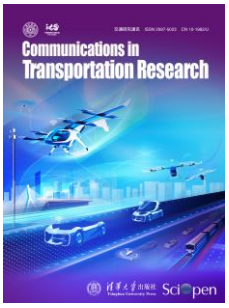
COMMTR

Communications in Transportation Research



RESEARCH ARTICLE

Driving risk emerges from the required two-dimensional joint evasive acceleration

Hao Cheng[1], Yanbo Jiang[1], Wenhao Yu[2], Rui Zhou[3], Jiang Bian[3], Keyu Chen[1], Zhiyuan Liu[1], Heye Huang[4], Hailun Zhang[5], Fang Zhang[2], Jianqiang Wang[1], Sifa Zheng[1,6,✉]

1 School of Vehicle and Mobility, Tsinghua University, Beijing 100084, China
2 State Key Laboratory of Intelligent Green Vehicle and Mobility, Beijing 100084, China
3 School of Traffic & Transportation Engineering, Central South University, Changsha 410000, China
4 Cho Chun Shik Graduate School of Mobility, Korea Advanced Institute of Science and Technology, Daejeon 34051, Republic of Korea
5 School of Automotive Engineering, Chang'an University, Xi'an 710064, China
6 Suzhou Automotive Research Institute, Tsinghua University, Suzhou 215134, China

✉ Corresponding author.

E-mail: zsf@tsinghua.edu.cn

Just Accepted

DOI 10.26599/COMMTR.2026.9640074

Received 29 April 2026

Revised 4 August 2026

Accepted 19 November 2026

COMMTR Homepage

Cite this article: https://doi.org/10.26599/COMMTR.2026.9640074

© The Author(s) 2026.

This is an open access article under the terms of the Creative Commons Attribution 4.0 International License (CC BY 4.0, http://creativecommons.org/licenses/by/4.0/), which permits use, distribution and reproduction in any medium, provided the original work is properly cited.

Published by **Tsinghua University Press**

https://www.sciopen.com/journal/2097-5023

清華大學出版社 Tsinghua University Press | SciOpen

Commun Transp Res

Research Article

Driving risk emerges from the required two-dimensional joint evasive acceleration

Hao Cheng[1], Yanbo Jiang[1], Wenhao Yu[2], Rui Zhou[3], Jiang Bian[3], Keyu Chen[1], Zhiyuan Liu[1], Heye Huang[4], Hailun Zhang[5], Fang Zhang[2], Jianqiang Wang[1], Sifa Zheng[1,6,✉]

Cite this article: xxx

ABSTRACT: Time-to-collision (TTC) is widely used in autonomous-driving safety evaluation. However, TTC and many of its extensions quantify risk using one-dimensional closing models, which may limit risk discrimination and misrepresent how risk evolves over time. We introduce evasive acceleration (EA), a physically interpretable two-dimensional (2D) framework for risk quantification. Considering all possible directions of relative acceleration, EA quantifies risk through the minimum magnitude of a constant 2D relative-acceleration vector required to render the extrapolated interaction collision-free. Across five open datasets and 630 reconstructed crash trajectories, EA provides earlier precrash warnings than all six baselines at matched noncrash percentile thresholds, with median lead times up to 2.0 s longer than TTC's. EA also provides stronger discrimination of eventual collision outcomes than the primary baselines and increases time-averaged retained predictive information about whether a crash occurs by 33.5–270.3% over the compared baselines. Across the evaluated EA–baseline pairs, adding EA yields 5.2–46.8 times as much time-averaged incremental predictive information as adding the corresponding baseline to EA, indicating greater predictive value beyond these baselines. These findings support 2D collision-avoidance demand as a physically interpretable basis for collision-risk quantification.

KEYWORDS: driving risk quantification, autonomous driving safety, traffic conflicts, early warning, collision avoidance

1 Introduction

Autonomous vehicles (AVs) are a major focus of mobility research and development. However, improving their safety remains a central challenge (Feng et al., 2026). Reliable quantification of driving risk is therefore essential for safety evaluation and system development.

Across academia and industry, time-to-collision (TTC) is widely used to quantify driving risk (Fig. 1a). AV benchmarks such as nuPlan (Caesar et al., 2021) and NAVSIM (Dauner et al., 2024) incorporate TTC into their safety evaluation metrics. TTC is also used to guide planner training and score candidate trajectories (Fan et al., 2024; Z. Li et al., 2024; H. Li et al., 2026). TTC thresholds are used in regulatory requirements and scenario-based safety assessments (de Gelder and Op den Camp, 2024; Nolte et al., 2025).

In its classical car-following formulation, TTC models collision imminence as a one-dimensional (1D) longitudinal closing problem. Extensions include anticipated collision time (ACT), proposed by Venthuruthiyil and Chunchu (2022), and the two-dimensional (2D) TTC formulation of Jiao (2023), hereafter TTC2D. These formulations account for collision geometry in the plane, including lateral and oblique encounters. Other developments include the 2D-TTC measure for safety-critical lane changes proposed by H. Guo et al. (2023), as well as adaptations for pedestrian–electric-scooter interactions and articulated vehicles (Jafari and Liu, 2024; Behera et al., 2026).

However, the suitability of TTC and its extensions for safety evaluation depends on the application (Westhofen et al., 2023; ElSamadisy et al., 2024). First, equal or similar TTC values need not indicate comparable interaction risk. Fig. 1b ii illustrates two longitudinal interactions with the same TTC. In Scene 1, lateral motion is about to resolve the conflict, whereas in Scene 2, the conflict is not being resolved laterally. These equal TTC values do not adequately distinguish the two states of conflict resolution. We refer to this limitation as limited risk informativeness. TTC-based warnings can also generate false alarms in resolving interactions (Choi and Kim, 2025).

Second, TTC-based measures can misrepresent how risk evolves over time. In the resolving interaction illustrated in Fig. 1b iii, evasive motion begins to reduce the conflict, while the extrapolated relative motion may still predict a collision. TTC can therefore continue to decrease while the indicated risk rises. Once the projected collision disappears, TTC becomes infinite and a TTC-derived risk representation can drop abruptly. The estimated risk may consequently peak near conflict resolution rather than reflect the earlier build-up and subsequent decline. We refer to this discrepancy as risk–time misalignment.

We argue that both limitations reflect a dimensionality mismatch between the underlying risk model and the 2D interaction it represents (Fig. 1b i). Although 2D TTC extensions can account for interactions from different directions, many still quantify risk by abstracting the interaction as an essentially 1D closing problem. This abstraction may not fully capture differences in risk between interaction states or the

[1] School of Vehicle and Mobility, Tsinghua University, Beijing 100084, China
[2] State Key Laboratory of Intelligent Green Vehicle and Mobility, Beijing 100084, China
[3] School of Traffic & Transportation Engineering, Central South University, Changsha 410000, China
[4] Cho Chun Shik Graduate School of Mobility, Korea Advanced Institute of Science and Technology, Daejeon 34051, Republic of Korea
[5] School of Automotive Engineering, Chang'an University, Xi'an 710064, China
[6] Suzhou Automotive Research Institute, Tsinghua University, Suzhou, Jiangsu 215134, China

✉ Corresponding author. E-mail: **zsf@tsinghua.edu.cn**

Received: xxx xx, 20xx; Revised: xxx xx, 20xx; Accepted: xxx xx, 20xx

© The Author(s) 20xx. This is an open access article under the terms of the Creative Commons Attribution 4.0 International License (CC BY 4.0, http://creativecommons.org/licenses/by/4.0/).

Just Accepted

temporal evolution of risk.

Beyond TTC-based measures, driving risk has been quantified through several complementary approaches. The Emergency Index (EI) quantifies the required intensity of evasive action by relating the maximum predicted intrusion depth into vehicle safety regions to the time available for avoidance (Cheng et al., 2025c). The Modified Emergency Index (MEI) refines the estimation of this available time (Cheng et al., 2025b). Traffic-safety-field methods represent longitudinal and lateral risk through spatially distributed fields (X. Zheng et al., 2018). Prediction-based probabilistic measures combine manoeuvre probabilities, collision probabilities, and expected crash severity over multimodal future trajectories (X. Wang et al., 2022). The generalized surrogate safety measure (GSSM) learns context-conditioned spacing distributions from naturalistic driving data without requiring crash or risk labels for training (Jiao et al., 2026). Reachability-based methods analyse safety under specified dynamics and admissible control and disturbance sets. Recent work combines backward reachability with stochastic forward reachability for probabilistic collision detection (X. Wang et al., 2024).

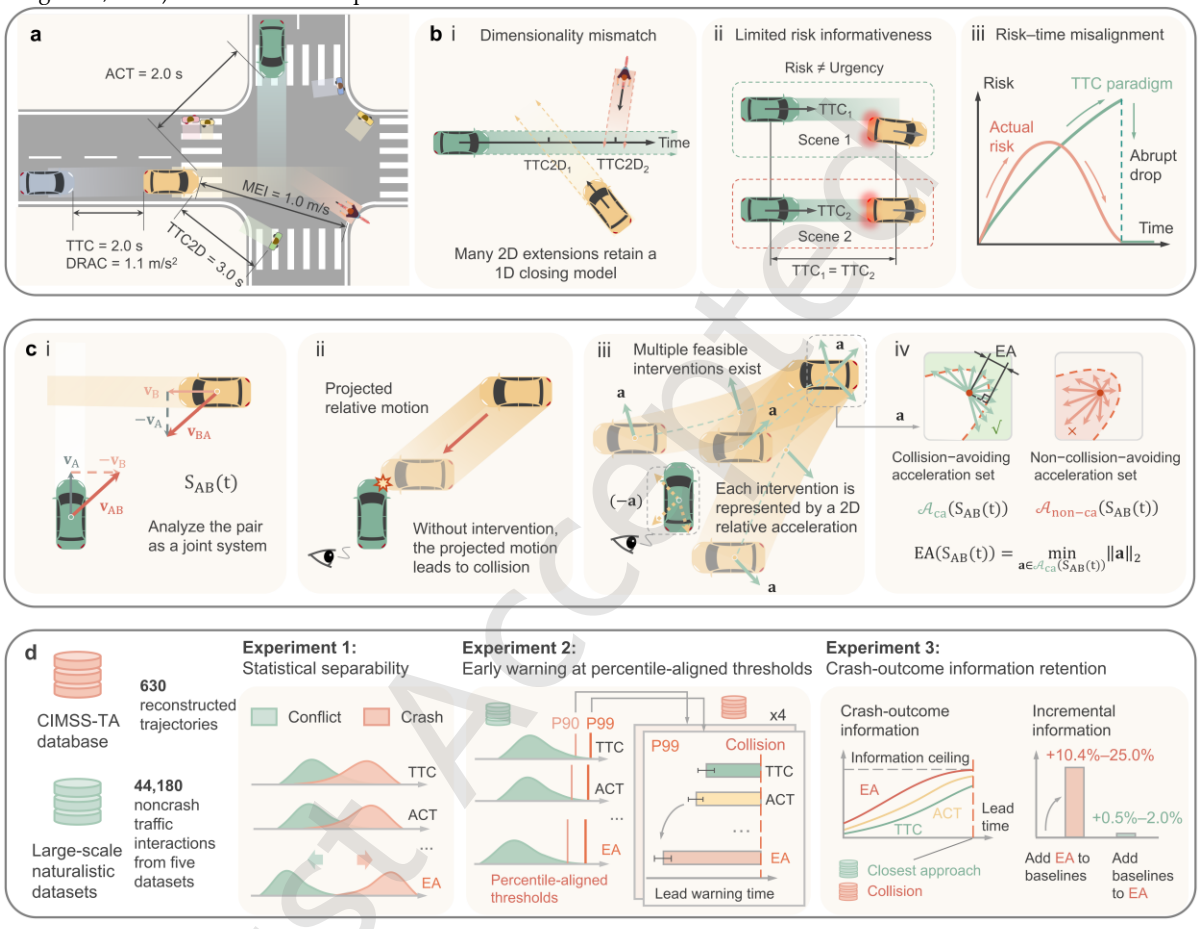


Fig. 1. Motivation, concept, and validation of the evasive acceleration (EA) framework. (a) Example traffic scene evaluated using time-to-collision (TTC), its extensions, and other physics-based risk measures. (b) Limitations of TTC-based risk representations. i, Dimensionality mismatch. Reducing two-dimensional (2D) interaction risk to a one-dimensional (1D) closing problem may impose a structural limitation on risk representation. ii, Limited risk informativeness. Equal TTC values can correspond to different collision-avoidance demands. iii, Risk–time misalignment. TTC-based risk may continue to rise and then drop abruptly to zero, failing to accurately reflect how risk evolves over time. (c) Conceptual illustration of EA. i, The interacting road users are analysed as a joint system through their relative motion. ii, The extrapolated relative motion predicts a collision without an evasive intervention. iii, Different constant 2D relative-acceleration vectors can make the extrapolated interaction collision-free. They describe interventions in the pair's relative motion and are shown on either vehicle to aid interpretation. iv, EA quantifies collision-avoidance demand through the minimum Euclidean magnitude in the collision-avoiding acceleration set under the specified motion extrapolation. (d) Validation using 630 reconstructed crash trajectories from CIMSS-TA and 44,180 noncrash traffic interactions from five open datasets, covering crash–noncrash discrimination, sustained precrash warning timeliness at matched noncrash percentile thresholds, and retained predictive information about crash outcomes.

We introduce evasive acceleration (EA), a physically interpretable framework for quantifying risk through 2D collision-avoidance demand. Fig. 1c i shows an interaction viewed as a joint system. Its extrapolated relative motion predicts a collision if no evasive change is introduced (Fig. 1c ii). We consider hypothetical constant relative-acceleration vectors with unrestricted directions in the plane (Fig. 1c iii). Collision-avoiding vectors form a feasible acceleration set, and EA quantifies risk through the minimum Euclidean magnitude in this set (Fig. 1c iv). EA quantifies avoidance demand in the pair's relative motion. Evaluating this demand at successive instants describes how risk builds up and subsides as an interaction evolves. A continuous-time implementation of EA averaged 0.154 ms per interaction frame in a single-core benchmark detailed in Section 3f of the Electronic Supplementary Material (ESM), supporting large-scale risk evaluation.

Just Accepted

We evaluate EA using 44,180 naturalistic noncrash traffic interactions from five open datasets collected in Germany, China, and the United States, together with 630 reconstructed crash trajectories from the CIMSS-TA database (Zhou et al., 2024). The evaluation comprises three complementary experiments (Fig. 1d). The primary comparisons cover widely used and recently proposed risk measures that quantify risk through directly interpretable physical quantities under simplified geometric or motion assumptions. Additional comparisons include time-to-react (TTR), which measures the time remaining to initiate a collision-avoiding emergency response (Tamke et al., 2011), and the parameterised Risk field model of Tan et al. (2024).

First, we assess how well the measures distinguish precrash states from noncrash interactions across four precrash lead windows. Second, we compare sustained precrash warning lead times at matched noncrash percentile thresholds, calibrated from the 90th, 95th, 99th, and 99.5th percentiles of eventwise maximum risk in the noncrash data. Third, we assess how much information each method's instantaneous risk value retains about whether a crash ultimately occurs. Across the primary evaluations, EA shows stronger overall crash–noncrash discrimination, longer median sustained precrash warning lead times at matched noncrash percentile thresholds, and greater time-averaged retained predictive information about crash outcomes than the evaluated primary baselines.

The remainder of this paper is organised as follows. Section 2 presents the EA formulation, computational method, and evaluation design. Section 3 reports the experimental results. Section 4 discusses additional comparisons, modelling limitations, and deployment considerations. Section 5 concludes the study.

2 Methods

2.1. Definition and modelling scope of evasive acceleration

EA measures the instantaneous pairwise collision-avoidance demand. By quantifying the relative acceleration required to make a predicted interaction collision-free, EA provides a physically grounded quantitative representation of collision risk under the specified motion extrapolation. EA does not estimate impact severity. Scenes with multiple road users can be represented as a set of pairwise interactions, each evaluated separately in the present formulation. A traffic conflict is an interaction for which extrapolation of the current motion state leads to a collision (Johnsson et al., 2018; Ray Sarkar et al., 2024; L. Zheng et al., 2014).

For two road users A and B at time t, let

$$S_{AB}(t) := \{S_A(t), S_B(t)\} \tag{1}$$

denote their joint interaction state, where $S_A(t)$ and $S_B(t)$ are the instantaneous states of the two road users. We evaluate the interacting pair jointly and define evasive acceleration (EA) as a pair-level quantity in relative-motion space. We consider planar interactions and represent each road user as a rigid body with a 2D occupied region. Unless otherwise specified, this occupied region is modelled as an oriented bounding box (OBB).

To separate the conceptual score from its computational truncation, first consider the extrapolated future motions under a specified motion model and without any additional interaction-induced evasive response. These yield trajectories $\boldsymbol{p}_A^0(t+s)$ and $\boldsymbol{p}_B^0(t+s)$ and the corresponding relative trajectory

$$\boldsymbol{r}^0(t+s) := \boldsymbol{p}_A^0(t+s) - \boldsymbol{p}_B^0(t+s), \qquad s \geq 0. \tag{2}$$

Let $\mathcal{C}(t+s) \subset \mathbb{R}^2$ denote the set of relative centre positions for which the interiors of the two occupied regions overlap, with their orientations given by the selected extrapolation model. Tangential contact is treated as the limiting boundary of collision avoidance. We then introduce a hypothetical constant relative acceleration

$$\boldsymbol{a} = \begin{bmatrix} a_x \\ a_y \end{bmatrix} \in \mathbb{R}^2, \tag{3}$$

defined directly in relative-motion space. Under this relative-acceleration perturbation, the relative trajectory becomes

$$\boldsymbol{r_a}(t+s) = \boldsymbol{r}^0(t+s) + \tfrac{1}{2}\boldsymbol{a}s^2, \qquad s \geq 0. \tag{4}$$

The ideal infinite-horizon collision-avoiding set and EA are

$$\begin{aligned} \mathcal{A}_\infty &:= \left\{ \boldsymbol{a} \in \mathbb{R}^2 : \boldsymbol{r_a}(t+s) \notin \mathcal{C}(t+s), \right. \\ &\qquad \left. \forall s > 0 \right\}, \\ \mathrm{EA}_\infty(S_{AB}(t)) &:= \inf_{\boldsymbol{a} \in \mathcal{A}_\infty} \| \boldsymbol{a} \|_2. \end{aligned} \tag{5}$$

For tractable computation, we truncate the extrapolation at a finite prediction (look-ahead) horizon $T > 0$ and evaluate

$$\begin{aligned} \mathcal{A}_T &:= \left\{ \boldsymbol{a} \in \mathbb{R}^2 : \boldsymbol{r_a}(t+s) \notin \mathcal{C}(t+s), \right. \\ &\qquad \left. \forall s \in (0, T] \right\}, \\ \mathrm{EA}_T(S_{AB}(t)) &:= \inf_{\boldsymbol{a} \in \mathcal{A}_T} \| \boldsymbol{a} \|_2. \end{aligned} \tag{6}$$

We use the convention $\inf \varnothing := +\infty$. If the occupied regions overlap at $s = 0$ (i.e., a collision has already occurred), we assign $\mathrm{EA}_T = \mathrm{EA}_\infty = +\infty$, because no finite acceleration applied from the current instant can change the current configuration.

For any $0 < T_1 < T_2$, the longer horizon imposes every constraint of the shorter horizon and additional future constraints. Hence,

$$\begin{aligned} &\mathcal{A}_\infty \subseteq \mathcal{A}_{T_2} \subseteq \mathcal{A}_{T_1} \\ \Rightarrow \quad &\mathrm{EA}_{T_1} \leq \mathrm{EA}_{T_2} \leq \mathrm{EA}_\infty. \end{aligned} \tag{7}$$

For compact rigid footprints with continuously extrapolated poses, the above boundary convention makes each $\mathcal{A}_T$ closed. Since $\mathcal{A}_\infty = \bigcap_{T>0} \mathcal{A}_T$ and $\| \boldsymbol{a} \|_2$ is coercive, the nested-set argument detailed in Section 1c of the ESM yields

$$\lim_{T \to \infty} \mathrm{EA}_T = \sup_{T>0} \mathrm{EA}_T = \mathrm{EA}_\infty. \tag{8}$$

Accordingly, the finite-horizon score converges monotonically from below to the infinite-horizon value. For the bounded-speed constant velocity (CV) and constant turn rate and velocity (CTRV) extrapolations used here, Section 1c of the ESM provides a tail bound and a sufficient condition for finite-horizon stabilisation. The finite T used in computation is therefore a reported truncation rather than a parameter fitted to outcome labels.

2.2. Unified numerical estimator under CV/CTRV extrapolation

Equation (6) defines a continuous-time, continuous-direction optimisation. Its numerical estimate depends on the motion extrapolation, prediction horizon, and discretisation used to test collision avoidance. For the empirical risk analyses, we use the reference estimator described below for all pairings of the two standard extrapolation models, CV and CTRV. CV is implemented as the zero-yaw-rate branch of CTRV. All four CV/CTRV pairings use the same optimisation with pairing-specific baseline trajectories.

Under the CV model, the future state of a road user satisfies

$$\boldsymbol{x}(t+s) = \boldsymbol{x}(t) + \boldsymbol{v}(t)s, \qquad s \in [0, T], \tag{9}$$

Just Accepted

where $\boldsymbol{x}$ and $\boldsymbol{v}$ denote centre position and velocity, respectively, and orientation is held fixed. Under CV extrapolation, baseline path curvature may be misattributed to additional avoidance demand in turning or merging interactions. We therefore also use CTRV extrapolation (Baek et al., 2020; Polychronopoulos et al., 2007; Schubert et al., 2008, 2011), under which

$$\begin{aligned} \dot{x}(t) &= v(t)\cos\theta(t), \\ \dot{y}(t) &= v(t)\sin\theta(t), \\ \dot{\theta}(t) &= \omega, \end{aligned} \tag{10}$$

where ω denotes the constant yaw rate, and speed is assumed constant over the interval of interest.

For road user $q \in \{A, B\}$ under hypothesis $i \in \{\mathrm{CV}, \mathrm{CTRV}\}$, the propagated yaw rate is

$$\omega_q^{(i)} = \begin{cases} 0, & i = \mathrm{CV}, \\ \widehat{\omega}_q, & i = \mathrm{CTRV}, \end{cases} \tag{11}$$

where $\widehat{\omega}_q$ is the yaw-rate input, taken from the dataset when available or otherwise estimated using only current and preceding trajectory samples (Section 1a in the ESM). An explicit straight-line, fixed-heading branch is used when $|\omega_q^{(i)}| < \varepsilon_\omega$, avoiding division by a near-zero yaw rate.

For every paired motion hypothesis, the baseline centre and orientation trajectories define time-varying oriented bounding boxes (OBBs). A hypothetical constant relative acceleration $\boldsymbol{a} = (a_x, a_y) \in \mathbb{R}^2$ is then applied to the predicted relative translation. This variable represents the joint intervention required by the pair but does not prescribe how that intervention is allocated between the two participants.

In implementation, $\boldsymbol{a}$ is realised as a quadratic translation of one predicted centre trajectory while the baseline orientation evolution is retained. If $\boldsymbol{p}_m^{(A)}(t+s)$ denotes the baseline predicted centre of road user A under paired hypothesis m, then

$$\widetilde{\boldsymbol{p}}_m^{(A)}(t+s; \boldsymbol{a}) = \boldsymbol{p}_m^{(A)}(t+s) + \tfrac{1}{2}\boldsymbol{a}\, s^2. \tag{12}$$

Let $\mathcal{O}_{q,m}(s; \boldsymbol{a})$ denote the OBB of participant q at time $t+s$, with Eq. (12) applied to the relative representation of A. On a numerical prediction grid $\mathcal{T}_{\Delta t} \subset (0, T]$, the grid-based model-specific feasible set is

$$\hat{\mathcal{S}}_{m,T} = \left\{ \begin{matrix} \boldsymbol{a} \in \mathbb{R}^2 : \mathrm{int}\mathcal{O}_{A,m}(s; \boldsymbol{a}) \\ \cap\, \mathrm{int}\mathcal{O}_{B,m}(s) = \varnothing \\ \forall s \in \mathcal{T}_{\Delta t} \end{matrix} \right\}. \tag{13}$$

After excluding states flagged as current contact or overlap, the numerical estimate is zero when the zero-acceleration baseline passes the fine-grid collision check.

To numerically approximate the minimisation over all possible directions in the EA definition, acceleration space is searched by direction. For $\boldsymbol{u}(\phi) = (\cos\phi, \sin\phi)^\top$,

$$\boldsymbol{a} = \rho\, \boldsymbol{u}(\phi), \qquad \rho \in [0, a_{\max}], \tag{14}$$

where $a_{\max}$ is a numerical search bound.

For each sampled direction, we use an interval search based on the Separating Axis Theorem (SAT) to find the smallest acceleration magnitude verified as collision-free on the numerical time grid. Repeating this calculation over a finite direction set Φ gives

$$\widehat{\mathrm{EA}}_{m,T} = \min_{\phi \in \Phi} \rho_m(\phi). \tag{15}$$

The hat distinguishes the computed value from the continuous-time, continuous-direction quantity in Eq. (6). We use coarse-to-fine angular and temporal grids. Temporal sampling can underestimate demand, whereas angular sampling can overestimate it. Section 1a of the ESM gives the SAT interval derivation, refinement algorithm, boundary conventions, and failure handling.

Unless otherwise stated, the prediction horizon is $T = 10.0\,\mathrm{s}$. Angular and temporal resolutions and search bounds are specified in Section 1c of the ESM. The magnitude bound is numerical rather than a participant capability limit. Current OBB interior overlap is assigned $+\infty$ because the intervention cannot change the current configuration. The numerical estimator also assigns $+\infty$ to current OBB contact. In the ideal formulation, tangential contact lies on the boundary of the collision-avoiding set. If no feasible magnitude is found within $a_{\max}$, the component and hence the aggregate are reported as $+\infty$ rather than clipped to the bound.

Each OBB is constructed directly from the reported centre position, heading, length, and width, without dimension scaling or an additional safety margin.

Because both CV and CTRV are simplified short-horizon motion hypotheses, either assumption alone may introduce systematic bias in some interaction patterns. We therefore allow each road user to be extrapolated using either CV or CTRV and average equally over the resulting paired hypotheses. Let

$$\mathcal{E}_2 = \{\mathrm{CV}, \mathrm{CTRV}\}, \qquad \mathcal{M} = \mathcal{E}_2 \times \mathcal{E}_2. \tag{16}$$

For each $m \in \mathcal{M}$, we compute $\widehat{\mathrm{EA}}_{m,T}(t)$ with the same numerical solver and define the reported EA as the arithmetic mean:

$$\mathrm{EA}(t) = \tfrac{1}{|\mathcal{M}|} \textstyle\sum_{m \in \mathcal{M}} \widehat{\mathrm{EA}}_{m,T}\,(t). \tag{17}$$

The paired extrapolations receive equal weights fixed independently of outcome labels. Because their minimisers can point in different directions, Eq. (17) aggregates four scalar minimum magnitudes and does not define a unique aggregate acceleration vector.

Sections 1c and 3d of the ESM report sensitivity to the prediction horizon ($T = 5$–$20\,\mathrm{s}$), angular and temporal resolutions, acceleration-search bound, and individual CV/CTRV pairings.

2.3. Crash–noncrash discrimination

The primary comparison includes EA, TTC, deceleration rate to avoid a crash (DRAC) (Westhofen et al., 2023), TTC2D (Jiao, 2023), anticipated collision time (ACT) (Venthuruthiyil and Chunchu, 2022), 2D DRAC (DRAC2D) (Huang and Chen, 2025), modified emergency index (MEI) (Cheng et al., 2025b), and bounding-box distance. From 658 archived reconstructed crash records, we retained one record from each of 28 identical-trajectory pairs, yielding 630 distinct trajectories for all crash-dependent analyses (Section 2b in the ESM).

Discrimination is quantified by the area under the precision–recall curve (AUPRC), the area under the receiver operating characteristic curve (AUROC), the Kolmogorov–Smirnov (KS) statistic, and recall at false-positive rates of 1%, 5%, and 10%. The value $+\infty$ for TTC, TTC2D, and ACT denotes the lowest-risk no-collision state, whereas $+\infty$ for EA, MEI, DRAC, and DRAC2D is retained at the high-risk end of the ordering, above all finite values. Zero values are also retained. Only NaN or parsing failures are treated as missing. We use a cluster bootstrap over crash cases and noncrash events to estimate 95% confidence intervals and paired method differences. A one-frame-per-crash analysis assesses whether the comparison is sensitive to using multiple frames from each crash. Sections 2a and 3a of the ESM provide

Just Accepted

event construction, resampling details, and complete results.

2.4. Precrash warning lead times and sensitivity analysis

After orienting outputs so that larger values indicate greater risk, we calibrate each method's warning thresholds using the same empirical percentiles of its eventwise maximum risk in the noncrash data. This procedure approximately aligns screened-event exceedance fractions across methods. Empirical percentile scores are compared with the thresholds without rounding. Risk ordering and missing-value handling follow Section 2.3. Here, time is measured relative to the crash reference time, set to zero. The primary sustained-warning lead time is defined as $L = -t_{\mathrm{onset}}$, where t_{onset} is the start of the final uninterrupted above-threshold episode that remains active through the final evaluated sample, with zero terminal tolerance (Section 2b in the ESM). Thus, $L \geq 0$, and larger values indicate earlier warning. A crash without a qualifying final episode is assigned $L = 0$, even if an earlier transient alert occurred. An inactive intermediate frame breaks an episode. Within each case, all methods use the same available history and final-episode criterion, with thresholds calibrated by the same percentile procedure. Sensitivity analyses use first-crossing and hysteresis-sustained warning definitions and additionally evaluate warning duration, episode count, missed-warning rate, and hysteresis-margin sensitivity. The three warning definitions use the same observation window, and the two sustained definitions use the same terminal criterion. Complete definitions and supplementary results are provided in Sections 2b and 3b of the ESM. A direct comparison with the full-history protocol allowing a terminal tolerance of 0.1 s is reported in Section 3b of the ESM.

The percentile calibration does not estimate a continuous-stream false-alarm rate.

2.5. Evaluation of retained predictive information

We assess how much information the instantaneous values of EA, ACT, TTC2D, TTC, DRAC, DRAC2D, and MEI retain about whether a crash ultimately occurs.

For crash episodes, the alignment anchor is the crash reference time, defined as $t = 0$ (Section 2b in the ESM). The reference time is reconstructed from the first sampled OBB contact where available. Otherwise, the final available trajectory frame is placed at -0.1 s. Independent impact-time annotations are unavailable to validate this reference time. For a noncrash episode, its anchor is the time of minimum bounding-box distance between the two interacting road users,

$$t_{\min} = \operatorname*{argmin}_{t} d_{\mathrm{bbox}}(t), \tag{18}$$

where $d_{\mathrm{bbox}}(t)$ is the instantaneous minimum distance between the occupied rigid-body bounding boxes. This anchor is shifted to $t = 0$. It provides a geometric reference for aligning all methods independently of their risk values.

After alignment, all crash and noncrash episodes are sampled on the same lead-time grid,

$$\tau \in \{-3.0, -2.9, \ldots, -0.1\}\ \mathrm{s}. \tag{19}$$

For each target lead time, the nearest available frame is selected within a tolerance of 0.03 s. The procedure is applied at the episode level, so each included episode contributes at most one sampled frame to each lead-time slice.

Output orientations are standardised so that larger transformed values denote greater risk. TTC, TTC2D, and ACT are sign-reversed, whereas EA, MEI, DRAC, and DRAC2D retain their original orientations.

For the standalone retained-information analysis, crash probabilities are estimated separately for each method and lead-time slice using five repetitions of five-fold stratified group out-of-fold prediction. Crash trajectory identifiers and naturalistic recording or scenario identifiers define the groups, so no group appears in both the training set and the held-out fold. Within each training fold, each oriented method value is transformed to its empirical percentile, and the fitted transformation is then applied to the held-out fold. A quadratic logistic regression ($C = 0.5$, L2-penalised liblinear solver, maximum 3,000 iterations) is fitted to the transformed training values and applied to the held-out fold. The out-of-fold probabilities are pooled across folds within each repetition and clipped to $[10^{-6}, 1 - 10^{-6}]$ before cross-entropy is calculated. The five resulting cross-entropies are then averaged.

For each lead time τ, let X_τ denote the interaction state and let $M_j(\tau) = f_j(X_\tau)$ denote the risk value returned by method j.

Let $Y \in \{0,1\}$ denote the eventual crash outcome. Within lead-time slice τ, the sample-label entropy is

$$H_\tau(Y) = -p_\tau \log_2 p_\tau - (1 - p_\tau)\log_2(1 - p_\tau), \tag{20}$$

where p_τ is the crash proportion in that slice. For method M_j, residual predictive uncertainty is estimated by the empirical held-out cross-entropy

$$\hat{H}_\tau^{\mathrm{pred}}(Y \mid M_j) = -\frac{1}{n_\tau}\sum_{i=1}^{n_\tau}\begin{bmatrix} y_i \log_2 \hat{p}_{ij,\tau} \\ +(1 - y_i)\log_2(1 - \hat{p}_{ij,\tau}) \end{bmatrix}, \tag{21}$$

where $\hat{p}_{ij,\tau}$ is the out-of-fold crash probability for episode i. Retained predictive information is defined as

$$\hat{\jmath}_j^{\mathrm{ret}}(\tau) = H_\tau(Y) - \hat{H}_\tau^{\mathrm{pred}}(Y \mid M_j). \tag{22}$$

The directional incremental-information analysis uses the same repeated group folds, training-fold percentile transformations, and quadratic-logistic model family. For each EA–baseline pair and lead-time slice, EA alone, the baseline alone, and their joint model are fitted within every training fold and evaluated only on held-out episodes. A single-input model contains its linear and quadratic terms, whereas a joint model contains the two linear terms, their quadratic terms, and their interaction. Thus, all three models entering a directional comparison have the same regularisation, feature degree, and held-out evaluation protocol.

For baseline M_k, the predictive information gained by adding EA is

$$\begin{aligned} \Delta\hat{\jmath}_{\mathrm{EA}|k}^{\mathrm{ret}}(\tau) = \hat{H}_\tau^{\mathrm{pred}}(Y \mid M_k) \\ -\hat{H}_\tau^{\mathrm{pred}}(Y \mid \mathrm{EA}, M_k), \end{aligned} \tag{23}$$

and the converse gain is

$$\begin{aligned} \Delta\hat{\jmath}_{k|\mathrm{EA}}^{\mathrm{ret}}(\tau) = \hat{H}_\tau^{\mathrm{pred}}(Y \mid \mathrm{EA}) \\ -\hat{H}_\tau^{\mathrm{pred}}(Y \mid \mathrm{EA}, M_k). \end{aligned} \tag{24}$$

A positive value indicates a reduction in held-out predictive cross-entropy after adding the second method. These model- and sample-dependent quantities measure reductions in held-out predictive cross-entropy rather than directly estimating population mutual information.

Sampling uncertainty for the standalone retained-information and residual-uncertainty curves is quantified using 1,000 case-level bootstrap replicates. Crash and noncrash episodes are resampled with replacement within their outcome groups, with the same resampled episodes used for every method in a replicate. The out-of-fold

Just Accepted

predictions are held fixed during resampling.

3 Results

The primary evaluation uses 630 reconstructed crash trajectories from the CIMSS-TA database and 44,180 naturalistic noncrash potential-conflict events extracted from five open datasets, highD, inD, exiD, SIND, and Argo2-LCR. Noncrash events were initially screened by retaining those with an eventwise minimum below 5 s for at least one of TTC, TTC2D, or ACT. This event set was held fixed when the risk measures were recomputed (Section 2a in the ESM). Using crash and noncrash interactions to evaluate risk measures complements collision-avoidability-based assessment (Cheng et al., 2025a). The three experiments assess crash–noncrash discrimination, sustained precrash warning timeliness, and retained predictive information about crash outcomes.

The trajectory datasets used in the primary analyses provide state and object-geometry records. EA is evaluated directly from the released trajectories at their native sampling rates of 10 Hz for Argo2-LCR, SIND, and CIMSS-TA, and 25 Hz for inD, highD, and exiD. No additional smoothing is applied to the released position, speed or heading trajectories in the primary analyses. Input sensitivity is assessed separately using offline state-input and temporal-data perturbations (Sections 1d and 3h in the ESM).

3.1. Risk evolution in representative cases

Fig. 2a shows a merging conflict between cars #266 and #267 that is eventually resolved. EA peaks at $1.18\ \mathrm{m/s^2}$ and then declines as car #266 decelerates to yield, capturing the build-up and subsequent dissipation of risk.

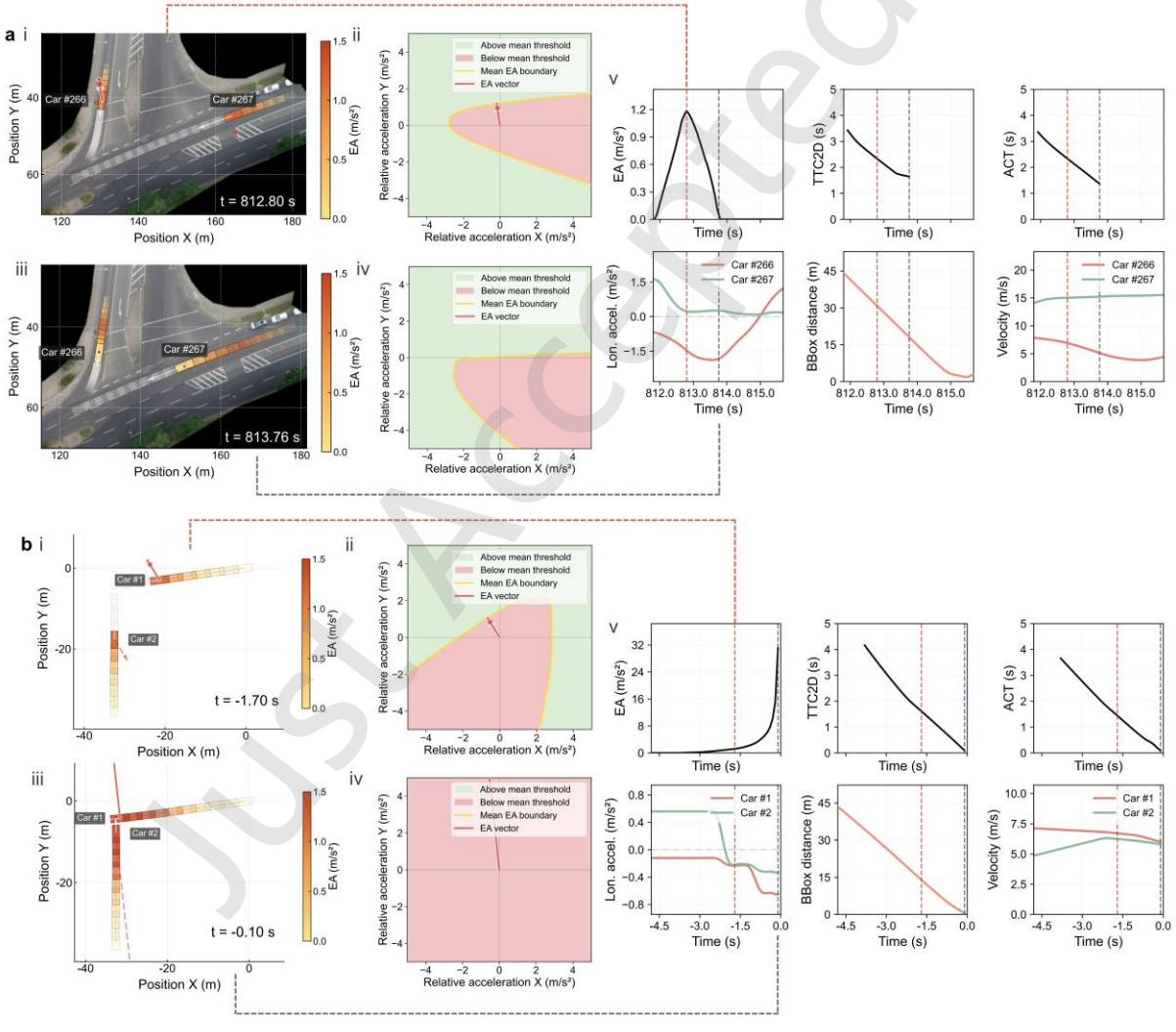


Fig. 2. Representative interaction cases showing how EA captures risk evolution and conflict resolution. In both cases, the trajectory colours in panels i and iii and the EA time series in panel v show the mean EA over the four CV/CTRV pairings. Corresponding trajectory and acceleration-space panels show the same EA vector. Its magnitude equals the mean EA, and its direction follows the sum of the four pairing-specific minimum vectors. The solid and dashed arrows show alternative representations of the same relative-acceleration intervention on the two vehicles. In panels ii and iv, the yellow curve shows the mean of the four minimum collision-avoiding acceleration magnitudes along each direction. The shaded regions distinguish magnitudes above and below this directional mean. (a) Merging interaction at an intersection. EA initially rises and then declines as car #266 decelerates to yield. (b) T-bone crash in Hunan, China. EA rises sharply as impact approaches. The arrows in b iii and b iv extend beyond their respective plot limits. The yellow curve in b iv lies outside the displayed range.

Just Accepted

TTC2D and ACT continue to indicate increasing risk until the projected collision disappears. They then jump to positive infinity, exemplifying risk–time misalignment (Fig. 2a v). At the later illustrated instant, ACT = 1.37 s and TTC2D = 1.65 s, while EA has fallen to 0.05 m/s^2 (Fig. 2a iii–iv).

Fig. 2b shows a T-bone crash that occurred in Hunan, China. At 1.70 s before the reconstructed crash reference time, EA has already reached 1.23 m/s^2 and continues to increase (see Fig. 2b i–ii). By 0.1 s before that reference time, EA has risen to 31.16 m/s^2, indicating sharply increasing collision-avoidance demand under the extrapolation (see Fig. 2b iii–iv). Over the interval in which TTC2D and ACT are finite, both decline approximately linearly, whereas EA rises sharply and nonlinearly (see Fig. 2b v).

The two cases have similar EA values at the earlier instants, then diverge as the merging conflict resolves and the crash conflict intensifies.

We next compare EA with the primary baselines in terms of crash discrimination, warning lead time, and retention of collision-relevant information.

3.2. Crash–noncrash discrimination

We compare EA with the seven primary baselines using the 630 reconstructed crash trajectories and 44,180 naturalistic potential-conflict events. All risk measures are computed frame by frame from the same interaction states. Each noncrash event contributes one observation equal to its eventwise maximum risk. Crash observations comprise all available frames in four precrash windows, $[-0.5, -0.1]$, $[-1.0, -0.1]$, $[-1.5, -0.1]$, and $[-2.0, -0.1]$ s. These nested windows test whether discrimination persists as earlier precrash states are included. Discrimination metrics and event-level inference procedures are specified in Section 2.3.

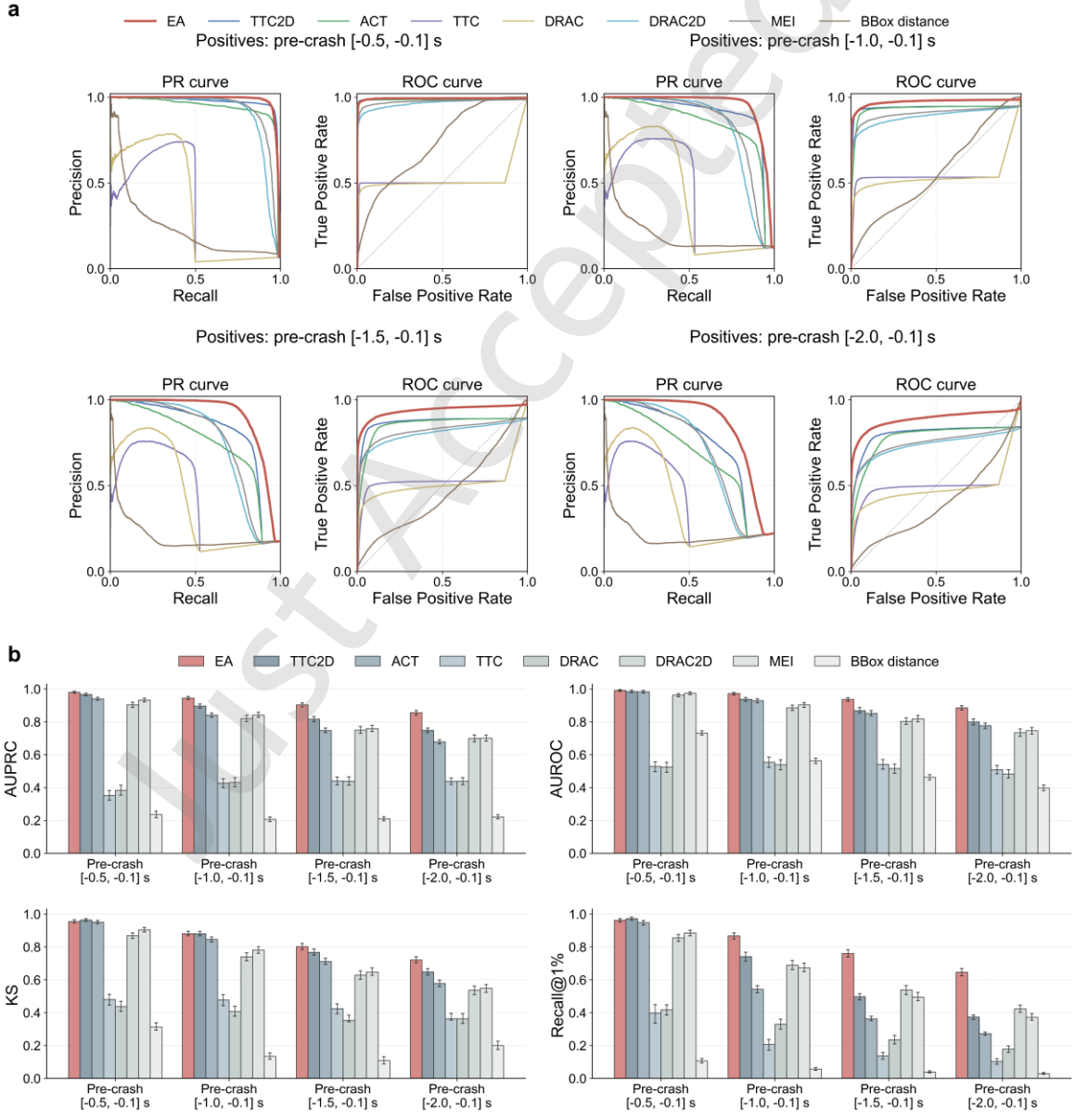


Fig. 3. Discrimination between precrash states and noncrash interactions. (a) Precision–recall (PR) and receiver operating characteristic (ROC) curves for the four precrash lead windows $[-0.5, -0.1]$, $[-1.0, -0.1]$, $[-1.5, -0.1]$, and $[-2.0, -0.1]$ s. Across all four lead windows, EA has higher AUPRC and AUROC than the evaluated baselines. (b) Summary discrimination metrics across the same four precrash lead windows, including AUPRC, AUROC, the Kolmogorov–Smirnov (KS) statistic, and recall at a 1% false-positive rate (FPR). Error bars show 95% cluster-bootstrap confidence intervals, with crash cases and noncrash events as the resampling units. In the [-0.5, -0.1] s window, TTC2D has slightly higher KS and Recall@1% FPR values than EA.

Just Accepted

Fig. 3 summarises the four lead windows, and panel b reports 95% event-level cluster-bootstrap confidence intervals. EA's AUPRC and AUROC exceed those of the seven primary baselines across all four precrash windows. In the nearest window, TTC2D is slightly higher for the KS statistic and Recall@1% FPR. In $[-1.5, -0.1]$ s, EA attains an AUPRC of 0.903 and an AUROC of 0.937, compared with baseline ranges of 0.210–0.816 and 0.463–0.869. The advantage persists in the longest precrash window, $[-2.0, -0.1]$ s, where EA attains an AUPRC of 0.856 and an AUROC of 0.886, compared with 0.748 and 0.800 for TTC2D. Event-level cluster-bootstrap and one-frame-per-crash analyses both retain EA's leading AUPRC and AUROC. Tables 3–6 of the ESM provide complete estimates, confidence intervals, paired comparisons, and sensitivity results.

Covariate matching retains 307, 374, 389, and 404 crash–noncrash pairs at 0.5, 1.0, 1.5, and 2.0 s, respectively. Within these common-support samples, EA's AUROC and event-weighted AUPRC exceed those of the compared methods at every lead time. In the strict-balance subsets, EA outperformed TTC2D on both AUROC and AUPRC. In the analysis using one score per event at the matched reference frames, EA ranked first on both metrics among all evaluated methods (Section 3g in the ESM).

3.3. Precrash warning lead times at matched percentile thresholds

We compare EA with ACT, TTC2D, TTC, DRAC, DRAC2D, and MEI at four percentile-aligned operating points. Thresholds are calibrated from the 90th, 95th, 99th, and 99.5th percentiles of the 44,180 noncrash eventwise maxima, approximately matching screened-event exceedance fractions across methods. Each of the 630 crash trajectories is evaluated over the same $[-3.0, -0.1]$ s precrash window, with zero terminal tolerance. Within each case, all methods use the same available history and final-episode criterion defined in Section 2.4 and Section 2b of the ESM.

Table 1. Median sustained precrash warning lead times at matched noncrash percentile thresholds. *Median WLT* (s) denotes the median sustained-warning lead time, reported as the nonnegative duration $L = -t_{\text{onset}}$. Larger values indicate earlier warning. ΔL (s) denotes $\text{Median}_{\text{EA}} - \text{Median}_{\text{other}}$, with positive values indicating an earlier warning by EA. *Impv.* (%) denotes the median lead-time difference as a percentage of the baseline median. It is shown as --- when the baseline median is zero.

Method	Risk percentile											
	90th			95th			99th			99.5th		
	Median WLT (s)	ΔL (s)	Impv. (%)	Median WLT (s)	ΔL (s)	Impv. (%)	Median WLT (s)	ΔL (s)	Impv. (%)	Median WLT (s)	ΔL (s)	Impv. (%)
EA	**2.00**	—	—	**1.70**	—	—	**1.30**	—	—	**1.20**	—	—
ACT	1.30	+0.70	+54	1.00	+0.70	+70	0.50	+0.80	+160	0.40	+0.80	+200
TTC2D	**1.70**	+0.30	+18	**1.40**	+0.30	+21	**0.80**	+0.50	+63	**0.60**	+0.60	+100
TTC	0.00	+2.00	—	0.00	+1.70	—	0.00	+1.30	—	0.00	+1.20	—
DRAC	0.00	+2.00	—	0.00	+1.70	—	0.00	+1.30	—	0.00	+1.20	—
DRAC2D	1.30	+0.70	+54	1.10	+0.60	+55	**0.80**	+0.50	+63	**0.60**	+0.60	+100
MEI	1.25	+0.75	+60	1.10	+0.60	+55	0.70	+0.60	+86	**0.60**	+0.60	+100

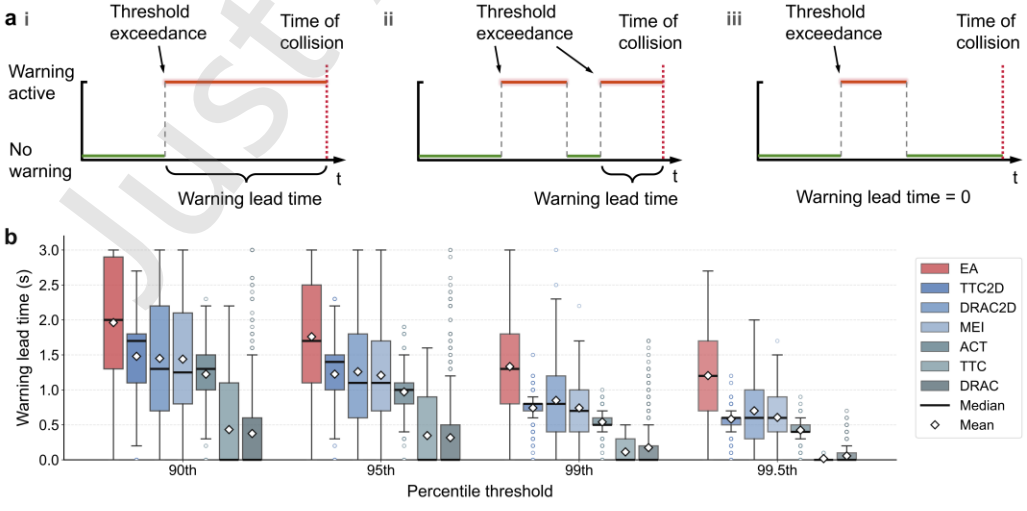


Fig. 4. Sustained precrash warning lead times at matched noncrash percentile thresholds. (a) Definition of warning lead time (WLT). i, WLT is measured from the onset of the final uninterrupted above-threshold episode that remains active through the last evaluated precrash sample. ii, An inactive frame breaks the episode. iii, WLT is zero if no episode meets this criterion. (b) Distributions of sustained-warning lead times across all 630 crash cases. Thresholds are set at the 90th, 95th, 99th, and 99.5th percentiles of each method's eventwise maximum risk across 44,180 noncrash interactions. Each case uses the [-3.0,-0.1] s precrash window (Section 2b in the ESM). Cases with zero WLT are included in all summary statistics. Boxes span the interquartile range, whiskers extend to the most extreme observations within 1.5 times this range of the box edges, and circles mark outliers. Black lines indicate medians and white diamonds indicate means.

Just Accepted

EA yields longer median sustained-warning lead times than the six evaluated baselines at every threshold (Fig. 4b and Table 1), with median lead times of 2.00, 1.70, 1.30, and 1.20 s from P90 to P99.5. Relative to TTC2D, EA advances the median by 0.30, 0.30, 0.50, and 0.60 s, respectively. The gains over ACT, DRAC2D, and MEI range from 0.50 to 0.80 s across the reported operating points.

EA yields longer median warning lead times than the six primary baselines under first-crossing, sustained, and hysteresis-sustained definitions. Its missed-warning rate remains 0.48–0.63% across these definitions. Detailed paired comparisons and protocol-sensitivity results are provided in Section 3b of the ESM.

3.4. Retained predictive information about crash outcomes

We next assess how much information instantaneous risk values retain about eventual crash outcomes.

We compare 630 crash episodes with 40,417 noncrash potential-conflict episodes that contain a complete preevent window from −3.0 s to −0.1 s after alignment, selected from the 44,180 extracted noncrash events. All methods are evaluated at the same relative lead times for each episode. Episode construction, temporal alignment, and nearest-frame matching are detailed in Section 2.5.

Retained predictive information is the reduction in outcome uncertainty defined in Eq. (22). Residual uncertainty is quantified by held-out predictive cross-entropy. For the 630 crash and 40,417 noncrash episodes, the baseline outcome entropy $H_\tau(Y)$ is about 0.1145 bits and sets the ceiling for retained predictive information. We report both residual uncertainty and retained information in bits. Table 2 summarises directional asymmetry. Tables 9 and 10 of the ESM provide time-resolved retained predictive information and EA's advantage over the baselines.

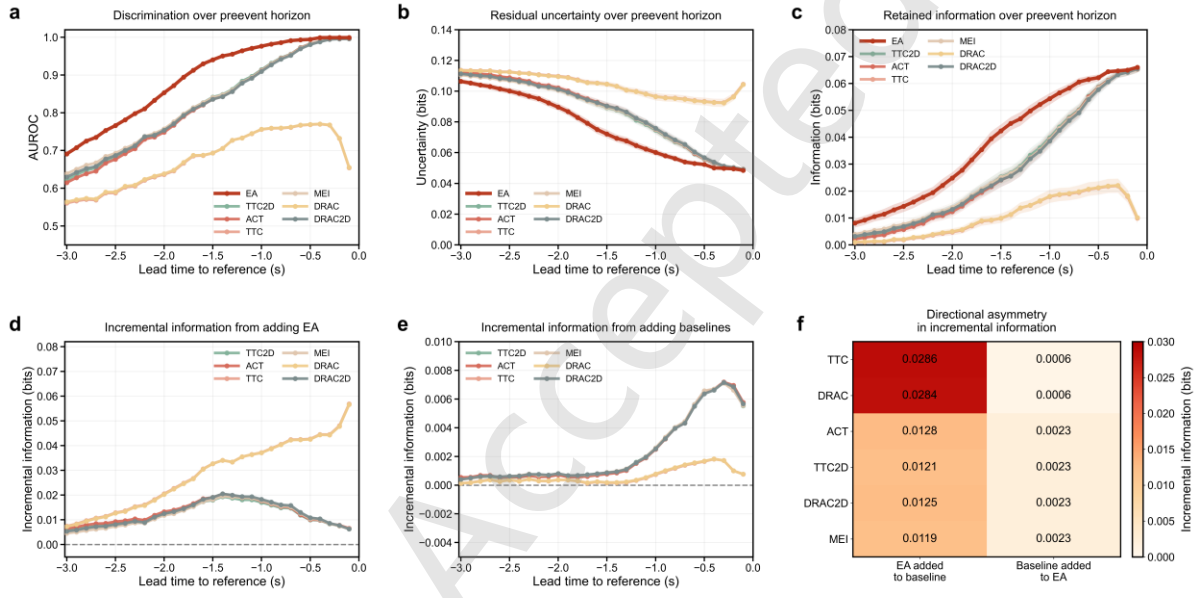


Fig. 5. Discrimination and predictive information about crash outcomes. Episodes are aligned to the crash reference time for crashes and the time of minimum bounding-box distance for noncrashes. The analysis includes 630 crashes and 40,417 complete noncrash episodes from −3.0 s to −0.1 s relative to these reference times. Shaded bands in panels b and c denote case-level bootstrap 95% confidence intervals. (a) Standalone discrimination of eventual crash outcomes across the preevent horizon, measured by AUROC. (b) Residual predictive uncertainty, measured by cross-entropy on episodes excluded from model fitting. (c) Cross-entropy-based retained predictive information in bits. (d) Incremental predictive information gained by adding EA to each baseline method. (e) Incremental predictive information gained by adding each baseline method on top of EA. (f) Incremental predictive information averaged across all sampled lead times for the two directions shown in d and e.

As shown in Fig. 5a, EA's AUROC exceeds those of the evaluated methods across the full preevent horizon. Averaged over the 3 s window, EA achieves a mean AUROC of 0.892, compared with 0.828 for MEI, 0.826 for TTC2D, and 0.677–0.679 for TTC and DRAC. At −3.0 s, EA reaches an AUROC of 0.691, compared with 0.615–0.638 for ACT, MEI, TTC2D, and DRAC2D and 0.561–0.564 for TTC and DRAC, showing that the separation is already evident at the earliest sampled lead time.

EA retains more information and leaves less residual uncertainty than the evaluated methods throughout the preevent horizon (Fig. 5b,c). At −3.0 s, it retains 0.0081 bits (7.1% of the ceiling), about twice MEI's value and more than twice the values of the other baselines. Averaged over the sampled lead times, EA retains 0.038 bits (33.5%), compared with 0.029 bits for MEI, 0.029 bits for TTC2D, and 0.010–0.028 bits for the other baselines. The mean retained predictive information across sampled lead times is 33.5–270.3% higher for EA than for the six baselines.

For EA, ACT, TTC2D, DRAC2D, and MEI, retained information generally increases as the evaluation time approaches each episode's reference time. At −0.1 s, EA reaches 57.6% of the ceiling, these four baselines reach 57.0–57.1%, and TTC and DRAC remain near 9%.

EA's information advantage over these four baselines is most pronounced in the early and intermediate preevent phases and narrows near each episode's reference time.

Across the six EA–baseline pairs, adding EA yields 5.2–46.8 times as much time-averaged incremental predictive information as adding the corresponding baseline to EA (Fig. 5d–f and Table 2).

Just Accepted

Table 2. Directional asymmetry in incremental predictive information between EA and the baselines. B denotes the baseline. Values are averaged across all sampled lead times. EA → B denotes the gain from adding EA to B, and B → EA denotes the gain from adding B to EA. Percentages are normalised by the outcome entropy of 0.1145 bits. The asymmetry ratio is defined as $(EA \rightarrow B)/(B \rightarrow EA)$. The asymmetry gap is defined as $(EA \rightarrow B) - (B \rightarrow EA)$.

Baseline	EA → B (bit)	EA → B (%)	B → EA (bit)	B → EA (%)	Asym. gap (bit)	Asym. ratio
ACT	0.0128	11.1	0.0023	2.0	0.0105	5.6
TTC2D	0.0121	10.6	0.0023	2.0	0.0098	5.4
TTC	0.0286	25.0	0.0006	0.5	0.0280	46.8
DRAC	0.0284	24.9	0.0006	0.5	0.0278	45.7
DRAC2D	0.0125	10.9	0.0023	2.0	0.0102	5.5
MEI	0.0119	10.4	0.0023	2.0	0.0096	5.2

4 Discussion

4.1. Comparison with additional risk measures

Beyond the primary baselines, we compare EA with TTR (Tamke et al., 2011; Wagner et al., 2018) and the Risk field model of Tan et al. (2024).

Paired event-level bootstrap analysis showed higher AUPRC for EA than for TTR in all four precrash windows. The paired AUROC differences had 95% confidence intervals containing zero in all four windows. TTR's AUROC point estimate was 0.009 higher in the longest window (Fig. 6a,b and Section 3e of the ESM). EA provides earlier median sustained warnings than TTR at all four matched noncrash percentile thresholds. With its published parameters, the road-user component of Risk field yielded lower AUPRC and AUROC and shorter median warning lead times than EA.

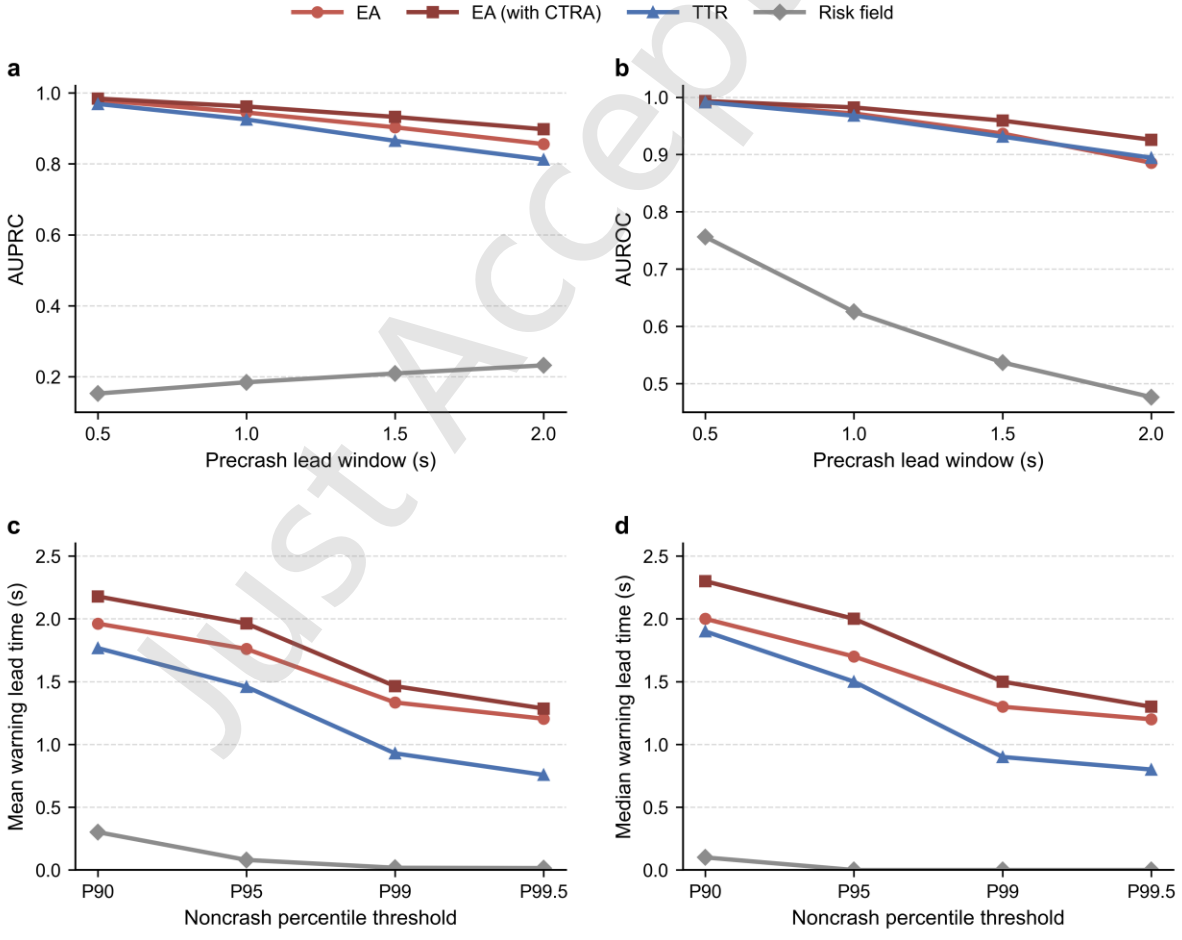


Fig. 6. Expanded comparison with TTR and Risk field, including EA with CTRA. (a, b) AUPRC and AUROC for distinguishing precrash frames in four nested windows ending at -0.1 s, with starting times of -0.5, -1.0, -1.5, and -2.0 s, from eventwise maximum-risk observations across 44,180 noncrash interactions. (c, d) Mean and median sustained-warning lead times across all 630 crash trajectories. Thresholds are set at the 90th, 95th, 99th, and 99.5th percentiles of each method's eventwise maximum risk across 44,180 noncrash interactions. All methods use the [-3.0,-0.1] s precrash window and the final-episode criterion defined in Section 2b of the ESM. Cases without a qualifying warning are assigned zero warning lead time. The default EA uses CV and CTRV extrapolation for each road user, whereas EA (with CTRA) additionally includes CTRA extrapolation. Expanded comparison methods and results are provided in Sections 2c and 3e of the ESM.

Just Accepted

We further examined whether richer motion extrapolation could improve EA by extending its CV/CTRV model set to include the constant turn-rate and acceleration (CTRA) model (Schubert et al., 2008), retaining the same collision-avoidance objective and numerical search. Within this expanded comparison, the resulting variant achieves the highest AUPRC and AUROC point estimates across the four precrash windows. In the [−2.0, −0.1] s window, AUPRC increases from 0.856 to 0.898 and AUROC from 0.886 to 0.926 relative to EA using CV/CTRV extrapolation. It also yields the longest median sustained-warning lead times at the four matched operating points. These results suggest that richer motion extrapolation can strengthen EA's risk representation (Section 3e of the ESM).

4.2. Robustness and practical considerations

We further assess EA's robustness with respect to numerical settings and input errors. EA's crash–noncrash discrimination varied little across the tested prediction horizons (5–20 s), angular resolutions and temporal resolutions (Section 3d of the ESM). Input-sensitivity tests showed stable frame-level EA rankings under position, speed and geometry perturbations. Among the inputs tested individually, heading error had the largest effect. At a heading noise standard deviation of 1°, the Spearman correlation between the original and perturbed rankings was 0.814, highlighting the importance of heading accuracy in practical applications (Sections 1d and 3h of the ESM).

To assess EA's computational efficiency, we benchmarked a continuous-time implementation on a single CPU core using over 850,000 real-world interaction frames. Computation time averaged 0.154 ms per interaction frame, with a 95th percentile of 0.619 ms and a maximum of 14.448 ms (Section 3f of the ESM).

We release the EA-P90 Pairwise Conflict Dataset[1] as a resource for AV testing. It comprises the top 10% of potential-conflict events in the naturalistic data, ranked by each event's maximum EA value. Construction and access details are provided in Section 2a of the ESM.

5 Conclusion

We introduce EA as a physically interpretable 2D framework for quantifying pairwise collision risk through collision-avoidance demand. In the primary comparisons, EA achieves higher AUPRC and AUROC, retains more time-averaged predictive information about crash outcomes, and provides longer median sustained precrash warning lead times at matched noncrash percentile thresholds. These findings support 2D collision-avoidance demand as a basis for quantifying collision risk in traffic.

Building on these findings, EA could be further refined and extended in several directions. The present formulation evaluates pairwise interactions without explicitly accounting for differences in manoeuvring capabilities across directions. Future work could incorporate these differences and extend EA to multi-agent interactions, including how avoidance actions are allocated among road users. EA's risk estimates also depend on the extrapolated baseline trajectories. Further research should explore richer extrapolation methods and trajectory prediction models to refine baseline interaction modelling and assess their effects on risk quantification.

[1] https://drive.google.com/drive/folders/1JGE76fxJyQwvCQTU5soV8YLZySyFQBv9?usp=sharing

Author contributions

H.C. and S.Z. conceived the EA framework and led the study. H.C., Y.J., and W.Y. developed the algorithms, processed the data, organised the results, and drafted the manuscript. Z.L. and K.C. supported the preliminary validation. R.Z. and J.B. contributed crash data. F.Z. supported the analysis of naturalistic driving data. H.H., H.Z., and J.W. provided feedback on algorithm development and experimental procedures. All authors contributed to manuscript revision and discussion of the results.

Replication and data sharing

The replication package for the 630-trajectory analysis has been prepared as an update to ETS-Data record ETS-Data-26-00091 (https://doi.org/10.26599/ETSD.2026.9190106). It contains aggregate results, noncrash eventwise risk values, timing observations, analysis code and a Replication Explanatory File. The public workflow verifies supplied summaries and reconstructs Figs. 5 and 6. Accident-level trajectories and per-case derived time series are excluded. Independent recalculation of accident-based statistics requires separately authorised inputs. Cohort selection is described in Section 2b of the ESM. Shared summaries, noncrash derived data and author-created documentation are licensed under CC BY-NC-SA 4.0, subject to source-dataset terms. Separately applicable software licences remain in force.

Data availability

The naturalistic driving datasets used in this study are highD, inD, exiD, SIND, and the Lateral Conflict Resolution Dataset (derived from Argoverse-2), which are available at https://www.highd-dataset.com, http://www.inD-dataset.com, https://levelxdata.com/exid-dataset/, https://github.com/SOTIF-AVLab/SinD, and https://github.com/RomainLITUD/conflict_resolution_dataset, respectively.

Code availability

The source code for EA is available at https://github.com/AutoChengh/evasive-acceleration.

Acknowledgements

This work was supported by the National Natural Science Foundation of China (Grant Nos. 52572497 and 52572477), the National Key R&D Program of China (Grant No. 2023YFB2504400), the Independent Research Project of the State Key Laboratory of Intelligent Green Vehicle and Mobility, Tsinghua University (Grant No. ZZ-PY-20250409), and the Open Fund Project of the State Key Laboratory of Intelligent Green Vehicle and Mobility under Grant KFZ260403. The authors also sincerely thank Dr. Yiru Jiao for her valuable guidance on the validation.

Declaration of competing interest

The authors declare no competing interests.

Declaration of generative AI and AI-assisted technologies in the writing process

During the preparation of this work, the authors used ChatGPT for English language polishing. The authors

Just Accepted

reviewed and edited the content as needed and take full responsibility for the final content of the article.

Electronic Supplementary Material

Supplementary theoretical derivations, experimental details, additional validation results and sensitivity analyses are provided in the Electronic Supplementary Material accompanying the online version of this article.

References

Baek, M., Jeong, D., Choi, D., Lee, S., 2020. Vehicle trajectory prediction and collision warning via fusion of multisensors and wireless vehicular communications. *Sensors*, *20*, 288.

Behera, A., Kharrazi, S., Frisk, E., Aramrattana, M., 2026. Two-dimensional time-to-collision measures for articulated vehicles: Predicting sideswipe and rear-end collisions. *Transportation Research Interdisciplinary Perspectives*, *39*, 102236.

Caesar, H., Kabzan, J., Tan, K. S., Fong, W. K., Wolff, E., Lang, A., et al., 2021. nuPlan: A closed-loop ML-based planning benchmark for autonomous vehicles. *arXiv preprint arXiv:2106.11810*. URL: https://arxiv.org/abs/2106.11810.

Cheng, H., Ge, Q., Jiang, Y., Li, H., Chen, K., Wang, J., et al., 2025a. A necessary criterion for evaluating scene-level criticality metrics in safety verification of autonomous driving. In *2025 IEEE Intelligent Vehicles Symposium (IV)* (pp. 722–727). IEEE.

Cheng, H., Jiang, Y., Shi, Q., Meng, Q., Chen, K., Yu, W., et al., 2025b. Modified-emergency index (MEI): A criticality metric for autonomous driving in lateral conflict. *arXiv preprint arXiv:2510.27333*.

Cheng, H., Jiang, Y., Zhang, H., Chen, K., Huang, H., Xu, S., et al., 2025c. Emergency index (EI): A two-dimensional surrogate safety measure considering vehicles' interaction depth. *Transportation Research Part C: Emerging Technologies*, *171*, 104981.

Choi, J., Kim, S., 2025. Predictive risk-aware reinforcement learning for autonomous vehicles using safety potential. *Electronics*, *14*, 4446.

Dauner, D., Hallgarten, M., Li, T., Weng, X., Huang, Z., Yang, Z., et al., 2024. NAVSIM: Data-driven non-reactive autonomous vehicle simulation and benchmarking. *Advances in Neural Information Processing Systems*, *37*, 28706–28719.

ElSamadisy, O., Shi, T., Smirnov, I., Abdulhai, B., 2024. Safe, efficient, and comfortable reinforcement-learning-based car-following for AVs with an analytic safety guarantee and dynamic target speed. *Transportation Research Record*, *2678*, 643–661.

Fan, Y., Li, Y., Wang, S., 2024. Risk-aware self-consistent imitation learning for trajectory planning in autonomous driving. In *European Conference on Computer Vision* (pp. 270–287). Springer.

Feng, S., Zhu, H., Sun, H., Yan, X., He, L., Yang, J., et al., 2026. Breaking through safety performance stagnation in autonomous vehicles with dense learning. *Nature Communications*, *17*, 3163.

de Gelder, E., Op den Camp, O., 2024. Scenario-based assessment of automated driving systems: How (not) to parameterize scenarios? In *2024 IEEE International Automated Vehicle Validation Conference (IAVVC)* (pp. 1–7). IEEE.

Guo, H., Xie, K., Keyvan-Ekbatani, M., 2023. Modeling driver's evasive behavior during safety–critical lane changes: Two-dimensional time-to-collision and deep reinforcement learning. *Accident Analysis & Prevention*, *186*, 107063.

Huang, Y.-L., Chen, Y.-H., 2025. Estimating intersections' near-crash conflicts with the drone-based image-recording data. *Traffic Safety Research*, *9*, e000084.

Jafari, A., Liu, Y.-C., 2024. Pedestrians' safety using projected time-to-collision to electric scooters. *Nature Communications*, *15*, 5701.

Jiao, Y., 2023. A fast calculation of two-dimensional Time-to-Collision. URL: https://github.com/Yiru-Jiao/Two-Dimensional-Time-To-Collision.

Jiao, Y., Calvert, S. C., van Cranenburgh, S., van Lint, H., 2026. Learning collision risk proactively from naturalistic driving data at scale. *Nature Machine Intelligence*, *8*, 337–350.

Johnsson, C., Laureshyn, A., De Ceunynck, T., 2018. In search of surrogate safety indicators for vulnerable road users: a review of surrogate safety indicators. *Transport Reviews*, *38*, 765–785.

Li, H., Li, T., Yang, J., Shang, M., Wu, G., Wang, C., et al., 2026. PlannerRFT: Reinforcing diffusion planners through closed-loop and sample-efficient fine-tuning. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)* (pp. 24929–24938).

Li, Z., Li, K., Wang, S., Lan, S., Yu, Z., Ji, Y., et al., 2024. Hydra-MDP: End-to-end multimodal planning with multi-target Hydra-Distillation. *arXiv preprint arXiv:2406.06978*.

Nolte, M., Brettin, L. J., Steege, H., Salem, N., Loba, M., Graubohm, R., et al., 2025. A review of conceptualizations of safety and risk in current automated driving regulation. *arXiv preprint arXiv:2502.06594*.

Polychronopoulos, A., Tsogas, M., Amditis, A. J., Andreone, L., 2007. Sensor fusion for predicting vehicles' path for collision avoidance systems. *IEEE Transactions on Intelligent Transportation Systems*, *8*, 549–562.

Ray Sarkar, D., Ramachandra Rao, K., Chatterjee, N., 2024. A review of surrogate safety measures on road safety at unsignalized intersections in developing countries. *Accident Analysis & Prevention*, *195*, 107380.

Schubert, R., Adam, C., Obst, M., Mattern, N., Leonhardt, V., Wanielik, G., 2011. Empirical evaluation of vehicular models for ego motion estimation. In *2011 IEEE Intelligent Vehicles Symposium (IV)* (pp. 534–539). IEEE.

Schubert, R., Richter, E., Wanielik, G., 2008. Comparison and evaluation of advanced motion models for vehicle tracking. In *2008 11th International Conference on Information Fusion* (pp. 1–6). IEEE.

Tamke, A., Dang, T., Breuel, G., 2011. A flexible method for criticality assessment in driver assistance systems. In *2011 IEEE Intelligent Vehicles Symposium (IV)* (pp. 697–702). IEEE.

Tan, H., Lu, G., Wang, Z., Hua, J., Liu, M., 2024. A unified risk field-based driving behavior model for car-following and lane-changing behaviors simulation. *Simulation Modelling Practice and Theory*, *136*, 102991.

Just Accepted

Venthuruthiyil, S. P., Chunchu, M., 2022. Anticipated collision time (ACT): A two-dimensional surrogate safety indicator for trajectory-based proactive safety assessment. *Transportation Research Part C: Emerging Technologies*, *139*, 103655.

Wagner, S., Groh, K., Kühbeck, T., Dörfel, M., Knoll, A., 2018. Using time-to-react based on naturalistic traffic object behavior for scenario-based risk assessment of automated driving. In *2018 IEEE Intelligent Vehicles Symposium (IV)* (pp. 1521–1528). IEEE.

Wang, X., Alonso-Mora, J., Wang, M., 2022. Probabilistic risk metric for highway driving leveraging multi-modal trajectory predictions. *IEEE Transactions on Intelligent Transportation Systems*, *23*, 19399–19412.

Wang, X., Li, Z., Alonso-Mora, J., Wang, M., 2024. Reachability-based confidence-aware probabilistic collision detection in highway driving. *Engineering*, *33*, 90–107.

Westhofen, L., Neurohr, C., Koopmann, T., Butz, M., Schütt, B., Utesch, F., et al., 2023. Criticality metrics for automated driving: A review and suitability analysis of the state of the art. *Archives of Computational Methods in Engineering*, *30*, 1–35.

Zheng, L., Ismail, K., Meng, X., 2014. Traffic conflict techniques for road safety analysis: open questions and some insights. *Canadian Journal of Civil Engineering*, *41*, 633–641.

Zheng, X., Huang, B., Ni, D., Xu, Q., 2018. A novel intelligent vehicle risk assessment method combined with multi-sensor fusion in dense traffic environment. *Journal of Intelligent and Connected Vehicles*, *1*, 41–54.

Zhou, R., Zhang, G., Huang, H., Wei, Z., Zhou, H., Jin, J., et al., 2024. How would autonomous vehicles behave in real-world crash scenarios? *Accident Analysis & Prevention*, *202*, 107572.

Author biography



Hao Cheng received the B.Eng. degree in automotive engineering from the School of Vehicle and Mobility, Tsinghua University, Beijing, China, in 2022, where he is currently pursuing the Ph.D. degree in mechanical engineering. His research interests include driving risk quantification and safety-oriented post-training for end-to-end autonomous driving.



Yanbo Jiang received the bachelor's degree in traffic engineering from the School of Civil Engineering, Tsinghua University, Beijing, China, in 2023. He is currently pursuing the Ph.D. degree in mechanical engineering with the School of Vehicle and Mobility, Tsinghua University. His research interests include autonomous vehicles and large language models.



Wenhao Yu received the B.S. degree in vehicle engineering and the Ph.D. degree in transportation engineering from Jiangsu University in 2014 and 2020, respectively. He was a Visiting Ph.D. Student at the University of Waterloo in 2018–2019. He is currently an Associate Research Fellow with the State Key Laboratory of Intelligent Green Vehicle and Mobility, Tsinghua University. His research interests include the safety of the intended functionality of autonomous vehicles, traffic-law-compliant driving, risk assessment, safety decision-making, and testing and evaluation of autonomous driving systems.



Rui Zhou received the master's degree in control engineering from Wuhan University of Technology, Wuhan, China, in 2020, and the Ph.D. degree in traffic and transportation engineering from Central South University, Changsha, China, in 2024. He is currently a Postdoctoral Researcher with the School of Traffic and Transportation Engineering, Central South University. His research interests include traffic-safety analysis and autonomous-vehicle testing.



Jiang Bian received the master's degree in traffic and transportation planning and management from Beijing Jiaotong University, Beijing, China, in 2024. He is currently pursuing the Ph.D. degree in traffic and transportation engineering with Central South University, Changsha, China. His research interests include traffic-safety analysis and autonomous-vehicle testing.



Keyu Chen received the B.Eng. degree in vehicle engineering from Nanjing University of Aeronautics and Astronautics, Nanjing, China, in 2023. He is currently working toward the Ph.D. degree in mechanical engineering with the School of Vehicle and Mobility, Tsinghua University, Beijing, China. His research focuses on reinforcement learning post-training for autonomous driving, with broader interests in interactive learning and simulation for physical AI.

Just Accepted



Zhiyuan Liu received his bachelor's degree in mechanics and vehicle engineering from Xingjian College, Tsinghua University, Beijing, China, in 2024. He is currently pursuing the Ph.D. degree in mechanical engineering with the School of Vehicle and Mobility, Tsinghua University, Beijing, China. His research interests include end-to-end autonomous driving and vision-language-action (VLA) models.



Heye Huang received the Ph.D. degree in mechanical engineering from Tsinghua University, Beijing, China, in 2023. She was a Research Associate at the University of Wisconsin–Madison and a Postdoctoral Associate at the MIT SMART Center. She is currently an Assistant Professor with the Cho Chun Shik Graduate School of Mobility, Korea Advanced Institute of Science and Technology, Daejeon, Republic of Korea. Her research interests include trustworthy autonomy, generative artificial intelligence, risk-sensitive decision-making, and human-centered autonomous systems.



Hailun Zhang received the Ph.D. degree in vehicle engineering from Chang'an University in 2022. He was a Visiting Scholar at Delft University of Technology in 2021–2022 and a Postdoctoral Research Fellow at Tsinghua University in 2023–2025. He is currently an Associate Professor with the School of Automotive Engineering, Chang'an University. His research interests include vehicle dynamics, driving behavior modeling, and intelligent connected vehicles.



Fang Zhang received the B.Eng. degree in vehicle engineering and the Ph.D. degree in mechanical engineering from Tsinghua University, Beijing, China, in 2012 and 2018, respectively. He is currently an Associate Research Fellow with the State Key Laboratory of Intelligent Green Vehicle and Mobility, Tsinghua University. His research interests include autonomous driving and vehicle–infrastructure cooperation.



Jianqiang Wang is currently a Professor with the School of Vehicle and Mobility, Tsinghua University, Beijing, China. His research interests include intelligent vehicles, driving-assistance systems, and driver behavior. He was named a Changjiang Distinguished Professor in 2017.



Sifa Zheng received the B.Eng. and Ph.D. degrees in automotive engineering from Tsinghua University, Beijing, China, in 1993 and 1997, respectively. He is currently a Professor with the School of Vehicle and Mobility and the State Key Laboratory of Intelligent Green Vehicle and Mobility, Tsinghua University. He is also the Executive Deputy Director of the Suzhou Automotive Research Institute, Tsinghua University. His research interests include autonomous driving, vehicle dynamics, and control.

Just Accepted