Proc. of the European Safety and Reliability Conference (ESREL 2026)
Edited by José C. Matos, Paulo B. Lourenço, Michael Beer, Daniel Oliveira, Edoardo Patelli, Hélder S. Sousa, Rui A. Silva and Monica Santamaria

doi: 10.3850/ESREL2026061419_esrel26-p26409-cd

From Risk Theory to Road Safety: Developing an Axiomatic Risk Measure for Connected and Automated Vehicles (CAVs).

Hishma M Shah
Centre for Future Transport and Cities, Coventry University, United Kingdom. Email: shahh24@coventry.ac.uk
School of Engineering, Deakin University, Waurn Ponds, Victoria, Australia.

Huw Davies
Centre for Future Transport and Cities, Coventry University, United Kingdom. Email: huw.davies@coventry.ac.uk

Christophe Bastien
Centre for Future Transport and Cities, Coventry University, United Kingdom. Email: aa3425@coventry.ac.uk

Ashim Kumar Debnath
School of Engineering, Deakin University, Waurn Ponds, Victoria, Australia. Email: ashim.debnath@deakin.edu.au

Existing risk assessment methods like rule, risk, judgement, event, and vulnerability-impact based offer valuable insights for assessment of CAV safety. However, risk and safety remain contested concepts: risk is measured quantitatively, while safety provides normative reference boundaries defined outside the risk model. This paper reviews different approaches to risk assessment, their methods and their numerical measures, including probabilistic distributions, quantile thresholds, impact-based scoring, and composite indices. Sectoral exemplars from aviation, finance, medicine, rail, and cybersecurity show how each discipline operationalizes risk and the principles or axioms governing those practices. Together, these perspectives reveal both the diversity of risk methods and their shared limitations when applied in dynamic traffic scenarios. Drawing on the strengths and limitations identified across the sectoral exemplars, the paper proposes an axiomatic, real-time, risk measure for CAVs based on the probability-of-exceedance. Denoted P_γ, the measure quantifies the proportion of safety scores (e.g., time-to-collision) that fall below a critical threshold γ, under a probabilistic model, i.e., behaviour of other actors, weather and road conditions. The measure links probabilistic risk estimation to normative safety boundaries (e.g., thresholds derived from road rules), supports composite safety scoring, and enables time-consistent updating as new information becomes available. The result, a conceptual framework founded on the axioms of monotonicity, subadditivity, translation invariance, positive homogeneity, time-consistency, and interpretability, and is demonstrated through a worked scenario. The framework provides a foundation for empirical estimation of the proposed risk measure using Monte Carlo simulation under dependent, multi-actor environment models, with future work using expert elicitation to support the axiomatic grounding of the framework.

Keywords: Connected and Automated Vehicles, Axiomatic Risk Measures, Probability of Exceedance, Real-Time Risk, Risk Theory

1. Introduction

Risk and safety are central but contested concepts in engineering and regulation. Risk is typically quantified as a function of probability and consequence, while safety is normatively defined as the absence of "unreasonable risk," reflecting societal and moral standards (Lowrance 1976; ISO 2018). Connected and Automated Vehicles (CAVs) operate in dynamic, uncertain environments requiring real-time decisions across continuous interactions with pedestrians, cyclists, and other vehicles. Unlike conventional vehicles, CAVs must encode risk assessment algorithmically, making the choice of risk measure a design decision with direct safety and regulatory consequences. Establishing a quantitative basis that can align risk assessment with normatively defined safety criteria is a prerequisite for credible CAV deployment.

Existing risk measures are diverse, spanning probabilistic distributions, threshold-based indices, impact scoring, and composite factors (Kaplan and Garrick 1981; Aven 2011; MacKenzie 2014). Sectoral domains provide exemplars: aviation relies on rule-based certification and probabilistic risk assessment, finance has developed axiomatic risk measures such as Value-at-Risk (VaR) and Conditional Value-at-Risk (CVaR), medicine employs judgment-based severity indices, rail adopts event-based precursors, and cybersecurity integrates vulnerability–impact scoring. Each reflects not only different measures but different underlying philosophies of risk assessment (Hansson 2004). However, no single method is directly transferable to the CAV domain, where decision-making must be both real-time and societally legitimate (Koopman and Wagner 2017).

This paper addresses this gap by synthesizing cross-sectoral risk philosophies and numerical measures, extracting their implicit axioms or governing principles, and adapting them to the requirements of CAV safety assessment. On this basis, an axiomatic, real-time risk measure is developed, within which a probability-of-exceedance risk measure, denoted P_γ, is derived. The measure quantifies the proportion of safety scores (e.g., time-to-collision) that fall below a critical threshold γ, under a probabilistic model of how surrounding actors and environmental conditions may behave in the current situation. Different scenario assumptions correspond to different probability distributions within the same underlying model. The measure links probabilistic risk estimation to normative safety thresholds, supports time-consistent updating, and enables composite safety scoring under uncertainty, making it suitable for dynamic, multi-actor driving environments. The contribution of this paper is threefold:

- **i.** A structured review of risk assessment philosophies and numerical measures drawn from safety-critical domains such as aviation, finance, medicine, rail, and cybersecurity.
- **ii.** An adapted axiomatic foundation for CAV risk measure, derived from classical risk theory and informed by

requirements observed in established safety critical sectors.

iii. A probability-of-exceedance risk measure that operationalizes these axioms for real-time CAV risk estimation.

The remainder of the paper is organized as follows. Section 2 outlines the methodological approach. Section 3 reviews risk assessment philosophies and sectoral exemplars. Section 4 presents the development of the proposed risk measure. Section 5 and 6 concludes with implications for CAV safety assurance and future empirical validation.

2. Methodology

This paper adopts a conceptual review and synthesis methodology, structured to connect risk theory with the operational requirements of CAV safety. Rather than performing an exhaustive systematic review, the approach is designed to extract and integrate foundational principles from established safety-critical domains to inform the development of a CAV-specific risk measure (Jesson, Matheson, and Lacey 2011; Snyder 2019). The methodology consists of four stages.

i. Conceptual framing: The analysis begins by distinguishing probabilistic risk quantification from normative safety judgements (Section 1), providing a clear basis for comparing how different sectors operationalise risk.

ii. Sectoral review: Five safety-critical domains aviation, finance, medicine, rail, and cybersecurity are examined in Section 3, selected for their relevance as exemplars of distinct risk assessment philosophies: rule-based, risk-based, judgement-based, event-based, and vulnerability–impact-based. These sectors are selected not only for their maturity, but because they represent distinct and complementary risk assessment philosophies that together span the principal approaches across safety-critical systems; additional sectors (e.g., process safety, which combines event-based fault tree analysis with probabilistic risk assessment) largely reflect combinations of these approaches rather than introducing a distinct philosophy.

iii. Typological mapping: Risk measures identified within each sector are classified in *Section 3.3* according to numerical typologies, including probabilistic measures, threshold-based indices, impact-based scoring, and composite measures. This dual mapping by philosophy and by numerical form enables systematic comparison across domains and highlights both commonalities and limitations when applied to dynamic, multi-actor traffic environments.

iv. Axiomatic synthesis: Drawing on Artzner et al.'s (1999) axioms from financial risk theory and the implicit governing principles observed in other domains, an adapted set of axioms for CAV risk assessment is proposed in *Section 4.2*: monotonicity, subadditivity, translation invariance, positive homogeneity, time-consistency, and interpretability. These axioms are treated as *design criteria* that constrain how multiple uncertain factors should be combined within a coherent risk measure (Riedel 2004). They subsequently guide the formulation of the proposed probability-of-exceedance risk measure P_γ in *Section 4.3*, defined as the probability that a safety threshold is violated under an assumed probabilistic environment model

By combining cross-sectoral review with axiomatic synthesis, this methodology ensures that the proposed framework is both theoretically grounded in established risk theory and practically relevant to real-time CAV operation. While the review is selective rather than exhaustive, it brings together risk philosophies, numerical measures, and guiding principles to inform the development of a CAV-specific risk measure.

3. Sectoral Exemplars of Risk Assessment

Risk assessment traditions differ not only in methodology but in their underlying philosophies what constitutes "risk," what evidence is admissible, and how uncertainty is translated into decisions. For CAVs, where real-time decision-making must coexist with public accountability and regulatory legitimacy, understanding established approaches to risk assessment is essential. This section examines five dominant risk assessment philosophies, each illustrated by a sector with mature and well-established practice.

3.1 Rule-based risk assessment: Aviation

Aviation provides a clear example of a rule-based approach to risk assessment. Safety is judged primarily through compliance with explicit rules, standards, and numerical thresholds, producing binary certification outcomes: acceptable or unacceptable. Regulatory authorities, like the U.S. Federal Aviation Administration (FAA) and the European Union Aviation Safety Agency (EASA), mandate prescriptive reliability targets, such as catastrophic failure probabilities below 10^{-9}per flight hour, alongside defined design and verification processes (Federal Aviation Administration 1988; Hobbs et al. 2023). To support these rule-based decisions, aviation also relies on probabilistic modelling frameworks. Probabilistic Risk Assessment (PRA) decomposes hazards into scenarios s_i, each associated with a probability $P(s_i)$ and consequence $C(s_i)$, with aggregate risk expressed as Eq. (1) (Kaplan and Garrick 1981):

$$\boldsymbol{R} = \sum_i \boldsymbol{P}(\boldsymbol{s_i})\boldsymbol{C}(\boldsymbol{s_i}) \qquad (1)$$

In practice, the scenario probabilities $P(s_i)$ are derived from a combination of historical accident and incident data, system reliability models, and structured expert judgement. Where dependencies between contributing factors are complex or empirical data are limited, Bayesian Belief Networks (BBNs) can be used to represent causal structure and update probabilities consistently (Pearl 1988). The governing principles in aviation include regulatory anchoring, interpretability, and traceability. For CAVs, the lesson is clear: safety thresholds should be anchored in normative road rules (e.g., the Highway Code), while probabilistic and Bayesian tools provide the evidential basis for demonstrating compliance.

3.2 Risk-based assessment: Finance

Finance has developed the most formalized axiomatic treatment of risk. Unlike aviation's rule-based compliance, where the focus is on binary pass/fail thresholds, finance embraces probabilistic modelling to evaluate how portfolios behave under uncertainty (Aven 2016). The dominant questions are: *What is the probability of large losses? and how bad could they be?* Risk is expressed as a function of the underlying distribution of returns or losses, and the goal is to manage exposure to extreme but plausible outcomes (McNeil et al. 2015). Measures such as Value-at-Risk (VaR), Eq. (2), reflect a risk-based philosophy: they are not rule-driven or binary, but instead use the distribution of outcomes to characterise decision-relevant thresholds (Jorion 1996). VaR, identifies a loss threshold associated with a specified tail probability. For a loss random variable X and confidence level $q \in (0,1)$, VaR is defined as:

$$VaR_q(X) = \inf\{x \in \mathbb{R}: P(X \le x) \ge q\} \quad (2)$$

where, $\mathrm{P}(\cdot)$denotes probability, $\mathbb{R}$ the real numbers, and inf selects the smallest threshold x such that at least a fraction q of losses satisfy $X \le x$. Equivalently, losses exceed $\mathrm{VaR}_q(X)$ with probability $1 - q$.

However, VaR is known to violate the axiom of subadditivity, meaning that the assessed risk of a combined position can exceed the sum of its parts (Acerbi and Tasche 2002). While this limitation is well understood in finance, its implications are more severe for embodied systems such as robots and CAVs. As argued by Majumdar and Pavone (2020), risk measures that fail subadditivity can behave unintuitively in sequential and interactive settings, where multiple hazards interact non-linearly over time. In such contexts, combining individually “low risk” components can lead to underestimation or misrepresentation of overall system risk, undermining safety assurance.

This motivates the need for coherent risk measures that remain stable under aggregation and interaction, particularly in safety-critical robotic systems operating under uncertainty. To address these weaknesses, finance introduced Conditional Value-at-Risk (CVaR) Eq. (3), defined as:

$$\mathrm{CVaR_q(X)} = \mathbb{E}[\mathrm{X}|\mathrm{X} \ge \mathrm{VaR_q(X)}] \quad (3)$$

where, $\mathbb{E}[\cdot]$denotes expectation thus, CVaR is the average loss conditional on losses being at least as large as the VaR threshold, i.e. the tail mean over the most adverse $1 - q$ proportion of outcomes (Rockafellar and Uryasev 2002). CVaR is coherent, satisfying all Artzner et al.'s (1999) axioms and promoting diversification as a risk-reducing strategy. For CAVs, the key lesson is that axioms provide a rigorous foundation for combining uncertain factors and ensuring predictable behaviour of the risk measure, particularly in rare but high-consequence scenarios.

3.3 Judgment-based assessment: Medicine

Medical risk assessment often operates under conditions of sparse data and high contextual variability. As a result, expert judgement plays a central role (Rosa 1998; Gigerenzer 2008; Aven and Renn 2009). Severity scoring systems such as the Abbreviated Injury Scale (AIS) and the Injury Severity Score (ISS) encode clinical expertise into structured, interpretable impact measures (Baker et al. 1974; Association for the Advancement of Automotive Medicine [AAAM] 2016). Clinical risk decisions, such as whether to proceed with surgery, prescribe medication, or classify trauma severity, therefore rely on an explicit translation step performed by a health professional. Observable clinical findings are first interpreted and encoded into structured severity categories (e.g. AIS or ISS), which then link to empirically established outcome curves such as mortality or complication probabilities (Greenhalgh et al. 2014). As a result, probabilistic risk estimates in medicine are conditional on expert judgement rather than derived directly from raw observations. This makes medicine a natural exemplar of a judgement-based approach to risk assessment, in which formal probability models are mediated by professional interpretation and normative standards of care.

These approaches prioritize interpretability, severity, and patient-centered outcomes. For CAVs, where empirical data may be limited for rare or novel scenarios, structured expert elicitation provides a principled way to inform probabilistic models without sacrificing transparency.

3.4 Event-based assessment: Rail

Rail safety is grounded in event-based analysis, focusing on accident precursors and causal chains from initiating events to outcomes (Hollnagel and Goteman 2004; Aven 2011). Event trees, fault trees, and precursor monitoring are used to assess the frequency and severity of hazardous sequences. This approach has long been applied in high-reliability sectors such as nuclear power and is particularly well-developed in rail safety, where system complexity, interaction effects analogous to those observed in road transport environments, and public exposure demand rigorous causal modelling (Evans 2013).

Risk is commonly represented through probability–consequence relationships, emphasizing causality and traceability. For CAVs, event-based reasoning supports interpretability and post-incident accountability, particularly in complex multi-actor interactions.

3.5 Vulnerability–Impact based assessment: Cybersecurity

Cybersecurity adopts a vulnerability–impact philosophy, modelling risk as a function of threat likelihood, system vulnerabilities, and potential consequences (Stoneburner et al. 2002). Frameworks such as NIST SP 800-30 and ISO 27005 express risk as:

$$\mathrm{R} = \text{Likelihood} \times \text{Vulnerability} \times \text{Impact}$$

Composite indices, such as the Common Vulnerability Scoring System (CVSS), integrate qualitative and quantitative factors to prioritize mitigation efforts (Mell et al. 2007). For CAVs, this highlights the necessity of composite risk representations that account for interacting behavioral, environmental, and technical vulnerabilities (Zadeh et al. 2023).

4. Developing an Axiomatic Risk Measure for CAVs

Across safety-critical domains, different components of risk assessment are emphasized. Aviation operationalises safety primarily through rule-based compliance, supported by probabilistic models to demonstrate acceptable levels of residual risk. Finance focuses on formally defined risk measures derived from probability distributions of loss. Medicine relies heavily on expert judgement and qualitative risk scoring to assess patient harm. Rail safety emphasizes causal traceability through event-based models, while cybersecurity adopts vulnerability–impact representations to reason about exposure. Despite these differing emphases, all domains ultimately combine three elements:

i. Probabilistic uncertainty,
ii. Threshold-based safety boundaries,

iii. Structured principles for aggregating risk.

The CAV context demands a framework that integrates these elements while remaining interpretable and normatively anchored in road rules and expected behaviour. These axioms motivate the probability-of-exceedance risk measure P_γ which is subsequently illustrated using a worked intersection example.

4.1 Classical axioms of risk

Early work in finance and decision theory formalized risk measures through a set of *axioms* that ensure consistency and rationality (Föllmer and Schied 2002). Artzner et al. (1999) defined "coherent risk measures" with four core axioms:

i. **Monotonicity:** If losses under X are always greater than or equal to losses under Y, then the risk of X must be greater than or equal to the risk of Y, i.e. $\rho(X) \geq \rho(Y)$.
ii. **Subadditivity:** Diversification should not increase risk; the joint risk of two financial positions (i.e., two sources of uncertain risk exposure) is no worse than the sum of each, i.e. $\rho(X+Y) \leq \rho(X) + \rho(Y)$.
iii. **Translation invariance**: Adding a sure loss c increases risk by exactly c, i.e. $\rho(X+c) = \rho(X) + c$
iv. **Positive homogeneity:** Scaling exposure by $\lambda > 0$ scales risk proportionally, i.e. $\rho(\lambda X) = \lambda\rho(X)$.

These axioms were intended for finance but have since become recognized in risk theory. While mathematically elegant, these axioms are static and domain-agnostic. According to Riedel (2004) they assume:

- Stationary distributions of risk.
- Decisions made at a single time horizon.
- Risk independence from context or norms.

Majumdar and Pavone 2020 critique these axioms as *too weak* for embodied agents: a measure can satisfy them yet produce unintuitive or unsafe results in real-time control. For example, subadditivity might hold in finance, but for CAVs two "low risk" maneuvers (accelerating + lane change) can combine into high risk if traffic context changes. They proposed extensions:

- **Time-Consistency:** Risk assessments must not reverse preference over time (a maneuver deemed safe at t should not suddenly become unsafe at t+1 without new evidence).
- **Non-Negativity:** No negative risk; risk $\rho(X) \geq 0$ for all scenarios.
- **Interpretability:** Risk measure must map to meaningful, explainable safety concepts.

Rather than assuming a single deterministic future, CAV risk assessment must account for uncertainty in how surrounding road users and environmental conditions may evolve over time. This uncertainty is represented probabilistically, reflecting behavioural variability, data limitations, and contextual factors. Within this framework, risk is defined relative to a normative safety threshold and evaluated as the probability that this threshold is breached under the assumed environment model.

This perspective is particularly relevant for CAV safety, where empirical data are sparse for rare events and edge cases dominate safety outcomes. Importantly, it motivates the move away from single-model risk assessment and anticipates the probability-of-exceedance formulation introduced later, which likewise evaluates risk under explicitly modelled uncertainty in surrounding actor behaviour.

4.2 Adapted axioms for real-time CAV risk

Drawing on the limitations of classical risk axioms and the cross-sector synthesis in Section 3, an adapted set of axioms is proposed that preserves coherence where appropriate while incorporating properties essential for real-time road safety.

i. **Monotonicity (adapted).** If changes in traffic conditions such as junction control, visibility, or uncertainty in other actors' behaviour increase the likelihood of unsafe outcomes, the assessed risk must increase. This includes conditions that shift the distribution of safety metrics (e.g. time-to-collision) toward unsafe values.
ii. **Subadditivity and compositionality (context-aware).** In road traffic, diversification is not the objective. Instead, this axiom is interpreted as coherent aggregation under dependence, ensuring that interacting hazards (such as pedestrians and cross-traffic) are evaluated jointly without artificial inflation or masking of risk.
iii. **Translation invariance (threshold interpretation).** For CAVs, translation invariance reflects shifts in normative or regulatory definitions of unacceptable states. Changes in safety standards, such as revised rules governing acceptable interactions between road users should translate the assessed risk consistently, without altering the underlying risk framework.
iv. **Positive homogeneity (exposure scaling).** Scaling key exposures, such as speed or traffic density, should lead to predictable changes in assessed risk. For example, higher approach speeds increase stopping distances and interaction severity, shifting outcome distributions toward unsafe regions and increasing the probability of unsafe events. The requirement is consistent directional scaling rather than strict linearity.
v. **Time-consistency.** A maneuver assessed as unsafe at time t should not be classified as safe at time $t+1$ in the absence of new evidence. Risk assessment must therefore update sequentially and coherently as information becomes available.
vi. **Interpretability.** Risk outputs must map to meaningful safety concepts that are intelligible to regulators, policymakers, and the public, not only to engineers. This includes explicit linkage to rule-like safety thresholds.
vii. **Uncertainty sensitivity.** Risk must respond monotonically to changes in uncertainty about surrounding actors and conditions, such that increased behavioural or environmental uncertainty leads to higher assessed risk when safety margins are reduced.

Together, these axioms define the properties that a viable CAV risk measure must satisfy. However, a distinction must be noted regarding vehicle centric and system centric application of axioms. Vehicle centric measures the risk to the vehicle i.e. the CAV. System centric measures the risk present in the system. For example, consider a three-lane motorway where a vehicle in the nearside lane and a vehicle in the offside lane simultaneously merge into the middle lane. At a vehicle level, subadditivity holds;

the risk to the ego vehicle in the middle lane is no greater than the sum of the individual merge risks. However, at a system level, the simultaneous merging creates additional collision risks that are not present in either individual assessment, including the risk of the two merging vehicles colliding with each other. If the system evaluates each merge independently and treats joint risk as simply the sum of individual risks, these interactions are missed entirely, and subadditivity may be violated. This suggests that the proposed axioms are most directly applicable to vehicle-centric risk assessment, and that system-centric evaluations may require additional consideration in future work.

4.2.1 Road rules as de facto axioms

Unlike finance, CAV risk operates under explicit normative constraints derived from traffic law and shared road-user expectations (Hale and Borys 2013). In the UK, the Highway Code encodes behavioural requirements such as maintaining safe following distance, giving priority at junctions, and complying with speed limits. Within the proposed framework, these rules do not constitute additional axioms. Instead, they operationalize the adapted axioms by constraining both:

- **The definition of unsafe states** (through thresholds γ, supporting *interpretability* and *translation invariance*)
- **Set of plausible environment evolutions** $\mathbb{P}_0$ (through regulatory constraints on behaviour).

For example, a headway threshold derived from *"maintain a safe following distance"* defines the boundary between safe and unsafe regions in $f(X)$, directly supporting *interpretability*. Similarly, *right-of-way rules at junctions* determine which interaction outcomes are normatively unacceptable, shaping how *monotonic* risk changes are interpreted across contexts. As a result, even a manoeuvre with a low statistical collision probability may still be classified as unsafe if it violates these norms. This illustrates that CAV risk measures must integrate *mathematical coherence* with *normative constraints*, rather than treating safety as a purely statistical construct.

4.2.2 CAV Specific Requirements

The *Statement of Safety Principles* (Department of Transport 2025) defines safe driving as behaviour consistent with a *"careful and competent driver."* Translating this into quantitative terms requires a risk measure that:

- Operates in real time (responsive to changing traffic conditions).
- Is interpretable (transparent to regulators, operators, and the public).
- Is scenario-sensitive (vehicle following, merging, lane changing).
- Responds to uncertainty (sensor noise, unpredictable human behaviour).

Thus, any viable risk measure for CAVs must satisfy both the adapted axioms and the operational requirements of real-world driving. This motivates the application of a probability-of-exceedance risk formulation within a CAV-specific framework introduced in Section 4.3, that integrates probabilistic environment models, normatively defined safety thresholds, and composite representations of driving context.

4.3 The proposed measure: P_γ

Building on the adapted axioms and the normative definition of unsafe states, we define CAV risk as the probability of entering an unsafe state under uncertainty about future behaviour. Future interactions are inherently probabilistic and are represented through a probabilistic model of how surrounding road users and environmental conditions may evolve from the current situation. The proposed probability-of-exceedance risk measure is defined as Eq. 4:

$$P_\gamma = \mathbb{P}_0(f(X) < \gamma) \tag{4}$$

where:

- X denotes the system variables that capture ego behaviour, surrounding actor behaviour and relevant environmental conditions,
- $f(X)$ is a deterministic transformation of system input variables into a single scalar safety score, such as minimum time-to-collision, headway, or braking margin,
- γ is a regulatory or normative safety threshold (e.g. TTC < 2s),
- $\mathbb{P}_0$ denotes the probability measure over the future states of the system, capturing the range and frequency of possible outcomes. The distributional character of the safety score arises from $\mathbb{P}_0$ varying across possible system states.

This formulation defines risk directly as the probability that a safety threshold is violated. The threshold γ defines what constitutes an unsafe state and is therefore a normative input to the risk definition rather than a parameter that scales the underlying physical danger of a situation. It follows that a more conservative γ, for example TTC < 4s rather than TTC < 2s, produces a higher exceedance probability for the same physical situation, reflecting a stricter definition of acceptable safety rather than greater underlying danger. Importantly, P_γ is a function of both γ and $\mathbb{P}_0$, and their relationship is non-linear, a step change in γ does not produce a proportionate change in P_γ, as the effect depends on the shape of the distribution at that threshold. The physical risk of a situation is therefore compared across scenarios by holding γ fixed and varying the environment model $\mathbb{P}_0$.

This definition yields several properties aligned with the requirements of CAV risk assessment:

- **Regulatory anchoring.** Road rules shape both the probabilistic model of future environment states (e.g. signalized versus unsignalized junctions lead to different behavioural likelihoods) and the normative reference points for interpreting which states are unacceptable.
- **Composite safety representation.** The safety score $f(X)$ combines multiple system variables. For a TTC score, these inputs include the ego vehicle's speed, position, and heading, together with the corresponding states of other system actors.
- **Time-consistent updating.** The measure may be indexed by time, $P_{\gamma,t}$, and recomputed as new sensor evidence updates the environment state.
- **Uncertainty reporting.** Confidence intervals or sensitivity bounds on P_γ may be reported to reflect estimation uncertainty where appropriate.

Therefore, P_γ is a *risk-theoretic measure intended for safety assessment and assurance*. It does not prescribe vehicle actions and is not a control objective but rather provides a transparent and interpretable basis for evaluating safety under uncertainty.

4.4 Worked example: Time-to-Collision at a signalized pedestrian crossing

This section presents a worked example demonstrating how the proposed probability-of-exceedance risk measure P_γ is constructed and evaluated for a simple but representative connected and automated vehicle (CAV) scenario. The purpose of the example is to show how uncertainty in the surrounding environment alters the assessed level of risk, rather than to model driver decision-making or vehicle control.

4.4.1 Baseline scenario

Consider a signalized urban intersection with a marked pedestrian crossing. At time t_0, the ego CAV approaches the stop line at approximately 40 km/h on a dry road surface. A pedestrian is present at the curb, and the signal state permits pedestrian crossing. The positions and velocities of all actors are known at t_0; however, future interactions are uncertain because the pedestrian's crossing decision and movement are not known in advance. Accordingly, the uncertain environment state is represented as Eq. (5):

$$X = \{T_{p,\text{start}}, v_p\}, \tag{5}$$

where $T_{p,start}$ denotes the pedestrian step-off time relative to t_0, and $v_p$denotes pedestrian walking speed during the crossing. Environmental conditions W(e.g., rain, reduced visibility, surface friction) are not included directly in X. Instead, they are treated as contextual factors that modify the probability distributions of $T_{p,start}$ and v_p, thereby reshaping the induced distribution of future interactions and TTC. *Pedestrian step-off time* is modelled using Liu et al.'s (2014) parametric survival distributions, which represent uncertainty in pedestrian crossing initiation. Empirical studies of signalised crossings show that waiting durations vary widely across individuals and contexts, and that the *probability of stepping off* changes continuously rather than at a fixed threshold. These models therefore provide plausible probabilistic representations of pedestrian entry into the conflict zone. Figure 1 shows probability density functions of pedestrian walking speed $\mathrm{v_p}$ once the pedestrian has stepped onto the crossing. The horizontal axis represents walking speed in metres per second, and the vertical axis represents probability density. They are illustrative parametric models, with parameters chosen to reflect ranges reported in Koh and Wong (2014) and Hashimoto et al. (2016), these will be replaced with observed distributions in future work.

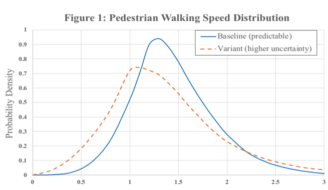

Figure 1: Pedestrian walking speed distributions. The baseline distribution reflects normal crossing behaviour under favourable conditions. The higher-uncertainty distribution reflects increased behavioural variability due to factors such as crowding, occlusion, or poor visibility (Koh and Wong 2014; Hashimoto et al. 2016).

The baseline distribution is narrow, reflecting relatively predictable walking behaviour under normal conditions. The higher-uncertainty distribution is wider, capturing greater variability in walking speed due to hesitation, rushing, or heterogeneous pedestrian responses. Together, uncertainty in pedestrian step-off timing and walking speed defines uncertainty in how long and when the pedestrian occupies the conflict zone.

4.4.2 Time-to-collision distribution arising from pedestrian uncertainty

A safety score function is defined as Eq. (6):

$$f(X) = \mathrm{TTC}\,, \tag{6}$$

Uncertainty in pedestrian step-off timing and walking speed jointly defines uncertainty in the future relative position and speed of the pedestrian with respect to the ego vehicle. Because relative position and speed are uncertain, TTC is not a single value, but a range of possible values $\mathbb{P}_0$ (Eq.4). Figure 2 shows the cumulative distribution function of TTC resulting from propagation of uncertainty in pedestrian behaviour, including uncertainty in crossing initiation and walking speed from Figure 1. The vertical dashed line indicates the safety threshold $\gamma = 2$s. A normative safety threshold $\gamma = 2$s is adopted, consistent with empirical road-safety guidance and the principle of a "careful and competent driver".

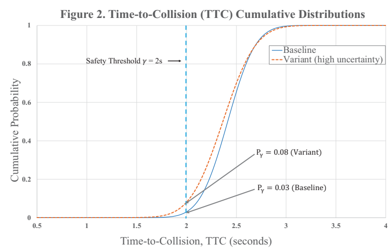

Figure 2: Cumulative distribution functions of minimum time-to-collision (TTC) between the ego vehicle and a pedestrian under baseline and higher-uncertainty environment models. The vertical dashed line indicates the safety threshold $\gamma = 2$ s. The value of the cumulative distribution at the threshold corresponds to the exceedance probability $\mathrm{P}_\gamma = \mathbb{P}_0\,[\mathrm{TTC} < \gamma]$.

The area to the left of the threshold corresponds to the probability that the safety threshold is breached. For the baseline uncertainty model $\mathbb{P}_{0\,base}$, the cumulative probability at γ is approximately 0.03, indicating a 3% probability of entering an unsafe state Eq. (7):

$$P_\gamma(base) = \mathbb{P}_{0\,base}[f(X) < \gamma] = 0.03. \tag{7}$$

Under the variant uncertainty model $\mathbb{P}_{0\,var}$, corresponding to increased uncertainty in pedestrian behaviour, the TTC distribution shifts such that a larger proportion of probability mass lies below the threshold. The exceedance probability increases to Eq. (8):

$$P_\gamma(\mathrm{var}) = \mathbb{P}_{0\,var}[f(X) < \gamma] = 0.08. \tag{8}$$

These values illustrate how increased behavioural uncertainty shifts the assessed level of risk from 3% to 8% under the same normative threshold.

4.4.3 Interpretation and axiomatic behaviour

This worked example demonstrates how the proposed probability-of-exceedance risk measure operates in a concrete CAV safety context. The increase in assessed risk from the baseline to the variant case illustrates *monotonicity*: worsening uncertainty in surrounding actor behaviour leads to a higher probability of breaching the safety threshold. The unsafe region remains explicitly defined by the same normative threshold $f(X) < \gamma$, preserving *interpretability* across scenarios.

By evaluating risk as the probability of violating a safety threshold under an uncertain environment, the formulation remains sensitive to changes in surrounding actor behaviour. Because the definition of risk does not change as uncertainty models are updated, the approach supports *time-consistent* reassessment as new information becomes available.

The example does not assume or prescribe any driver or vehicle response. Instead, it demonstrates that the assessed level of risk changes solely as a function of uncertainty in the surrounding environment, providing a principled foundation for later analysis of behavioral adaptation and decision-making.

5. Implications for CAV Risk Measure

5.1 Risk as exceedance of unsafe states

A central implication of the proposed framework is the reframing of CAV risk in terms of exceedance of unsafe states rather than expected loss or severity-weighted outcomes. In road traffic, safety is governed by normative boundaries such as minimum headway, stopping distance, or time-to-collision, beyond which behaviour is considered unacceptable regardless of how seldom violations occur. Defining risk as the probability of breaching such boundaries therefore aligns naturally with regulatory logic, which assesses safety relative to expected, rule-conforming behaviour rather than worst-case adversarial actions.

This supports comparison across scenarios and conditions and enables consistent interpretation of safety performance under variability in behaviour and environment.

5.2 Relationship to existing risk measures

The proposed probability-of-exceedance measure improves on existing sectoral risk measures by combining their strengths while addressing their limitations in dynamic, multi-actor road environments. Unlike VaR measures commonly used in finance, P_γ preserves tail sensitivity while maintaining subadditivity and interpretability. In contrast to purely rule-based thresholds employed in aviation, the measure adapts continuously as conditions evolve, rather than operating as a static pass–fail criterion.

Compared to judgement-based approaches in medicine, the framework provides a mathematical risk measure while still accommodating expert elicitation through the specification of probabilistic environment models. Unlike event-based approaches in rail safety, which focus on discrete failure modes, the formulation generalizes naturally to continuous, time-varying, and interactive traffic scenarios. Finally, in contrast to vulnerability–impact models used in cybersecurity, uncertainty is embedded directly through probabilistic modelling of surrounding actors and conditions rather than relying on scenario enumeration alone.

5.3 *Estimation and practical application*

While the contribution of this paper is conceptual rather than algorithmic, the proposed measure is amenable to practical estimation. In applied settings, P_γ can be estimated using Monte Carlo simulation or related rare event sampling techniques, drawing samples from the joint distribution of the environment state. The probabilistic environment model $\mathbb{P}_0$ is specified through distributions over uncertain elements such as road-user behaviour and environmental conditions, as illustrated in the pedestrian-crossing example.

The formulation also supports time-consistent updating. As new sensor data are observed, the distribution over the environment state can be refined and the exceedance probability recomputed, yielding a rolling estimate of risk that remains consistent with the underlying axioms.

5.4 Implications for regulation and safety assurance

Because the proposed risk measure is explicitly anchored to normative safety thresholds, it provides a transparent basis for regulatory interpretation and for supporting safety assurance arguments. Evaluation under explicitly modelled uncertainty in surrounding behaviour aligns with conservative decision-making practices in other safety-critical domains and supports structured reasoning about risk under uncertainty. Rather than replacing existing rules or safety metrics, the framework complements them by quantifying the likelihood of threshold violations under uncertainty. This makes it suitable for use in assurance cases, comparative safety assessment, and regulatory discussions of acceptable risk in the context of CAV deployment.

6. Conclusion

This paper has argued that the safe deployment of CAVs requires a risk framework that is both mathematically coherent and normatively interpretable. By synthesising risk philosophies and numerical measures across aviation, finance, medicine, rail, and cybersecurity, the paper identified a set of axioms and operational requirements that are particularly relevant to dynamic, multi-actor road environments.

Building on this foundation, the paper introduced P_γ, a probability-of-exceedance formulation for CAV safety risk. The contribution does not lie in the exceedance equation itself but, in how it is instantiated for road traffic: through the identification of traffic-relevant axioms, the explicit grounding of safety thresholds in road rules and behavioural norms, and the construction of safety score functions that encode real interaction constraints between multiple system actors.

The formulation embeds regulatory anchoring, time-consistent updating, and structured interpretation within a single risk measure. The worked intersection example demonstrated how uncertainty in surrounding system actors behaviour can be propagated through this framework to yield an interpretable probability of entering an unsafe state.

Future work will focus on empirical estimation and validation through expert elicitation, Monte Carlo simulation, and scenario-based testing. These methods are not add-ons but necessary consequences of the framework's complexity: once behaviour, interaction, and normative constraints are represented explicitly, analytical solutions are no longer sufficient. Together, these steps will support both real-time risk assessment and regulatory safety assurance and provide a foundation for the development and evaluation of interpretable safety metrics for CAVs.

References

Acerbi, Carlo, and Dirk Tasche. 2002. “On the Coherence of Expected Shortfall.” *Journal of Banking & Finance* 26 (7): 1487–1503. https://doi.org/10.1016/S0378-4266(02)00283-2.

Artzner, Philippe, Freddy Delbaen, Jean-Marc Eber, and David Heath. 1999. “Coherent Measures of Risk.” *Mathematical Finance* 9 (3): 203–28. https://doi.org/10.1111/1467-9965.00068.

Aven, Terje, and Ortwin Renn. 2009. “On Risk Defined as an Event Where the Outcome Is Uncertain.” *Journal of Risk Research* 12 (1): 1–11. https://doi.org/10.1080/13669870802488883.

Aven, Terje. 2011. *Quantitative Risk Assessment: The Scientific Platform*. Cambridge: Cambridge University Press. https://doi.org/10.1017/cbo9780511974120

Aven, Terje. 2016. “Risk Assessment and Risk Management: Review of Recent Advances on Their Foundation.” *European Journal of Operational Research* 253 (1): 1–13. https://doi.org/10.1016/j.ejor.2015.12.023.

Baker, Susan P., Brian O’Neill, William Haddon Jr., and William B. Long. 1974. “The Injury Severity Score: A Method for Describing Patients with Multiple Injuries and Evaluating Emergency Care.” *Journal of Trauma* 14 (3): 187–96.

Department for Transport. 2025. *Automated Vehicles: Statement of Safety Principles*. London: UK Government. https://www.gov.uk/government/calls-for-evidence/automated-vehicles-statement-of-safety-principles

Evans, Andrew W. 2013. “The Economics of Railway Safety.” *Research in Transportation Economics* 43 (1): 137–47. https://doi.org/10.1016/j.retrec.2012.12.003.

Federal Aviation Administration. 1988. *AC 25.1309-1A: System Design and Analysis*. Washington, DC: FAA. https://doi.org/10.21949/1514499.

Föllmer, Hans, and Alexander Schied. 2002. “Convex Measures of Risk and Trading Constraints.” *Finance and Stochastics* 6 (4): 429–47. https://doi.org/10.1007/s007800200072.

Gigerenzer, Gerd. 2008. “Why Heuristics Work.” *Perspectives on Psychological Science* 3 (1): 20–29. https://doi.org/10.1111/j.1745-6916.2008.00058.x.

Greenhalgh, Trisha, Jeremy Howick, and Neal Maskrey. 2014. “Evidence Based Medicine: A Movement in Crisis?” *BMJ* 348: g3725. https://doi.org/10.1136/bmj.g3725.

Hale, Andrew, and David Borys. 2013. “Working to Rule, or Working Safely? Part 1: A State of the Art Review.” *Safety Science* 55: 207–21. https://doi.org/10.1016/j.ssci.2012.05.011.

Hansson, Sven Ove. 2004. “Philosophical Perspectives on Risk.” *Techné* 8 (1): 10–35. https://doi.org/10.5840/techne2004818

Hashimoto, Yoriyoshi, Yanlei Gu, Li-Ta Hsu, Miho Iryo-Asano, and Shunsuke Kamijo. "A probabilistic model of pedestrian crossing behavior at signalized intersections for connected vehicles." *Transportation research part C: emerging technologies* 71 (2016): 164-181. https://doi.org/10.1016/j.trc.2016.07.011

Hobbs, Kathryn L., Matthew L. Mote, Michael C. Abate, Samuel D. Coogan, and Eric M. Feron. 2023. “Runtime Assurance for Safety-Critical Systems.” *IEEE Control Systems Magazine* 43 (2): 28–65. https://doi.org/10.1109/MCS.2023.3234380.

Hollnagel, Erik, and Olof Goteman. 2004. “The Functional Resonance Accident Model.” In *Proceedings of Cognitive System Engineering in Process Plant*, 155–61.

ISO. 2018. *ISO 31000:2018 Risk Management—Guidelines*. Geneva: International Organization for Standardization.

Jesson, Jill, Fiona M. Lacey, and Lydia Matheson. "Doing your literature review: Traditional and systematic techniques." (2011): 1-192. https://doi.org/10.1080/09500790.2011.581509

Jorion, P. (1996). Risk2: Measuring the risk in value at risk. *Financial analysts journal*, *52*(6), 47-56. https://doi.org/10.2469/faj.v52.n6.2039

Kaplan, Stanley, and B. John Garrick. 1981. “On the Quantitative Definition of Risk.” *Risk Analysis* 1 (1): 11–27. https://doi.org/10.1111/j.1539-6924.1981.tb01350.x.

Koh, P. P., and Y. D. Wong. "Gap acceptance of violators at signalised pedestrian crossings." *Accident Analysis & Prevention* 62 (2014): 178-185. https://doi.org/10.1016/j.aap.2013.09.020

Koopman, Philip, and Michael Wagner. 2017. “Autonomous Vehicle Safety: An Interdisciplinary Challenge.” *IEEE Intelligent Transportation Systems Magazine* 9 (1): 90–96. https://doi.org/10.1109/MITS.2016.2583491.

Liu, Yan, Rushdi Alsaleh, and Tarek Sayed. "Modeling pedestrian temporal violations at signalized crosswalks: A random intercept parametric survival model." *Transportation research record* 2676, no. 6 (2022): 707-720. https://doi.org/10.1177/03611981221076119

Lowrance, William W. 1976. *Of Acceptable Risk: Science and the Determination of Safety*. Los Altos, CA: William Kaufmann. http://hdl.handle.net/10822/772475

MacKenzie, Cameron A. 2014. “Summarizing Risk Using Risk Measures and Risk Indices.” *Risk Analysis* 34 (12): 2143–62. https://doi.org/10.1111/risa.12220.

Majumdar, Anirudha, and Marco Pavone. 2020. “How Should a Robot Assess Risk?” In *Proceedings of Robotics Research*, 75–84. Springer. https://doi.org/10.1007/978-3-030-28619-4_10.

McNeil, Alexander J., Rüdiger Frey, and Paul Embrechts. *Quantitative risk management: concepts, techniques and tools-revised edition*. Princeton university press, 2015.

Mell, Peter, Karen Scarfone, and Sasha Romanosky. 2007. “Common Vulnerability Scoring System.” *IEEE Security & Privacy* 4 (6): 85–89. https://doi.org/10.1109/MSP.2006.145

Pearl, Judea. 1988. *Probabilistic Reasoning in Intelligent Systems*. San Francisco: Morgan Kaufmann.

Pritsker, Matthew. 1997. “Evaluating Value at Risk Methodologies.” *Journal of Financial Services Research* 12 (2): 201–42. https://doi.org/10.1023/A:1007978820465.

Riedel, Frank. 2004. “Dynamic Coherent Risk Measures.” *Stochastic Processes and Their Applications* 112 (2): 185–200. https://doi.org/10.1016/j.spa.2004.03.004

Rockafellar, R. Tyrrell, and Stanislav Uryasev. 2002. “Conditional Value-at-Risk for General Loss Distributions.” *Journal of Banking & Finance* 26 (7): 1443–71. https://doi.org/10.1016/S0378-4266(02)00271-6.

Rosa, Eugene A. 1998. “Metatheoretical Foundations for Post-Normal Risk.” *Journal of Risk Research* 1 (1): 15–44. https://doi.org/10.1080/136698798377303.

Snyder, Hannah. "Literature review as a research methodology: An overview and guidelines." *Journal of business research* 104 (2019): 333-339. https://doi.org/10.1016/j.jbusres.2019.07.039

Stoneburner, Gary, Alice Goguen, and Alexis Feringa. 2002. *Risk Management Guide for Information Technology Systems*. NIST SP 800-30. https://doi.org/10.6028/NIST.SP.800-30.

Zadeh, Amin, Benjamin Lavine, Hamed Zolbanin, and David Hopkins. 2023. “A Cybersecurity Risk Quantification and Classification Framework.” *Decision Analytics Journal* 9: 100328. https://doi.org/10.1016/j.dajour.2023.100328.