\mu_{\text{MC}} - R(x_j)\big)^2}. \tag{7}$$

2) NONPARAMETRIC IMPORTANCE SAMPLING

In general, it can be expected that the probability of collision, $\mathbb{E}_{S\in\mathcal{C}}[R(x_S)]$, is small. As a result, none or few of the N_{MC} scenario simulations may end with a collision and the relative uncertainty, i.e., $\sigma_{\text{MC}}/\mu_{\text{MC}}$, will be high or undefined (in case none of the scenario simulations ends with a collision). With importance sampling, the scenario parameters are sampled from a different distribution — the so-called importance density — such that the simulation runs focus more on scenarios in which the probability of collision is high. This will lead to a lower relative uncertainty of the estimated probability of collision. We use nonparametric importance sampling, which means that a nonparametric method is used to estimate the unknown optimal importance density [71]. More specifically, as proposed in [71], the nonparametric importance sampling method employs KDE to construct the importance density.

Let $g(\cdot)$ denote the importance density for sampling the scenario parameters. If N_{NIS} scenarios are simulated in nonparametric importance sampling, the estimated probability of collision is

$$\mu_{\text{NIS}} = \frac{1}{N_{\text{NIS}}}\sum_{j=1}^{N_{\text{NIS}}} R(x_j)\frac{\hat{f}_{\mathcal{C}}(x_j)}{g(x_j)}, \quad x_j \sim g. \tag{8}$$

The weight $\hat{f}_{\mathcal{C}}(x_j)/g(x_j)$ is used to correct for the bias introduced by sampling from $g(\cdot)$ instead of $\hat{f}_{\mathcal{C}}(\cdot)$. If $g(x) > 0$ whenever $R(x)\hat{f}_{\mathcal{C}}(x) > 0$, then (8) gives unbiased results [72]. The estimated standard deviation of (8) is

$$\sigma_{\text{NIS}} = \frac{1}{N_{\text{NIS}}}\sqrt{\sum_{j=1}^{N_{\text{NIS}}}\left(\frac{R(x_j)\hat{f}_{\mathcal{C}}(x_j)}{g(x_j)} - \mu_{\text{NIS}}\right)^2}. \tag{9}$$

In the ideal case, we choose $g(\cdot)$ such that the actual standard deviation of μ_{NIS} is minimized. This ideal $g(\cdot)$ is, however, unknown because it depends on $R(\cdot)$, for which no functional form is available, and on the unknown $\mathbb{E}_{S\in\mathcal{C}}[R(x_S)]$. Therefore, we propose to first conduct crude Monte Carlo sampling and to base $g(\cdot)$ on this result. Let $\{x_{(1)}, \ldots, x_{(N_{\text{MC}})}\}$ denote the ordered set of scenario parameters from the crude Monte Carlo sampling, such that $x_{(1)}$ and $x_{(N_{\text{MC}})}$ are the parameter vectors corresponding to the "most critical" scenario and "least critical" scenario, respectively. Then, we use the KDE technique described earlier to construct $g(\cdot)$ using the $N_{\text{C}} < N_{\text{MC}}$ "most critical" scenarios:

$$g(x) = \frac{1}{N_{\text{C}}h_{\text{NIS}}^d}\sum_{i=1}^{N_{\text{C}}} K\left(\frac{1}{h_{\text{NIS}}}\big(x - x_{(i)}\big)\right). \tag{10}$$

Using the Gaussian kernel of (4), it follows that $g(x) > 0$ for all x, such that (8) gives an unbiased result. The bandwidth h_{NIS} is estimated using leave-one-out cross validation, similar to h in (5).

To order the scenarios from "most critical" to "least critical", a metric for quantifying the (maximum) risk during a single simulation run can be used. For illustration purposes, this work uses the minimum absolute Time to Collision (TTC) [73] during each simulation run. Note that the presented method can easily be applied with other metrics. The TTC is defined as the ratio of the distance toward an approaching object and the speed difference with that object. In case the simulation of a scenario ends in a collision, the minimum TTC is 0. Note that N_{C} must be larger than the number of simulation runs of the crude Monte Carlo sampling that ended in a collision and N_{C} must be substantially smaller than N_{MC}, e.g., an order of magnitude, to really benefit from the nonparametric importance sampling.

E. CALCULATION OF SEVERITY

Besides the probability of a collision, risk also includes the extent of the harm in a potential collision. We express the severity as the expectation of the probability of a moderate injury or worse, corresponding to a Maximum Abbreviated Injury Scale (MAIS) [74] level of 2 or higher: $\mathbb{E}_{S\in\mathcal{C}}[\mathbb{P}(I(x_S))]$.

Typically, two different approaches are considered for predicting the extent of harm in a crash. The first approach involves simulation of the crash. The simulations as explained in Section III-C and as used in Section III-D consider the pre-crash phase and are used to determine initial conditions and boundary conditions for the simulation of the in-crash phase. The extent of harm is predicted by simulations of the in-crash phase using biomechanical models; see [75]–[79] for an overview. The second approach makes use of phenomenological injury risk functions based on the research of the relationships between collision parameters and the probability of an injury [80]. The most commonly used collision parameter is the impact velocity change [81]–[85]. Other factors for determining the probability of an injury are belt usage [81]–[83], occupant age [81], [84], peak acceleration during the crash [86], airbag deployment [84], and seating position of the occupants [84]. Typically, logistic regression is used to model the relationship between collision parameters and the probability of an injury.

In this article, we assume that a method to estimate the probability of an injury given a parameterized scenario, i.e., $\mathbb{P}(I(x_S))$, is known. In the case study in Section V, an example is provided. To estimate $\mathbb{E}_{S\in\mathcal{C}}[\mathbb{P}(I(x_S))]$, the same approach for estimating the expectation of $R(x)$ in (8) is used:

$$\mathbb{E}_{S\in\mathcal{C}}[\mathbb{P}(I(x_S))] \approx \mu_{\text{injury}} = \frac{1}{N_{\text{NIS}}}\sum_{j=1}^{N_{\text{NIS}}} w(x_j), \; x_j \sim g, \tag{11}$$

with

$$w(x_j) = \mathbb{P}(I(x_S))R(x_j)\frac{\hat{f}_{\mathcal{C}}(x_j)}{g(x_j)}. \tag{12}$$

The estimated standard deviation of μ_{injury} is

$$\sigma_{\text{injury}} = \frac{1}{N_{\text{NIS}}}\sqrt{\sum_{j=1}^{N_{\text{NIS}}}\left(w(x_j) - \mu_{\text{injury}}\right)^2}. \tag{13}$$

F. RISK QUANTIFICATION

The risk associated with a scenario category $\mathcal{C}$ can be defined as the combination of the probability of occurrence of a scenario of $\mathcal{C}$ and the probability of an injury given such a scenario. Thus, the risk is mathematically defined as:

$$\text{Risk}(\mathcal{C}) = \mathbb{E}[n_{\mathcal{C}}] \cdot \mathbb{E}_{S\in\mathcal{C}}[\mathbb{P}(I(S))], \tag{14}$$

where $\mathbb{E}[n_{\mathcal{C}}]$ is defined in (1) and $\mathbb{E}_{S\in\mathcal{C}}[\mathbb{P}(I(S))]$ is estimated in (11).

Remark 3: The risks of two scenario categories can be combined as follows:

$$\text{Risk}(\mathcal{C}_1 \cup \mathcal{C}_2) = \text{Risk}(\mathcal{C}_1) + \text{Risk}(\mathcal{C}_2) - \text{Risk}(\mathcal{C}_1 \cap \mathcal{C}_2). \tag{15}$$

In general, it is sufficient to estimate the upper bound of the risk, so in case it is practically difficult to evaluate $\text{Risk}(\mathcal{C}_1 \cap \mathcal{C}_2)$, one can use

$$\text{Risk}(\mathcal{C}_1 \cup \mathcal{C}_2) \leq \text{Risk}(\mathcal{C}_1) + \text{Risk}(\mathcal{C}_2), \tag{16}$$

where equality applies if two scenario categories, $\mathcal{C}_1$ and $\mathcal{C}_2$, do not overlap. ◊

IV. RELATION WITH ISO 26262 AND ISO/DIS 21448

In this section, we propose how to quantify the exposure, severity, and controllability in Sections IV-A to IV-C, respectively. Finally, in Section IV-D, we show that combining these aspects results in the earlier calculated risk of (14).

A. EXPOSURE

We consider the likelihood of being in a scenario that is comprised by the scenario category $\mathcal{C}$. Hence, the exposure is similar to (1):

$$\text{Exposure}(\mathcal{C}) = \mathbb{E}[n_{\mathcal{C}}]. \tag{17}$$

Note that the ISO 26262 standard [35] also mentions the "failure mode", see Table 1. The ISO/DIS 21448 standard [36], however, does not consider a specific "failure mode", which is why we focus on the likelihood of "being in an operational situation". Here, the "operational situation" is described by the scenario category $\mathcal{C}$.

B. SEVERITY

In Section III-E, a method to compute the severity has been proposed. Following the reasoning of the ISO 26262 standard, however, the severity is defined slightly differently, as it is assumed that there is no operator that can avoid harm or damage. Therefore, we define severity as follows:

$$\text{Severity}(\mathcal{C}) = \mathbb{E}_{S\in\mathcal{C}}\left[\mathbb{P}(I(S))|\text{no operator}\right], \tag{18}$$

where "no operator" indicates that the backup function of the operator is not considered in the simulations.

Remark 4: An operator might still be considered in the simulation. For example, if the ADS only controls the vehicle in the longitudinal direction, an operator is still in charge of lateral control. In this example, the notation "no operator" indicates that the driver is not a backup for the longitudinal control. ◊

Note that similar to the exposure, the severity in (18) is calculated with respect to a scenario category, whereas the ISO 26262 standard determines the severity with respect to a "hazardous event". Our method still allows to evaluate the severity considering such malfunctioning behavior by simply injecting this malfunctioning behavior as part of the sensor(s), ADS(s), and/or vehicle(s). In a similar manner, any triggering conditions (see Remark 1) that may cause hazardous behavior can be included into the simulations.

C. CONTROLLABILITY

To determine the controllability, we compare the probability of an injury with and without a backup operator. Hence,

$$\text{Controllability}(\mathcal{C}) = \frac{\mathbb{E}_{S\in\mathcal{C}}[\mathbb{P}(I(S))]}{\mathbb{E}_{S\in\mathcal{C}}\left[\mathbb{P}(I(S))|\text{no operator}\right]}. \tag{19}$$

If $\text{Controllability}(\mathcal{C}) = 1$, then a backup operator cannot avoid any potential harm. The ISO 26262 standard calls this "difficult to control or uncontrollable". If, on the other hand, $\text{Controllability}(\mathcal{C}) = 0$, then a backup operator is able to avoid any potential harm.

Remark 5: It might be counterintuitive that a higher score for $\text{Controllability}(\mathcal{C})$ indicates that the situation is less controllable. The ISO 26262 standard also assigns higher scores for the controllability in case the situation is less controllable, see Section II-A. We have chosen to prefer consistency with respect to the ISO 26262 standard, which is why $\text{Controllability}(\mathcal{C}) = 1$ means that the situation is difficult to control and $\text{Controllability}(\mathcal{C}) = 0$ means that the situation is easy to control. ◊

D. COMBINING THE RISK ASPECTS TO COMPUTE THE RISK

To compute the risk, the scores for the exposure, severity, and controllability are multiplied:

$$\text{Risk}(\mathcal{C}) = \text{Exposure}(\mathcal{C}) \cdot \text{Severity}(\mathcal{C}) \cdot \text{Controllability}(\mathcal{C}). \tag{20}$$

Substituting (17) to (19) into (20) results in the risk of (14).

Remark 6: The ASIL ranking is obtained by summing the scores for the exposure, severity, and controllability, while the risk in (20) is obtained by multiplying the scores for the exposure, severity, and controllability. Following the ISO 26262 standard, the scoring for the exposure, severity, and controllability is such that one point difference corresponds to an order in magnitude. Therefore, loosely said, the ASIL is proportional to the log of the risk of (20) while the individual scores of the exposure, severity, and controllability are proportional to the log of (17), (18), and (19), respectively. Since we can rewrite (20) as

$$\begin{aligned}\log \text{Risk}(\mathcal{C}) &= \log \text{Exposure}(\mathcal{C}) \\ &\quad + \log \text{Severity}(\mathcal{C}) + \log \text{Controllability}(\mathcal{C}),\end{aligned} \tag{21}$$

(20) is consistent with the way the risk is determined in the ISO 26262 standard. ◊

V. CASE STUDY

This section explains the details of the case study; the results are reported in the next section. In Section V-A, the models for the ADS under test and the human driver that serves as a backup operator are described. The scenario categories and triggering conditions are listed in Sections V-B and V-C, respectively. Section V-D describes the data set. This section is ended with details on the simulation (Section V-E) and the severity estimation (Section V-F).

A. AUTOMATED DRIVING SYSTEM UNDER TEST

For the ADS under test, we consider the ACC model described by Xiao *et al.* [87]. This ACC model is based on the ACC model proposed by Milanés and Shladover [88] but also includes a safety distance d_0. The ACC function adjusts the ego vehicle speed such that the ego vehicle maintains a safe distance from a leading vehicle. If there is no vehicle ahead or the distance between the ego vehicle and the leading vehicle is large, then the ACC acts like a Cruise Control (CC) that maintains a constant speed set by the driver. The acceleration of the ACC is based on the speed of the ego vehicle, $v_e(t)$, the speed of the leading vehicle, $v_l(t)$, and the gap between the leading vehicle's back and the ego vehicle's front, $g(t)$, at time t. If the ACC function is active, the acceleration of the ego vehicle at time t is described by the following equations [87]:

$$a_e(t) = \max(\min(a_{ACC}(t), a_{CC}(t)), -d_{max}), \tag{22}$$

$$a_{ACC}(t) = \begin{cases} k_1 g_{error}(t) + k_2 v_{error}(t) & \text{if } g(t) < d_{ACC}, \\ a_{CC}(t) & \text{otherwise}, \end{cases} \tag{23}$$

$$g_{error}(t) = g(t) - d_0(v_e(t)) - \tau_h v_e(t), \tag{24}$$

$$v_{error}(t) = v_l(t) - v_e(t), \tag{25}$$

$$d_0(u) = \begin{cases} 5\,\text{m} & \text{if } u \geq 15\,\text{m/s}, \\ 7\,\text{m} & \text{if } u < 10.8\,\text{m/s}, \\ \dfrac{75\,\text{m}^2/\text{s}}{u} & \text{otherwise}, \end{cases} \tag{26}$$

$$a_{CC}(t) = k_{CC}(v_{set} - v_e(t)). \tag{27}$$

The values and descriptions of the parameters d_{max}, d_{ACC}, k_1, k_2, τ_h, and k_{CC} are provided in Table 3. The parameter v_{set} is the desired speed, which is assumed to be the same as the initial speed of the ego vehicle in each simulation run, i.e., $v_{set} = v_e(0)$.

Following Xiao *et al.* [87], the human driver takes over control from the ACC in the following circumstances:

- In case of a Forward Collision Warning (FCW). If the FCW is at time t, then the human driver takes over at $t + t_r$, where t_r is the reaction time of the human driver.
- In case the ego vehicle approaches the leading vehicle with a relative speed of 15 m/s and the gap between the ego vehicle and the leading vehicle is less than the perception range, d_{view}.

Similar as in [87], the FCW model is taken from [89] and is triggered when $1/(1 + \exp\{-\beta\}) > 0.75$, with

$$\beta = \begin{cases} -6.09 + 18.82 \dfrac{v_e - v_l}{g} + 0.12 v_e & \text{if } v_l > 0 \wedge a_l < 0, \\ -6.09 + 12.58 \dfrac{v_e - v_l}{g} + 0.12 v_e & \text{if } v_l > 0 \wedge a_l \geq 0, \\ -9.07 + 24.23 \dfrac{v_e - v_l}{g} + 0.12 v_e & \text{otherwise}, \end{cases} \tag{28}$$

where a_l denotes the acceleration of the leading vehicle. The reaction time of the driver, t_r, is distributed according to the log-normal distribution with a mean of 0.92 s and a standard deviation of 0.28 s [90].

Similar to Xiao *et al.* [87], we use the Intelligent Driver Model plus (IDM+) [92] to model the human driver behavior. If the human driver is controlling the vehicle and $g(t) \leq d_{view}$, then the acceleration of the ego vehicle at time $t + t_r$ is described by the following equations [92]:

$$a_e(t + t_r) = \max(a_{IDM}(t), -d_{max}), \tag{29}$$

$$a_{IDM}(t) = k_a \min\left(1 - \left(\frac{v_e(t)}{v_{set}}\right)^{\delta}, 1 - \left(\frac{g^*(v_e(t), v_e(t) - v_l(t))}{g(t)}\right)^2\right), \tag{30}$$

$$g^*(u, v) = s_0 + \tau_h u + \frac{uv}{2\sqrt{k_a k_d}}. \tag{31}$$

Table 3 provides the descriptions of the constants k_a, k_d, δ, s_0, and τ_h. If $g(t) > d_{view}$, then $a_e(t + t_r) = 0\,\text{m/s}^2$.

B. SCENARIO CATEGORIES

This case study considers 3 scenario categories named "leading vehicle decelerating (LVD)" ($\mathcal{C}_{LVD}$), "cut-in" ($\mathcal{C}_{cut-in}$), and "approaching slower vehicle (ASV)" ($\mathcal{C}_{ASV}$), see Figure 3. The relevance of the first 2 scenario categories is exemplified by the proposed regulation for automated lane-keeping systems in which these two scenario categories are mentioned as "critical scenarios" [93]. The scenario category ASV, which also includes scenarios in which the leading vehicle is stationary, accounts for more than 25 % of all

TABLE 3. Parameters of the system under test and human driver behavior model. If the parameter value is taken from literature, the corresponding reference is mentioned after the value. For the parameters $d_{\max}$ and d_{view}, different values are considered so-called triggering conditions are included in the scenarios, see Section V-C.

Parameter	Description	Value
$d_{\max}$	Maximum deceleration (value with triggering condition "limited braking capacity")	$6\,\text{m/s}^2$ ($3\,\text{m/s}^2$)
d_{ACC}	Maximum sensor range of ACC	150 m
k_1	Distance gain of ACC	$0.23\,\text{s}^{-2}$ [87]
k_2	Speed gain of ACC	$0.07\,\text{s}^{-1}$ [87]
τ_{h}	Time-gap setting, also known as desired time headway	1.1 s [87]
k_{CC}	Speed gain of CC	$0.4\,\text{s}^{-1}$ [87]
v_{set}	Desired speed	Variable
t_{r}	Reaction time	Variable [90]
d_{view}	Perception range of human driver (value with triggering condition "poor visibility")	150 m (60 m) [87]
k_{a}	Maximum acceleration for human driver	$0.73\,\text{m/s}^2$ [91]
k_{d}	Maximum deceleration for human driver	$1.67\,\text{m/s}^2$ [91]
δ	Acceleration exponent for human driver	4 [91]
s_0	Safety distance for human driver	2 m [91]

crashes that involve two vehicles [94]. For each scenario category, we will list the parameters that describe the scenarios. Furthermore, we will describe how the speed of the leading vehicle is modeled. Since the ego vehicle is controlled by the ACC or the human driver, only its initial state at $t = 0$ is provided.

1) $\mathcal{C}_{\text{LVD}}$: LEADING VEHICLE DECELERATING

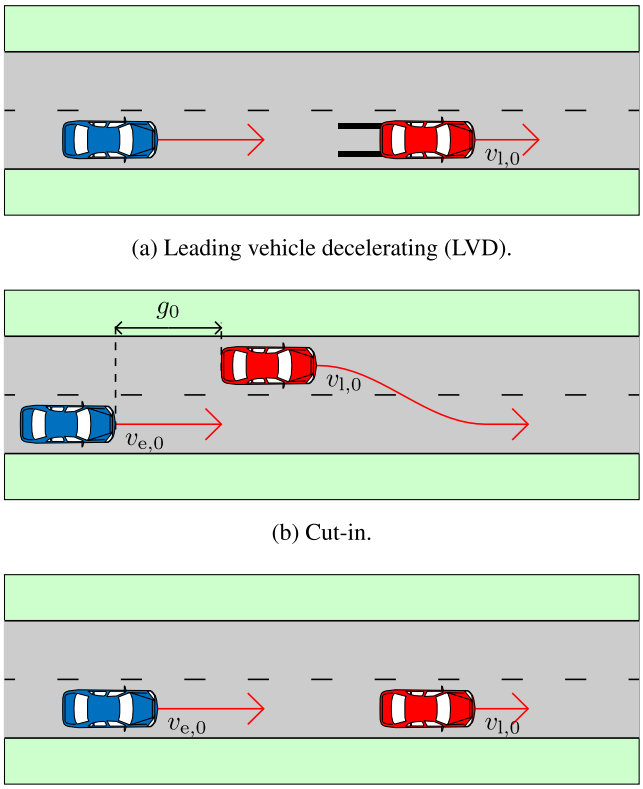
In an LVD scenario, the ego vehicle is following another vehicle, which is referred to as the leading vehicle. The LVD scenario starts as soon as the leading vehicle starts to decelerate. The simulation of LVD scenario ends if the distance between the ego vehicle and the leading vehicle is no longer decreasing or if the ego vehicle and the leading vehicle collide. To describe an LVD scenario, 3 parameters are used:

- $v_{\text{l},0} > 0$: The initial speed of the leading vehicle;
- $\Delta_v \in (0, v_{\text{l},0}]$: The speed difference of the leading vehicle obtained through decelerating; and
- $\bar{a} > 0$: The average deceleration of the leading vehicle.

To model the speed of the leading vehicle during its deceleration activity, a sinusoidal shape is used:

$$v_{\text{l}}(t) = \begin{cases} v_{\text{l},0} - \dfrac{\Delta_v}{2}\left(1 - \cos\left(\dfrac{\pi \bar{a} t}{\Delta_v}\right)\right) & \text{if } t < \dfrac{\Delta_v}{\bar{a}}, \\ v_{\text{l},0} - \Delta_v & \text{otherwise.} \end{cases} \tag{32}$$

It is assumed that the ego vehicle is following the leading vehicle at the same speed, i.e., $v_{\text{e}}(0) = v_{\text{l},0}$. This is also the desired speed, thus $v_{\text{set}} = v_{\text{l},0}$. The initial gap is such that the ego vehicle is initially driving at a constant speed, so $g(0) = d_0(v_{\text{e}}(0)) + \tau_{\text{h}} v_{\text{e}}(0)$, such that the initial distance error, $g_{\text{error}}(0)$, is zero, see (24).



(a) Leading vehicle decelerating (LVD).

(b) Cut-in.

(c) Approaching slower vehicle (ASV).

FIGURE 3. Schematic representation of the scenario categories considered in the case study. The left vehicle is the ego vehicle. All parameters are shown except Δ_v and $\bar{a}$ of the LVD scenario.

2) $\mathcal{C}_{\text{cut-in}}$: CUT-IN

In a cut-in scenario, another vehicle changes lane such that it becomes the leading vehicle of the ego vehicle. The reason for the other vehicle to change lane is not considered, i.e., it may change lane because the driver of the vehicle assumes it is safe and appropriate to change lane or because a lane change is mandatory. A cut-in scenario starts as soon as the other vehicle enters the lane of the ego vehicle. Similarly as for an LVD scenario, the simulation of a cut-in scenario ends as soon as the distance between the ego vehicle and the leading vehicle is no longer decreasing or if there is a collision. To describe a cut-in scenario, 3 parameters are used:

- $g_0 > 0$: The gap between the leading vehicle and the ego vehicle at the moment of the cut-in;
- $v_{\text{l},0} > 0$: The initial speed of the leading vehicle; and
- $v_{\text{e},0} > 0$: The initial speed of the ego vehicle.

It is assumed that the leading vehicle drives at a constant speed, thus $v_{\text{l}}(t) = v_{\text{l},0}$. The initial gap and the initial speed of the ego vehicle are provided by the parameters: $g(0) = g_0$ and $v_{\text{e}}(0) = v_{\text{e},0}$. The initial speed of the ego vehicle is also the desired speed: $v_{\text{set}} = v_{\text{e},0}$.

3) $\mathcal{C}_{\text{ASV}}$: APPROACHING SLOWER VEHICLE

In an ASV scenario, another vehicle, referred to as the leading vehicle, is driving in front of the ego vehicle. Furthermore, the

leading vehicle is driving slower than the ego vehicle, such that the ego vehicle is approaching this vehicle. The ASV scenario starts if the ego vehicle is at a safe distance and ends if the distance between the two vehicle is no longer decreasing or if the two vehicles collide. To describe an ASV scenario, 2 parameters are used:

- $v_{l,0} \geq 0$: The initial speed of the leading vehicle; and
- $v_{e,0} > 0$: The initial speed of the ego vehicle.

It is assumed that the leading vehicle drives at a constant speed, thus $v_l(t) = v_{l,0}$. The initial speed of the ego vehicle is $v_e(0) = v_{e,0}$, which is also the desired speed: $v_{set} = v_{e,0}$. The initial gap is $g(0) = \tau_{h,0} v_e(0)$ with $\tau_{h,0} = 4$ s, such that the initial distance is safe, considering a time headway of 4 s.

C. TRIGGERING CONDITIONS

To illustrate the application of the proposed method for the risk quantification for the validation of the SOTIF, triggering conditions are included in the scenarios that may trigger hazardous situations, see Remark 1. For comparison, the risk is calculated without a triggering condition and with a triggering condition.

The first triggering condition is a limited braking capacity of the ego vehicle. As a result, the maximum deceleration of the ego vehicle is $d_{max} = 3\,\text{m/s}^2$. The reason for a limited braking capacity is not further specified here, but it could be caused by a loss of road friction, e.g., due to an adverse road condition.

The second triggering condition is a poor visibility. It is assumed that the human driver can only see up to $d_{view} = 60$ m. It is further assumed that the maximum sensor range of the ACC, d_{ACC}, is unaffected. This assumption is reasonable in case the poor visibility is caused by fog, because fog has a limited effect on automotive radars that are typically used for an ACC.

D. DATA SET

To estimate the exposure and the pdfs of the scenario parameters, the data set described in [95] is used. The data were recorded from a single vehicle in which 20 experienced drivers were asked to drive a prescribed route. Table 4 lists more information about the drivers. Each driver drove the 50 km route, as shown in Figure 4, 6 times, which resulted in 63 h of data. The route contains urban roads, rural roads, and highways. To measure the surrounding traffic, the vehicle was equipped with three radars and one camera. The surrounding traffic was measured by fusing the data of the radars and the camera as described in [96].

To extract the scenarios from the data set, the approach described in [55] is used. LVD scenarios are found by querying for a vehicle that has the tags "leading vehicle" and "braking" at the same time. Cut-in scenarios are found by querying for a vehicle that initially has the tags "changing lane" and "no leading vehicle" which changes into the tags "changing lane" and "leading vehicle" [55]. ASV scenarios are found by querying for a vehicle that has the tags "leading vehicle" and "driving slower", where the tag "driving slower" indicates that the vehicle has at most 90 % of the speed of the ego vehicle. In 63 hours of driving, 1300 LVD scenarios, 297 cut-in scenario, and 291 ASV scenarios has been found.

TABLE 4. Information of the participants that drove the 50 km route 6 times.

	Minimum	Maximum	Mean
Age [years]	29	60	52.3
Experience [years]	8	42	32.4
Mileage [km/year]	$10 \cdot 10^3$	$40 \cdot 10^3$	$16.1 \cdot 10^3$

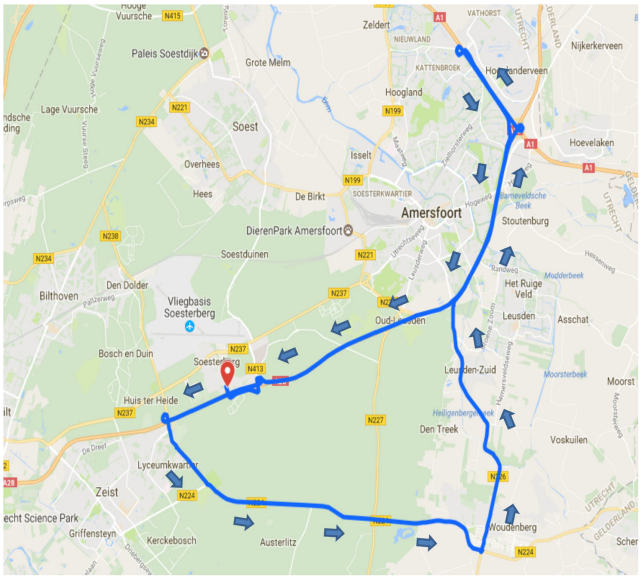


FIGURE 4. The route, including the driving direction, that was driven by the drivers.

E. SIMULATIONS

For the simulation, we use the forward Euler method with a step size of 0.01 s to compute the positions of the ego vehicle and the leading vehicle. We used Python as the coding language. The code is available on a public repository.[3] Initially, the crude Monte Carlo sampling with $N_{MC} = 10000$ simulation runs is performed where the scenario parameters are drawn from $\hat{f}_{\mathcal{C}}(\cdot)$ of (3). Additionally, if applicable, the driver reaction time, t_r, is sampled from the log-normal distribution described in Section V-A.

The importance density $g(\cdot)$ of (10) is constructed using the $N_C = 200$ "most critical" scenarios of the crude Monte Carlo simulation: the scenarios with the lowest TTC. If applicable, the driver reaction time is also part of the multivariate pdf $g(\cdot)$. Note that these N_C "most critical" scenarios always include scenarios that ended with a collision, because $\mu_{MC} < N_C/N_{MC} = 0.02$.

[3] https://github.com/ErwindeGelder/ScenarioRiskQuantification

F. PROBABILITY OF AN INJURY

For the calculation of the injury rate, see (11), it has been assumed that the probability of an injury given a parameterized scenario, $\mathbb{P}(I(S))$, is known. The presented case study uses the model from Kusano & Gabler [83] to determine $\mathbb{P}(I(x_S))$, since this model is also used in [57], [97]. Kusano & Gabler [83] model the relationship between impact velocity change and belt usage with the probability of an injury with MAIS≥2 in rear-end crashes:

$$\mathbb{P}(I(x_S)) = \frac{1}{1+\exp\{-(\beta_0+\beta_1\Delta v(x_S)+\beta_2 b)\}}. \quad (33)$$

Here, $\beta_0 = -6.068$, $\beta_1 = 0.1000$ s/m, and $\beta_2 = 0.6234$ are parameters that are fitted to data of 1406 rear-end crashes [83]. For the velocity change during a crash, $\Delta v(x_S)$, which depends on the masses of the two objects colliding, we assume equal masses such that $\Delta v(x_S)$ is half of the impact speed (i.e., the speed difference at the start of the crash). If the belt is not used, then $b = -1$. In this case study, it is assumed that the belt is always used, so $b = 1$.

VI. RESULTS

This section provides the results of the case study described in Section V. First, the exposures of the scenarios are listed. Next, the severity and controllability are reported. Finally, we explain the results of the simulations that include the triggering conditions.

A. EXPOSURE

The bar graph in Figure 5 shows the estimated probability that the respective scenario is found n times in 1 hour of driving. Given the number of encounters of the scenarios, we have the following exposures:

$$\text{Exposure}(\mathcal{C}_{\text{LVD}}) = \mathbb{E}\left[n_{\mathcal{C}_{\text{LVD}}}\right] = 20.6\,\text{h}^{-1}, \quad (34)$$

$$\text{Exposure}(\mathcal{C}_{\text{cut-in}}) = \mathbb{E}\left[n_{\mathcal{C}_{\text{cut-in}}}\right] = 4.71\,\text{h}^{-1}, \quad (35)$$

$$\text{Exposure}(\mathcal{C}_{\text{ASV}}) = \mathbb{E}\left[n_{\mathcal{C}_{\text{ASV}}}\right] = 4.62\,\text{h}^{-1}. \quad (36)$$

As mentioned in Remark 2, it is not uncommon to assume that the probability of encountering n scenarios belonging to a specific scenario category is Poisson-distributed. To compare the results with the Poisson distribution of (2), the probability of the Poisson distribution is also shown in Figure 5. The parameter λ is estimated by maximizing the likelihood. This gives the same results as (34) to (36), i.e., $\lambda = 20.6\,\text{h}^{-1}$ for the LVD scenarios, $\lambda = 4.71\,\text{h}^{-1}$ for the cut-in scenarios, and $\lambda = 4.62\,\text{h}^{-1}$ for the ASV scenarios. According to the Chi-square goodness of test, the probability that $\mathbb{P}\left(n_{\mathcal{C}_{\text{LVD}}} = n\right)$, $\mathbb{P}\left(n_{\mathcal{C}_{\text{cut-in}}} = n\right)$, and $\mathbb{P}\left(n_{\mathcal{C}_{\text{ASV}}} = n\right)$ are distributed according to the Poisson distribution is very small ($p < 0.001$). Hence, it can be concluded that the assumption that the number of encounters of the LVD, cut-in, and ASV scenarios is Poisson-distributed is unrealistic.

B. SEVERITY AND CONTROLLABILITY

As explained in Section III-D, to estimate the probability of a collision, the pdfs of the scenario parameters need to be estimated. In Figures 6 to 8, the results of the pdf estimation are shown. The histograms show the original data that are used to estimate the pdfs. The multivariate pdfs are estimated using KDE, see (3). To account for the infinite support of the Gaussian kernel of (4), the resulting pdfs are set to 0 if the parameters are outside the valid range of parameters as mentioned in Section V-B. Next, the pdfs are scaled such that they integrate to 1. The solid lines in Figures 6 to 8 represent the marginal distributions of the resulting pdfs.

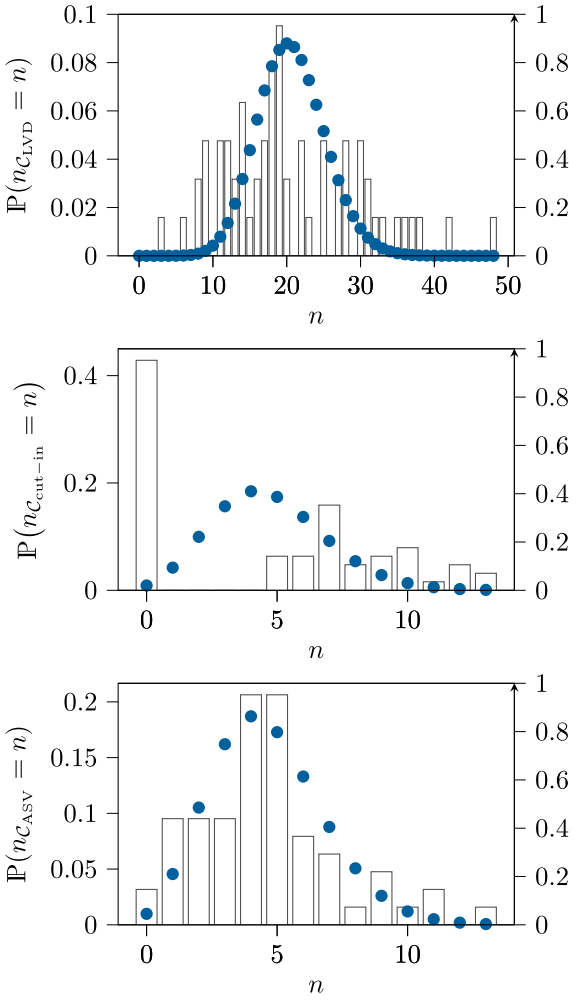


FIGURE 5. The bar graph shows how often a scenario of scenario category $\mathcal{C}_{\text{LVD}}$, $\mathcal{C}_{\text{cut-in}}$, or $\mathcal{C}_{\text{ASV}}$ is encountered n times in 1 hour of driving divided over the total number of hours, 63. For comparison, the dots denote the Poisson distribution of (2) with the maximum likelihood estimate of λ.

Table 5 shows the results of the simulation runs. It shows that the estimated probability of an injury with MAIS≥2 in LVD scenarios is $2.77 \cdot 10^{-7}$ with an estimated uncertainty of $1.27 \cdot 10^{-7}$. If, however, there would be no backup from an operator, then this probability is substantially higher: $\text{Severity}(\mathcal{C}_{\text{LVD}}) = 1.12 \cdot 10^{-5}$. Therefore, the controllability score is small ($\text{Controllability}(\mathcal{C}_{\text{LVD}}) = 2.47 \cdot 10^{-2}$), resulting in $\text{Risk}(\mathcal{C}_{\text{LVD}}) = 5.72 \cdot 10^{-6}\,\text{h}^{-1}$. In other words, it is expected that, on average, one collision in an LVD scenario with a moderate injury or worse happens in $1.75 \cdot 10^5$ h of driving.

TABLE 5. Results of the case study. μ_{NIS} of (8): estimated probability of collision. σ_{NIS} of (9): estimated standard deviation of μ_{NIS}. μ_{injury} of (11): estimated probability of an injury with MAIS≥2. σ_{injury} of (13): estimated standard deviation of μ_{NIS}. Severity: estimation of (18). Controllability: estimation of (19). Risk: estimation of (20). Note that for the risk estimation, the exposure of the triggering condition is not considered.

$\mathcal{C}$	Trig. cond.	μ_{NIS}	σ_{NIS}	μ_{injury}	σ_{injury}	Severity	Controllability	Risk
$\mathcal{C}_{\text{LVD}}$	None	$1.61 \cdot 10^{-4}$	$5.95 \cdot 10^{-5}$	$2.77 \cdot 10^{-7}$	$1.27 \cdot 10^{-7}$	$1.12 \cdot 10^{-5}$	$2.47 \cdot 10^{-2}$	$5.72 \cdot 10^{-6}\,\text{h}^{-1}$
$\mathcal{C}_{\text{LVD}}$	Limited braking	$7.22 \cdot 10^{-3}$	$1.58 \cdot 10^{-4}$	$1.16 \cdot 10^{-5}$	$2.70 \cdot 10^{-7}$	$1.52 \cdot 10^{-5}$	0.763	$2.39 \cdot 10^{-4}\,\text{h}^{-1}$
$\mathcal{C}_{\text{LVD}}$	Poor visibility	$1.61 \cdot 10^{-4}$	$5.95 \cdot 10^{-5}$	$2.77 \cdot 10^{-7}$	$1.27 \cdot 10^{-7}$	$1.12 \cdot 10^{-5}$	$2.47 \cdot 10^{-2}$	$5.72 \cdot 10^{-6}\,\text{h}^{-1}$
$\mathcal{C}_{\text{cut-in}}$	None	$1.95 \cdot 10^{-3}$	$1.32 \cdot 10^{-4}$	$3.71 \cdot 10^{-6}$	$2.71 \cdot 10^{-7}$	$3.92 \cdot 10^{-6}$	0.947	$1.75 \cdot 10^{-5}\,\text{h}^{-1}$
$\mathcal{C}_{\text{cut-in}}$	Limited braking	$2.20 \cdot 10^{-3}$	$1.35 \cdot 10^{-4}$	$4.79 \cdot 10^{-6}$	$3.23 \cdot 10^{-7}$	$4.54 \cdot 10^{-6}$	1.06	$2.26 \cdot 10^{-5}\,\text{h}^{-1}$
$\mathcal{C}_{\text{cut-in}}$	Poor visibility	$1.95 \cdot 10^{-3}$	$1.32 \cdot 10^{-4}$	$3.71 \cdot 10^{-6}$	$2.71 \cdot 10^{-7}$	$3.92 \cdot 10^{-6}$	0.947	$1.75 \cdot 10^{-5}\,\text{h}^{-1}$
$\mathcal{C}_{\text{ASV}}$	None	$1.03 \cdot 10^{-7}$	$1.02 \cdot 10^{-7}$	$1.66 \cdot 10^{-10}$	$1.64 \cdot 10^{-10}$	$2.12 \cdot 10^{-5}$	$7.83 \cdot 10^{-6}$	$7.68 \cdot 10^{-10}\,\text{h}^{-1}$
$\mathcal{C}_{\text{ASV}}$	Limited braking	$4.51 \cdot 10^{-3}$	$1.72 \cdot 10^{-4}$	$9.78 \cdot 10^{-6}$	$4.15 \cdot 10^{-7}$	$3.70 \cdot 10^{-5}$	0.265	$4.52 \cdot 10^{-5}\,\text{h}^{-1}$
$\mathcal{C}_{\text{ASV}}$	Poor visibility	$3.57 \cdot 10^{-3}$	$1.41 \cdot 10^{-4}$	$1.01 \cdot 10^{-5}$	$3.97 \cdot 10^{-7}$	$2.12 \cdot 10^{-5}$	0.476	$4.66 \cdot 10^{-5}\,\text{h}^{-1}$

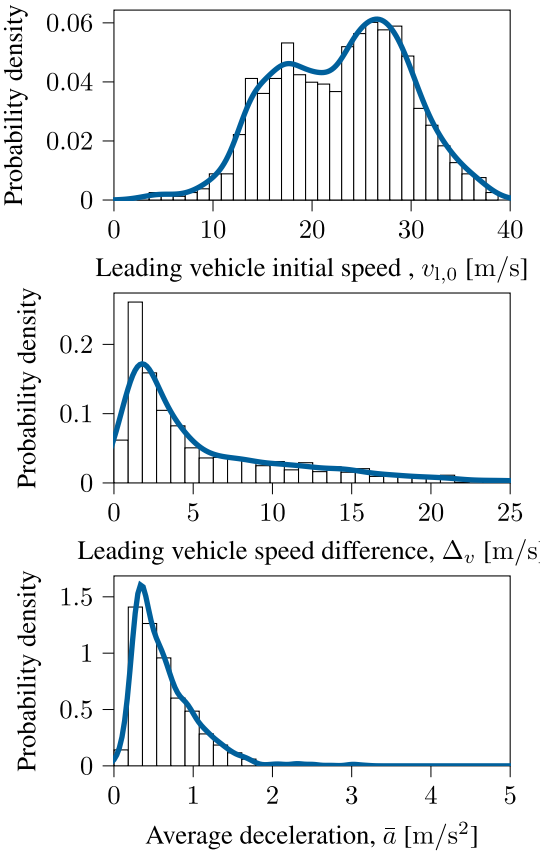


FIGURE 6. Three parameters of the LVD scenarios. The histograms show the original data and the solid lines represent the marginal probability distributions of the estimated pdf.

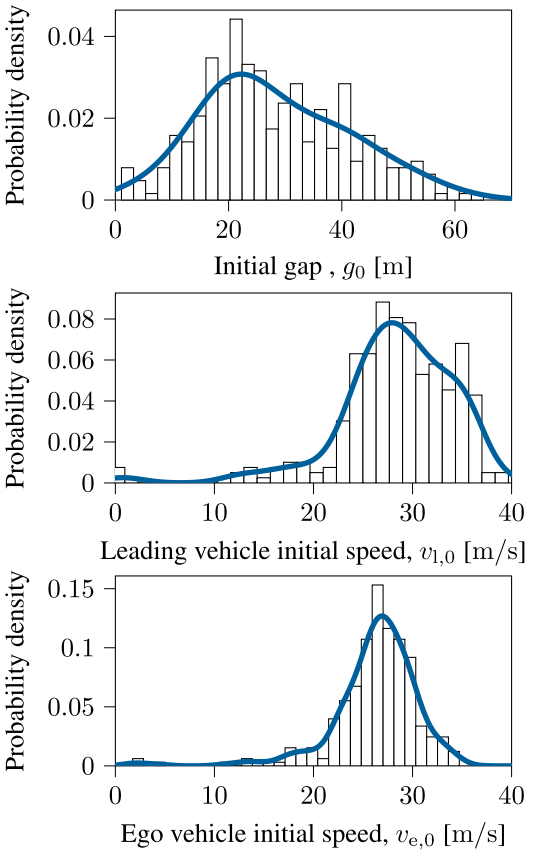


FIGURE 7. Three parameters of the cut-in scenarios. The histograms show the original data and the solid lines represent the marginal probability distributions of the estimated pdf.

For the cut-in scenario category, the controllability score is almost 1, indicating that it is unlikely that a human driver can avoid any potential harm. Although the severity and exposure are lower than for the LVD scenario category, the higher controllability score results in a higher overall risk ($\text{Risk}(\mathcal{C}_{\text{cut-in}}) = 1.75 \cdot 10^{-5}\,\text{h}^{-1}$).

Out of the 3 scenario categories, the ASV scenario category has the highest severity ($\text{Severity}(\mathcal{C}_{\text{ASV}}) = 2.12 \cdot 10^{-5}$). According to the simulations, however, the human driver is able to avoid any harm in almost all cases. Therefore, the controllability score is low ($\text{Controllability}(\mathcal{C}_{\text{ASV}}) = 7.83 \cdot 10^{-6}$) and the estimated risk is also low ($\text{Risk}(\mathcal{C}_{\text{ASV}}) = 7.68 \cdot 10^{-10}\,\text{h}^{-1}$). Note that for the ASV scenario category, σ_{injury} is almost equal to μ_{injury}, indicating that the relative uncertainty is high. This indicates that the actual risk could be an order of magnitude lower than the estimated risk.

C. TRIGGERING CONDITIONS

Table 5 also shows the results of the simulations that include a triggering condition. The risk in LVD scenarios is

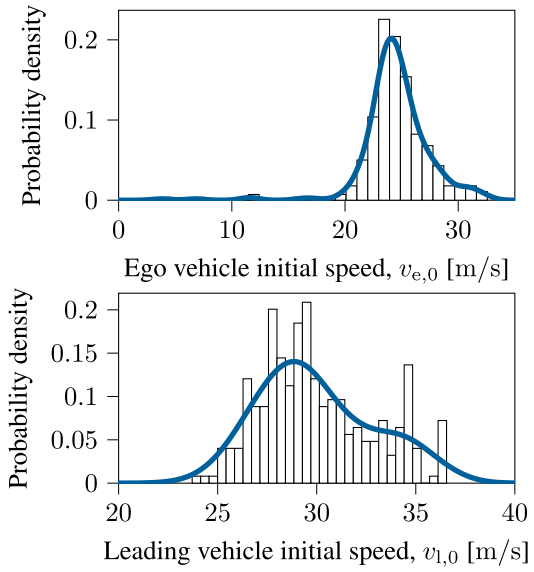


FIGURE 8. Three parameters of the ASV scenarios. The histograms show the original data and the solid lines represent the marginal probability distributions of the estimated pdf.

approximately 50 times higher when considering the triggering condition "limited braking capacity"; so this triggering condition has a significant impact on the safety during such scenarios. For cut-in scenarios, the risk is of the same order of magnitude when considering the triggering condition "limited braking capacity". The estimated risk in ASV scenarios is the highest when considering a limited braking capacity.

The triggering condition "poor visibility" does not influence the risk in LVD scenarios. This is because the leading vehicle is always within the viewing range of the driver, i.e., $s_0 + v_e\tau_h < d_{view}$. The risk in cut-in scenarios is also not influenced by this triggering condition. It may be possible that the vehicle cutting in is at a larger distance than d_{view}, but in potential risky scenarios that may result in harm, the vehicle cutting in is at a smaller distance than d_{view}. Hence, the vehicle cutting in is still in the limited viewing range of the human driver. For ASV scenarios, the risk is significantly higher when considering the triggering condition "poor visibility". The severity is the same because the limited viewing range of the human driver does not affect the ACC, but the controllability score is significantly higher. This indicates that it is less likely that the human driver will prevent harm in ASV scenarios when its view toward the leading vehicle is limited.

VII. DISCUSSION

Quantifying the risk in driving scenarios is an important component of the overall risk assessment of an ADS. Research question 1 addresses this by asking how to quantify the risk of an ADS. The proposed method in Section III answers this question. To answer Research question 2, we have also shown how the proposed method for risk quantification contributes to the risk assessment as proposed in the ISO 26262 [35] and ISO/DIS 21448 [36] standards by decomposing the risk into the components exposure, severity, and controllability. In this section, we provide further interpretations of the results and we discuss limitations of the presented research that are to be addressed in future research.

Whereas the ISO 26262 standard addresses functional safety and the ISO/DIS 21448 standard addressed SOTIF, for the safe deployment and operation of an ADS, cybersecurity [98] needs to be addressed as well. The cybersecurity ensures that an ADS is well protected against security attacks [99], [100]. This includes, e.g., mitigating privacy-related risks [101]–[107]. Note that cybersecurity is out-of-scope of the ISO 26262 and ISO/DIS 21448 standards and, thus, it is also out-of-scope of the current work. We refer to the ISO 21434 standard [108] for more information regarding cybersecurity of ADSs.

The proposed risk quantification serves two purposes. One purpose is to support the evaluation of whether it is safe to actually deploy an ADS in real-world traffic. In fact, the output of the proposed method, i.e., the expected number of injuries per hour of driving, can be compared with road crash statistics. Another purpose is to facilitate the design decisions during the development of an ADS. For example, based on the high controllability score of the ACC in ASV scenarios and poor visibility conditions (Table 5), it might be decided that a subsystem needs to be in place to detect poor visibility conditions such that the speed is reduced under these conditions. Also, the severity score might be lowered by adapting the control logic of the ADS.

The exposure of a scenario category is estimated by counting the number of occurrences of the corresponding scenarios and dividing this number by the number of hours of driving. As pointed out in Remark 2, it is not uncommon to assume that the rate of occurrence is constant such that the occurrence is Poisson-distributed. Looking at the results in Figure 5, however, the data do not support the assumption that the occurrence is Poisson-distributed. This suggests that the rate of occurrence is not constant and depends on other factors, e.g., the road layout, the environment (a cut-in is less likely on a rural road than on the highway), the driver, and the time of the day. If such factors are different during the deployment of the ADS, the estimated exposure needs to be reconsidered.

The presented method for risk quantification comes with limitations that are to be addressed in future research. While we advertised the use of data such that the risk estimation relies less on possibly subjective opinions from safety experts, it might be difficult to justify the adequacy of the data. First, it is important that we have enough data to accurately determine the pdf of the scenario parameters.[4] Second, the data need to match the Operational Design Domain (ODD) (see Table 2 for the definition) of the ADS. For example, the data that have been used in the case study were obtained from driving a specific route multiple times during daytime and good weather conditions. If the ODD

[4]To determine whether enough data have been collected to estimate the pdf accurately, the metric proposed in [65] can be used.

of the ADS considers the same route during daytime and good weather conditions, then the data are representing this ODD. If, however, the ODD is substantially different, extra arguments are needed to justify that the data still represent the ODD. The estimated exposure of the scenarios and the estimated parameter pdfs might not be accurate for the specified ODD in case the data have been recorded under different circumstances. As a result, the estimated risk might not be accurate enough.

Another difficulty involves accounting for the exposure of the triggering conditions. If the occurrence rate of the triggering condition can be assumed to be independent of the scenario category, the calculated risk of (20) — which does not consider the exposure of the triggering condition — can simply be multiplied with the estimated exposure of the triggering condition. It may be difficult, however, to justify this independence. More research is needed to investigate the possible triggering conditions, their occurrence rates, their relations with the occurrence of different scenarios, and the (possible) dependency between the scenario parameter values and the triggering conditions.

The proposed method for risk quantification employs simulations of the ADS response in driving scenarios. As a proof of concept, we have implemented simulations using simple models for the system-under-test and the driver behavior model. On the one hand, using these simple models contributes to the explainability of the results, ensures short simulation run times, and facilitates the reproducibility of the case study. On the other hand, the fidelity of the simulation results may be compromised by the simplicity of the simulations. When using the proposed method to assess the risk of deploying an ADS in the real world, evidence is needed to justify the fidelity of the simulation results. Therefore, the development of high-fidelity simulators for ADSs is an important research topic; see [27], [30] for an overview. More research is needed to actually verify the fidelity of such simulators. Note that it might be possible to combine virtual simulations with, e.g., hardware-in-the-loop tests, vehicle-in-the-loop tests, and proving ground tests [109]. For example, proving ground tests may be used to verify the virtual simulation results and virtual simulations are used to alleviate the required resources as no longer all tests have to be performed physically.

The proposed risk quantification is performed with regards to a given scenario category, possibly including one or more triggering conditions. To thoroughly review the risk of deploying an ADS in real-world traffic, many scenario categories and triggering conditions need to be considered. It remains an open question how many scenario categories and triggering conditions are to be considered. The 67 scenario categories described in [54] might be a good starting point, but it is expected that more scenario categories are needed for an ADS that aims for deployment in a complex ODD. The ISO/CD 34502 standard [110], currently under development, provides a process to determine the relevant scenario categories and triggering conditions for the safety validation of an ADS that is active on highways. Thus, once published, this standard may help to answer the question of which and how many scenario categories and triggering conditions are needed for a thorough risk assessment.

VIII. CONCLUSION

Validating the safety of an Automated Driving System (ADS) is essential to enable the safe deployment of an ADS. Part of the safety validation is the estimation of the risk of an ADS when dealing with real-world driving scenarios. The current work has presented a method to quantify this risk given a set of driving scenarios that are comprised by a scenario category. Since a data-driven approach has been proposed, the method relies less on the possibly subjective inputs from safety experts. Simulations are employed such that risks can be assessed prospectively, i.e., before real-life testing. The method supports the decomposition of the risk into the 3 aspects of risk, i.e., exposure, severity, and controllability, as mentioned by the leading standards in automotive safety, i.e., the ISO 26262 and ISO/DIS 21448 standards. Thus, the current article provides a method that can help engineers when designing an ADS in compliance with these standards and when evaluating the compliance of the ADS to these standards.

The risk quantification method has been illustrated by means of a case study. In the case study, the risk of a moderate injury or worse per hour of driving of an ADS is estimated for 3 scenario categories: leading vehicle decelerating (LVD), cut-in, and approaching slower vehicle (ASV). In addition to these scenarios, the scenarios have been complemented with so-called triggering conditions, such as a poor visibility conditions. The results have indicated, e.g., that the role of a human driver as a back-up is essential to reduce the risk in ASV scenarios and that, therefore, conditions that cause a poor visibility of the driver's environment have a significant impact on the risk.

Future work involves determining the statistics of the triggering conditions and the dependencies between triggering conditions and the different scenarios. Furthermore, verification methods for the fidelity of the simulation results is a topic of ongoing research. Finally, more research is needed to determine the relevant scenario categories that are to be considered for a full risk assessment of an ADS.

REFERENCES

[1] C.-Y. Chan, "Advancements, prospects, and impacts of automated driving systems," *Int. J. Transp. Sci. Technol.*, vol. 6, no. 3, pp. 208–216, 2017.

[2] I. Mahdinia, R. Arvin, A. J. Khattak, and A. Ghiasi, "Safety, energy, and emissions impacts of adaptive cruise control and cooperative adaptive cruise control," *Transp. Res. Rec., J. Transp. Res. Board*, vol. 2674, no. 6, pp. 253–267, Jun. 2020.

[3] A. Mammeri, G. Lu, and A. Boukerche, "Design of lane keeping assist system for autonomous vehicles," in *Proc. 7th Int. Conf. New Technol., Mobility Secur. (NTMS)*, Jul. 2015, pp. 1–5.

[4] "Taxonomy and definitions for terms related to driving automation systems for on-road motor vehicles," SAE Int., Warrendale, PA, USA, Tech. Rep. J3016, Apr. 2021.

[5] K. Bimbraw, "Autonomous cars: Past, present and future a review of the developments in the last century, the present scenario and the expected future of autonomous vehicle technology," in *Proc. 12th Int. Conf. Inform. Control, Automat. Robot. (ICINCO)*, vol. 1, Jul. 2015, pp. 191–198.

[6] D. Milakis, M. Snelder, B. van Arem, B. van Wee, and A. H. D. A. Correia, "Scenarios about development and implications of automated vehicles in The Netherlands," in *Proc. 95th Annu. Meeting Transp. Res. Board*, 2016, pp. 1–17.
[7] A. M. Madni, "Autonomous system-of-systems," in *Transdisciplinary Systems Engineering*. Cham, Switzerland: Springer, 2018, pp. 161–186.
[8] P. Liu, R. Yang, and Z. Xu, "How safe is safe enough for self-driving vehicles?" *Risk Anal.*, vol. 39, no. 2, pp. 315–325, Feb. 2019.
[9] N. Kalra and S. M. Paddock, "Driving to safety: How many miles of driving would it take to demonstrate autonomous vehicle reliability?" *Transp. Res. A, Policy Pract.*, vol. 94, pp. 182–193, Dec. 2016.
[10] C. Roesener, J. Sauerbie, A. Zlocki, F. Fahrenkrog, L. Wang, A. Várhelyi, E. de Gelder, J. Dufils, S. Breunig, P. Mejuto, and F. Tango, "A comprehensive evaluation approach for highly automated driving," in *Proc. 25th Int. Tech. Conf. Enhanced Saf. Veh.*, 2017, pp. 1–13.
[11] A. Pütz, A. Zlocki, J. Bock, and L. Eckstein, "System validation of highly automated vehicles with a database of relevant traffic scenarios," in *Proc. 12th ITS Eur. Congr.*, 2017, pp. 1–8.
[12] M. Althoff, M. Koschi, and S. Manzinger, "CommonRoad: Composable benchmarks for motion planning on roads," in *Proc. IEEE Intell. Veh. Symp. (IV)*, Jun. 2017, pp. 719–726.
[13] H. Elrofai, J.-P. Paardekooper, E. de Gelder, S. Kalisvaart, and O. Op den Camp, "Scenario-based safety validation of connected and automated driving," Netherlands Org. Appl. Sci. Res., TNO, The Hague, The Netherlands, Tech. Rep., 2018. [Online]. Available: http://publications.tno.nl/publication/34626550/AyT8Zc/TNO-2018-streetwise.pdf
[14] J. Ploeg, E. de Gelder, M. Slavík, E. Querner, T. Webster, and N. de Boer, "Scenario-based safety assessment framework for automated vehicles," in *Proc. 16th ITS Asia–Pacific Forum*, 2018, pp. 713–726.
[15] T. Menzel, G. Bagschik, and M. Maurer, "Scenarios for development, test and validation of automated vehicles," in *Proc. IEEE Intell. Veh. Symp. (IV)*, Jun. 2018, pp. 1821–1827.
[16] P. Nitsche, R. H. Welsh, A. Genser, and P. D. Thomas, "A novel, modular validation framework for collision avoidance of automated vehicles at road junctions," in *Proc. 21st Int. Conf. Intell. Transp. Syst. (ITSC)*, Nov. 2018, pp. 90–97.
[17] E. Thorn, S. Kimmel, and M. Chaka, "A framework for automated driving system testable cases and scenarios," Nat. Highway Traffic Saf. Admin., The Hague, The Netherlands, Tech. Rep. DOT HS 812623, 2018.
[18] J. Antona-Makoshi, U. Nobuyuki, Y. Kunio, O. Koichiro, K. Eiichi, and T. Satoshi, "Development of a safety assurance process for autonomous vehicles in Japan," in *Proc. 26th Int. Tech. Conf. Enhanced Saf. Veh. (ESV)*, 2019, pp. 1–18.
[19] R. Myers and Z. Saigol, "Pass-fail criteria for scenario-based testing of automated driving systems," 2020, *arXiv:2005.09417*.
[20] S. Riedmaier, T. Ponn, D. Ludwig, B. Schick, and F. Diermeyer, "Survey on scenario-based safety assessment of automated vehicles," *IEEE Access*, vol. 8, pp. 87456–87477, 2020.
[21] M. Scholtes, L. Westhofen, L. R. Turner, K. Lotto, M. Schuldes, H. Weber, N. Wagener, C. Neurohr, M. H. Bollmann, F. Kortke, J. Hiller, M. Hoss, J. Bock, and L. Eckstein, "6-layer model for a structured description and categorization of urban traffic and environment," *IEEE Access*, vol. 9, pp. 59131–59147, 2021.
[22] C. Neurohr, L. Westhofen, M. Butz, M. H. Bollmann, U. Eberle, and R. Galbas, "Criticality analysis for the verification and validation of automated vehicles," *IEEE Access*, vol. 9, pp. 18016–18041, 2021.
[23] S. Ulbrich, T. Menzel, A. Reschka, F. Schuldt, and M. Maurer, "Defining and substantiating the terms scene, situation, and scenario for automated driving," in *Proc. IEEE 18th Int. Conf. Intell. Transp. Syst.*, Sep. 2015, pp. 982–988.
[24] E. de Gelder, J.-P. Paardekooper, A. Khabbaz Saberi, H. Elrofai, O. Op den Camp, S. Kraines, J. Ploeg, and B. De Schutter, "Towards an ontology for scenario definition for the assessment of automated vehicles: An object-oriented framework," 2021, *arXiv:2001.11507*.
[25] J. Kapinski, J. V. Deshmukh, X. Jin, H. Ito, and K. Butts, "Simulation-based approaches for verification of embedded control systems: An overview of traditional and advanced modeling, testing, and verification techniques," *IEEE Control Syst.*, vol. 36, no. 6, pp. 45–64, Dec. 2016.
[26] A. Dosovitskiy, G. Ros, F. Codevilla, A. Lopez, and V. Koltun, "CARLA: An open urban driving simulator," in *Proc. Conf. Robot Learn.*, 2017, pp. 1–16.
[27] 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, p. 648, 2019.
[28] P. Wimmer, M. Düring, H. Chajmowicz, F. Granum, J. King, H. Kolk, O. Op den Camp, P. Scognamiglio, and M. Wagner, "Toward harmonizing prospective effectiveness assessment for road safety: Comparing tools in standard test case simulations," *Traffic Injury Prevention*, vol. 20, no. 1, pp. S139–S145, Jun. 2019.
[29] G. Rong, B. H. Shin, H. Tabatabaee, Q. Lu, S. Lemke, M. Mozeiko, E. Boise, G. Uhm, M. Gerow, S. Mehta, E. Agafonov, T. H. Kim, E. Sterner, K. Ushiroda, M. Reyes, D. Zelenkovsky, and S. Kim, "LGSVL simulator: A high fidelity simulator for autonomous driving," in *Proc. IEEE 23rd Int. Conf. Intell. Transp. Syst. (ITSC)*, Sep. 2020, pp. 1–6.
[30] P. Kaur, S. Taghavi, Z. Tian, and W. Shi, "A survey on simulators for testing self-driving cars," 2021, *arXiv:2101.05337*.
[31] F. Schuldt, A. Reschka, and M. Maurer, "A method for an efficient, systematic test case generation for advanced driver assistance systems in virtual environments," in *Automotive Systems Engineering II*. Cham, Switzerland: Springer, 2018, pp. 147–175.
[32] S. Feng, Y. Feng, C. Yu, Y. Zhang, and H. X. Liu, "Testing scenario library generation for connected and automated vehicles—Part I: Methodology," *IEEE Trans. Intell. Transp. Syst.*, vol. 22, no. 3, pp. 1573–1582, Mar. 2021.
[33] J. Spooner, V. Palade, M. Cheah, S. Kanarachos, and A. Daneshkhah, "Generation of pedestrian crossing scenarios using ped-cross generative adversarial network," *Appl. Sci.*, vol. 11, no. 2, p. 471, Jan. 2021.
[34] E. de Gelder, J. Hof, E. Cator, J.-P. Paardekooper, O. Op den Camp, J. Ploeg, and B. De Schutter, "Scenario parameter generation method and scenario representativeness metric for scenario-based assessment of automated vehicles," 2021.
[35] *Road Vehicles—Functional Safety*, Standard ISO 26262, International Organization for Standardization, 2018.
[36] *Road Vehicles—Safety of the Intended Functionality*, Standard ISO/DIS 21448, International Organization for Standardization, 2019.
[37] C. M. Hruschka, D. Topfer, and S. Zug, "Risk assessment for integral safety in automated driving," in *Proc. 2nd Int. Conf. Intell. Auton. Syst. (ICoIAS)*, Feb. 2019, pp. 102–109.
[38] R. V. Cowlagi, R. C. Debski, and A. M. Wyglinski, "Risk quantification for automated driving using information from V2V basic safety messages," in *Proc. IEEE 93rd Veh. Technol. Conf. (VTC-Spring)*, Apr. 2021, pp. 1–5.
[39] W. M. D. Chia, S. L. Keoh, A. L. Michala, and C. Goh, "Real-time recursive risk assessment framework for autonomous vehicle operations," in *Proc. IEEE 93rd Veh. Technol. Conf. (VTC-Spring)*, Apr. 2021, pp. 1–7.
[40] G. Li, Y. Yang, T. Zhang, X. Qu, D. Cao, B. Cheng, and K. Li, "Risk assessment based collision avoidance decision-making for autonomous vehicles in multi-scenarios," *Transp. Res. C, Emerg. Technol.*, vol. 122, Jan. 2021, Art. no. 102820.
[41] G. Li, Y. Yang, S. Li, X. Qu, N. Lyu, and S. E. Li, "Decision making of autonomous vehicles in lane change scenarios: Deep reinforcement learning approaches with risk awareness," *Transp. Res. C, Emerg. Technol.*, Nov. 2021, Art. no. 103452.
[42] B. Kramer, C. Neurohr, M. Büker, E. Böde, M. Fränzle, and W. Damm, "Identification and quantification of hazardous scenarios for automated driving," in *Proc. Int. Symp. Model-Based Saf. Assessment*, 2020, pp. 163–178.
[43] S. Wagner, K. Groh, T. Kuhbeck, M. Dorfel, and A. Knoll, "Using time-to-react based on naturalistic traffic object behavior for scenario-based risk assessment of automated driving," in *Proc. IEEE Intell. Veh. Symp. (IV)*, Jun. 2018, pp. 1521–1528.
[44] F. Warg, M. Skoglund, A. Thorsen, R. Johansson, M. Brannstrom, M. Gyllenhammar, and M. Sanfridson, "The quantitative risk norm—A proposed tailoring of HARA for ADS," in *Proc. 50th Annu. IEEE/IFIP Int. Conf. Dependable Syst. Netw. Workshops (DSN-W)*, Jun. 2020, pp. 86–93.
[45] C. Bergenhem, M. Majdandzic, and S. Ursing, "Concepts and risk analysis for a cooperative and automated highway platooning system," in *Proc. Eur. Dependable Comput. Conf.*, 2020, pp. 200–213.
[46] S. Teuchert, "ISO 26262—Blessing or curse?" *ATZelektronik worldwide*, vol. 7, no. 6, pp. 4–9, Oct. 2012.
[47] S. Khastgir, S. Birrell, G. Dhadyalla, H. Sivencrona, and P. Jennings, "Towards increased reliability by objectification of hazard analysis and risk assessment (HARA) of automated automotive systems," *Saf. Sci.*, vol. 99, pp. 166–177, Nov. 2017.

[48] A. Abdulkhaleq, D. Lammering, S. Wagner, J. Röder, N. Balbierer, L. Ramsauer, T. Raste, and H. Boehmert, "A systematic approach based on STPA for developing a dependable architecture for fully automated driving vehicles," *Proc. Eng.*, vol. 179, pp. 41–51, Jan. 2017.
[49] D. Tokody, I. J. Mezei, and G. Schuster, "An overview of autonomous intelligent vehicle systems," in *Vehicle and Automotive Engineering*, K. Jármai and B. Bolló, Eds. Cham, Switzerland: Springer, 2017, pp. 287–307.
[50] *Self-Driving Safety Report*, General Motors, Detroit, MI, USA, 2018.
[51] *The New Era of Mobility*, Aurora, Pittsburgh, PA, USA, 2019.
[52] *Waymo Safety Report*, Waymo, Mountain View, CA, USA, 2021.
[53] M. Gyllenhammar, R. Johansson, F. Warg, D. Chen, H.-M. Heyn, M. Sanfridson, J. Söderberg, A. Thorsén, and S. Ursing, "Towards an operational design domain that supports the safety argumentation of an automated driving system," in *Proc. 10th Eur. Congr. Embedded Real Time Syst. (ERTS)*, 2020, pp. 1–10.
[54] E. de Gelder, O. Op den Camp, and N. de Boer, "Scenario categories for the assessment of automated vehicles," CETRAN, Singapore, Version 1.7, Tech. Rep., 2020. [Online]. Available: http://cetran.sg/wp-content/uploads/2020/01/REP200121_Scenario_Categories_v1.7.pdf
[55] E. de Gelder, J. Manders, C. Grappiolo, J.-P. Paardekooper, O. Op den Camp, and B. De Schutter, "Real-world scenario mining for the assessment of automated vehicles," in *Proc. IEEE 23rd Int. Conf. Intell. Transp. Syst. (ITSC)*, Sep. 2020, pp. 1073–1080.
[56] R. Krajewski, J. Bock, L. Kloeker, and L. Eckstein, "The highD dataset: A drone dataset of naturalistic vehicle trajectories on German highways for validation of highly automated driving systems," in *Proc. 21st Int. Conf. Intell. Transp. Syst. (ITSC)*, Nov. 2018, pp. 2118–2125.
[57] D. Zhao, X. Huang, H. Peng, H. Lam, and D. J. LeBlanc, "Accelerated evaluation of automated vehicles in car-following maneuvers," *IEEE Trans. Intell. Transp. Syst.*, vol. 19, no. 3, pp. 733–744, Mar. 2018.
[58] S. Jesenski, N. Tiemann, J. E. Stellet, and J. M. Zollner, "Scalable generation of statistical evidence for the safety of automated vehicles by the use of importance sampling," in *Proc. IEEE 23rd Int. Conf. Intell. Transp. Syst. (ITSC)*, Sep. 2020, pp. 1–8.
[59] E. de Gelder and J.-P. Paardekooper, "Assessment of automated driving systems using real-life scenarios," in *Proc. IEEE Intell. Veh. Symp. (IV)*, Jun. 2017, pp. 589–594.
[60] Y. Xu, Y. Zou, and J. Sun, "Accelerated testing for automated vehicles safety evaluation in cut-in scenarios based on importance sampling, genetic algorithm and simulation applications," *J. Intell. Connected Veh.*, vol. 1, no. 1, pp. 28–38, Oct. 2018.
[61] M. Rosenblatt, "Remarks on some nonparametric estimates of a density function," *Ann. Math. Statist.*, vol. 27, no. 3, pp. 832–837, 1956.
[62] E. Parzen, "On estimation of a probability density function and mode," *Ann. Math. Statist.*, vol. 33, no. 3, pp. 1065–1076, Sep. 1962.
[63] A. B. Turlach, "Bandwidth selection in kernel density estimation: A review," Institut für Statistik und Ökonometrie, Humboldt-Universität zu Berlin, Berin, Germany, Tech. Rep., 1993. [Online]. Available: http://citeseerx.ist.psu.edu/viewdoc/download;jsessionid=74BAC9974E24193267F74312E1D2969A?doi=10.1.1.44.6770&rep=rep1&type=pdf
[64] T. Duong, "KS: Kernel density estimation and kernel discriminant analysis for multivariate data in R," *J. Stat. Softw.*, vol. 21, no. 7, pp. 1–16, 2007.
[65] E. de Gelder, J.-P. Paardekooper, O. Op den Camp, and B. De Schutter, "Safety assessment of automated vehicles: How to determine whether we have collected enough field data?" *Traffic Injury Prevention*, vol. 20, no. S1, pp. S162–S170, Jun. 2019.
[66] E. de Gelder, E. Cator, J.-P. Paardekooper, O. Op den Camp, and B. De Schutter, "Constrained sampling from a kernel density estimator to generate scenarios for the assessment of automated vehicles," in *Proc. IEEE Intell. Veh. Symp. Workshops (IV Workshop)*, Jul. 2021, pp. 203–208.
[67] B. W. Silverman, *Density Estimation for Statistics and Data Analysis*. Boca Raton, FL, USA: CRC Press, 1986.
[68] D. M. Bashtannyk and R. J. Hyndman, "Bandwidth selection for kernel conditional density estimation," *Comput. Statist. Data Anal.*, vol. 36, no. 3, pp. 279–298, May 2001.
[69] M. C. Jones, J. S. Marron, and S. J. Sheather, "A brief survey of bandwidth selection for density estimation," *J. Amer. Stat. Assoc.*, vol. 91, no. 433, pp. 401–407, 1996.
[70] A. Gramacki and J. Gramacki, "FFT-based fast bandwidth selector for multivariate kernel density estimation," *Comput. Statist. Data Anal.*, vol. 106, pp. 27–45, Feb. 2017.
[71] P. Zhang, "Nonparametric importance sampling," *J. Amer. Stat. Assoc.*, vol. 91, no. 435, pp. 1245–1253, Sep. 1996.
[72] A. B. Owen, *Monte–Carlo Theory, Methods Examples*, 2013.
[73] J. C. Hayward, "Near miss determination through use of a scale of danger," Pennsylvania State Univ., State College, PA, USA, Tech. Rep. TTSC-7115, 1972.
[74] E. Petrucelli, J. D. States, and L. N. Hames, "The abbreviated injury scale: Evolution, usage and future adaptability," *Accident Anal. Prevention*, vol. 13, no. 1, pp. 29–35, Mar. 1981.
[75] J. C. Brewer, "Effects of angles and offsets in crash simulations of automobiles with light trucks," in *Proc. 17th Int. Tech. Conf. Enhances Saf. Veh.*, 2001, pp. 1–8.
[76] K. H. Yang, J. Hu, N. A. White, A. I. King, C. C. Chou, and A. Prasad, "Development of numerical models for injury biomechanics research: A review of 50 years of publications in the stapp car crash conference," *Stapp Car Crash J.*, vol. 50, pp. 429–490, Nov. 2006.
[77] I. A. Rafukka, B. B. Sahari, A. A. Nuraini, and A. Manohar, "Child dummy finite element models development: A review," *ARPN J. Eng. Appl. Sci.*, vol. 11, no. 10, pp. 6649–6656, 2006.
[78] Y. Peng, Y. Chen, J. Yang, D. Otte, and R. Willinger, "A study of pedestrian and bicyclist exposure to head injury in passenger car collisions based on accident data and simulations," *Saf. Sci.*, vol. 50, no. 9, pp. 1749–1759, Nov. 2012.
[79] F. Carollo, G. Virzi-Mariotti, and E. Scalici, "Injury evaluation in teenage cyclist-vehicle crash by multibody simulation," *WSEAS Trans. Biol. Biomed.*, vol. 11, no. 11, pp. 203–217, 2014.
[80] C. Jurewicz, A. Sobhani, J. Woolley, J. Dutschke, and B. Corben, "Exploration of vehicle impact speed–injury severity relationships for application in safer road design," *Transp. Res. Proc.*, vol. 14, no. 2352-1465, pp. 4247–4256, 2016.
[81] S. Yoshida, T. Hasegawa, S. Tominaga, and T. Nishimoto, "Development of injury prediction models for advanced automatic collision notification based on Japanese accident data," *Int. J. Crashworthiness*, vol. 21, no. 2, pp. 112–119, Mar. 2016.
[82] D. J. Gabauer and H. C. Gabler, "Comparison of delta-v and occupant impact velocity crash severity metrics using event data recorders," *Assoc. Advancement Automot. Med.*, vol. 50, pp. 57–71, Oct. 2006.
[83] K. D. Kusano and H. C. Gabler, "Potential occupant injury reduction in pre-crash system equipped vehicles in the striking vehicle of rear-end crashes," in *Proc. Ann. Adv. Automot. Med./Annu. Sci. Conf.*, vol. 54, 2010, p. 203.
[84] J. Augenstein, E. Perdeck, J. Stratton, K. Digges, G. Bahouth, N. Borchers, and P. Baur, "Methodology for the development and validation of injury predicting algorithms," in *Proc. 18th ESV Conf.*, vol. 467, 2003, pp. 1–13.
[85] D. J. Gabauer, "Predicting occupant injury with vehicle-based injury criteria roadside crashes," Ph.D. dissertation, Virginia Tech, Blacksburg, VA, USA, 2008.
[86] D. J. Gabauer and H. C. Gabler, "Acceleration-based occupant injury criteria for predicting injury in real-world crashes," *Biomed. Sci. Instrum.*, vol. 44, pp. 213–218, Jan. 2008.
[87] L. Xiao, M. Wang, and B. van Arem, "Realistic car-following models for microscopic simulation of adaptive and cooperative adaptive cruise control vehicles," *Transp. Res. Rec.*, vol. 2623, no. 1, pp. 1–9, Jan. 2017.
[88] V. Milanés and S. E. Shladover, "Modeling cooperative and autonomous adaptive cruise control dynamic responses using experimental data," *Transp. Res. C, Emerg. Technol.*, vol. 48, pp. 285–300, Nov. 2014.
[89] R. J. Kiefer, D. J. LeBlanc, and C. A. Flannagan, "Developing an inverse time-to-collision crash alert timing approach based on drivers' last-second braking and steering judgments," *Accident Anal. Prevention*, vol. 37, no. 2, pp. 295–303, Mar. 2005.
[90] Marc Green, "How long does it take to stop? Methodological analysis of driver perception-brake times," *Transp. Hum. Factors*, vol. 2, no. 3, pp. 195–216, 2000.
[91] M. Treiber, A. Hennecke, and D. Helbing, "Congested traffic states in empirical observations and microscopic simulations," *Phys. Rev. E, Stat. Phys. Plasmas Fluids Relat. Interdiscip. Top.*, vol. 62, no. 2, pp. 1805–1824, 2000.
[92] W. J. Schakel, B. van Arem, and B. D. Netten, "Effects of cooperative adaptive cruise control on traffic flow stability," in *Proc. 13th Int. IEEE Conf. Intell. Transp. Syst.*, Sep. 2010, pp. 759–764.
[93] *Proposal for a New UN Regulation on Uniform Provisions Concerning the Approval of Vehicles With Regards to Automated Lane Keeping System*, Standard ECE/TRANS/WP.29/2020/81, World Forum for Harmonization of Vehicle Regulations, 2020.

[94] G. W. Najm, D. J. Smith, and M. Yanagisawa, "Pre-crash scenario typology for crash avoidance research," U.S. Dept. Transp. Res. Innov. Technol. Admin., Tech. Rep. DOT HS 810 767, Apr. 2007.
[95] J.-P. Paardekooper, S. Montfort, J. Manders, J. Goos, E. de Gelder, O. Op den Camp, A. Bracquemond, and G. Thiolon, "Automatic identification of critical scenarios in a public dataset of 6000 km of public-road driving," in *Proc. 26th Int. Tech. Conf. Enhanced Saf. Veh. (ESV)*, 2019, pp. 1–8.
[96] J. Elfring, R. Appeldoorn, S. van den Dries, and M. Kwakkernaat, "Effective world modeling: Multisensor data fusion methodology for automated driving," *Sensors*, vol. 16, no. 10, pp. 1–27, 2016.
[97] K. D. Kusano and H. C. Gabler, "Safety benefits of forward collision warning, brake assist, and autonomous braking systems in rear-end collisions," *IEEE Trans. Intell. Transp. Syst.*, vol. 13, no. 4, pp. 1546–1555, Dec. 2012.
[98] D. Craigen, N. Diakun-Thibault, and R. Purse, "Defining cybersecurity," *Technol. Innov. Manage. Rev.*, vol. 4, no. 10, pp. 13–21, Oct. 2014.
[99] F. Sakiz and S. Sen, "A survey of attacks and detection mechanisms on intelligent transportation systems: VANETs and IoV," *Ad Hoc Netw.*, vol. 61, pp. 33–50, Jun. 2017.
[100] J. Cui, L. S. Liew, G. Sabaliauskaite, and F. Zhou, "A review on safety failures, security attacks, and available countermeasures for autonomous vehicles," *Ad Hoc Netw.*, vol. 90, Jul. 2019, Art. no. 101823.
[101] J. Zhou, Z. Cao, X. Dong, and A. V. Vasilakos, "Security and privacy for cloud-based IoT: Challenges," *IEEE Commun. Mag.*, vol. 55, no. 1, pp. 26–33, Jan. 2017.
[102] J. Liu, J. Yu, and S. Shen, "Energy-efficient two-layer cooperative defense scheme to secure sensor-clouds," *IEEE Trans. Inf. Forensics Security*, vol. 13, no. 2, pp. 408–420, Feb. 2018.
[103] M. A. Khan and K. Salah, "IoT security: Review, blockchain solutions, and open challenges," *Future Gener. Comput. Syst.*, vol. 82, pp. 395–411, May 2018.
[104] T. Wang, M. Z. A. Bhuiyan, G. Wang, L. Qi, J. Wu, and T. Hayajneh, "Preserving balance between privacy and data integrity in edge-assisted Internet of Things," *IEEE Internet Things J.*, vol. 7, no. 4, pp. 2679–2689, Apr. 2020.
[105] Z. Wu, R. Wang, Q. Li, X. Lian, G. Xu, E. Chen, and X. Liu, "A location privacy-preserving system based on query range cover-up or location-based services," *IEEE Trans. Veh. Technol.*, vol. 69, no. 5, pp. 5244–5254, May 2020.
[106] Z. Wu, S. Shen, X. Lian, X. Su, and E. Chen, "A dummy-based user privacy protection approach for text information retrieval," *Knowl.-Based Syst.*, vol. 195, May 2020, Art. no. 105679.
[107] Z. Wu, G. Li, S. Shen, X. Lian, E. Chen, and G. Xu, "Constructing dummy query sequences to protect location privacy and query privacy in location-based services," *World Wide Web*, vol. 24, no. 1, pp. 25–49, Jan. 2021.
[108] *Road Vehicles—Cybersecurity Engineering*, Standard ISO 21434, International Organization for Standardization, 2021.
[109] S. Riedmaier, J. Nesensohn, C. Gutenkunst, T. Düser, B. Schick, and H. Abdellatif, "Validation of x-in-the-loop approaches for virtual homologation of automated driving functions," in *Proc. 11th Graz Symp. Virtual Veh.*, 2018, pp. 1–12.
[110] *Road Vehicles—Engineering Framework and Process of Scenario-Based Safety Evaluation*, Standard ISO/CD 34502, Standard at Committee Draft (CD) Stage, International Organization for Standardization, 2020.

ERWIN DE GELDER received the M.Sc. degree *(cum laude)* in systems and control from the Delft University of Technology, Delft, The Netherlands, in 2014, where he is currently pursuing the Ph.D. degree. Since 2014, he has been with the Netherlands Organization for Applied Scientific Research (TNO). From 2017 to 2019, he worked with Nanyang Technological University, Singapore, to apply his research for an assessment procedure for automated vehicles in Singapore. His current research interest includes quantitative assessment methodology for automated vehicles using scenarios obtained from real-world data.

HALA ELROFAI received the M.Sc. degree in applied mathematics from the University of Twente, in 2003, and the Ph.D. degree in applied mathematics from the VU Amsterdam, in 2008. Following her Ph.D., she worked as a Postdoctoral Fellow at the Eindhoven University of Technology. After that, she joined the Research and Development Department, ASML. She currently works as a Senior Research Scientist at TNO. Her research interests include artificial intelligence and data analytics for development of safety assessment methods and perception and decision-making models for automated driving systems.

ARASH KHABBAZ SABERI received the M.Sc. degree in embedded systems and the Professional Doctorate in Engineering (P.D.Eng.) degree in automotive system design from the Eindhoven University of Technology (TU/e), in 2013 and 2015, respectively. Finally, in 2020, he defended his Ph.D. thesis on application of model-based software engineering for functional safety of automated driving. He is currently the Product Safety Engineering Manager at TomTom responsible for product safety of the entire TomTom product portfolio.

JAN-PIETER PAARDEKOOPER received the M.Sc. degree in astrophysics and the Ph.D. degree in theoretical astrophysics from Leiden University, Leiden, The Netherlands, in 2006 and 2010, respectively. He is currently a Researcher at TNO, The Netherlands, on the topic of artificial intelligence (AI) in connected and automated vehicles. Since 2018, he has been a part-time Research Fellow with the Department of Artificial Intelligence, Donders Institute for Brain, Cognition and Behaviour, Radboud University, Nijmegen, The Netherlands. His research interests include AI safety, neuro-symbolic AI, and situation awareness for AI.

OLAF OP DEN CAMP received the Engineering degree *(cum laude)* in mechanical engineering and the Engineering degree *(cum laude)* in biomedical engineering from the Eindhoven University of Technology, and the Ph.D. degree with a study into the identification of material parameters of biological tissue, in 1996. In 1995, he joined the Netherlands Organization for Applied Scientific Research (TNO). Since 2018, he has been working as a part-time at Nanyang Technological University, Singapore, to develop and implement a scenario-based safety assessment framework to support safe deployment of autonomous vehicles within the CETRAN program. He is a Senior Consultant and a Technical Lead of the StreetWise scenario database development.

BART DE SCHUTTER (Fellow, IEEE) received the Ph.D. degree *(summa cum laude with congratulations of the examination jury)* in applied sciences from the KU Leuven, Belgium, in 1996.

He is currently a Full Professor and the Head of the Department, Delft Center for Systems and Control, Delft University of Technology, Delft, The Netherlands. His current research interests include multi-level and multi-agent control, learning-based control, and control of hybrid systems with applications in intelligent transportation systems and smart energy systems. He is a Senior Editor of the IEEE TRANSACTIONS ON INTELLIGENT TRANSPORTATION SYSTEMS and an Associate Editor of the IEEE TRANSACTIONS ON AUTOMATIC CONTROL.