Natural Resources for Human Health
eISSN: 2583-1194

Original Research

Received 11 July 2026
Revised 12 Aug 2026
Accepted 26 Aug 2026

Natural Resources for Human Health 2026, Vol. 6 (19s): 1279–1298

KEYWORDS:
ISO 26262, ISO 21448, SOTIF, ASIL D, functional safety, sensor fusion, voting, watchdog, autonomous vehicles, work-zone safety, Monte Carlo simulation.

Confidence-Based Sensor Fusion and Watchdog Fallback for an ASIL D Construction-Zone Overspeed Hazard: A Simulation-Based Study

Srikanth Annamareddy[1], Gummadi Govardhan Reddy[2], Sony Karra[3], Sonti Thirupathamma[4], Dr. M. V. D. Prasad[5]*

[1]Professor of Practice, Department of ECE, K. L. University, India
[2]M. Tech Student, Department of ECE, K. L. University, India
[3]Assistant Professor, Department of ECE, K. L. University, India
[4]Ph.D. Research Scholar, Department of ECE, SRM University–AP, Amaravati, India
[5]Associate Professor, Department of ECE, KL University, Vaddeswaram, India
*Corresponding Author: Dr. M. V. D. Prasad
Email: mvd_ece@kluniversity.in

ABSTRACT: Construction zones ask an automated vehicle to drop from highway speed to a temporary limit, often marked by only a portable sign and a few cones. An earlier ISO 26262 case study rated this hazard ASIL D. In six runs, a camera-only design reached the zone boundary at about 99 km/h against a 40 km/h limit, while a three-path fusion design stayed near 40 km/h, but that study gave no decision rule. This paper defines one. Each sensing path (camera, LiDAR, HD map with GPS) reports a confidence score. One path above the threshold starts braking, and braking is released only when two healthy paths that can see the zone both read low. A watchdog forces emergency braking if a weak cue is not confirmed within 1.5 s. We tested the rule in a simulator we wrote in Python 3.12 using NumPy and SciPy. It is a 20 Hz kinematic and probabilistic model, so we did not use CARLA or CarMaker, and no sensor data is rendered. The simulator covers twelve operating conditions with 1,000 trials each and compares the rule with a camera-only baseline and three other voting rules on identical random draws. No measured or third-party dataset was used, and every input distribution is our own assumption. Under these assumptions, 97.0% of proposed-system trials cross the boundary at or below 42 km/h, against 57.8% for the baseline, and faults in all three paths remain the weak case at 75.3%. In 20,000 hazard-free approaches, the system issues a brief brake command in 1.25% and leaves a brake unreleased in 0.02%. Because the inputs are assumed, we report sensitivity sweeps and a 200-draw perturbation ensemble instead of significance tests. The results are simulation evidence for a proposed design and not a safety case.

I. INTRODUCTION

A highway work zone changes the rules quickly. A moment ago the road allowed 100 km/h, and now it allows 40 km/h, with only a portable sign and a row of cones to say so. A human driver slows down because the road looks different. An automated vehicle slows down only if its sensors pick up the sign and its software acts on what they report. If any link in that chain fails, the vehicle carries its cruise speed straight into the zone.

An earlier ISO 26262 case study [1] looked at this hazard. It defined the item as adaptive cruise control with traffic sign recognition and speed adaptation, carried out a hazard analysis and risk assessment, and rated the event "vehicle exceeds safe speed in a construction zone" as ASIL D. The study then wrote a safety concept around three diverse sensing paths, built an

FMEA and a fault tree, and checked the design in six closed-loop runs in CARLA [2], with CarMaker as a cross-check. The camera-only design reached the zone boundary at about 99 km/h, while the fusion design stayed near 40 km/h in every run.

Six runs show that the mechanism can work, but they say little about how often it works. They leave out rain, fog, night driving, a dirty lens, a damaged sign, a sign missing from the map, and faults that arrive together. The safety concept also says only that the item should cross-check the paths, without giving a rule for how the cross-check decides.

We take the architecture of that study as given and add what it lacks. This paper defines a decision rule, builds a simulation that exercises the rule across twelve conditions, and reports where the rule works and where it does not. Section IV gives the rule, Section V describes the simulation, and Section VI reports the results. Sections VII to X cover the discussion, limitations, future work, and conclusions.

The paper contributes the following.

• A confidence score for each of three diverse sensing paths (camera sign recognition, LiDAR cone and barrier clustering, HD-map geofence), including a self-diagnostic health factor.

• An asymmetric voting rule. Any one path above the threshold starts braking, and releasing the brake takes two healthy paths that can see the zone and both read low. This is a 1-out-of-3 rule to act combined with a 2-out-of-3 rule to clear. Voting of this family is long established in functional safety [3]–[5]. What we add is a version for confidence-scored perception paths, together with its evaluation on this hazard.

• A watchdog with a defined trigger. A weak cue from any path arms it. If no path reaches the act threshold within the 1.5 s fault tolerant time interval, the watchdog forces emergency braking.

• A 20 Hz simulation in which a camera-only baseline, plain 1-of-3 fusion, symmetric 2-of-3 fusion, the asymmetric rule without the watchdog, and the full proposal all run on identical random draws, so the contribution of each component can be seen on its own.

• A sensitivity study that varies the threshold, the fault tolerant time interval, the actuator and watchdog fault rates, a common-cause factor, and the set of available paths, together with a 200-draw perturbation ensemble over the input distributions.

II. BACKGROUND AND RELATED WORK

This section covers standards, validation practice for automated vehicles, redundancy and voting, and simulation platforms and work-zone material.

A. Functional Safety Standards

ISO 26262 [6], [7] governs the functional safety of electrical and electronic systems in road vehicles. The first edition appeared in 2011 and the second in 2018. The standard is built around a V-shaped development process, hazard analysis and risk assessment (HARA), automotive safety integrity levels (ASIL), and a chain that runs from safety goals to functional and technical safety concepts. It derives from IEC 61508 [3], and MISRA C [8] supplies the coding rules usually cited with its software part. ISO 21448 (SOTIF) [9] covers hazards that occur even when every component behaves as specified. A camera that cannot see an occluded sign is a SOTIF case. The earlier study treated occlusion that way, and so do we. The item sits at the driver-assistance end of the SAE J3016 scale [10], although the same fusion logic could serve higher automation levels.

FMEA scores severity, occurrence, and detection bottom-up into a risk priority number (RPN) [11]. Fault tree analysis works top-down from the hazardous event through gates to basic events [12]. Both describe design-time risk. Neither says what the vehicle does at runtime.

B. Automated Vehicle Testing and Validation

Koopman and Wagner [13] list five challenges that strain classical test-to-requirement mapping for automated vehicles: the driver is out of the loop, requirements are complex and incomplete, algorithms are non-deterministic, some components are learned, and fail-operational behavior is required. Their later papers [14], [15] argue for validation frameworks and safety standards suited to fully autonomous vehicles. Kalra and Paddock [16] estimate that road mileage alone would need hundreds of millions to hundreds of billions of miles to show reliability, which is the usual argument for adding simulation and scenario-based evidence. Wachenfeld and Winner [17] reach a similar conclusion about releasing autonomous vehicles. Junietz et al. [18]

compare validation approaches for automated driving, and Amersbach and Winner [19] propose decomposing approval into functions to cut effort.

The Responsibility-Sensitive Safety model [20] gives formal definitions of a dangerous situation and a proper response. Adaptive stress testing [21], [22] and rare-event simulation [23] search adversarially for failure-inducing scenarios. The twelve-condition sweep in this paper is a simpler relative of those methods. It uses a fixed set of conditions that we chose by engineering judgment.

C. Redundancy, Voting, and Sensor Fusion

Redundancy with voting is older than automotive safety. Von Neumann [24] analyzed majority voting over unreliable components, and Lyons and Vanderkulk [25] applied triple modular redundancy to computers. Functional-safety practice describes the result as k-out-of-n architectures. IEC 61508-6 [3] gives reliability models for 1oo1, 1oo2, 2oo2, 1oo2D, and 2oo3 arrangements, and Goble [4] and Rausand and Høyland [5] treat them in detail. In a 1oo2 arrangement, either channel can demand the safe state. In a 2oo3 arrangement, two channels must agree. The choice of k moves a design between missed demands and spurious trips. Distributed detection theory contains the same idea as counting rules for fusing local decisions [26]. The rule in this paper combines two of these choices, 1-out-of-3 for braking and 2-out-of-3 for release, and applies them to health-scored perception paths. IEC 61508-6 also supplies the common-cause beta factor that we use in Section VI-F.

Sensor fusion for perception grew from the Kalman filter [27] and its nonlinear extensions [28]. Bar-Shalom et al. [29] cover estimation for tracking, and Elmenreich [30] gives an introduction to sensor fusion. Camera-based sign recognition usually relies on convolutional detectors such as YOLO [31], and LiDAR scene representation draws on occupancy-grid methods [32]. These methods produce estimates. The rule in this paper decides what to do with the outputs of several detectors that can fail independently.

D. Simulation Platforms and Work Zones

CARLA [2] is an open-source simulator with configurable sensors and weather, and the earlier study used it with CarMaker. That study also used IBM DOORS Next [33] for traceability and Ansys Medini Analyse [34] for the FMEA and fault tree. Work-zone signing practice is set out in the MUTCD [35], and FHWA statistics [36] show work zones as a high-risk highway setting. Waymo's safety methodology [37] and NHTSA's ADS 2.0 guidance [38] show the mix of evidence a mature program combines: simulation, closed-course testing, on-road testing, and a safety argument that ties them together.

III. RESEARCH GAP

Three gaps come out of the material above.

1) Rates. Six runs cannot estimate a detection rate, a false-alarm rate, or the spread of residual risk. Koopman and Wagner [13] make this point in general terms but give no worked example for a specific ASIL D hazard.

2) Decision logic. The FMEA and fault tree give static numbers, an RPN and a per-hour probability. The safety concept describes runtime behavior in words such as "cross-check" and "plausibility voter". Nothing connects the two quantitatively, so it is hard to ask what happens to the false-alarm rate when the rule asks for two agreeing paths and what happens when it asks for one.

3) Independence and conditions. The earlier study ran clear weather, a camera failure with and without fallback, a GPS error, a partly occluded sign, and a camera failure in rain. Night, fog, a damaged sign, an unmapped sign, LiDAR degradation, and combined faults were not covered. The benefit of fusion also depends on the paths failing independently, and six runs cannot test that.

Sections IV to VI address these gaps.

IV. METHOD

Fig. 1 shows the architecture. It keeps the three sensing paths of the earlier study and its chain from sensing to a fusion and diagnostic unit, then to a decision and fallback unit, then to the actuators. Fig. 2 places the work of this paper on the ISO 26262 V-model. The simulation is a software-level fault-injection exercise. Hardware-in-the-loop testing, system validation, and the safety case review all remain open.

Natural Resources for Human Health
eISSN: 2583-1194

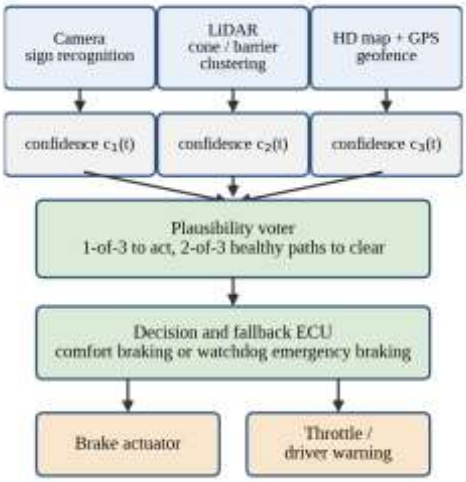


Fig. 1. Proposed architecture: three sensing paths feed per-path confidence scores, a plausibility voter, and a decision and fallback unit.

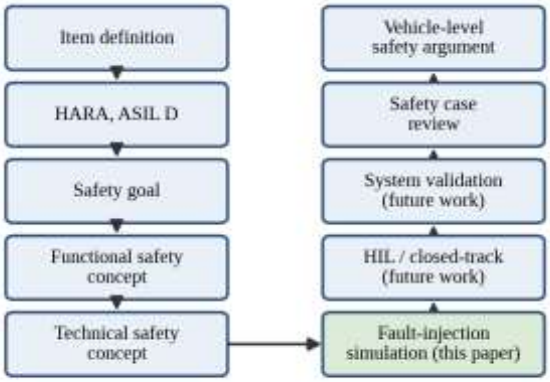


Fig. 2. Work in this paper placed on the ISO 26262 V-model.

A. Path Confidence

Each path i (camera, LiDAR, map) reports a confidence $c_i(t)$ between 0 and 1 at every 20 Hz cycle. The value combines the filtered plausibility signal $s_i(t)$ that the path sees a construction-zone restriction, a self-diagnostic health factor d_i between 0 and 1, and a fixed prior weight w_i:

$$c_i(t) = w_i \cdot s_i(t) \cdot d_i \qquad (1)$$

The health factor drops when a path flags contamination, low light, a poor GPS fix, or a fault code. The weights are 1.0 for camera and LiDAR and 0.85 for the map path. We gave the map a lower weight because the hazard is a physical, in-the-moment observation and a geofence is a prior. A path with nothing to report has c close to 0.

Each path also has a detection range D_i. Until the vehicle is within D_i of the zone boundary, the path is out of range and cannot see the zone. A healthy path in range that reads low is evidence that nothing is there. A path that is out of range or unhealthy is silent, and silence is no evidence at all.

Natural Resources for Human Health
eISSN: 2583-1194

B. Asymmetric Voting

A missed detection can end in an S3 outcome, with workers or barriers struck at highway speed. A false trigger costs a brief and unnecessary deceleration. The rule therefore acts on weak evidence and asks for strong evidence before it stands down. With act threshold τ_{thr} and clear threshold τ_{clr}, the rule is:

$$ACT \Leftrightarrow |\{ i : c_i(t) > \tau_{thr} \}| \geq 1 \quad (2)$$

$$CLEAR \Leftrightarrow |\{ i : r_i(t) = 1,\ d_i \geq d_{min},\ c_i(t) < \tau_{clr} \}| \geq 2 \quad (3)$$

Here $r_i(t) = 1$ when path i is in range. Eq. (2) starts nominal braking as soon as one path crosses τ_{thr}. Braking is confirmed, and can no longer be released, once two paths are above τ_{thr}. Eq. (3) releases braking and returns the vehicle to nominal driving when at least two healthy in-range paths have read below τ_{clr} for $T_{hold} = 0.5$ s. A path that triggered a released event is muted until its confidence falls below τ_{arm}, so a persistent false cue cannot restart braking every frame. Fig. 3 shows the states.

The values are $\tau_{thr} = 0.55$, $\tau_{clr} = 0.20$, $\tau_{arm} = 0.25$, and $d_{min} = 0.6$. We fixed them before running any experiment. Section VI-F shows how far the results move when τ_{thr} changes.

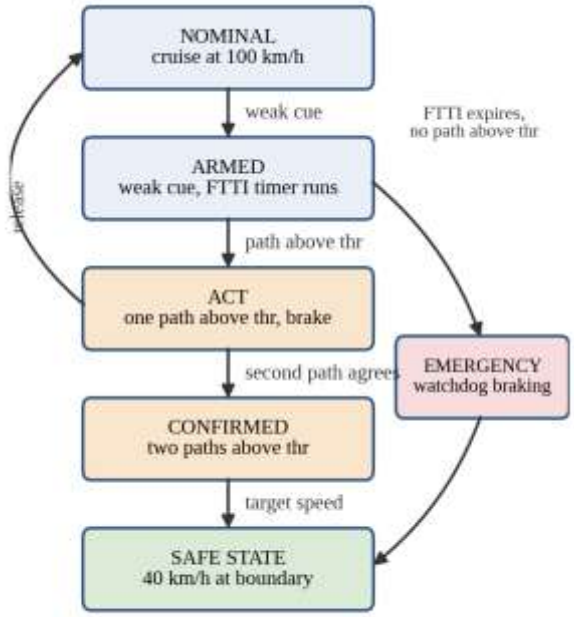


Fig. 3. Runtime states of the voter and watchdog. Release from ACT to NOMINAL follows Eq. (3). "thr" stands for τ_{thr}.

C. Watchdog Fallback

Eqs. (2) and (3) need at least one path to produce a signal. The watchdog covers the case where no path reaches τ_{thr}. It is armed the first time any path exceeds τ_{arm}, which is a weaker cue than the act threshold. If no path has reached τ_{thr} when the fault tolerant time interval (FTTI) of 1.5 s has run since arming, the watchdog forces emergency braking at 5.5 m/s². Braking already under way makes the watchdog unnecessary, and a release under Eq. (3) disarms it.

The watchdog cannot help if every path stays below τ_{arm}, because it never arms. We also do not treat it as a guaranteed save. Braking capability depends on friction and on the actuator. The simulation gives the watchdog timer a failure probability of 0.18% per approach and gives the brake actuator an uncovered fault probability of 0.18%. ISO 26262 Part 5 likewise treats the diagnostic coverage of any real mechanism as below 100%.

Natural Resources for Human Health
eISSN: 2583-1194

D. Inputs from the Earlier Study

Tables I and II reproduce the FMEA and fault tree figures of the earlier study [1]. Fig. 4 shows the fault tree and Fig. 5 the share of RPN by failure mode. We use these figures for context and for the discussion in Section VII. They enter the simulation through the conditions in Table IV and through nothing else.

TABLE I FMEA OF THE EARLIER CASE STUDY

Failure mode	S	O	D	RPN
Sign partially occluded (foliage, truck)	8	6	5.6	270
Sensor fusion disagreement unresolved	8	3	5.6	135
The camera misses or misreads the sign	8	3	4.5	108
Software crash in recognition module	9	2	5.0	90
HD map / GPS position error	7	4	3.0	84
LiDAR sensor failure	6	2	3.0	36

S = severity, O = occurrence, D = detection difficulty (each 1–10). RPN = S × O × D, from the earlier study [1].

TABLE II FAULT TREE BASIC EVENTS OF THE EARLIER CASE STUDY

Basic event	Probability (per hour)
Traffic sign occluded	1.0×10^{-3}
HD map outdated / sign unmapped	1.0×10^{-4}
Camera sensor failed	1.0×10^{-5}
Image processing error	5.0×10^{-6}
Sensor fusion logic failed	3.0×10^{-6}
Software bug in recognition pipeline	2.0×10^{-6}

OR-gate logic: any basic event alone reaches the top event (Fig. 4).

Natural Resources for Human Health
eISSN: 2583-1194

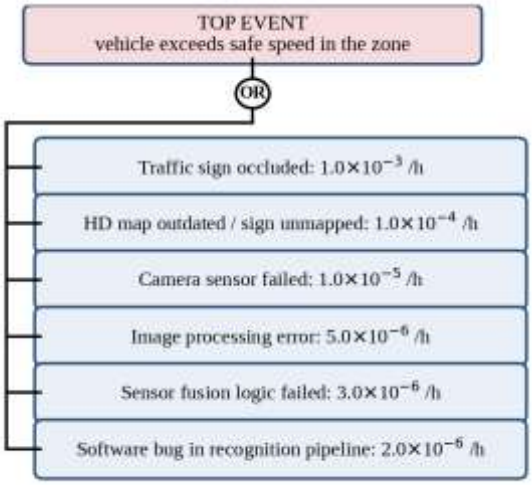


Fig. 4. Fault tree for the overspeed hazard from the earlier case study [1].

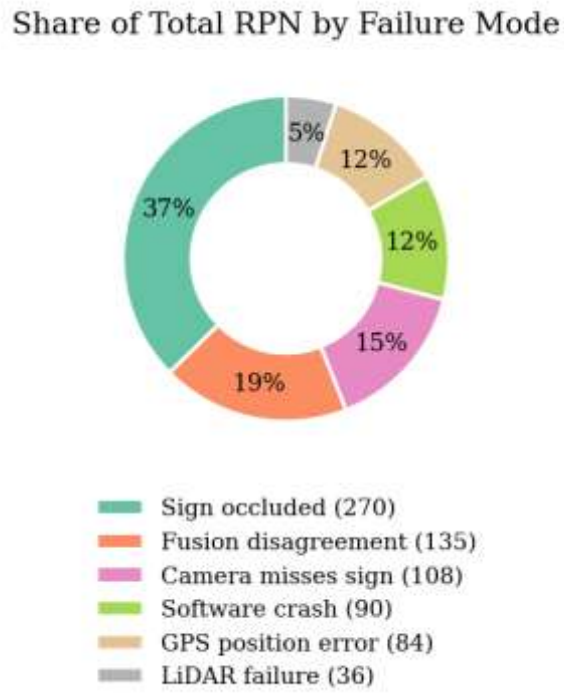


Fig. 5. Share of total RPN by failure mode (values from Table I).

V. SIMULATION SETUP

The earlier study ran CARLA and CarMaker. We ran neither. The simulator in this paper is a Python program that we wrote for this study. It is a time-stepped model of the approach and of the three sensing paths, and it produces no rendered sensor data. We chose this form because it lets us run 12,000 hazard trials and 20,000 control trials for every voting rule in a few minutes, all on common random numbers.

A. Approach and Vehicle Model

Fig. 6 shows the geometry. The ego vehicle cruises at 100 km/h and starts 250 m before the zone boundary, where a 40 km/h limit applies. The model steps at 20 Hz. The vehicle holds cruise speed until braking starts, then decelerates at 4.0 m/s² (nominal) or 5.5 m/s² (watchdog). Each rate is multiplied by a friction factor μ for the condition and by a per-trial factor with mean 1 and

SD 0.06. Braking starts after an actuator delay of 150 ms (SD 30 ms). Once at 40 km/h the vehicle holds that speed with Gaussian tracking noise (SD 0.55 km/h). After a release the vehicle accelerates again at 1.0 m/s². We record the speed at the boundary. With $\mu = 1$, braking must start within about 5.9 s of the start of the approach to reach 40 km/h at the boundary.

A brake actuator fault with probability 0.18% per approach makes every braking command ineffective, for both systems. A watchdog timer fault with the same probability disables the watchdog of the proposed system.



Fig. 6. Approach geometry: 250 m of warning before the zone boundary.

B. Sensing-Path Model

Each path has a steady signal level a_i, a detection range D_i, a rise time T_i, a frame noise σ_i, and a health d_i. Once the vehicle is within range, the raw signal is $a_i(1 - \exp(-t_{in}/T_i))$ plus Gaussian frame noise, where t_{in} is the time spent in range. An exponential filter with coefficient 0.3 turns the raw signal into $s_i(t)$. For each trial, a_i and d_i are drawn from normal distributions around the condition means (SD 0.04 to 0.06 for a_i, 0.05 for d_i), and D_i has an SD of 12% of its mean. Table III lists the clear-road values. The HD map contains the zone with probability 0.95. When the zone is missing, the map path is healthy but silent, with a_{map} drawn between 0.05 and 0.15. Every number in this subsection is our assumption. None was fitted to measured sensor data.

TABLE III CLEAR-ROAD PATH PARAMETERS

Path	a (SD)	D (m)	d	σ	T (s)	w
Camera	0.92 (0.04)	200	0.98	0.05	0.6	1.00
LiDAR	0.85 (0.06)	120	0.98	0.05	0.6	1.00
HD map / GPS	0.90 (0.04)	250	0.95	0.03	0.3	0.85

a = steady signal level, D = detection range, d = health factor, σ = frame noise, T = rise time, w = prior weight.

C. Operating Conditions

Table IV lists the twelve conditions. Each changes some of the clear-road values. The multiplier μ scales the achievable deceleration. Condition 8 sets the probability that the zone appears in the map to 0.05. Conditions 5 and 7 lower the camera signal level while the health factor stays high, because a camera does not know that it is looking at an occluded or damaged sign. This is the SOTIF case. Condition 9 removes the camera. A thirteenth case, camera failure in rain, is used only for the comparison in Section VI-A.

TABLE IV Operating Conditions

#	Condition	μ	Changes from clear-road values
1	Clear / Sunny	1.00	none

Natural Resources for Human Health
eISSN: 2583-1194

#	Condition	μ	Changes from clear-road values
2	Rain	0.78	camera a 0.8, D 150, d 0.85, σ 0.08; LiDAR a 0.78, D 100, d 0.85
3	Night	0.96	camera a 0.78, D 140, d 0.8, σ 0.08
4	Fog	0.92	camera a 0.65, D 105, d 0.6, σ 0.1; LiDAR a 0.6, D 75, d 0.6, σ 0.09
5	Occluded sign	1.00	camera a 0.45, D 130, d 0.9, σ 0.07
6	Dirty lens	1.00	camera a 0.6, D 170, d 0.5, σ 0.1
7	Damaged sign	1.00	camera a 0.45, D 150, d 0.9, σ 0.08
8	Unmapped temp. sign	1.00	zone in map with probability 0.05
9	Camera failure	1.00	camera a 0, d 0
10	LiDAR degradation	1.00	LiDAR a 0.5, D 70, d 0.5, σ 0.09
11	GPS inaccuracy	1.00	map a 0.55, d 0.55, σ 0.05
12	Multiple faults	0.80	camera a 0.6, D 90, d 0.55, σ 0.1; LiDAR a 0.55, D 60, d 0.55, σ 0.09; map a 0.55, d 0.55, σ 0.05

Units follow Table III. Unlisted values keep their clear-road setting.

D. Systems Compared

Five decision rules run on identical random draws. The baseline is a camera-only design that brakes when the raw camera signal exceeds τ_{thr}. It has no health factor, no second opinion, and no watchdog. The 1oo3 rule uses the fused confidences of Eq. (1) and Eq. (2) with no release and no watchdog. The 2oo3 rule is the same but needs two paths above τ_{thr} before it brakes. The asymmetric rule adds the release of Eq. (3). The proposed system adds the watchdog.

E. Hazard-Free Control Trials

To study false alarms we run 20,000 approaches with no work zone. In each one, every path raises a spurious candidate with probability 0.8% (camera), 0.3% (LiDAR), or 0.2% (map). A candidate is a burst of signal with amplitude between 0.55 and 0.95 that lasts 0.3 to 1.5 s and starts at a random time between 0.5 s and 7 s. These rates are our assumptions. We report two outcomes: a brake command, meaning any braking command at all, and an unreleased brake, meaning one that is still active when the vehicle reaches the boundary.

F. Metrics

For every condition and system we report the share of trials that command braking before the boundary, the mean and SD of boundary speed, the share at or below 42 km/h (the 40 km/h limit plus a 2 km/h operating margin, following the reduced-margin pass in the earlier study), the share above 70 km/h (a coarse indicator of a high-energy entry, not a crash simulation), and the mean decision time. For pooled comparisons we add the compliance of the worst single condition. Differences in mean boundary speed carry Welch 95% intervals [39], and proportions carry Wilson 95% intervals [40]. Every rule sees the same draws, so a paired analysis would give tighter intervals than Welch; we use Welch as the more conservative choice. The ROC analysis scores each unbraked approach by the maximum confidence it reaches.

G. Reproducibility and Statistical Reporting

All random draws come from the NumPy PCG64 generator [41] with seed 20260725, and every rule sees the same draws. The code runs in about five minutes on one CPU core with Python 3.12.3 [42], NumPy 2.4.4 [43], and SciPy 1.17.1 [44]. The code and seeds are supplied as supplementary material. We report no p-values. Every input is an assumption, so extra trials only narrow the intervals around a model quantity, and a p-value below 0.001 would say nothing about a real vehicle. The sensitivity runs of Section VI-F check how far the conclusions depend on the inputs.

H. Software, Data and Provenance

The simulator is written in Python 3.12.3 [42] with NumPy 2.4.4 [43] and SciPy 1.17.1 [44], all of which are open source. No measured sensor recordings, map data, or third-party datasets were used. Every sensor, fault, friction, and braking value in Tables III and IV is our own engineering assumption. It was chosen to be plausible and was not fitted to data. The only numbers taken from earlier work are the six case-study results in Table V and the FMEA and fault tree values in Tables I and II, all of which come from the earlier study [1]. CARLA [2], CarMaker, IBM DOORS Next [33], and Ansys medini analyze [34] belong to that earlier study, and none of them was used here.

VI. RESULTS

A. Comparison with the Earlier Case Study

Table V sets the six results of the earlier study beside the matching cases in this simulation. The parameters in Tables III and IV were fixed before this comparison and we did not adjust them afterwards. The baseline reaches the boundary at 40.1 km/h in clear conditions and at 100.0 km/h with the camera failed. The proposed system stays between 40.1 and 41.4 km/h in the four fusion cases. These values sit in the range of the earlier results, but the model gives about 40 km/h whenever braking succeeds, so this is a plausibility check and not a validation. They say nothing about the other conditions, which the earlier study never ran.

TABLE V Earlier Case Study And This Simulation

ID	Case in earlier study	Earlier result	This model
TC01	Clear weather, camera-only	~40 km/h, pass	40.1 km/h
TC02	Camera failure, no fallback	~99 km/h, fail	100.0 km/h
TC03	Camera failure, fusion and fallback	~40 km/h, pass	40.2 km/h
TC04	GPS position error, fusion	~40–46 km/h, pass	40.1 km/h
TC05	Sign partly occluded, fusion	~41 km/h, pass, reduced margin	40.2 km/h

Natural Resources for Human Health
eISSN: 2583-1194

ID	Case in earlier study	Earlier result	This model
TC06	Camera failure with rain, fusion	~42 km/h, pass, reduced margin	41.4 km/h

This model = mean boundary speed over 1,000 trials (baseline for TC01 and TC02, proposed system for TC03 to TC06).

B. Boundary Speed

Table VI gives the per-condition results and Fig. 7 plots them. The baseline fails in the conditions that hide the sign from the camera. It commands braking in only 15.3% of occluded-sign trials and 26.2% of damaged-sign trials, and never with the camera failed. Fog delays its decision enough that only 9.7% of fog trials reach the boundary at or below 42 km/h. Pooled over the twelve conditions the baseline complies in 57.8% of trials, with 32.2% above 70 km/h. The pooled figure gives each of the twelve conditions equal weight, so it depends on which conditions we chose, and the per-condition results in Table VI are the more reliable comparison.

The proposed system complies in 97.0% of pooled trials, with 2.3% above 70 km/h. It commands braking in at least 99.0% of trials in every condition. Two conditions stand out. In fog it complies in 95.3% of trials. With faults in all three paths it complies in 75.3%, and 22.8% of trials arrive above 70 km/h, because braking starts late (mean decision time 4.25 s). Fig. 8 shows the full distributions for six conditions

TABLE VI Boundary Speed By Condition, 1,000 Trials Each

Condition	Base: brake %	Base: mean (SD)	Base: ≤ 42 %	Base: > 70 %	Prop: brake %	Prop: mean (SD)	Prop: ≤ 42 %	Prop: > 70 %	Mean difference [95% CI]
Clear / Sunny	100.0	40.1 (2.0)	99.9	0.1	100.0	40.1 (2.0)	99.9	0.1	0.00 [-0.17, 0.17]
Rain	100.0	41.5 (5.2)	86.5	0.4	100.0	40.3 (3.2)	98.5	0.3	1.29 [0.91, 1.67]
Night	100.0	40.7 (4.8)	95.0	0.5	100.0	40.3 (4.3)	99.3	0.5	0.40 [0.01, 0.80]
Fog	100.0	59.4 (11.4)	9.7	17.9	100.0	41.7 (8.6)	95.3	3.1	17.68 [16.79, 18.56]
Occluded sign	15.3	95.8 (12.1)	1.4	92.9	100.0	40.2 (1.6)	98.7	0.0	55.66 [54.90, 56.41]
Dirty lens	100.0	40.8 (4.9)	95.7	0.9	100.0	40.2 (2.8)	99.4	0.2	0.63 [0.28, 0.98]
Damaged sign	26.2	91.0 (18.2)	5.7	84.9	100.0	40.0 (2.0)	99.8	0.1	50.94 [49.80, 52.07]
Unmapped temp. sign	100.0	40.1 (2.0)	99.9	0.1	100.0	40.1 (2.0)	99.9	0.1	0.00 [-0.17, 0.17]
Camera failure	0.0	100.0 (0.0)	0.0	100.0	100.0	40.2 (3.0)	98.7	0.2	59.75 [59.56, 59.94]
LiDAR degradation	100.0	40.1 (2.0)	99.8	0.1	100.0	40.1 (2.0)	99.8	0.1	0.00 [-0.17, 0.17]
GPS inaccuracy	100.0	40.1 (2.7)	99.8	0.2	100.0	40.1 (2.7)	99.8	0.2	0.00 [-0.24, 0.24]

Natural Resources for Human Health
eISSN: 2583-1194

Condition	Base: brake %	Base: mean (SD)	Base: ≤ 42 %	Base: > 70 %	Prop: brake %	Prop: mean (SD)	Prop: ≤ 42 %	Prop: > 70 %	Mean difference [95% CI]
Multiple faults	98.9	79.2 (8.0)	0.0	88.7	99.0	51.5 (20.9)	75.3	22.8	27.70 [26.31, 29.08]

Speeds in km/h. Brake % = share of trials with braking commanded when the vehicle reaches the boundary. Mean difference = baseline minus proposed.

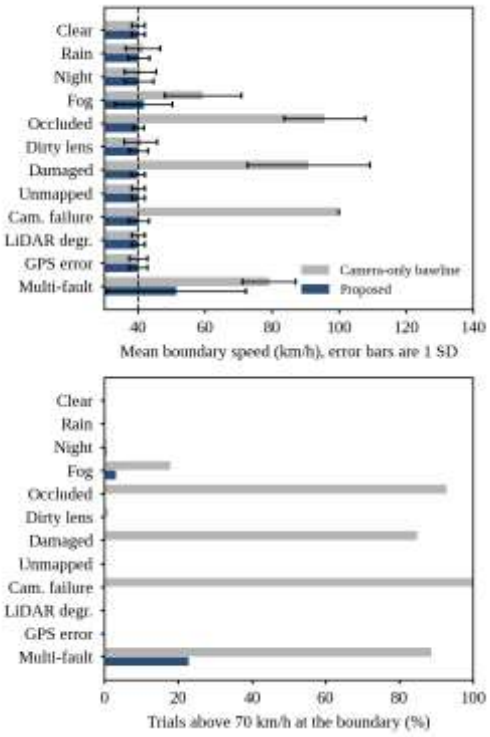


Fig. 7. Boundary speed (top) and share of trials above 70 km/h (bottom) by condition, 1,000 trials per bar.

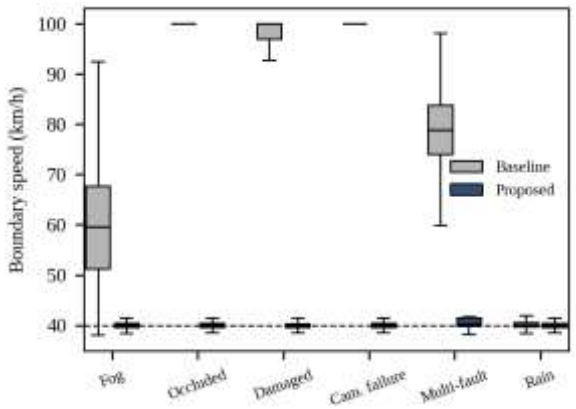


Fig. 8. Boundary speed distributions for six conditions. Boxes show quartiles and the median. Whiskers follow the 1.5 × IQR rule and outliers are omitted.

Natural Resources for Human Health
eISSN: 2583-1194

C. Decision Time

Fig. 9 shows when each system decides. In most conditions the proposed system decides after a mean of 1.0 to 1.3 s. That early decision comes from the map path, which is in range from the first frame, so it rests on the assumption that the zone is in the map (Section VI-F). Where the map is silent (condition 8) the decision comes from the camera at about 2.5 s, the same time as the baseline. With GPS error the map confidence is weak, and in 27.4% of trials the watchdog fires before the camera confirms. With faults in all three paths the mean decision time is 4.25 s and the watchdog acts in 99.0% of trials. Fig. 10 shows one such trial.

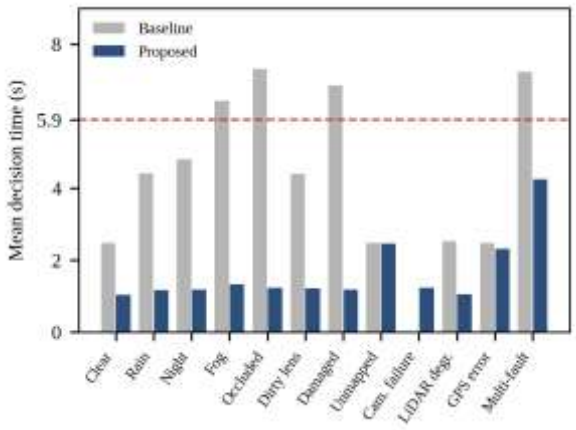


Fig. 9. Mean decision time by condition. The dashed line marks the latest braking start (5.9 s) that still gives 40 km/h at the boundary with $\mu = 1$. The baseline never decides with the camera failed.

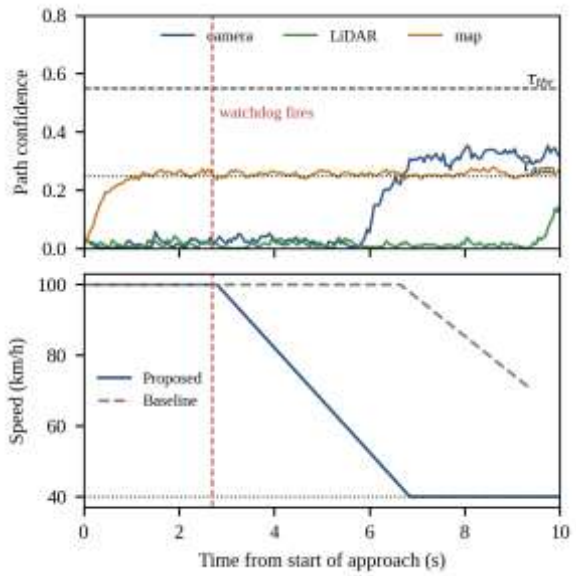


Fig. 10. One trial with faults in all three paths. Top: path confidences and the two thresholds. Bottom: speed of the proposed system and of the baseline on the same random draws. The map path exceeds τ_{arm} at 1.2 s and the watchdog fires 1.5 s later.

D. Voting Rules and False Alarms

Table VII compares the five rules. The camera-only baseline complies in 57.8% of hazard trials. Plain 1oo3 fusion raises that to 90.6%, but every spurious candidate becomes braking that nothing releases: 1.25% of control approaches, with a mean speed loss of 56 km/h. Symmetric 2oo3 fusion produces no brake commands in the control trials and complies in only 67.9% of hazard trials, because a hazard seen by a single path is ignored. The asymmetric rule keeps the 1oo3 hazard performance (90.6%)

and releases every false command, after a mean of 0.93 s and with a mean speed dip of 10.8 km/h. Adding the watchdog lifts pooled compliance to 97.0% and worst-condition compliance from 0.0% to 75.3%.

The asymmetric rule has a price. It issues more brake commands than the baseline (1.25% of control approaches for the proposed system, 0.87% for the baseline), each one short-lived. The proposed system leaves a brake unreleased in 0.02% of control approaches, against 0.87% for the baseline.

TABLE VII VOTING RULES ON IDENTICAL RANDOM DRAWS

Rule	Hazard ≤ 42 %	Hazard > 70 %	Worst cond. ≤ 42 %	Ctrl: brake cmd %	Ctrl: unrel. %
Camera-only baseline	57.8	32.2	0.0	0.87	0.87
Fusion 1oo3, no release	90.6	8.9	0.0	1.25	1.25
Fusion 2oo3	67.9	19.2	0.0	0.00	0.00
Asymmetric, no watchdog	90.6	8.9	0.0	1.25	0.00
Proposed (with watchdog)	**97.0**	**2.3**	**75.3**	**1.25**	**0.02**

Hazard columns pool 12,000 trials (12 conditions). Ctrl = control trials: 20,000 hazard-free approaches, brake cmd = any brake command, unrel. = brake unreleased at the boundary.

E. Detection Quality

Fig. 11 shows ROC curves for the maximum confidence reached during an unbraked approach. The fused score has an AUC of 0.997 and the camera-only score 0.955. Neither curve separates the classes perfectly. Some hazard approaches, mostly in the fog and multiple-fault conditions, stay below the confidence of a spurious candidate. At τ_{thr} = 0.55 the fused score detects 91.3% of hazard approaches at a false-alarm rate of 1.25%, and the camera-only score detects 78.4% at 0.87%. The camera-only AUC in the camera-failure condition is 0.50, as it should be.

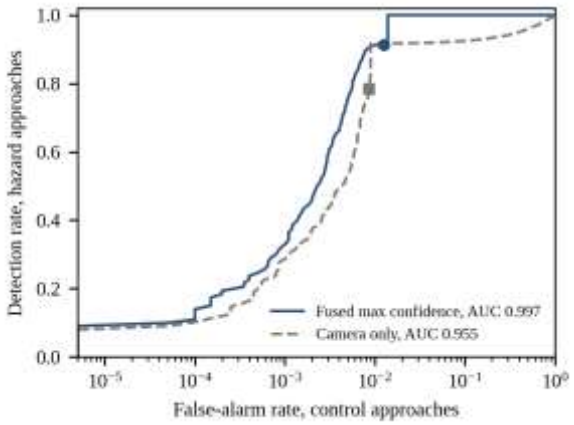


Fig. 11. ROC curves of the maximum confidence reached during an unbraked approach. Markers show τ_{thr} = 0.55.

Natural Resources for Human Health
eISSN: 2583-1194

F. Sensitivity to Assumptions

Threshold. Fig. 12 varies τ_{thr} from 0.35 to 0.75. Pooled compliance stays between 96.6% and 97.2%, and worst-condition compliance changes little (Fig. 12). The control brake-command rate falls from 1.31% to 0.61% as the threshold rises, while the share of approaches with an unreleased brake rises from 0.02% to 0.16%. Over this range the threshold trades nuisance activations against unreleased false brakes and leaves hazard performance almost unchanged.

Fault tolerant time interval. Shortening the FTTI from 3.0 s to 1.0 s raises pooled compliance from 96.6% to 97.5% and worst-condition compliance from 73.2% to 78.6%. Control behavior barely moves. The "All paths present" and "FTTI 1.5 s" rows of Table VIII describe the same setting and differ by 5.0 points in worst-condition compliance, so we read the FTTI trend in that column as within run-to-run noise.

Fault rates. Raising the brake actuator fault probability from 0 to 5% lowers proposed-system compliance from 96.9% to 92.2% and baseline compliance from 57.8% to 54.8%. The watchdog timer fault probability matters less: 96.7% at 0 and 96.3% at 5%.

Common cause. A beta factor β is the probability that one event degrades all three paths at once (signal levels times 0.4, health times 0.6). At $\beta = 0.10$ proposed compliance falls to 88.1% and high-energy entries rise to 11.1%. The baseline falls from 57.6% to 52.2%. The redundancy gain survives but shrinks as the paths share failures.

Path removal. Removing the camera path gives 93.6% compliance, removing LiDAR gives 96.4%, and removing the map path gives 76.8% with 13.2% high-energy entries. In this model the map path supplies most of the early warning, because it is in range from the first frame and holds the zone in 95% of trials. No other single input has this much influence.

Perturbation ensemble. We drew 200 sets of inputs: signal levels and detection ranges within ±15%, health factors within ±10%, frame noise times 0.7 to 1.5, spurious rates times 0.5 to 2, and both fault probabilities log-uniform between 0.05% and 1%. The proposed system beat the baseline in pooled compliance in 100% of the draws (Fig. 13). Its pooled compliance had a median of 96.7% (5th to 95th percentile 92.7% to 98.2%) against 57.8% (53.3% to 62.4%) for the baseline. Across all 200 draws the proposed system never fell below 90.5% and the baseline never rose above 68.1%. The worst-condition compliance of the proposed system ranged from 10.0% to 92.8% with a median of 73.8%, so the multiple-fault result depends heavily on the assumed inputs, and we do not conclude its value. The median control brake-command rate was 1.31% for the proposed system and 0.78% for the baseline, and the median unreleased rate for the proposed system was 0.04%.

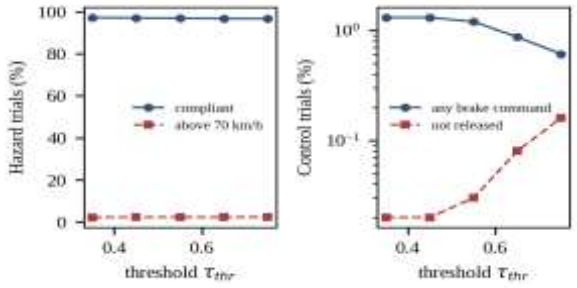


Fig. 12. Effect of the act threshold τ_{thr}. Left: hazard trials. Right: control trials (log scale).

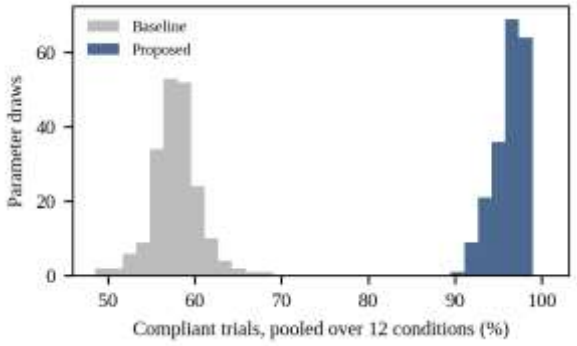


Fig. 13. Pooled compliance across 200 draws of the input distributions.

Natural Resources for Human Health
eISSN: 2583-1194

TABLE VIII ONE-AT-A-TIME SENSITIVITY RUNS

Setting	Prop. ≤ 42 %	Prop. > 70 %	Worst cond. ≤ 42 %	Base ≤ 42 %
All paths present	96.8	2.4	73.2	–
FTTI 1.0 s	97.5	1.7	78.6	–
FTTI 1.5 s	97.2	2.1	78.2	–
FTTI 2.0 s	97.1	2.3	77.4	–
FTTI 3.0 s	96.6	2.5	73.2	–
Actuator fault 0%	96.9	2.2	74.6	57.8
Actuator fault 0.18%	96.7	2.4	74.6	57.6
Actuator fault 1%	95.9	3.2	74.0	57.1
Actuator fault 5%	92.2	6.9	70.6	54.8
$\beta = 0$	96.9	2.4	74.0	57.6
$\beta = 0.02$	95.3	4.0	72.6	56.6
$\beta = 0.05$	92.6	6.7	70.4	54.9
$\beta = 0.10$	88.1	11.1	67.6	52.2
Camera removed	93.6	3.4	73.2	–
LiDAR removed	96.4	2.7	73.2	–
Map path removed	76.8	13.2	0.0	–

Each run uses 500 hazard trials per condition and 10,000 control trials, so the base case differs slightly from Table VII.

VII. DISCUSSION

A. What the Results Show

The proposed rule helps in every draw we tried, and the size of the help depends on the inputs. Across the perturbation ensemble the pooled compliance of the proposed system never fell below 90.5%, while the baseline never rose above 68.1%. The gap is large enough that its direction is safe to state for this model. The exact percentages need more evidence than this model can give.

Two design pieces do most of the work. The asymmetric release keeps the hazard performance of plain 1oo3 fusion and removes the sustained false braking that plain 1oo3 causes. The watchdog handles conditions in which no path is confident. With faults in all three paths the asymmetric rule alone complies in 0.0% of trials and the full proposal in 75.3%.

B. Cost of the Asymmetry

The proposed system issues brake commands in 1.25% of hazard-free approaches, against 0.87% for the baseline. Nearly all of them are released within about a second, with a mean speed dip of 10 km/h. That seems acceptable for an S3 hazard, but whether it is acceptable in a product depends on a nuisance-braking requirement, and this study does not have one.

C. Dependence on the Map Path

Removing the map path drops pooled compliance from about 97% to 76.8%. Two assumptions drive this. The map path is in range from the first frame, because its 250 m range equals the approach length, and it holds the zone in 95% of trials. A temporary work zone is often missing from a real map, so the 95% availability we assumed deserves the closest scrutiny in any follow-up. Condition 8 shows that camera and LiDAR keep the system compliant (99.9%) when the map is silent in that one condition, while the ablation shows the cost of losing the map everywhere.

D. Simultaneous Faults

The design does not solve the case of faults in all three paths. It only bounds the outcome in some trials by forcing emergency braking after the FTTI. Compliance there is 75.3% with 22.8% of trials above 70 km/h, and the ensemble shows the value swinging between 10.0% and 92.8% as the assumptions move. A real claim for this condition needs measured data.

E. Relation to the FMEA and Fault Tree

The FMEA ranks sign occlusion highest (RPN 270) and the fault tree gives it the highest basic-event probability (about 1.0×10^{-3} per hour). In the simulation, occlusion and sign damage are where the camera-only baseline fails most, with braking commanded in only 15.3% and 26.2% of trials. The proposed system reaches the boundary at or below 42 km/h in 98.7% and 99.8% of those trials, because the map and LiDAR paths see the zone. The static analysis and the simulation point at the same condition. The simulation takes its severity from assumptions and draws nothing from the FMEA numbers, so we treat the agreement as a plausibility check only.

F. Relation to the Literature

Kalra and Paddock [16] argue against mileage as the only source of evidence. A disclosed simulation can add to a safety case next to the on-road and closed-track evidence that Waymo and NHTSA describe [37], [38]. It cannot replace that evidence. The results reproduce a trade-off that the voting literature already knows [4], [5]. A 1oo3 arrangement detects more and trips more often. A 2oo3 arrangement trips rarely and misses more hazards. The asymmetric release detects as well as 1oo3 and its false trips are short. Relative to RSS [20], the rule makes the trade-off between missed detections and false alarms an explicit parameter for a single hazard.

VIII. LIMITATIONS

The simulator is a kinematic and probabilistic model. It renders no camera images, simulates no LiDAR point clouds, and models no GPS multipath. Every input distribution is our assumption, and only six earlier results anchor any of them.

- The three paths fail independently except through the common-cause factor of Section VI-F and through the condition table. Real weather affects several paths together in ways this model does not capture.
- The watchdog needs at least one path to exceed τ_{arm}. A zone that produces no cue at all is invisible to the system.
- Braking is constant deceleration. The model has no ACC dynamics, no traffic, and no driver.
- Spurious candidates in the control trials are synthetic bursts at fixed rates. Real false alarms cluster around objects such as other signs and parked trucks.
- The thresholds were fixed and varied one at a time. Interactions between them were explored only through the perturbation ensemble.
- The map path decides many outcomes. Its range equals the approach length and it holds the zone in 95% of trials, and both of these are our assumptions.
- Release needs two healthy in-range paths that both read low. A healthy but blind camera, a missing map entry, and weak LiDAR cues together could therefore release braking during a real hazard. We did not test that combination.
- The pooled percentages give the twelve conditions equal weight, so they depend on which conditions we chose. The per-condition results are the safer comparison.

• The sensitivity runs use 500 hazard trials per condition, so small differences in worst-condition compliance are within run-to-run noise.

• The study covers one hazardous event. We have not tested the architecture on other ASIL D hazards.

• No hardware-in-the-loop, closed-track, or on-road testing was done, and no timing was measured on automotive hardware. The results are simulation evidence for a proposed design and form no certified safety argument.

IX. FUTURE WORK

• Replace the assumed sensor distributions with recorded camera, LiDAR, and GPS data in hardware-in-the-loop tests.

• Measure brake and actuator behavior on a closed track and replace the assumed friction, noise, and fault values.

• Run the same rule inside CARLA so that the paths produce confidences from rendered sensor data.

• Search adversarially [21]–[23] for conditions outside the twelve that drive the system toward baseline behavior.

• Optimise the thresholds jointly and study the map availability with real work-zone map data.

• Test the release rule against a blind camera, a missing map entry, and weak LiDAR cues at the same time.

• Extend the architecture to other ASIL D and SOTIF hazards and to radar and V2X paths.

X. CONCLUSION

This paper turns the fusion concept of an earlier six-run ISO 26262 case study into a defined rule: a confidence score per sensing path, one path to act, two healthy paths to release, and a watchdog that forces emergency braking when a weak cue is never confirmed. In a 20 Hz simulation over twelve conditions, the rule brings 97.0% of trials to the boundary at or below 42 km/h, against 57.8% for a camera-only baseline. In hazard-free approaches it commands brief brakes in 1.25% and leaves a brake unreleased in 0.02%.

The results depend on assumed inputs. The perturbation ensemble supports the direction of the improvement in every draw and leaves the exact percentages open. Faults in all three paths remain the weak case, and the map path carries a large share of the benefit. Hardware-in-the-loop and closed-track testing are the next steps.

REFERENCES

[1] S. Vishnu and M. Anji Reddy, "ISO 26262-Based Safety Analysis of Autonomous Emergency Braking in Construction Zones," unpublished manuscript, 2025.
[2] Dosovitskiy, G. Ros, F. Codevilla, A. López, and V. Koltun, "CARLA: An open urban driving simulator," in Proc. 1st Conf. Robot Learning (CoRL), 2017, pp. 1–16.
[3] IEC 61508:2010, Functional Safety of Electrical/Electronic/Programmable Electronic Safety-Related Systems, Parts 1–7, Int. Electrotechnical Commission, Geneva, Switzerland, 2010.
[4] W. M. Goble, Control Systems Safety Evaluation and Reliability, 3rd ed. Research Triangle Park, NC, USA: ISA, 2010.
[5] M. Rausand and A. Høyland, System Reliability Theory: Models, Statistical Methods, and Applications, 2nd ed. Hoboken, NJ, USA: Wiley, 2004.
[6] ISO 26262:2018, Road Vehicles — Functional Safety, Parts 1–12, Int. Org. for Standardization, Geneva, Switzerland, 2018.
[7] ISO 26262:2011, Road Vehicles — Functional Safety, 1st ed., Int. Org. for Standardization, Geneva, Switzerland, 2011.
[8] MISRA, MISRA C:2012 Guidelines for the Use of the C Language in Critical Systems, MIRA Ltd, Nuneaton, U.K., 2013.
[9] ISO 21448:2022, Road Vehicles — Safety of the Intended Functionality, Int. Org. for Standardization, Geneva, Switzerland, 2022.
[10] SAE International, J3016_202104, Taxonomy and Definitions for Terms Related to Driving Automation Systems for On-Road Motor Vehicles, SAE Int., Warrendale, PA, USA, 2021.
[11] D. H. Stamatis, Failure Mode and Effect Analysis: FMEA from Theory to Execution, 2nd ed. Milwaukee, WI, USA: ASQ Quality Press, 2003.
[12] W. E. Vesely, F. F. Goldberg, N. H. Roberts, and D. F. Haasl, Fault Tree Handbook, U.S. Nuclear Regulatory Commission, Washington, DC, USA, Rep. NUREG-0492, 1981.

[13] P. Koopman and M. Wagner, "Challenges in autonomous vehicle testing and validation," SAE Int. J. Transp. Safety, vol. 4, no. 1, pp. 15–24, 2016.
[14] P. Koopman and M. Wagner, "Autonomous vehicle safety: An interdisciplinary challenge," IEEE Intell. Transp. Syst. Mag., vol. 9, no. 1, pp. 90–96, 2017.
[15] P. Koopman, U. Ferrell, F. Fratrik, and M. Wagner, "A safety standard approach for fully autonomous vehicles," in Proc. SAFECOMP Workshops, 2019, pp. 326–332.
[16] N. Kalra and S. M. Paddock, "Driving to safety: How many miles of driving would it take to demonstrate autonomous vehicle reliability?," RAND Corp., Santa Monica, CA, USA, Rep. RR-1478-RC, 2016.
[17] W. Wachenfeld and H. Winner, "The release of autonomous vehicles," in Autonomous Driving: Technical, Legal and Social Aspects, M. Maurer, J. C. Gerdes, B. Lenz, and H. Winner, Eds. Berlin, Germany: Springer, 2016, pp. 425–449.
[18] H. Junietz, W. Wachenfeld, K. Klonecki, and H. Winner, "Evaluation of different approaches to address safety validation of automated driving," in Proc. 21st Int. Conf. Intell. Transp. Syst. (ITSC), 2018, pp. 491–497.
[19] C. Amersbach and H. Winner, "Functional decomposition — An approach to reduce the approval effort for highly automated driving," in Proc. 12. Workshop Fahrerassistenzsysteme, 2019.
[20] S. Shalev-Shwartz, S. Shammah, and A. Shashua, "On a formal model of safe and scalable self-driving cars," arXiv:1708.06374, 2017.
[21] R. Lee, O. J. Mengshoel, A. Saksena, R. Gardner, D. Genin, J. Silbermann, M. Owen, and M. Kochenderfer, "Adaptive stress testing: Finding likely failure events with reinforcement learning," J. Artif. Intell. Res., vol. 69, pp. 1165–1201, 2020.
[22] M. Koren, S. Alsaif, R. Lee, and M. J. Kochenderfer, "Adaptive stress testing for autonomous vehicles," in Proc. IEEE Intell. Vehicles Symp. (IV), 2018, pp. 1898–1904.
[23] M. O'Kelly, A. Sinha, H. Namkoong, R. Tedrake, and J. C. Duchi, "Scalable end-to-end autonomous vehicle testing via rare-event simulation," in Proc. Adv. Neural Inf. Process. Syst. (NeurIPS), 2018, pp. 9827–9838.
[24] J. von Neumann, "Probabilistic logics and the synthesis of reliable organisms from unreliable components," in Automata Studies, C. E. Shannon and J. McCarthy, Eds. Princeton, NJ, USA: Princeton Univ. Press, 1956, pp. 43–98.
[25] R. E. Lyons and W. Vanderkulk, "The use of triple-modular redundancy to improve computer reliability," IBM J. Res. Develop., vol. 6, no. 2, pp. 200–209, 1962.
[26] P. K. Varshney, Distributed Detection and Data Fusion. New York, NY, USA: Springer, 1997.
[27] R. E. Kalman, "A new approach to linear filtering and prediction problems," J. Basic Eng., vol. 82, no. 1, pp. 35–45, 1960.
[28] S. J. Julier and J. K. Uhlmann, "New extension of the Kalman filter to nonlinear systems," in Proc. SPIE Signal Process., Sensor Fusion, Target Recognit. VI, vol. 3068, 1997, pp. 182–193.
[29] Y. Bar-Shalom, X. R. Li, and T. Kirubarajan, Estimation with Applications to Tracking and Navigation: Theory, Algorithms and Software. New York, NY, USA: Wiley, 2001.
[30] W. Elmenreich, "An introduction to sensor fusion," Vienna Univ. Technol., Vienna, Austria, Res. Rep. 47/2001, 2001.
[31] J. Redmon, S. Divvala, R. Girshick, and A. Farhadi, "You only look once: Unified, real-time object detection," in Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR), 2016, pp. 779–788.
[32] Elfes, "Sonar-based real-world mapping and navigation," IEEE J. Robot. Autom., vol. 3, no. 3, pp. 249–265, 1987.
[33] IBM, "IBM Engineering Requirements Management DOORS Next," IBM Corp., Armonk, NY, USA, product documentation, accessed 2026.
[34] Ansys, Inc., "Ansys medini analyze for functional safety and cybersecurity," product documentation, accessed 2026.
[35] Federal Highway Administration, Manual on Uniform Traffic Control Devices for Streets and Highways, U.S. Dept. Transportation, Washington, DC, USA, 2023.
[36] Federal Highway Administration, "Work zone facts and statistics," U.S. Dept. Transportation, Washington, DC, USA, 2023. [Online]. Available: https://ops.fhwa.dot.gov/wz/resources/facts_stats.htm
[37] Waymo LLC, "Waymo safety methodologies and readiness determinations," Waymo, Mountain View, CA, USA, White Paper, 2021.
[38] National Highway Traffic Safety Administration, "Automated driving systems 2.0: A vision for safety," U.S. Dept. Transportation, Washington, DC, USA, DOT HS 812 442, 2017.
[39] B. L. Welch, "The generalization of 'Student's' problem when several different population variances are involved," Biometrika, vol. 34, no. 1–2, pp. 28–35, 1947.
[40] E. B. Wilson, "Probable inference, the law of succession, and statistical inference," J. Amer. Statist. Assoc., vol. 22, no. 158, pp. 209–212, 1927.

[41] M. E. O'Neill, "PCG: A family of simple fast space-efficient statistically good algorithms for random number generation," Harvey Mudd College, Claremont, CA, USA, Tech. Rep. HMC-CS-2014-0905, 2014.
[42] G. Van Rossum and F. L. Drake, Python 3 Reference Manual. Scotts Valley, CA, USA: CreateSpace, 2009.
[43] C. R. Harris et al., "Array programming with NumPy," Nature, vol. 585, pp. 357–362, 2020.
[44] P. Virtanen et al., "SciPy 1.0: Fundamental algorithms for scientific computing in Python," Nat. Methods, vol. 17, pp. 261–272, 2020.