Preprint version of:
Gyllenhammar et al. "Efficient Injury Risk Assessment for Automated Driving Systems Using Subset Simulation" in B. Gallina et al. (Eds.): SAFECOMP 2025, LNCS 15954, 2025.
https://doi.org/10.1007/978-3-032-01241-8_6

Efficient Injury Risk Assessment for Automated Driving Systems Using Subset Simulation

Magnus Gyllenhammar[1,2] [0000-0001-9020-6501], Daniel Åsljung[1] [0000-0003-1670-6078], Majid Khorsand Vakilzadeh[1] [0000-0003-0019-7547], and Gabriel Rodrigues de Campos[1] [0000-0002-8218-6915]

[1] Zenseact, Lindholmspiren 2, 417 56 Gothenburg, Sweden
magnus.gyllenhammar@zenseact.com
[2] KTH Royal Institute of Technology

Abstract. Assessing safety and reliability of Automated Driving Systems (ADSs) is a crucial activity before deployment on public roads. However, determining the absence of risk is more than estimating the frequency of collisions, it also entails understanding the likelihood of collisions with different injury levels. In this paper we present a new method adapting a well-known approach for estimating probability of rare events; namely, Subset Simulation (SuS). Our adaptation leverages all generated samples and enables simultaneous estimation of the probability of multiple injury levels and not just the collision rate. A new composite metric, incorporating both the brake threat number as well as the collision velocity of an eventual collision, is used to guide SuS toward high injury risk scenarios. The usefulness of the proposed method is demonstrated through simulations on a degraded Automated Driving (AD) functionality – emulated in the form of an Adaptive Cruise Control (ACC) function with emergency braking capabilities and limited sensing horizon – exposed to cut-in scenarios. For benchmark purposes, we consider a Monte-Carlo reference and the results show that the proposed method finds comparable injury probabilities with respect to the reference, but with more than three orders of magnitude fewer samples, therefore representing a very efficient alternative.

Keywords: Injury probability · Safety · Subset simulation · Injury levels · Safety evaluation · Automated Driving System.

1 Introduction

Automated Driving Systems (ADSs), aimed to operate in open and uncertain environments, require a thorough qualitative and quantitative assessment of the system's safety before deployment on public roads. From a quantitative perspective, such an assessment involves estimating both the probability of different outcomes, as well as the associated injury level(s) (i.e. severity). Methods for (virtual) scenario-based testing have been introduced to provide efficient assessment as part of the Verification and Validation (V&V) strategies for ADSs [8,11], but these methods struggle with exploring the "infinitely many characteristics" of the scenario-space [14], and especially to find rare cases with severe consequences. While methods for critical scenario identification can be used to guide the assessment towards critical regions [18], many of these methods do not provide quantitative estimates of the failure probability. Furthermore, methods that do, such as importance sampling [19] or Subset Simulation (SuS) [21,17,20], only provide quantitative estimates for one specific failure case, such as, e.g., collisions.

Therefore, in this paper we propose a new method adapting SuS [20] by leveraging a new composite metric that captures the proximity to an accident as well as the eventual collision velocity, which enables an efficient exploration beyond mere collisions. Furthermore, by using continuous injury risk functions [9], we elicit accurate estimates of the injury risk for each of the generated samples and also propose a method for simultaneously estimating the probability of multiple injury levels from one single run of SuS.

The proposed system can be used for three distinct purposes:

1. To efficiently (time- and cost-wise) provide virtual assessment of a specific (ADS) feature (as part of the V&V) enabling falsification of safety requirements spanning several injury levels;
2. To enable efficient assessment and falsification of different system configurations or failure modes – thereby understanding the (safety) implications of different system faults; and
3. To derive more detailed situation awareness [16] and subsequently support the derivation of, e.g., a precautionary safety driving policy [2,12], as described in [4].

The usefulness and merits of the proposed method are demonstrated through a simulation-based case study of a degraded Automated Driving (AD) functionality – in the form of an Adaptive Cruise Control (ACC) function with emergency braking capabilities with limited sensing horizon – faced with cut-in scenarios drawn from a 15-dimensional statistical model created from 12 443 recorded cut-in scenarios. We show that the proposed method is able to estimate multiple injury probabilities from one single run of SuS. Furthermore, we compare the results to a Monte-Carlo (MC) based sampling approach – and show that the proposed method achieves similar estimates of the injury probabilities with more than three orders of magnitude fewer samples, highlighting the efficiency properties of the proposed method.

The contributions of the paper can be summarised as follows:

- A **method**, leveraging SuS to simultaneously and efficiently **determine the probabilities of multiple injury levels**. The method includes:
  - a **composite metric** for SuS that incorporates i) the Brake Threat Number (BTN) and ii) the collision velocity;
  - a **procedure** to determine the probabilities of the individual injury levels from the generated chain of samples from SuS; and
  - the procedure further **incorporating continuous injury risk models**, i.e. risk curves.
- A **case study** of the approach, evaluating an AD function based on data from 12 443 recorded real cut-in scenarios.

The paper is organised as follows. In Sec. 2 the steps used to adapt SuS for estimating multiple injury probabilities from continuous risk curves is described. The details of the particular case study investigated, are given in Sec. 3 and the results are presented in Sec. 4. A discussion is provided in Sec. 5 and conclusions are given in Sec. 6.

2 Subset Simulation for Injury Risk Estimation

To elicit the quantitative likelihood of an injury, given a specific scenario type, we require not only a method for finding critical scenarios but also need to ground this in the probability of the scenario occurring in the first place. Furthermore, since the final risk of an injury might be stochastic – even given a specific outcome in terms of scenario parameters – there is a need to understand the transfer function from a scenario into the different injury levels.

In this section, we give an overview of our proposed method that provides quantitative estimates of multiple injury probabilities simultaneously. First, the transfer function, i.e. the function for estimating the injury risk probabilities of a collision, is elaborated on. Second, we describe SuS and its ability to explore critical scenarios and ultimately provide an estimate for the probability of failure – or, as in our case, the probability of multiple injury levels. Third, we present our new composite metric for evaluating the function in order to allow SuS to find not only collisions but also collisions with high collision velocities. Lastly, we present the algorithm for determining the probability of multiple injury levels simultaneously from SuS, using continuous injury risk functions.

2.1 Injury Risk Probabilities

There are a lot of factors that can impact the resulting injury level of a collision but only some of those factors are measurable or observable during (virtual) evaluation of, e.g., an ADS. This means that for any given collision there is a probability of different injury levels occurring. [9] presents a statistical model that is able to predict the injury risk for all foreseeable planar collisions. In practice, it entails one function for the probability of a collision resulting in a maximum abbreviated injury scale (MAIS) 3 and higher (3+), and one function for the probability of MAIS5+. A MAIS3+ corresponds to a maximum injury that is *serious or worse*, whereas MAIS5+ denotes one that is *critical or worse*. The Abbreviated Injury Scale (AIS), created for, and broadly used within, the automotive domain, is an anatomical-based coding system to classify and describe the severity of injuries. We use these injury risk functions to estimate the probability of each injury level for a collision scenario.

Without loss of generality, for the purpose of this paper, we limit ourselves to estimating the probability of MAIS3+ and MAIS5+ injuries. However, this is a quite reasonable choice, which is also aligned with the severity classifications of, e.g., ISO 26262 [7]. Furthermore, we assume that the principal direction of force for all collisions is aligned with the direction of travel. This means that we set the collision angle to zero and, as suggested by the statistical models of [9], this entails that increased collision velocities result in an increased probability of injury risk following an "S-curve". With the curve being shifted to higher velocities for MAIS5+ when compared to MAIS3+.

2.2 Subset Simulation

SuS is proposed to selectively explore intermediate regions of the search space to progressively move closer to the failure region with each step, also denoted level [20].

Through a sequence of subsets, that the levels of SuS correspond to, it is possible to estimate the probability of reaching each level through the sequence of conditional probabilities of all levels before. Traditionally, in SuS with a binary failure region, the failure probability can then be estimated through:

$$\mathbb{P}(F) = \mathbb{P}(F_m) = \mathbb{P}(F_0)\mathbb{P}(F_1|F_0)\mathbb{P}(F_2|F_1)\cdots\mathbb{P}(F_m|F_{m-1}), \tag{1}$$

where $\mathbb{P}(F_k|F_{k-1})$ is the conditional probability of event F_k given the occurrence of F_{k-1}, and m corresponds to the final level of SuS.

In order to determine the closeness to a failure, one makes use of a performance evaluation function, or Limit-State Function (LSF) as it is called in SuS terminology. Each level of SuS further explores the scenarios from the previous level that are closest to a failure. This is achieved through a sampling method called Markov Chain Monte Carlo (MCMC) and the specific algorithm, used within this paper, is the Modified Metropolis Algorithm (MMA), see [10] and [15] for further details. The intermediate failure region consists of samples, θ, fulfilling:

$$LSF(\theta) \leq y^*, \tag{2}$$

where y^* is the threshold which captures p_0 percent (usually set to 10%, to ensure that intermediate regions can be easily found) of the samples in the level. The samples from the intermediate failure region is subsequently used as a starting point for the next level of SuS. This level-wise exploration stops once more than $p_0\%$ of the samples of the final level fulfil (2) with $y^* = 0$. An LSF of less than zero is the definition of the *failure region* which SuS is tasked with estimating.

2.3 Brake Threat Number

In order to guide SuS towards collision scenarios there is a need to use some kind of threat metric. How SuS is able to find collision scenarios using different threat metrics is investigated in [21]. Based on the analysis presented therein we will use Brake Threat Number (BTN) as one of the factors for our composite metric described below. In fact, we will use a version of the BTN, the BTN_{post}, which looks at the scenario retrospectively to find the most critical time instance. The BTN_{post} for that instance captures the needed share of available braking capability that the ego vehicle would need to use to avoid a collision. The main difference between BTN and BTN_{post} is that the latter only uses the relative positions and velocities, and no prediction of states is needed. For further details, see [22, Eqs. (13) - (16)]. For the sake of notation clarity, though, in the remainder of this paper we will use BTN to signify BTN_{post}.

2.4 A Composite Failure Metric

Since we are looking for characterising injury risks and not only a collision rate, we introduce here a new metric for the LSF that is able to capture both of these properties. For that purpose, we propose the function:

$$g(\theta) = (1 + \tau_{\Delta v}) - \text{SEVBTN} = (1 + \tau_{\Delta v}) - (min(\text{BTN}, 1) + \Delta v)\,. \tag{3}$$

Here, τ_{Δ_v} is a constant related to the maximum collision velocity (Δv) that should represent a failure case when guiding SuS and BTN is the maximum BTN of the considered scenario. To ensure that the value of BTN does not influence the SuS exploration after a collision has been identified, we cap it using $min(\text{BTN}, 1)$, after which (3) is only reduced by the increase in collision velocity, Δv. The termination criterion for SuS is defined as when LSF ≤ 0, which in our case is determined by the leading factor of $(1 + \tau_{\Delta v})$, such that $g(\theta) \leq 0$ when there is a collision with $\Delta v \geq \tau_{\Delta v}$. The factor following the minus sign of (3) we coin severity-BTN (SEVBTN).

Note that for the purpose of this paper, we take the relative velocity difference at the moment of impact (Δv) as a proxy for the real collision velocity. While this is a simplification that might impact the final probability estimates, it does not impact the applicability of the methods or the results provided in this paper.

2.5 Estimating Injury Probabilities from SuS

Given our goal of using injury risk functions to estimate multiple injury levels using SuS, there is one obstacle remaining. Namely, the fact that we have a non-binary failure region when estimating the probability of MAIS3+ and MAIS5+, respectively. In particular, any sample of enough collision velocity can contribute with some (albeit small) probability of a MAISX+ injury. This is different from the classical binary failure assessment used for SuS.

To accommodate this we need to modify the classical failure estimation of SuS, as described in (1). This modification is twofold:

(i) We need an additional function to determine the injury risks from a given sample and cannot directly use the values from the performance evaluation function, $g(\theta)$; and

(ii) We need to consider samples generated also in earlier levels of SuS when determining the final probability of injury.

To address the former, (i), we need simply add an evaluation of the injury risk functions of [9] of each sample. However, to address the latter, i.e. (ii), we need to modify the way the failure probability from SuS is estimated.

The modified algorithm for estimating the injury level probabilities include the following steps:

1. Remove the $p_0\%$ samples closest to a failure from each level (i.e. those samples used as the starting point for the next level of SuS);
2. Calculate the average injury probability of the remaining samples in the level;
3. Weight each level with the probability of getting to that level while accounting for the removed samples; and
4. Sum the weighted contributions.

While the first two steps are to a certain extent straightforward, the third step above might require some elaboration. Since we remove the seed samples from each level, this also means that the corresponding probability of getting into that level is no longer p_0 to the power of the level number, but we need to remove the probability of us getting into the next level, i.e.

$$w_i = p_0^i - p_0^{i+1}, \quad \forall\, i \in [1, m), \tag{4}$$

where m is the number of levels of SuS. With these details in mind we can write the injury level probability as:

$$R_{\text{MAISX+}} = w_m \sum_{j}^{N} \left(\mathbb{P}_{\text{MAISX+}}(\theta_m^j)\right) + \sum_{i}^{m-1} w_i \sum_{j}^{N(1-p_0)} \left(\mathbb{P}_{\text{MAISX+}}(\theta_i^j)\right), \quad (5)$$

where $\mathbb{P}_{\text{MAISX+}}(\cdot)$ is the injury probability function for MAISX+ injuries; w_i is the weight capturing the probability of getting into level i as given in (4); θ_i^j is the jth ordered sample from level i; N is the number of samples in each level; and m is the number of levels of SuS. The limit of j to $N(1-p_0)$ corresponds to the exclusion of the $p_0\%$ samples closest to a failure (following the description above). The first term of (5), captures the last level of SuS from which all samples should be included. This approach provides a correct unbiased estimation since all samples in each level are drawn from the intermediate failure distributions [20].

3 Case Study

In this section, we present a case study to evaluate the applicability and usefulness of the proposed performance evaluation function of (3) and the described algorithm, of Sec. 2.5, for estimating the injury level probabilities. The case study involves a capable ACC function exposed to cut-in scenarios. It should be noted that the use of SuS for the evaluation of ADS functions have previously been investigated for lane change [17] and cut-in [21] scenarios. However, these publications evaluate the AD feature for collisions only. Furthermore, they do neither cover multiple injury levels nor account for continuous risk functions as we do in this paper.

3.1 Scenario Modelling: Cut-In

The cut-in used within this case study is defined, following [3], as a lane-change manoeuvre by other traffic participants that starts in the adjacent lane and ends in front of, and in the same lane as, the ADS (i.e. ego vehicle). To statistically model these cut-in scenarios we fit a Gaussian Mixture Model (GMM) to 12 443 recorded cut-ins. The trajectory of each of the recorded cut-ins is modelled by a five-degree polynomial for the longitudinal and lateral movements, respectively. In addition to the 12 parameters resulting from these two polynomials, the modelling also includes the duration of the scenario and the initial lateral position and longitudinal velocity of the ego vehicle – resulting in a total of 15 dimensions. These trajectories are subsequently split into whether they occur from the left or right lane adjacent of the ego vehicle and a 6 component GMM is used to fit each cut-in direction. The component numbers were selected following an evaluation of the Bayesian Information Criteria (BIC) and selected due to a plateau starting from 6 components.

In the evaluations of this paper, we only sample from the one component of the GMM that rendered the highest failure rate. In fact, this component resulted in collision frequencies one order of magnitude larger than the second most salient component, therefore suggesting that the total failure frequency would be dominated by the contributions of the selected component – partly motivating our choice of evaluating one component. This choice is also motivated by our ambition to showcase the method proposed in this paper, rather than provide a full evaluation of the considered AD function.

3.2 Software Under Test: ACC with Limited Sensing Horizon

For the purpose of this case study, we analyse an Adaptive Cruise Control (ACC) based on the "Adaptive Cruise Control with Sensor Fusion" example from the Automated Driving Toolbox in Mathworks Matlab [13]. The core ACC functionality tries to maintain a set speed while trying to keep distance and minimising the relative velocity to a possible target (i.e. vehicle in front). The original ACC functionality has been modified in five ways: (1) the deceleration (min_ac) was increased to -10m/s^2 to enable full braking capability; (2) the position error gain, $xerr\_gain$, was increased to 0.3 – to be able to reach full braking even in low relative velocity cases; (3) the longitudinal time constant, tau, was reduced to 0.25 – to give more realistic brake force build up; (4) the Tracking and Sensor Fusion block was replaced by a simpler lead vehicle selection assuming perfect sensing, where the target is selected if it is predicted to enter a (virtual) corridor/lane extending 0.3 m to the side of the ego vehicle; (5) the sensing horizon was capped to 30 m – beyond which point the vehicle assumes empty space.

While the former three modifications render a more capable ACC function, the last one would correspond to some degraded mode of an ADS. For example, the limited sensing horizon could be due to a perception system error. The low failure rate of the full ACC function (as estimated by initial runs of our SuS approach) made it infeasible to provide an MC reference, requiring multiple billions of samples. Hence, for the sake of this paper, we introduce the limited sensing horizon as a plausible failure mode for which it is now feasible to produce samples for an MC reference. The limitations to the ACC functionality analysed do, however, not impact the ability to showcase the applicability and usefulness of our SuS approach. Furthermore, to ensure safety of a system, it is crucial to understand the failure rates when operating in degraded modes, of which the selected ACC function here could be one.

3.3 Delimitations to the Case Study

When estimating the injury risks, by using the risk curves of [9], the case study assumes straight frontal collisions for all the collision samples generated, i.e. with zero angle. This restriction makes sense for the purpose of evaluating the approach proposed in this paper, since the considered performance evaluation function of (3) only extends the BTN with the relative velocity at the collision. The risk curves of [9] also include values for the full spectrum of collision angles, and the proposed approach could be extended by a reformulation of the considered performance evaluation function. This extension will be considered in future research.

Furthermore, the collision velocities used to calculate the injury probabilities are estimated using the relative velocity at the time of collision. This means that we do not consider any deformations of the two cars nor do we consider the case where the two vehicles might continue travelling forward after the collision.

While it could be argued that the above two simplifications reduce the realism of the simulation, they do not diminish the results nor the pertinence of the proposed SuS approach, especially considering that the samples from both the MC reference and SuS use the same limitations/assumptions. Furthermore, future extensions can be made to tackle those specific limitations.

3.4 Settings for SuS

For the results presented in the next section we generated 2×10^4 MC samples at the first level and selected 5% of those samples as seeds for the next level of SuS. The reason for this was that only about 6% of the generated samples resulted in a meaningful value of BTN, and consequently also SEVBTN. Thus, to provide informative seeds for SuS we needed to generate more MC samples in the first level.

For the levels in SuS we used a conditional level probability $p_0 = 0.1$; the number of samples at each level were set to $N = 10^4$; and the collision velocity threshold, $\tau_{\Delta v} \approx 26.2$m/s, used in (3), was set to correspond to a 40% probability of a MAIS5+.

4 Results

In the following section, the estimated injury probabilities of the ACC function faced with cut-ins are presented. The results are based on 36 independent runs of SuS in order to estimate the standard deviation. As explained before, the composite metric of (3) is used to guide the SuS search. Furthermore, for benchmark purposes, we conducted an MC-based assessment with a total of $193 \cdot 10^6$ simulations. The number of samples in the MC reference was selected to enable a reliable estimate of the MAIS5+ probability.

4.1 Injury Probability Estimates

In Table 1, the average injury probability for MAIS$\{$X+$\}_{X=\{3,5\}}$ from the 36 independent runs of SuS are presented, alongside the standard deviations of these estimates, as well as the MC reference. We can see that the estimates from SuS are in the same order of magnitude as compared to the MC reference. However, the SuS estimates have a bias towards overestimation of both of the injury levels. Yet, this is a trade-off between accuracy and efficiency. Here, it is relevant to note, that a single run of SuS requires either $90\ 000$ or $100\ 000$ simulations (depending on the number of levels required to meet the target failure rate), compared to the almost 200 million samples used to construct the MC reference. This fact alone makes the proposed SuS approach significantly more efficient, and therefore attractive for evaluation, assessment and falsification purposes for ADS development.

As a further reference, Fig. 1 presents histograms of the estimated injury level probabilities across all the 36 runs of SuS, along with the MC reference. Here, it is clear that there is an overestimation bias in SuS. The overestimation is also consistent between the two severity levels.

Injury level	Subset Simulation		Monte-Carlo
	Average	Standard deviation	
$\mathbb{P}$(MAIS3+)	21×10^{-6}	2.5×10^{-6}	6.6×10^{-6}
$\mathbb{P}$(MAIS5+)	22×10^{-7}	2.8×10^{-7}	5.3×10^{-7}
Number of samples	96 389	4 871	193×10^{6}

Table 1. Shows the average estimated injury probabilities from 36 runs of SuS compared to the estimates from the MC reference.

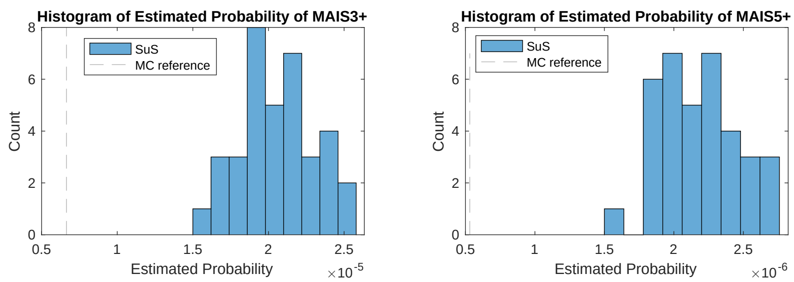

Fig. 1. Provides a histogram of the injury level probabilities from the 36 runs of SuS. The MC reference is depicted as the vertical gray dashed line.

4.2 High Collision Velocity Trajectories

To showcase the ability of SuS to guide us towards severe injury scenarios, Fig. 2 presents the trajectories from three different levels of a particular run of SuS. The trajectories are coloured according to the SEVBTN value of each trajectory, with blue representing no collision to low collision velocities, green relating to collisions with moderate collision velocities and red corresponding to high collision velocities.

There are just a few collisions at the first level (left panel), whereas at level 4 (middle) the collision velocities start approaching 15m/s. Note that, as SuS progresses through the levels, the trajectories start to look more and more similar. Interestingly, the results shown in the right panel indicate three distinct groups of high collision velocity cut-ins in level 7. Note that we do not restrict the trajectories to start in the adjacent lane. A loophole that SuS is apt at leveraging when generating high severity collisions, as can be seen from the trajectories with a starting lateral position of close to zero.

4.3 Exploring Critical Scenarios

In Fig. 3, the relative frequencies of SEVBTN values for a few different levels of SuS are presented. As SuS progresses, one can see that the collision velocities steadily increase with each level. At level 4 and onwards, we can see that SuS has found a reliable set of samples to continue the exploration towards high collision velocities. Furthermore, note that the values for SEVBTN at the later levels are rather homogeneous within each level, which can be explained by the smaller and smaller subsets of the parameter space used as a starting point for each consecutive level of SuS.

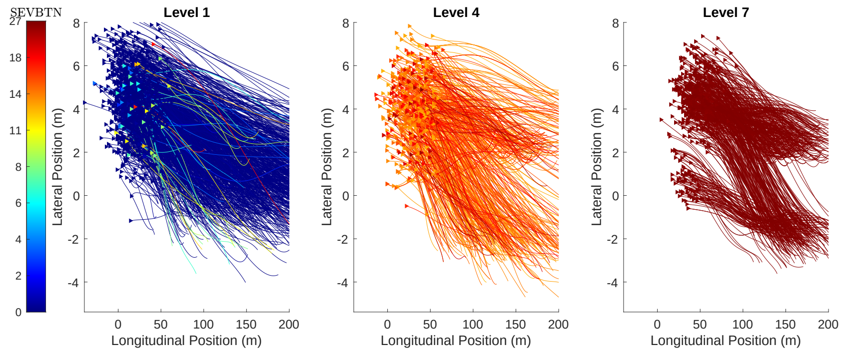

Fig. 2. Illustrates the trajectories of the generated samples at different levels of a specific run of SuS. The colour indicates the value of SEVBTN (see (3)). The longitudinal and lateral positions are relative to the starting point of the scenario not relative to the position of the ego vehicle.

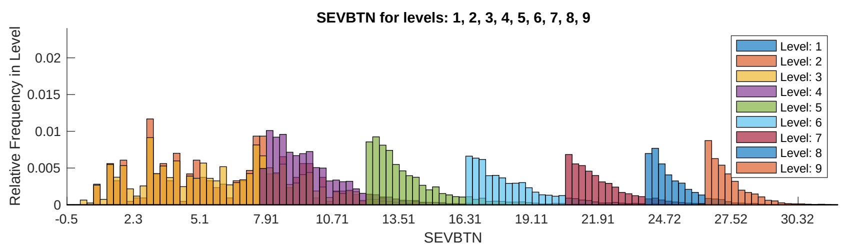

Fig. 3. Depicts the within-level-relative frequencies of SEVBTN values at different levels of SuS.

4.4 Marginal Distributions

The minimum longitudinal acceleration and the minimum lateral velocity of different sampled scenarios are shown in Fig. 4. To the left, samples from the second level of SuS are shown, in the middle samples from level 3 and to the right samples from level 4. The yellow samples and marginal distributions correspond to collision samples from the MC reference, the blue relate to collision samples from SuS, whereas red presents all samples from SuS for that level. Already at level 2 the collision samples from SuS start to resemble the marginal distributions of the MC-reference. Proceeding to level 3 we note that there is still a significant difference between the collision samples and the SuS samples at that level. This difference is completely gone at level 4 where all samples generated by SuS are collisions. This progression of samples is also reflected in the increasing value of SEVBTN as shown in Fig. 3.

In Fig. 5, we continue to explore even more critical scenarios and the associated parameters. Here the samples with a collision velocity $\Delta v \geq 15$m/s are highlighted. Continuing from level 5 (left panel of Fig. 5), we note that there is a significant difference between the samples generally generated from SuS and those with $\Delta v \geq 15$m/s, corresponding to the red and blue samples respectively. At level 6, SuS only generates samples fulfilling $\Delta v \geq 15$m/s and the generated samples at this level of SuS neatly follow the samples from the MC reference. At the last level of SuS (level 9), the generated

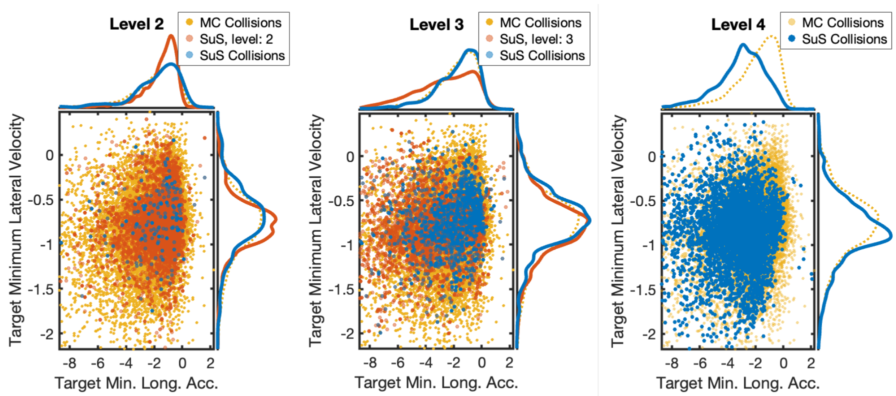

Fig. 4. Displays the relative longitudinal velocity and target longitudinal position at the beginning of the scenario for samples from SuS levels 2, 3 and 4. Collision samples from the MC reference are also depicted, in yellow. The red samples and line, correspond to all samples of SuS in that level whereas blue represent generated collision samples.

samples have been shifted to even more aggressive longitudinal accelerations. It seems as if the lateral velocity has less and less impact on generating high collision velocities in the context of the evaluated ACC function. Note that the difference between the SuS samples and the MC reference in the right-hand plot comes from the fact that SuS here only accepts samples related to the level threshold $\Delta v \geq 26.7$m/s.

5 Discussion

In this section we present some discussions on the presented results. We also highlight the limitations of the present work and provide avenues for future work and extensions to the proposed SuS approach.

5.1 Overestimation of the Injury Probabilities

The method, presented in this paper, concretises upon the approach outlined in [5] and shows how to use SuS to estimate severe injury collisions while also enabling estimation of several injury levels from a single run. The use of continuous risk curves might, however, be a reason for the overestimation of the injury probabilities compared to the MC reference. Considering the S-shaped risk curves given in [9], there are small but non-zero contributions for collisions at low velocities and similarly large contributions at the end of the curve. The question is how reliable the risk curves are at the fringes of the ”S” and how well the evaluation made corresponds to a smooth and well estimated distribution across the entire region of this S-curve. Through SuS, we are able to efficiently sample at the end of the S-curve. Something that the MC reference is not able to do. This is especially true for MAIS5+ events, for which the MC reference includes zero samples fulfilling the termination criteria we set for SuS, namely $\mathbb{P}$(MAIS5+)≥ 0.4.

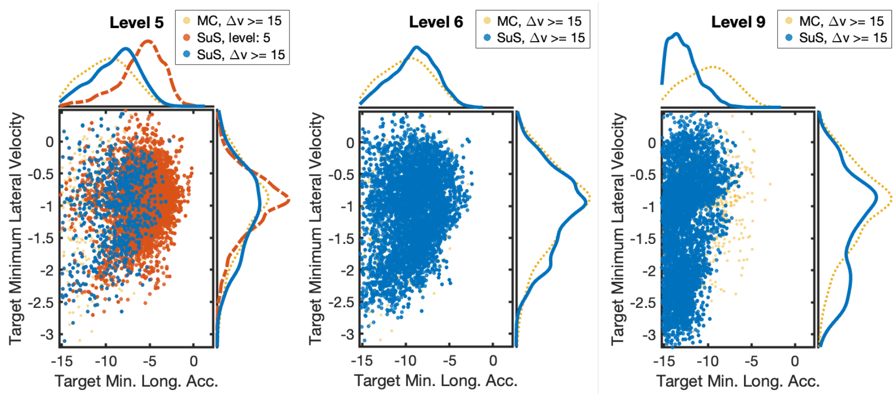

Fig. 5. Shows the same parameters as Fig. 4 but for samples with $\Delta v \geq 15$m/s. Starting from the left, the panels display samples from SuS levels 5, 6 and 9, respectively. In all three panels, samples from the MC references are also shown. Note that all samples in SuS level 9 have $\Delta v \geq 26.7$m/s.

5.2 Metric Selection

When using SuS, the selection of metrics to guide the exploration is very crucial, as shown in, e.g., [21]. In our case, we found that the BTN value was non-existent for many of the samples generated through pure MC. Consequently, in order for SuS to have a meaningful transition to the next level, we needed to produce more samples in the first level and also have a correspondingly lower conditional level probability. One other solution to this would be adding an additional term to the proposed composite metrics of (3) to discern the closeness of the scenario in producing a BTN value in the first place. However, again care should be taken in making this extension, such that the added factor helps guide SuS rather than set it off in the wrong direction. By merging several metrics in the same LSF there is of course the risk of producing conflicting directions for SuS when searching for the failure region. This needs to be carefully considered.

5.3 Efficiency

Even though SuS provides an overestimation of the injury probabilities it produces consistent results with three orders of magnitude fewer samples compared to the MC references. Furthermore, by the proposed approach of estimating multiple injury levels simultaneously we further increase the efficiency of each used sample compared to having dedicated SuS runs for each such estimate.

5.4 Limitations

For this paper we have focused on a single degraded mode of an ACC function as the software under test. While this might seem like an arbitrary choice, it is not unreasonable considering the importance of estimating safety properties also for different degraded modes of the full function. Considering the efficiency gain garnered from the proposed SuS approach it might be possible to assess the injury probabilities of not just the full function but also several of the main degraded modes. Consequently, this approach might efficiently support the use of Restricted Operational Domains (RODs) [1] as well as enable optimal performance also in degraded modes of the system.

By restricting the simulation analysis to only involve straight frontal collisions, i.e. at angle zero, one limits the number of collision variations and, to a certain extent, the realism of the simulation. In the process, we might also underestimate the total injury risk probability since the risk curves of [9] have their minima at angle zero. Including the angle of collision would therefore be a relevant and interesting extension to the present work. Such an extension would include not only the implementation of angle estimation from simulation but also include this as a factor within the composite metric to enable SuS to be guided by also this factor.

Another extension to further increase the realism of the simulations would be to consider the deformation of the two vehicles during collisions. This extension would entail the inclusion of a factor, i.e. a coefficient for restitution, as explored in [6], when estimating the collision velocity in the sampled collision scenarios. The exclusion of this factor makes the present simulations overestimate the severity of the collisions, since the collision velocities are overestimated.

Note, however, that the underestimation of the injury risk, from restricting the simulation analysis to frontal collisions only, and the overestimation of the injury risk, stemming from the simplified approximation of the collision velocity, do not impact the validity of the presented results with respect to showcasing the applicability and usefulness of the proposed SuS approach. Nevertheless, these limitations should be addressed and the proposed approach extended for improved benefits in terms of ADS feature validation. And certainly before the output of such methods is used to motivate the deployment of said features.

5.5 Future Work

As outlined above, we suggest expanding the simulations and metrics to include the collision angle as well as considering the influence from deformation on the collision velocity to improve the realism of the simulations. This would also allow for a more accurate estimate of the collision velocity compared to the velocity difference at impact, as used for our case study, and, as a consequence, would also allow for a more accurate estimate of the evaluated feature's injury risks.

It would be valuable to apply this method to other domains where there is a need to estimate probabilities of rare events while also considering multiple injury levels. Finally, how to use this method in conjunction with dynamic risk assessment would be a very interesting avenue for future work. This would include providing a probability estimate of detailed risks in the scene of the ADS to be used to improve the situation awareness beyond what is produced by currently available methods.

6 Conclusions

This paper presents an approach to use SuS for estimating the probability of collisions with multiple severity levels, suitable for supporting efficient safety evaluation of ADSs. A composite metric, capturing both the closeness to collision, as well as the collision velocity, is proposed. Furthermore, the injury probabilities are estimated from continuous risk curves. To investigate the proposed method, we estimate the injury probabilities for an ACC function faced with cut-in scenarios, emulating the functionalities of an degraded ADS. A statistical model is constructed based on 12 443 recorded cut-ins and the estimated injury probabilities from the proposed SuS method is compared to a MC reference. The results show that the SuS method provides consistent injury probabilities estimates that fall within the correct order of magnitude, when compared to the MC reference, but with three orders of magnitude fewer samples. The presented method therefore enables more efficient safety evaluation and falsification of automated driving features throughout the development phase and, as a consequence, would also enable more frequent iteration cycles. Finally, the proposed method opens up for new novel methods for dynamic risk assessment, leading to enhanced performance and safety properties.

Acknowledgments. This research has been partially supported by the Wallenberg AI, Autonomous Systems and Software Program (WASP), funded by the Knut and Alice Wallenberg Foundation.

References

1. Colwell, I., Phan, B., Saleem, S., Salay, R., Czarnecki, K.: An automated vehicle safety concept based on runtime restriction of the operational design domain. In: Intelligent Vehicles Symposium (IV). IEEE (2018)
2. Gyllenhammar, M., R. de Campos, G., Sandblom, F., Törngren, M., Sivencrona, H.: Uncertainty aware data driven precautionary safety for automated driving systems considering perception failures and event exposure. In: Intelligent Vehicles Symposium (IV). IEEE (2022)
3. Gyllenhammar, M., Zandén, C., Törngren, M.: Defining fundamental vehicle actions for the development of automated driving systems. In: World Congress (WCX). SAE (2020)
4. Gyllenhammar, M., Zandén, C., Vakilzadeh, M.K.: Estimation of risk exposure for autonomous vehicles (Aug 02 2023), EU Pat. EP4219261A1
5. Gyllenhammar, M., Zandén, C., Vakilzadeh, M.K., Falkovén, A., Olsson, J., Zwirglmaier, K., Ulas, C.: Estimation of probability of collision with increasing severity level for autonomous vehicles (Jan 19 2022), EU Pat. EP3940487A1
6. Ishikawa, H.: Impact center and restitution coefficients for accident reconstruction. SAE transactions pp. 657–671 (1994)
7. ISO: ISO 26262:2018 Road vehicles – Functional safety (2018)
8. Jesenski, S., Stellet, J.E., Branz, W., Zöllner, J.M.: Simulation-based methods for validation of automated driving: A model-based analysis and an overview about methods for implementation. In: Intelligent Transportation Systems Conference (ITSC). pp. 1914–1921. IEEE (2019)
9. McMurry, T.L., Cormier, J.M., Daniel, T., Scanlon, J.M., Crandall, J.R.: An omni-directional model of injury risk in planar crashes with application for autonomous vehicles. Traffic injury prevention **22**(sup1), S122–S127 (2021)
10. Papaioannou, I., Betz, W., Zwirglmaier, K., Straub, D.: Mcmc algorithms for subset simulation. Probabilistic Engineering Mechanics **41**, 89–103 (2015)

11. Riedmaier, S., Ponn, T., Ludwig, D., Schick, B., Diermeyer, F.: Survey on scenario-based safety assessment of automated vehicles. IEEE access **8**, 87456–87477 (2020)
12. Rodrigues de Campos, G., Kianfar, R., Brännström, M.: Precautionary safety for autonomous driving systems: Adapting driving policies to satisfy quantitative risk norms. In: Intelligent Transportation Systems Conf. (ITSC). IEEE (2021)
13. Simulink 10.1 (R2020a): Adaptive cruise control with sensor fusion (2020), [Online]. https://se.mathworks.com/help/driving/ug/adaptive-cruise-control-with-sensor-fusion.html
14. Stellet, J.E., Brade, T., Poddey, A., Jesenski, S., Branz, W.: Formalisation and algorithmic approach to the automated driving validation problem. In: 2019 IEEE Intelligent Vehicles Symposium (IV). pp. 45–51. IEEE (2019)
15. Vakilzadeh, M.K., Huang, Y., Beck, J.L., Abrahamsson, T.: Approximate bayesian computation by subset simulation using hierarchical state-space models. Mechanical Systems and Signal Processing **84**, 2–20 (2017)
16. Wardziński, A.: Safety assurance strategies for autonomous vehicles. In: Harrison, M.D., Sujan, M.A. (eds.) Computer Safety, Reliability, and Security. pp. 277–290. Springer Berlin Heidelberg, Berlin, Heidelberg (2008)
17. Zhang, S., Peng, H., Zhao, D., Tseng, H.E.: Accelerated evaluation of autonomous vehicles in the lane change scenario based on subset simulation technique. In: Intelligent Transportation Systems Conf. (ITSC). IEEE (2018)
18. Zhang, X., Tao, J., Tan, K., Törngren, M., Sánchez, J.M.G., Ramli, M.R., Tao, X., Gyllenhammar, M., Wotawa, F., Mohan, N., Nica, M., Felbinger, H.: Finding critical scenarios for automated driving systems: A systematic mapping study. IEEE Transactions on Software Engineering **49**(3), 991–1026 (2023). https://doi.org/10.1109/TSE.2022.3170122
19. Zhao, D., Lam, H., Peng, H., Bao, S., LeBlanc, D.J., Nobukawa, K., Pan, C.S.: Accelerated evaluation of automated vehicles safety in lane-change scenarios based on importance sampling techniques. IEEE Transactions on Intelligent Transportation Systems **18**(3), 595–607 (2016)
20. Zuev, K.M.: Subset Simulation Method for Rare Event Estimation: An Introduction, pp. 1–25. Springer Berlin Heidelberg, Berlin, Heidelberg (2021). https://doi.org/10.1007/978-3-642-36197-5_165-1, https://doi.org/10.1007/978-3-642-36197-5_165-1
21. Åsljung, D., Zandén, C., Fredriksson, J., Vakilzadeh, M.K.: On automated vehicle collision risk estimation using threat metrics in subset simulation. In: Intelligent Transportation Systems Conference (ITSC). pp. 58–63. IEEE (2021)
22. Åsljung, D., Zandén, C., Fredriksson, J.: A Risk Reducing Fleet Monitor for Automated Vehicles Based on Extreme Value Theory. techrxiv (5 2022). https://doi.org/10.36227/techrxiv.19780081.v1