Delft University of Technology

Delft Center for Systems and Control

Technical report 04-029

A probabilistic approach for validation of advanced driver assistance systems*

O.J. Gietelink, B. De Schutter, and M. Verhaegen

If you want to cite this report, please use the following reference instead:
O.J. Gietelink, B. De Schutter, and M. Verhaegen, "A probabilistic approach for validation of advanced driver assistance systems," *Proceedings of the 84th Annual Meeting of the Transportation Research Board*, Washington, DC, 24 pp., Jan. 2005. Paper 05-1524.

Delft Center for Systems and Control
Delft University of Technology
Mekelweg 2, 2628 CD Delft
The Netherlands
phone: +31-15-278.24.73 (secretary)
URL: https://www.dcsc.tudelft.nl

A Probabilistic Approach for Validation of Advanced Driver Assistance Systems

Authors:
O.J. Gietelink (corresponding author), TNO Automotive,
P.O. Box 756, 5700 AT Helmond, The Netherlands,
phone: +31 492 566507, fax: +31 492 566566, gietelink@wt.tno.nl

B. De Schutter, Delft Center for Systems and Control, Delft University of Technology,
Mekelweg 2, 2628 CD Delft, The Netherlands,
phone: +31 15 2785113, fax: +31 15 2786679, b.deschutter@dcsc.tudelft.nl

M. Verhaegen, Delft Center for Systems and Control, Delft University of Technology,
Mekelweg 2, 2628 CD Delft, The Netherlands,
phone: +31 15 2785119, fax: +31 15 2786679, m.verhaegen@dcsc.tudelft.nl

Abstract

This paper presents a methodological approach for validation of advanced driver assistance systems. The methodology relies on the use of randomized algorithms that are more efficient than conventional validation using simulations and field tests, especially with increasing complexity of the system. The methodology consists of first specifying the perturbation space and performance criteria. Then a minimum number of samples and a relevant sampling space is selected. Next an iterative randomized simulation is executed, followed by validation of the simulation model by hardware tests, in order to increase the reliability of the estimated performance. The proof of concept is illustrated with some examples of a case study involving an adaptive cruise control system. The case study also points out some characteristic properties of randomized algorithms regarding the necessary sample complexity, and the sensitivity to model uncertainty. Solutions for these issues are proposed as well as corresponding recommendations for future research.

1 INTRODUCTION

1.1 State-of-the-Art of Advanced Driver Assistance Systems

The increasing demand for safer passenger vehicles has stimulated the research and development of advanced driver assistance systems (ADASs) over the past decade. An ADAS typically consists of environment sensors (e.g. radar, laser, and vision sensors) and controllers to improve driving comfort and traffic safety by warning the driver, or even autonomous control of actuators. State-of-the-art examples of ADASs that have already been introduced by the automotive industry are adaptive cruise control (ACC) (*1*) (*2*), collision warning systems (*3*), and pre-crash systems (*4*).

1.2 Challenges in Design and Validation of ADASs

The demand for safety and reliability naturally increases with this increasing automation of the vehicle's driving task, since the driver must fully rely on a flawless operation of the ADAS. For instance, autonomous braking in an ACC should be executed only if the distance to the preceding vehicle will otherwise become unacceptably small, or even results in a collision. In order to handle a large variety of complex traffic scenarios and disturbances, redundancy and fault-tolerance measures are often implemented. In practice, it is however difficult to choose the right control measures and to validate their effectiveness. Manufacturers thus face an increasing effort in the design and validation of ADASs in contrast to a desired shorter time-to-market.

Currently, an iterative process of simulations and prototype test drives on a test track is used for validation purposes. The simulation effort can however be inefficient, since a large number of simulations are necessary. Instead, a worst-case analysis is often used, which leads to a conservative control system design, thus limiting the functional performance of the ADAS. On the other hand, test drives are more reliable, but can never cover the entire set of operating conditions, due to time and cost constraints. Furthermore, test results can be difficult to analyze for controller validation and benchmarking, because traffic scenarios cannot be exactly reproduced during a test drive. It may therefore become impossible to evaluate an ADAS with guaranteed measures for the level of performance, safety, and reliability.

1.3 Need for New Tools and Methods

The objective of this paper is therefore to present a methodological approach to provide an *efficient* test program in order to cover the entire set of operating conditions. Here we define efficiency in terms of a minimum number of experiments to be performed, in order to reduce the costs for validation. A simplified case study will be used as an illustration of this methodology for reasons of transparency. Note that the emphasis of this paper is therefore on the *validation* process, and not on control system *design*.

One validation tool in this methodology is the software tool PRESCAN, which is used in a probabilistic simulation strategy with the operating conditions chosen to be representative for the 'real' conditions. The simulation models can subsequently be validated with another tool: the VEhicle-Hardware-In-the-Loop (VEHIL) facility, TNO's tailor-made laboratory for testing ADASs. In VEHIL a real ADAS equipped vehicle is tested in a hardware-in-the-loop simulation, as shown by the working principle in Figure 1. With VEHIL the development process and more specifically the validation phase of ADASs can be carried out safer, cheaper, more manageable, and more reliable. However, before VEHIL testing takes place, the simulation phase should be able to provide a reliable estimate of the performance. Both VEHIL and PRESCAN are fully described in (*5*), and in this paper we restrict the discussion to the underlying validation methodology, as outlined below.

1.4 Outline of this Paper

In Section 2 a simplified model of an ACC system is presented, together with its performance measures and the perturbations acting on the system. Section 3 treats the background theory of a randomized algorithm (RA) that is subsequently applied to the ACC case study in a number of examples. These examples illustrate the advantages of using RAs, but also highlight some points for improvement, especially regarding the required number of samples. Section 4 discusses some possibilities for addressing this issue. Subsequently, Section 5 presents an improved methodological approach for the validation of ADASs. Finally, Section 6 summarizes the validation approach, and discusses ongoing research activities.

2 A CASE STUDY: ADAPTIVE CRUISE CONTROL

2.1 A Simplified Model for Adaptive Cruise Control

The ACC longitudinal control problem consists of two vehicles, as shown in Figure 2. This figure also indicates the position x and the velocity v, where the subscripts 'l' and 'f' denote leader and follower respectively. Further defined are the distance between the two vehicles (the *headway*) $x_{\mathrm{r}} = x_{\mathrm{l}} - x_{\mathrm{f}}$, the relative velocity $v_{\mathrm{r}} = v_{\mathrm{l}} - v_{\mathrm{f}}$, the desired distance x_{d}, and the headway separation error $e_x = x_{\mathrm{d}} - x_{\mathrm{r}}$.

ACC tries to maintain a pre-defined velocity set-point v_{cc}, unless a slower vehicle is detected ahead. In that case vehicle 'f' is controlled to follow vehicle 'l' with equal velocity $v_{\mathrm{f}} = v_{\mathrm{l}}$ at a desired distance x_{d}. Since the ACC objective is to control the motion of a vehicle *relative* to a preceding vehicle, the vehicle state is chosen as $\mathbf{x} = \begin{bmatrix} x_{\mathrm{r}} & v_{\mathrm{r}} \end{bmatrix}^T$. The initial condition of $\mathbf{x}$ is defined as $\mathbf{x}(0) = \begin{bmatrix} x_{\mathrm{r,0}} & v_{\mathrm{r,0}} \end{bmatrix}^T$, i.e. every time that the sensor detects a new lead vehicle that ACC should follow. We can then write the state space representation as

$$\dot{\mathbf{x}} = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} x_{\mathrm{r}} \\ v_{\mathrm{r}} \end{bmatrix} + \begin{bmatrix} 0 \\ -1 \end{bmatrix} a_{\mathrm{f}} + \begin{bmatrix} 0 \\ 1 \end{bmatrix} a_{\mathrm{l}}, \tag{1}$$

where the acceleration of the following vehicle a_{f} is the control input, and the acceleration of the leading vehicle a_{l} forms the disturbance to the system.

2.2 Constant Spacing Control Law

In velocity control mode, the ACC operates as a conventional cruise control, where the desired acceleration a_{d} is given by a simple proportional controller

$$a_{\mathrm{d}} = k_{\mathrm{cc}}(v_{\mathrm{cc}} - v_{\mathrm{f}}), \qquad k_{\mathrm{cc}} > 0. \tag{2}$$

In distance control mode, a_{d} is given by proportional feedback control of the distance separation error $e_x = x_{\mathrm{d}} - x_{\mathrm{r}}$ and its derivative $e_v = \dot{e}_x = v_{\mathrm{d}} - v_{\mathrm{r}}$

$$a_{\mathrm{d}} = -k_2 e_v - k_1 e_x, \qquad k_1,\ k_2 > 0. \tag{3}$$

The distance x_{d} and the feedback gains k_i can be a function of a number of state variables (usually x_{r}, v_{f}, and v_{r}), tuned in order to achieve a natural following behavior. However, here we use a constant spacing control law, where x_{d} is equal to a constant value $s_0 = 40\,\mathrm{m}$, $k_1 = 1.2$, and $k_2 = 1.7$. Since the desired relative velocity is obviously equal to zero, Eq. (3) can be rewritten as

$$a_{\mathrm{d}} = k_2 v_{\mathrm{r}} + k_1(x_{\mathrm{r}} - s_0). \tag{4}$$

Although control law Eq. (4) is not used for ACC systems because of string stability issues (*6*) and unnatural following behavior, we will use it in this paper for reasons of simplicity. The control law is however proven to be asymptotically stable, such that both e_x and e_v are always regulated to zero, provided $k_1,\ k_2 > 0$ (*6*).

In this paper we also neglect sensor processing delay and vehicle dynamics by assuming that the desired acceleration is realized at the input of the controlled system without any time lag, such that $a_{\mathrm{f}} = a_{\mathrm{d}}$. However, we do introduce an actuator saturation, since ACC systems usually restrict the minimum and maximum control input for safety reasons. In this case study we use the restriction that a_{f} is bounded between -2.5 and 2.5 m/s^2.

2.3 Performance Criteria for ACC

The performance of an ACC can be quantified in a number of measures ρ_i, e.g. overshoot, tracking error, time response, control effort, ride comfort, and string stability. Here we restrict the controller validation to the measure of *safety*, expressed as the probability p that no collision will occur for a whole range of traffic situations. The safety measure for a single experiment is denoted by $\rho_\mathrm{s} \in \{0, 1\}$, where $\rho_\mathrm{s} = 1$ means that the ACC system manages to follow the preceding vehicle at a safe distance, and $\rho_\mathrm{s} = 0$ means that the traffic scenario would require a brake intervention by the driver to prevent a collision[1]. Depending on the resulting value for p, the feedback gains k_1 and k_2 can be optimized. In practice, it is not desirable to achieve $p = 1$ for all traffic scenarios, since this would necessitate a very conservative controller with high autonomous braking capacity. ACC field test results suggest that $p = 0.95$ is more or less appropriate for highway cruising (*7*)[2].

2.4 Perturbation Space

The value of ρ_s for a particular traffic scenario obviously depends on the perturbations imposed by that scenario. The disturbance to the ACC system is formed by the motion of other vehicles that are detected by the environment sensors (which form the interface between the environment and the system). Apart from the acceleration of the preceding vehicle a_l, also the initial conditions $\mathbf{x}(0)$ determine the probability of a collision. These scenario parameters, together with measurement noise, process noise, unmodeled dynamics and various types of faults construct an n-dimensional *perturbation space* Δ. It is then of interest to evaluate the function ρ_Δ that reflects the dependency of ρ on the structure of Δ. In this paper we first consider the situation where Δ is limited to the subset $\mathcal{S} = \{a_\mathrm{l},\ x_{\mathrm{r},0},\ v_{\mathrm{r},0}\}$ of uncertain traffic scenario parameters and where the resulting safety measure ρ_s is non-decreasing.

2.5 The Control System Validation Problem

Designing a simple stable controller that meets the ACC performance requirements is not difficult. However, validation of this controller with respect to these requirements and subsequent tuning requires a lot of effort. For low-order systems, controller validation can still be solved in a deterministic way or by using iterative algorithms. But when the dimension of Δ increases and ρ_Δ becomes non-convex, the problem will eventually become more difficult to solve, i.e. it becomes *intractable*, as shown by the tutorial paper (*8*).

Currently, controller validation is therefore often performed by a *grid search*, where all parameters are varied through their operating range, as applied in the clearance of flight control laws (*9*). However, an exhaustive grid search may require an inefficient large number of experiments, perhaps even too large to be feasible. Alternatively, a worst-case analysis can be performed by lumping together a combination of disturbances, each with the direction and magnitude that has the worst impact on the system performance ρ. However, a worst-case control system analysis is often not realistic, since it may result in a conservative controller that is tuned to a specific non-reachable combination of operating conditions. In the next sections we therefore discuss and further extend an approach to address this problem, using a simple case study for proof of concept.

[1] Please note the difference between the performance level ρ for a particular experiment and its probability p for a whole range of experiments.

[2] This means that in 5 % of the scenarios the driver has to intervene, which is a reasonable value for ACC. Note that ACC is designed as a comfort system, not as a collision avoidance system.

3 A RANDOMIZED ALGORITHM FOR CONTROL SYSTEM VALIDATION

3.1 Motivation for a Probabilistic Approach

An alternative approach for solving a complex problem exactly, is to solve it *approximately* by using a *randomized algorithm* (RA). An RA is an algorithm that makes random choices during its execution, and covers both Monte Carlo, sequential and other probabilistic algorithms (*10*).

The use of an RA can turn an intractable problem into a *tractable* one, but at the cost that the algorithm may fail to give a correct solution (without the user knowing). The probability that the RA fails can be made arbitrarily close to zero, but never exactly equal to zero. A popular example is a Monte Carlo simulation strategy, where this probability depends on the *sample complexity*, i.e. the number of simulations performed.

An important issue is therefore the necessary sample complexity that guarantees a certain level of confidence for the simulation outcome. In this section we show that this sample complexity is bounded, depending on the desired level of accuracy and reliability, but also that these bounds are rather conservative.

3.2 Bounded Sample Complexity for RAs

The use of a randomized approach for controller validation can be illustrated as follows. Consider an arbitrary process with only two possible outcomes: 'failure' ($\rho = 0$) and 'success' ($\rho = 1$). Suppose that we wish to determine the probability p for a successful outcome of this process. If N denotes the number of experiments with this process and N_{S} the number of experiments with successful results, then the ratio N_{S}/N is called the *empirical probability* or *empirical mean* $\hat{p}_N$ for a successful result of the process.

However, $\hat{p}_N$ is unlikely to be exactly equal to the real probability p, although it is reasonable to expect that $\hat{p}_N$ will approach p, as $N \to \infty$, and as long as the samples are chosen to be representative of the set Δ. The question thus arises in what sense $\hat{p}_N$ converges to p, and how many experiments N have to be performed to give a *reasonable* estimate of p. An estimate $\hat{p}_N$ can be called reasonable if it differs from the real (unknown) value p by no more than $\epsilon > 0$, such that

$$|p - \hat{p}_N| \leq \epsilon, \tag{5}$$

where ϵ is called the *accuracy* of the estimate.

Since $\hat{p}_N$ is a random variable depending on the particular realization of the N samples, the outcome of the inequality Eq. (5) is a random variable as well with a certain probability of realization. Therefore, we cannot always guarantee that $|p - \hat{p}_N| \leq \epsilon$, even with a large number of experiments. This means that if the experiments are performed another N times, the estimate $\hat{p}_N$ will probably have another value. The probability that $|p - \hat{p}_{N_j}| > \epsilon$ for a particular set of simulations N_j is then the *unreliability* δ of the estimate $\hat{p}_N$. We would therefore like to know the level of accuracy ϵ, and reliability $1 - \delta$, that can be obtained with one particular set of N simulations. In practice, often the reverse problem is considered: how many experiments N are necessary to achieve a desired level of confidence, in terms of ϵ and δ?

To answer this question, first define the desired accuracy for determining the unknown quantity p by some number $\epsilon > 0$. Then, in order to know the reliability of our set of experiments, it is necessary to determine the probability that performing N experiments will generate an estimate $\hat{p}_N$ for which $|p - \hat{p}_N| > \epsilon$. A popular measure available is the Chernoff bound (*11*), which states that this probability, defined as $\delta > 0$, is no larger than $2e^{-2N\epsilon^2}$. This means that, after performing the experiment N times, we can state with a confidence of at least $1 - 2e^{-2N\epsilon^2}$ that the empirical probability $\hat{p}_N$ is no more than ϵ different from the true but unknown probability p. Therefore, in order to estimate the unknown quantity p to an accuracy of ϵ and with a confidence of $1 - \delta$, the number of experiments N should be chosen such that

$$2e^{-2N\epsilon^2} \leq \delta. \tag{6}$$

Eq. (6) can also be written as

$$N_{\mathrm{ch}} \geq \frac{1}{2\epsilon^2} \ln \frac{2}{\delta}, \tag{7}$$

which is known as the *additive Chernoff bound*. Table 1 presents the necessary sample complexity to calculate $\hat{p}_N$ for some values of ϵ and δ. Since $\hat{p}_N$ has a confidence interval, the upper bound for N is called a *soft* bound, as opposed to *hard* bounds given by a deterministic algorithm.

3.3 Formulation of an RA

The Chernoff bound forms the basis for the application of Monte Carlo methods, used widely in engineering for design and analysis of control system performance. Several references give a clear tutorial introduction to the use of RAs (*12, 13, 14, 8*). Here we summarize the main steps to be taken in control system analysis using RAs.

Suppose that the closed-loop system (in our case the ACC equipped vehicle and its controller) must be verified for a certain performance level ρ. It is then the goal to estimate the probability p that this performance ρ lies above a pre-specified threshold value γ. In order to compute $\hat{p}_N(\gamma)$, we generate N independent identically distributed (iid) samples [3]

$$\Delta_1, \Delta_2, \ldots, \Delta_N$$

in the perturbation space Δ according to its probability density function (pdf) f_Δ. The outcome of every i-th experiment is represented by an indicator function J, given by

$$J(\Delta_i) = \begin{cases} 0, & \text{if } \rho < \gamma \\ 1, & \text{if } \rho \geq \gamma \end{cases} \tag{8}$$

The empirical probability $\hat{p}_N$ can then be estimated as

$$\hat{p}_N = \frac{1}{N} \sum_{i=1}^{N} J(\Delta_i), \tag{9}$$

which is known as the *simple sampling* estimator.

Since the outcome of the inequality $|p - \hat{p}_N| \leq \epsilon$ is a random variable, it has a certain probability of realization. By introducing a confidence degree $1 - \delta$, this probability is defined as

$$\Pr\{|p(\gamma) - \hat{p}_N(\gamma)| \leq \epsilon\} \geq 1 - \delta, \qquad \forall\, \gamma \geq 0,\ \delta \in (0,1),\ \epsilon \in (0,1) \tag{10}$$

The sample complexity N that is required for Eq. (10) to hold, can then be calculated from Eq. (7). This method for probabilistic control system analysis is formalized as follows.

[3] In the following we will also use $\hat{p}_N$ instead of $\hat{p}_N(\gamma)$ for reasons of clarity.

Algorithm 3.1 (Probabilistic performance verification (*14*))
Given $\epsilon > 0$, $\delta \in (0,1)$ and $\gamma \geq 0$, this RA returns with probability at least $1-\delta$ an estimate $\hat{p}_N(\gamma)$ for the probability $p(\gamma)$, such that $|p(\gamma) - \hat{p}_N(\gamma)| \leq \epsilon$.

1. *Determine the necessary sample size*
   $N = N_{\mathrm{ch}} \geq \frac{1}{2\epsilon^2} \ln \frac{2}{\delta}$;
2. *Draw N samples $\Delta_1, \ldots, \Delta_N$ according to f_Δ;*
3. *Return the empirical probability*
   $\hat{p}_N(\gamma) = \frac{1}{N} \sum_{i=1}^{N} J\left(\Delta^{(i)}\right)$
   where $J(\cdot)$ is the indicator function of the set $\mathcal{B} = \{\Delta : \rho(\Delta) \geq \gamma\}$.

3.4 Examples of Application of an RA

In this section we will apply Algorithm 3.1 to the simplified ACC case study, of which the 'true' outcome is exactly known. If we are interested in knowing the number of traffic situations that would become dangerous, given a stochastic distribution of traffic scenarios, we can use Algorithm 3.1 to give an estimate $\hat{p}_N$ of this probability p with an accuracy level ϵ and a confidence level δ, based on N simulations.

3.4.1 Example 1: Uniformly Distributed Disturbance

Scenario Definition In this first example we choose a highway cruising scenario where the lead vehicle approaches a traffic jam and brakes to a full stop. The initial conditions are $x_{r,0} = x_{\mathrm{d}} = 40\,\mathrm{m}$, and $v_{\mathrm{l}}(0) = v_{\mathrm{f}}(0) = 30\,\mathrm{m/s}$. We assume that the deceleration of the preceding vehicle is the only disturbance with a uniform distribution between -10 and $0\,m/s^2$, denoted as $f^{\mathrm{U}}(a_{\mathrm{l}}) \sim \mathcal{U}[-10,\,0]$. We also restrict the analysis to the measure of *safety* ρ_{s}.

In situations when the lead vehicle brakes hard, the ACC vehicle cannot obtain the necessary deceleration a_{d}, since the actuator saturates at -2.5 m/s^2. Now, for fine-tuning the controller parameters, we would like to know the percentage of brake situations for which ρ_{s} meets a pre-defined threshold. The threshold γ is simply set at 1 in this case (no collision).

The safety of the system obviously decreases with a stronger deceleration a_{l}, such that the function $\rho_{\mathrm{s}}(a_{\mathrm{l}})$ is non-decreasing. Therefore, the boundary value Δ_γ, for which only just a collision is prevented, can easily be calculated using an iterative algorithm as $-3.015\,\mathrm{m/s}^2$. Below this value (the scenario will always result in a collision, above this value the ACC vehicle will be able to stop autonomously and prevent a collision. Since a_{l} is uniformly distributed on the interval $[-10,\,0]$, the real value for p is thus equal to $(3.015/10) = 0.3015$. The probability that the estimated value $\hat{p}_N$ differs from this real value by an accuracy smaller than ϵ is then larger than $1-\delta$, as defined by Eq. (10).

Randomization of the Problem Although for this example it is feasible to formulate a deterministic algorithm, in practice it can be difficult or even impossible to determine p in a deterministic way, when the dimension of Δ increases and the relation ρ_Δ becomes non-convex. So instead of calculating p explicitly in deterministic sense, the function is randomized in such a way that it takes a random input Δ_i from its distribution function $f^{\mathrm{U}}(a_{\mathrm{l}})$, according to Algorithm 3.1.

In order to verify the performance of Algorithm 3.1, we execute it 500 times (each with $N = 100$). So every j-th simulation set, for $j = 1, \ldots, 500$, gives us an estimate $\hat{p}_{N_j}$. The distribution of the estimate is shown in the histogram in Figure 3(a). With this example, the accuracy and reliability for a single simulation set (each consisting of 100 simulations), can be estimated as $\hat{\epsilon}$ and $\hat{\delta}$ respectively.

The empirical mean $\hat{p}_M$ of the probability of a collision-free scenario is 0.30714, based on all $M \cdot N = 50\,000$ simulations, which is very close to $p = 0.3015$, as could be expected. The variance of each individual estimate $\hat{p}_{N_j}$ can be found by the unbiased estimator for the variance

$$\sigma^2_{N_j-1} = \frac{1}{N_j - 1} \sum_{j=1}^{N_j} (\hat{p}_{N_j} - \hat{p}_M)^2, \tag{11}$$

which gives $\sigma^2_{N_j-1} = 0.0021042$.

Analysis of the Simulation Results Suppose we desire that $\delta = 0.10$ and $\epsilon = 0.10$. Eq. (7) then gives $N_{\text{ch}} = 150$ as an upper bound on the sample complexity. The desired values for δ and ϵ imply that the empirical mean $\hat{p}_{N_j}$ for the j-th set of simulations should lie in the interval $[0.2015, 0.4015]$ for every 9 out of 10 simulation sets. However, from Figure 3(a) it can be observed that in total only 15 out of the 500 simulation sets result in $\hat{p}_{N_j} \notin [0.2015, 0.4015]$. This means that δ is empirically determined at $\hat{\delta} = 15/500 = 0.03$, which is smaller than the desired δ. The estimate $\hat{p}_{N_j}$ for any simulation set j is thus more reliable than desired, and we only used 100 experiments, whereas Eq. (7) requires $N_{\text{ch}} = 150$.

This example shows that a certain value for δ and ϵ can be achieved with a lower number of samples than deemed necessary by the Chernoff bound. Correspondingly, by choosing $N = N_{\text{ch}}$, a higher level of accuracy and reliability can be obtained. Although the degree of conservatism for this example is rather limited, it can be shown that this conservatism increases with smaller values for δ and ϵ (*8*).

3.4.2 Example 2: Gaussian Distributed Disturbance

In the previous example, $f^{\text{U}}(a_{\text{l}})$ may not be representative for the entire operating range, whereas $\hat{p}_N$ greatly depends on the correctness of the underlying pdf. Let us therefore assume a more general acceleration profile for highway cruising. Typical vehicle measurements during ACC field testing (*7*) suggest that the acceleration profile can be roughly described as a random signal with a Gaussian distribution f^{N} with mean $\mu = 0$ and standard deviation $\sigma = 1.5$, denoted as $\mathcal{N}(0, 1.5)$, truncated on the interval $[-10, 10]$ m/s^2.

We now execute Algorithm 3.1 again with $a_{\text{l}} \sim \mathcal{N}(0, 1.5)$. The resulting histogram with the estimated collision probability $\hat{p}_{N_j}$ for every simulation set is shown in Figure 3(b). The empirical mean $\hat{p}_M$ is now 0.97752, with a variance of $\sigma^2_{N_j-1} = 0.00020506$. This variance is much smaller than in Example 1, due to the lower occurrence rate of dangerous situations.

Suppose that $\delta = 0.02$ and $\epsilon = 0.03$ are desired, Eq. (7) then gives $N_{\text{ch}} = 2559$. From Figure 3(b) can be seen that only 10 out of 500 estimates $\hat{p}_{N_j}$ fall outside the interval $[p - \epsilon, p + \epsilon]$, which indicates an estimated reliability $\hat{\delta} = 0.02$. So the same level of accuracy and reliability has been achieved with only $N = 100$ instead of $N_{\text{ch}} = 2559$ samples.

3.4.3 Example 3: Multi-Dimensional Disturbance

In the previous two examples a_{l} was considered the only disturbance. However, in practice p depends also on the initial distance and velocities of both vehicles at the moment of first detection by the sensor. Consider for instance a cut-in situation at close distance x_{r}, which is more dangerous than a vehicle cutting-in at larger distance (with identical a_{l} and v_{r}). The crash probability thus depends on the entire set $\mathcal{S} = \{v_{\text{l}},\ v_{\text{f}},\ x_{\text{r}},\ a_{\text{l}}\}$ that defines the longitudinal scenario, as discussed in Section 2.4. In this example we assume Gaussian distributions for both the velocities (with $\mu = 30$ m/s and $\sigma = 5$ m/s), and distance x_{r} (with $\mu = 60$ m, $\sigma = 20$ m, and truncated on $[10, 150]$ m).

Figure 3(c) shows the resulting histogram after execution of Algorithm 3.1 with $\hat{p}_M = 0.86848$ and $\sigma^2_{N_j-1} = 0.0010378$. Compared to Example 2 it can be seen that the variance is much larger, since the outcome of the experiment now depends on four independent pdf's. The sample complexity N is again conservative, since Figure 3(c) shows that only one estimate falls outside the interval $[p-\epsilon, p+\epsilon]$, indicating $\hat{\delta} = 0.002$. So the same level of accuracy and reliability has been achieved with only $N = 100$ instead of $N_{\text{ch}} = 346$ samples.

3.5 Characteristics Properties of RAs

The examples and key references (*8, 12, 13, 14*) point out the following characteristic properties:

- A limited number of test runs is necessary as compared to deterministic algorithms. However, the sample complexity to achieve a reasonable ϵ and δ, as given by the Chernoff bound, is quite conservative.

- An RA is very simple, since the Chernoff bound is completely independent of the nature of the underlying process and the perturbation space Δ. However, this also means that no advantage is taken of any a priori knowledge of the structure of Δ. It is therefore desired to modify the simulation approach such that knowledge of the system is applied.

- The reliability of the simulation outcome strongly depends on the reliability of the pre-defined pdf f_Δ. In the examples we have used simplified pdf's, which in practice have to be improved by empirical data. The outcome of the simulation approach also greatly depends on the modeling effort.

- The empirical mean $\hat{p}_N$ does not say anything about the *minimum* or *maximum* level of performance that can be expected. It may well be that a control system has a good *average* performance, but also a poor *worst-case* performance.

In the next section methods used in conventional control system validation will be applied to address these issues. These methods will be integrated with the RA approach, and illustrated with the ACC case study. The resulting methodology will be presented in Section 5.

4 REDUCTION OF THE SAMPLE COMPLEXITY

As mentioned in the previous section, the sample complexity N given by the Chernoff bound is too conservative. Reduction of N for ADAS validation is therefore an important challenge. One way is to reformulate the problem and use a different test objective. In addition, the sampling space Δ can be reduced by neglecting certain subsets that are impossible to occur. Subsequently, N can be further reduced by using a priori knowledge on which samples will be more interesting than others.

4.1 A Reduced Worst-Case Bound

The examples in Section 3 were based on estimation of the *mean* performance p. Alternatively, we can estimate the *worst-case* performance $\rho_{\max}$ by

$$\hat{\rho}_N = \max_{i=1,2,\ldots,N} \rho(\Delta_i). \tag{12}$$

Tempo (*15*) proves that the minimum number of samples N which guarantees that

$$\Pr\left\{\Pr\left\{\rho\left(\Delta\right) > \hat{\rho}_N\right\} \leq \epsilon\right\} \geq 1-\delta \tag{13}$$

is then given by

$$N \geq \frac{\ln(1/\delta)}{\ln(1/\left(1-\epsilon\right))}. \tag{14}$$

The sample complexity given by Eq. (14) is much lower than for Eq. (7), as shown in Table 1, and is more suitable in case of a worst-case analysis.

4.2 Reduction of the Sampling Space

Figure 4 shows a scatter plot with the occurrence of collisions in the perturbation space Δ, corresponding to Example 3. Obviously, a collision is more likely with lower values for $x_{\mathrm{r},0}$, $v_{\mathrm{r},0}$, and a_{l}, such that the collision occurrences are clustered in a specific subspace Δ_{F}.

This means that there is structure in the perturbation space Δ and in the function ρ_Δ that can be used to reduce the necessary perturbation space of interest. The sampling space can then be reduced by disregarding specific subsets of Δ, of which the outcome is a priori known. An example is the subset Δ_{S} with combinations of positive acceleration and positive relative velocity that will never result in a potential collision.

When this dependency ρ_Δ can be proven to be convex or non-decreasing, the sample complexity N can be significantly reduced, as illustrated in Figure 5. In case of a convex performance function, an RA can then iteratively search for the boundary value $\rho(\Delta_\gamma) = \gamma$, and neglect any Δ_i larger than Δ_γ, since the outcome of those samples can be predicted in advance.

4.3 Importance Sampling

As stated above, it makes sense to give more attention to operating conditions that are more likely to cause a collision than others.

4.3.1 A Randomized Algorithm for Importance Sampling

Suppose that we want to estimate a probability p, given a perturbation Δ. If f_Δ is a pdf $\sim \mathcal{U}\left[0,1\right]$ on the interval $S=[0,1]$, our goal is then to estimate

$$p = \int_S J(\Delta) f_\Delta(\Delta) d\Delta = E[J(\Delta)] \tag{15}$$

where J is the performance function, and $\Delta \sim f_\Delta$. *Importance sampling* is a sampling technique to increase the number of occurrences of the event of which the probability p should be estimated (*16*). This estimation is corrected by dividing it by the increased probability of the occurrence of the event. In order to highlight the interesting subset Δ_F it thus makes sense not to sample from the original pdf f_Δ, but instead use an artificial pdf, reflecting the importance of the events, and then reweighing the observations to get an unbiased estimate.

We can now define an importance sampling pdf φ that is strictly positive on S. We can then write

$$p = \int_S \left(\frac{J(\Delta) f_\Delta(\Delta)}{\varphi(\Delta)} \right) \varphi(\Delta) d\Delta = E\left[\frac{J(\Phi) f_\Delta(\Phi)}{\varphi(\Phi)} \right], \tag{16}$$

where $\Phi \sim \varphi$. The importance sampling estimator based on φ is

$$\hat{p}[\varphi]_N = \frac{1}{N} \sum_{i=1}^{N} \frac{J(\Phi_i) f_\Delta(\Phi_i)}{\varphi(\Phi_i)} \tag{17}$$

where $\Phi_1, \ldots, \Phi_N$ are iid with pdf φ. Its variance is

$$\mathrm{var}\left(\hat{p}\left[\varphi\right]_N\right) = \frac{1}{N} \left[\int_S \frac{J(\Delta)^2 f_\Delta(\Delta)^2}{\varphi(\Delta)} d\Delta - p^2 \right] \tag{18}$$

An efficient estimator $\hat{p}[\varphi]_N$ is obtained by choosing φ proportional to the 'importance' of the individual samples, where importance is defined as $|J(\Delta) f_\Delta(\Delta)|$. A rare but dangerous event can thus be equally important as a more frequent but less dangerous event. A randomized algorithm can then be formulated as follows.

Algorithm 4.1 (Importance Sampling)
Given ϵ, $\delta \in (0,1)$, γ, and the true distribution function f_Δ, this RA returns with probability at least $1-\delta$ an estimate $\hat{p}_N$ for the probability p, such that $|p-\hat{p}_N| < \epsilon$.

1. *Determine a strictly positive importance sampling pdf φ that emphasizes the interesting events;*

   WHILE $N_{\text{is}} < \lambda_{\text{is}} N_{\text{ch}}$;

2. Draw $N_{\text{is}} = \lambda_{\text{is}} N_{\text{ch}}$ samples $\Delta_1, \ldots, \Delta_N$ according to φ;

3. Return the empirical probability

   $\hat{p}[\varphi]_N = \frac{1}{N_{\text{is}}} \sum_{i=1}^{N_{\text{is}}} \frac{J(\Phi_i) f(\Phi_i)}{\varphi(\Phi_i)}$

4. Determine the importance sampling variance

   $\sigma^2_{\text{is}} = \frac{1}{N_{\text{is}}-1} \sum_{i=1}^{N_{\text{is}}} (\hat{p}[\varphi]_N - \hat{p}_M)^2$

5. Determine the importance sampling reduction factor

   $\lambda_{\text{is}} = \frac{\sigma^2_{\text{is}}}{\sigma^2_{\text{ss}}}$

   where σ^2_{ss} is the variance for the simple sampling technique;

6. Check if $N_{\text{is}} \geq \lambda_{\text{is}} N_{\text{ch}}$;

   IF yes, THEN end algorithm;

   IF no, THEN increase N_{is} and return to step 2;

4.4 Example 4: Importance Sampling

Suppose that we again want to estimate the probability p of a collision-free ACC system for the given perturbation, as specified by Example 2. The goal is then to estimate

$$p = \int_S J(a_{\text{l}}) f^{\text{N}}(a_{\text{l}}) da_{\text{l}} = E[J(\Delta)] \tag{19}$$

where $\Delta \sim f^{\text{N}}$. Assuming that this Gaussian distribution function $f^{\text{N}}(a_{\text{l}})$ is correct, then with a simple sampling method relatively few samples will lie in the interval of interest, i.e. [-10, -3.015], as was observed in Example 2.

We therefore define a more suitable φ to sample from. We therefore choose the linear pdf $\varphi(a_{\text{l}}) = -0.005a_{\text{l}} + 0.05$ with a_{l} bounded on the interval $S = [-10, 10]$. Note that $\int_{-10}^{10} \varphi(a_{\text{l}}) = 1$, as is the integral of the Gaussian distribution (neglecting the tails of the distribution), such that the reweighting process gives an unbiased estimator.

With Algorithm 4.1 more 'important' samples will be generated for every j-th simulation set, thus decreasing the variance of $\hat{p}_N$. This result can be seen from the histogram in Figure 3(d), where the empirical mean $\hat{p}_M = 0.97751$ and the variance $\sigma^2_{\text{is}} = 5.92 \cdot 10^{-5}$, based on 50 000 simulations. Note that $\hat{p}_M$ is approximately equal to the situation in Example 2. However, the variance of a particular realization $\hat{p}_{N_j}$ has decreased by a factor of 3.5! The inverse of this factor is called the importance sampling reduction factor λ_{is}, which indicates the reduction in the sample complexity necessary to achieve the same level of accuracy and reliability as the Chernoff bound (*17*).

However, the performance of the importance sampling method heavily depends on the reliability of the pdf φ to generate random variables, and of the models used in the simulation. In addition, the sample complexity cannot be determined a priori, thus requiring an iterative loop in Algorithm 4.1 and the need for a suitable stopping criterion.

5 METHODOLOGICAL APPROACH

A major problem with ADAS controller validation is that the system cannot be tested *exhaustively* for every disturbance under every operating condition. A validation methodology should therefore provide a suitable test program in order to sufficiently (but also efficiently) cover the entire perturbation space. To this aim we propose a generic methodological approach, consisting of the steps described in the following sections.

5.1 Specification

Firstly, define performance measures ρ and corresponding evaluation criterion γ. In addition, the desired δ and ϵ must be defined. Correspondingly, select the test objective in order to determine the type of bound for N:

- *Probability of performance satisfaction*: for a given δ, ϵ, check if the performance measure ρ is below threshold γ with a certain probability level p for the whole perturbation space Δ.
- *Worst-case performance*: check if the worst-case performance $\rho_{\max}$ is within ϵ of $\hat{\rho}_N$ with a certain probability $1 - \delta$.

Then, identify the perturbation space Δ and its pdf f_Δ by using preliminary field test results. Using knowledge on the structure of Δ or the function ρ_Δ, determine subsets Δ_{F} and Δ_{S}, of which the outcome is a priori known (either failure or success). Furthermore, a simulation model of the vehicle, its sensor system, and its control system is designed using the dedicated simulation tool PRESCAN (*18*).

5.2 Simulation

Then execute Algorithm 4.1 to cover the important part of the perturbation space to estimate the performance $\hat{p}_N$ with respect to the criteria defined earlier. In general ρ can be a continuous value, although we used a discrete value in the examples.

The performance of Algorithm 4.1 depends heavily on the reliability of the models and the pdf's used in the simulation phase. The robustness of $\hat{p}_N$ to model uncertainty should therefore be considered when validating an ADAS in a randomized approach. The experimental relation between ρ and Δ from the simulations is then bounded between $\rho_{\max}$ and $\rho_{\min}$, as illustrated in Figure 5. This means that the estimated boundary value lies within the interval $[\Delta_{\min},\ \Delta_{\max}]$, provided that ρ_Δ is a non-decreasing relation.

5.3 Model Validation

Therefore, the most interesting samples of the perturbation space Δ_i are chosen to be reproduced in the VEHIL facility, also in a randomized approach to efficiently cover Δ. These particular Δ_i are selected to lie within the interval $[\Delta_{\min},\ \Delta_{\max}]$. In VEHIL disturbances can be introduced in a controlled and accurate way, thereby achieving a more reliable estimate than $\rho_{\mathrm{sim}}(\Delta)$. In this way the model uncertainty can be reduced, because of the replacement of a vehicle and sensor model by real hardware. The corresponding test program can be formalized as follows

Algorithm 5.1 (Probabilistic model validation)

1. *Choose initial values for $\Delta_{\min}$ and $\Delta_{\max}$ in accordance with the uncertainty size, and a suitable N.*
2. *Test ρ at $\Delta_{\min}$*
3. IF $\rho \geq \gamma$, THEN decrease $\Delta_{\min}$ and GOTO 2.
   IF $\rho < \gamma$, THEN decrease $\Delta_{\max}$ and GOTO 4.
4. Test ρ at $\Delta_{\max}$
5. IF $\rho \geq \gamma$, THEN increase $\Delta_{\min}$ and GOTO 2.
   IF $\rho < \gamma$, THEN increase $\Delta_{\max}$ and GOTO 4.
6. Return the empirical probability
   $\hat{p}_N = \frac{1}{N}\sum_{i=1}^{N} J\left(\Delta^{(i)}\right)$

The sample complexity N for VEHIL is calculated as follows. Suppose that the requirement is that $\rho \leq \gamma$, where γ is a small number, and we want to determine N in the case that all tests have been successful, i.e. $\hat{p}_N = 1$. Then we want to check if $\Pr\{|\hat{p}_N - p| \leq \epsilon\} \geq 1 - \delta$, where ϵ is set equal to γ. From ϵ and the desired confidence level δ, the necessary number N can then be derived with the Chernoff bound.

Apart from obtaining ρ_{VEHIL}, the test results can also be used for model validation. The estimate $\hat{p}_N$ may indicate necessary improvements in the system design regarding fine-tuning of the controller parameters. Improve the simulation model using the VEHIL test results until the simulation model proves to provide adequate performance, convergence in $\hat{p}_N$ and sufficient samples N.

5.4 Performance Measure

In an iterative process the simulation results in step 5.2 and thus the estimate $\hat{p}_N$ can be improved. Subsequently, the VEHIL test program in step 5.3 can be better optimized by choosing a smaller interval $[\Delta_{\min}, \Delta_{\max}]$. From the combination of simulation and VEHIL results the performance $\hat{p}_N$ of the ADAS can then be estimated with a high level of reliability, and the controller design can be improved. Finally, determine ρ_{VEHIL} with Eq. (17), correcting it for the higher occurrence rate in the interval $[\Delta_{\min}, \Delta_{\max}]$.

6 CONCLUSIONS

We have presented a methodological approach for probabilistic performance validation of advanced driver assistance systems, and applied a randomized algorithm (RA) to a simple adaptive cruise control problem. This probabilistic approach cannot *prove* that the system is safe or reliable. However, we accept a certain risk of failure (though small), since any other conventional validation process (e.g. test drives) is also based on a probabilistic analysis. Furthermore, use can be made of a priori information on the system, thereby emphasizing interesting samples.

However, the resulting probabilistic measure $\hat{p}_N$ is always associated with a certain level of confidence, depending greatly on the reliability of the simulation models, the perturbation space Δ and its pdf f_Δ. The use of VEHIL hardware tests with the real vehicle and sensor can reduce this model uncertainty and improve the estimate $\hat{p}_N$ in an iterative approach.

Ongoing research is focused on extension of this methodological approach to more complex ADAS models, with non-convex performance functions ρ_Δ, a reliable pdf f_Δ formed by empirical data, and multiple performance criteria (safety, stability, and driving comfort). Finally, the presented methodology will be integrated with the use of PRESCAN and VEHIL.

ACKNOWLEDGMENTS

Research supported by TNO Automotive and TRAIL Research School.

REFERENCES

[1] W. Prestl. The BMW active cruise control ACC. *SAE Technical Paper Series*, (2000-01-0344), 2000.

[2] T. Iijima, A. Higashimata, S. Tange, K. Mizoguchiand, H. Kamiyama, K. Iwasaki, and K. Egawa. Development of an adaptive cruise control system with brake actuation. *SAE Technical Paper Series*, (2000-01-1353), 2000.

[3] R. Kiefer, D. LeBlanc, M. Palmer, J. Salinger, R. Deering, and M. Shulman. Development and validation of functional definitions and evaluation procedures for collision Warning/Avoidance systems. Technical Report DOT HS 808 964 NHTSA, DOT/National HighwayTraffic Safety Administration, 1999.

[4] Honda. Honda develops world's first 'collision mitigation brake system' (CMS) for predicting rear-end collisions and controlling brake operations, 2003. http://world.honda.com/news/2003/4030520.html.

[5] O.J. Gietelink, D.J. Verburg, K. Labibes, and A.F. Oostendorp. Pre-crash system validation with PRESCAN and VEHIL. In *Proc. of the IEEE Intelligent Vehicles Symposium (IV)*, Parma, Italy, June 14-17 2004.

[6] D.V.A.H.G. Swaroop. *String Stability of Interconnected Systems: An Application to Platooning in Automated Highway Systems*. PhD thesis, University of California at Berkeley, 1994.

[7] P. Fancher, R. Ervin, J. Sayer, M. Hagan, S. Bogard, Z. Bareket, M. Mefford, and J. Haugen. Intelligent cruise control field operational test. Technical Report DOT HS 808 849, DOT/National Highway Traffic Safety Administration, 1998. Final report.

[8] M. Vidyasagar. Statistical learning theory and randomized algorithms for control. *IEEE Control Systems Magazine*, 18(6):69–85, December 1998.

[9] C. Fielding, A. Vargas, S. Bennani, and M. Selier, editors. *Advanced Techniques for Clearance of Flight Control Laws*. Springer, Berlin, Germany, 2002.

[10] R. Motwani. *Randomized Algorithms*. Cambridge University Press, New York, 1995.

[11] H. Chernoff. A measure of asymptotic efficiency for tests of a hypothesis based on the sum of observations. *Annals of Mathematical Statistics*, 23:493–507, 1952.

[12] C. Alippi. Randomized algorithms: A system-level, poly-time analysis of robust computation. *IEEE Trans. on Computers*, 51(7):740–749, July 2002.

[13] G. Calafiore, F. Dabbene, and R. Tempo. Randomized algorithms in robust control. In *Proc. of the 42nd IEEE Conference on Decision and Control*, pages 1908–1913, Maui, Hawaii, USA, 2003.

[14] R. Tempo, G. Calafiore, and F. Dabbene. *Randomized Algorithms for Analysis and Control of Uncertain Systems*. Springer-Verlag, Berlin, Germany, 2004.

[15] R. Tempo, E.W. Bai, and F. Dabbene. Probabilistic robustness analysis: Explicit bounds for the minimum number of samples. *Systems & Control Letters*, 30:237–242, 1997.

[16] N. Madras. *Lectures on Monte Carlo Methods*. American Mathematical Society, Providence, Rhode Island, USA, 2002.

[17] J.S. Stadler. Adaptive importance sampling. *IEEE Journal on Selected Areas in Communications*, 11(3):309–316, April 1993.

[18] K. Labibes, Z. Papp, A.H.C. Thean, P.P.M. Lemmen, M. Dorrepaal, and F.J.W. Leneman. An integrated design and validation environment for intelligent vehicle safety systems (IVSS). In *Proc. of the 10th World Congress on Intelligent Transport Systems and Services (ITS)*, Madrid, Spain, November 16–20 2003. Paper 2731.

LIST OF TABLES

LIST OF FIGURES

δ	ϵ	N_{ch}	N_{wc}
0.10	0.10	150	22
0.05	0.10	185	29
0.03	0.10	210	34
0.02	0.10	231	38
0.01	0.10	265	44
0.002	0.10	346	59
0.05	0.05	738	59
0.02	0.05	922	77
0.01	0.05	1060	90
0.02	0.03	2559	129
0.01	0.01	26492	459
0.001	0.001	3800452	6905

TABLE 1 Values for ϵ, δ, N_{ch}, and N_{wc}, according to the Chernoff bound Eq. (7) and the worst-case bound Eq. (14).

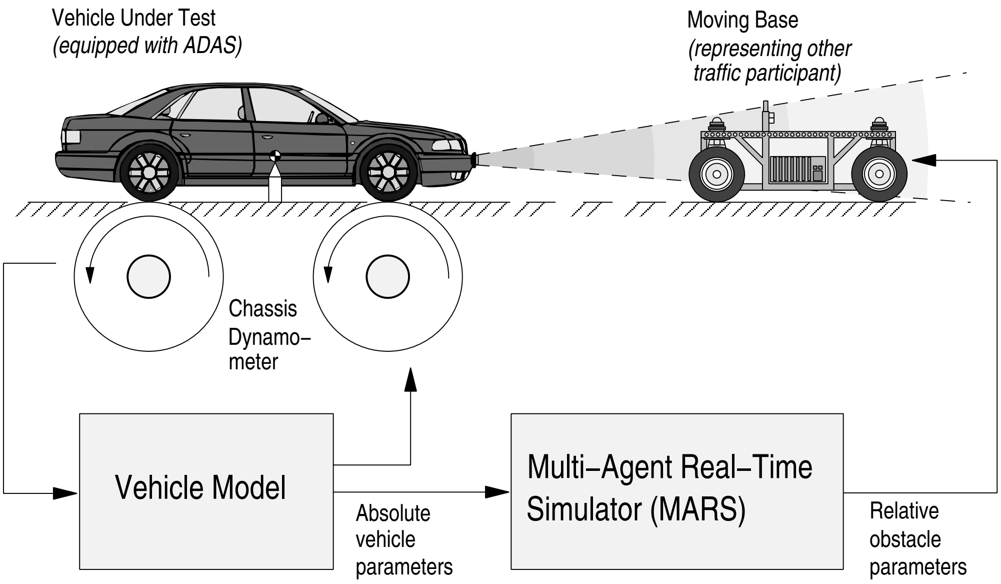


FIGURE 1 VEHIL working principle.

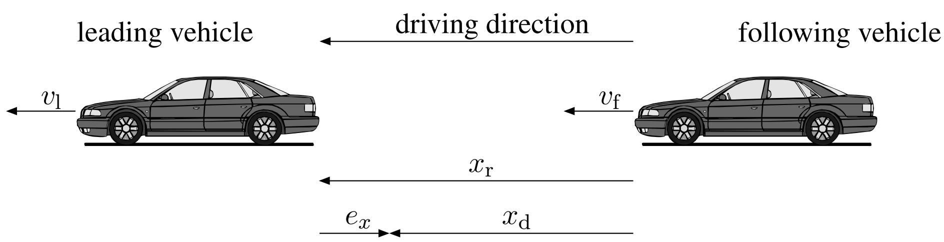
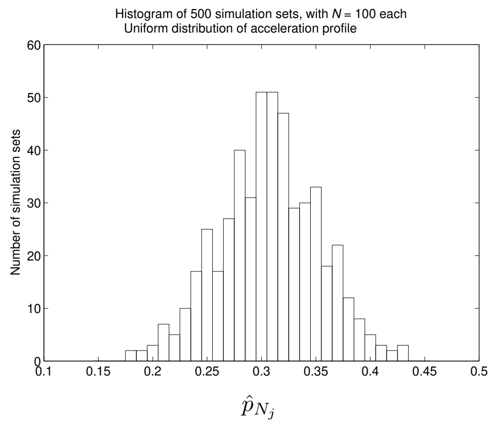
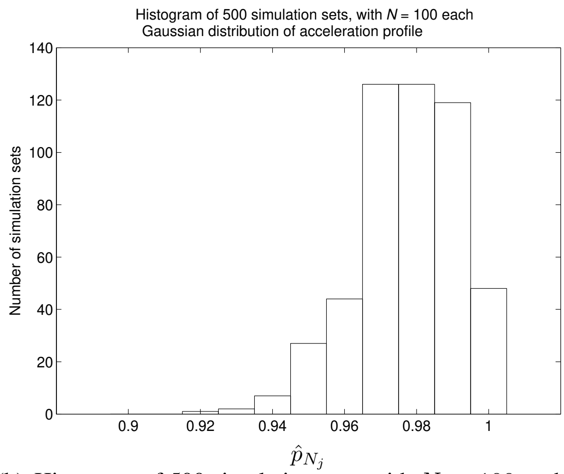


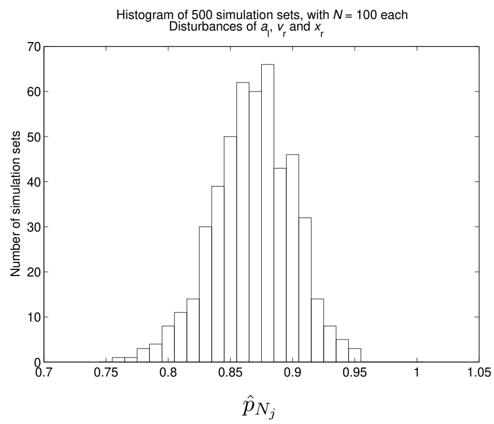
FIGURE 2 Two vehicles in ACC mode: one leader and one follower.



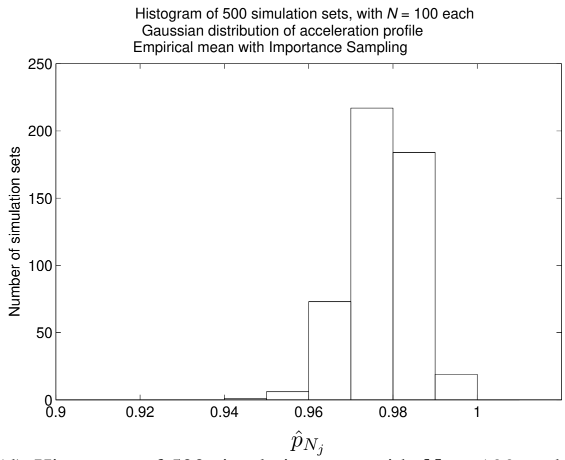
(a) Histogram of 500 simulation sets, with $N = 100$ each, where $a_l \sim \mathcal{U}(-10, 0)$.



(b) Histogram of 500 simulation sets, with $N = 100$ each, where the acceleration profile is sampled from a Gaussian distribution $\mathcal{N}(0, 1.5)$.



(c) Histogram of 500 simulation sets, with $N = 100$ each, with a multi-dimensional disturbance for $\mathcal{S} = \{v_l,\ v_f,\ x_r,\ a_l\}$.



(d) Histogram of 500 simulation sets, with $N = 100$ each, where the acceleration profile is sampled from an importance sampling distribution $\varphi = -0.005a_l + 0.05$.

FIGURE 3 Histograms for the estimate $\hat{p}_{N_j}$.

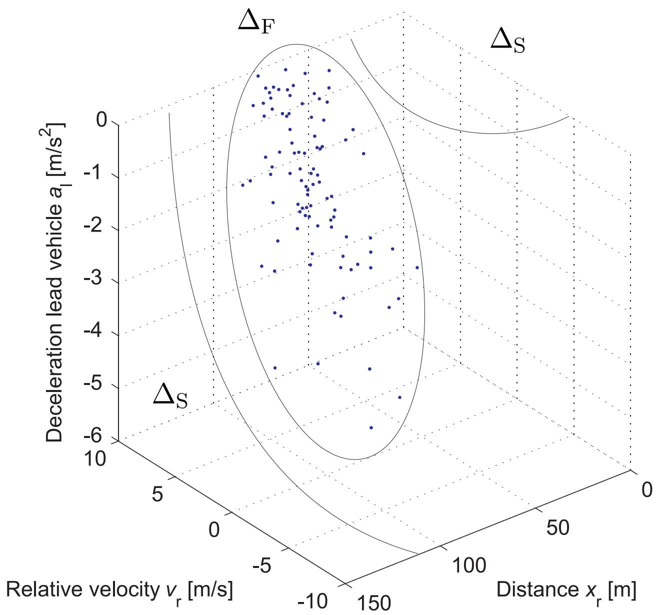


FIGURE 4 Scatter plot with the occurrence of collisions. Furthermore, the subspaces Δ_S and Δ_F of which the outcome can be a priori established as success or failures.

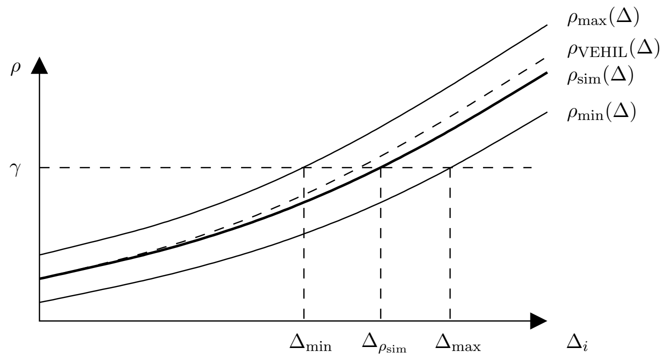


FIGURE 5 Dependency between the performance characteristic ρ and the specific value of Δ.