Proceedings of the 35th European Safety and Reliability & the 33rd Society for Risk Analysis Europe Conference
Edited by Eirik Bjorheim Abrahamsen, Terje Aven, Frederic Bouder, Roger Flage, Marja Ylönen
 Published by Research Publishing, Singapore.
doi: 10.3850/978-981-94-3281-3_ESREL-SRA-E2025-P8451-cd

From Uncertainty Representation to Safety Performance Monitoring for Operational Safety Assurance - A Systematic Approach

Nishanth Laxman
Fraunhofer IESE, Germany. E-mail: nishanth.laxman@iese.fraunhofer.de

Joel Varghese Jiby
RPTU Kaiserslautern-Landau, Germany. E-mail: jiby@rhrk.uni-kl.de

Ioannis Sorokos
Fraunhofer IESE, Germany. E-mail: ioannis.sorokos@iese.fraunhofer.de

Joshua Frey
Fraunhofer IESE, Germany. E-mail: joshua.frey@iese.fraunhofer.de

Jan Reich
Fraunhofer IESE, Germany. E-mail: jan.reich@iese.fraunhofer.de

Recent advancements in Automated Driving Systems (ADS), driven by substantial investments, have significantly enhanced ADS technologies. However, traditional methods for the design, development, verification, and validation of safety-critical automotive systems are inadequate for managing the increased complexity and operational uncertainties of ADS, making the assurance of their operational safety in dynamic environments an unresolved challenge. Current operational safety approaches use various approaches to incrementally challenge the validity of assurance cases but lack the integration of field data. The increasing availability of real-time vehicle data presents an opportunity to identify potential runtime uncertainties affecting safety assurance cases. By continuously refining and expanding assurance cases with field data, additional evidence or counter-evidence, and other relevant information through a DevSafeOps process, the safe operation of ADS can be assured.
A crucial aspect of operational safety assurance is Safety Performance Monitoring (SPM) using Safety Performance Indicators (SPIs), which are essential for both operational safety and compliance with standards such as UL 4600 and BSI PAS 1881 for the deployment of ADS. SPIs quantify safety performance and can be used to monitor the validity of safety arguments during operation. SPIs at sufficiently detailed sub-claim levels can proactively identify potential violations of safety case claims in a 'leading' manner, before safety-critical events occur. Additionally, they can provide supplementary evidence to address residual uncertainties after deployment.
This paper primarily addresses SPM for operational safety, presenting a novel systematic approach that spans from uncertainty representation in assurance cases using Dempster-Shafer theory to employing dialectics and argument defeaters, ultimately defining useful SPIs related to various claims in an assurance case. This approach aids in concretely identifying and defining SPIs based on an assurance case and facilitates the runtime field data-based validation of assurance cases, additionally aiding in standards conformance. The approach is demonstrated through a construction zone assist case study for ADS.

Keywords: Operational Safety Assurance, Safety Performance Monitoring, Autonomous Driving Systems, DevSafeOps

1. Introduction

Automated Driving Systems (ADS) represent a promising solution to key challenges in modern mobility, such as reduced emission, effective time management, and enhanced comfort for drivers. ADS are envisioned to operate in unpredictable and open environments; therefore it is essential for them to have robust and intelligent system architectures that can manage runtime uncertainties, adapt to changing conditions, and ensure safe operation in the inherently uncertain and complex nature of realworld driving. The emergence of

ADS is linked to substantial research in areas such as sensing, on-board processing, AI and machine learning, and object recognition and localization. The complexities associated with these fields, which are characterized by inherent uncertainties, complicate the realization of the benefits offered by ADS and ensuring the notion of public acceptance.

Safety assurance is paramount for ADS and it is essential to demonstrate safety by providing compelling evidence that supports the system's safety claims. The use of Safety Assurance Cases (SACs) has become a standard practice to show various stakeholders that these systems are safe [Ayoub et al. (2013)]. SACs present *"structured argument, supported by evidence, intended to justify that a system and activity is acceptably safe for a specific application in a specific operating environment and reflects recognized good practices"* [The British Standards Institution (2022)]. SACs and structured arguments are intended to be "living documents" but are treated closer to static ones in conventional systems. They are generally valid for a specific system under consideration at the time of submission to a certification authority, based on the evidence available during its development. For ADS, it can be the case that after deployment, new or contradictory information could emerge, potentially undermining the validity of SAC due to unexpected system properties, environmental uncertainties, or insufficiency in system performance. Even a well-defined SAC cannot fully address the uncertainties inherent in the subjective nature of safety argumentation, raising questions about its final conclusions. Furthermore, the evaluation done by experts is also subjective, as the evidence presented in SACs is manually analyzed and therefore the result could be prone to biases and prejudices [Gyllenhammar et al. (2023)].

Operational safety assurance methods are helpful to maintain confidence in the validity of SACs post-deployment, especially when ADS are in operation and the increasing availability of real-time vehicle data offers a valuable opportunity to detect potential runtime uncertainties that could affect this validity [Laxman et al. (2024)]. Research in this direction is slowly gaining traction, especially for runtime model-based assurance cases. R. Wei et al. in [Wei et al. (2024)] emphasizes the need to transition from conventional design-time processes to runtime processes using SACM [OMG (OMG)]. E. Asaadi et al. in [Asaadi et al. (2020)] introduce dynamic assurance cases specifically in the context of AI and machine learning systems. C.Carlan et al. in [Cârlan et al. (2024)] propose a Dynamic Safety Case Management System (DSCMS) to address the challenges of frontier AI systems.

Safety Performance Indicators (SPIs) are designed to provide predictive safety insights, assess the effectiveness of safety measures, enable continuous improvement, support risk management decision-making, and advocate Safety Performance Monitoring (SPM), to achieve operational safety. Through continuous monitoring and analysis of SPIs, SACs can be validated, thereby reinforcing safety arguments or identifying and mitigating runtime uncertainties before they lead to accidents [Koopman (2022)]. Currently, there is no standard/systematic approach widely used for specifying SPIs, to the best of our knowledge.

This paper presents a novel systematic approach implemented during the development phase to tackle the challenges posed by uncertainties related to SACs. It delineates a method for resolving the claims through Eliminative Argumentation (EA) and effectively identifies specific claims within SACs where valuable SPIs can be defined to validate certain claims using field data. This process not only facilitates safety performance monitoring but also enhances operational safety. The presented approach will also assist in compliance for standards like UL 4600 [Underwriter Laboratories (2022)] and BSI PAS 1881 [The British Standards Institution (2022)].

The rest of the paper is organized as follows; Background and related work are presented in Chapter 2. The systematic approach is presented in Chapter 3 and exemplarily illustrated with an use case from ADS domain in Chapter 4. Chapter 5 presents our conclusions and future planned work.

2. Background and Related Work

2.1. *Dempster-Shafer Theory (DST)*

Dempster-Shafer Theory (DST) differs from probability theory by representing belief, disbelief, and uncertainty regarding an event through beliefs. It accommodates ambiguity and encompasses all aspects of assumptions. DST is particularly valuable when multiple evidence sources are involved, ensuring that all conflicts and uncertainties are addressed [Shafer (1992)]. In DST, *Frame of discernment* refers to the set of all possible outcomes of an event, while the *Power set* describes all the possible subsets derived from the frame of discernment. *Mass function* assigns probabilities to the elements within the power set. To further elaborate, consider an event A: its *Belief* is defined as the sum of all masses that support the assertion that A is true, whereas its *Disbelief* represents the sum of masses that contradict or negate the occurrence of A. *Plausibility* is the sum of all masses that may support the truth of A, indicating that some uncertainty remains, with belief lying within the range of plausibility. *Conflict*, or *conflict measure (K)*, quantifies the degree of contradiction among the sources of evidence, calculated as the sum of the products of mass functions where the subsets have no overlap. Lastly, *Normalizing factor*, represented as "1-K" describes the non-contradicting portion of the evidence [Yager (1987)].

2.2. *Operational Safety Assurance*

Eliminative argumentation (EA) is a framework for constructing and assessing confidence in assurance cases, which explain why a system possesses desired properties, such as safety. The approach emphasizes the iterative nature of argumentation, where confidence in claims of a SAC can be incrementally increased as doubts or "defeaters" are systematically identified and eliminated. Defeaters are the reasons for doubting a particular claim. Confidence is quantified using a Baconian probability notation, where the proportion of eliminated doubts to total doubts reflects one's confidence in a claim [Goodenough et al. (2015)].

EA considers three types of defeaters: Rebutting, Undermining, and Undercutting. A *Rebutting* defeater counters a claim. An *Undermining* defeater questions the validity of evidence provided. An *Undercutting* defeater identifies specific conditions under which the validity of a conclusion is uncertain, when the premises of an *Inference Rule* are true. The interplay of arguments, defeaters, inference rules, and evidence creates a confidence map, which visually represents reasons to doubt a claim. *Rebutting* and *Undercutting* defeaters are introduced with "unless," while *Undermining* defeaters start with "but," aiding clarity and readability. As more defeaters are identified and resolved with appropriate evidence, the assurance level (confidence level) increases incrementally [Goodenough et al. (2013)].

Dialectic argumentation (DA) is a systematic approach for conducting dialectic analysis of safety cases. It is employed prior to the deployment of a system to identify potential runtime challenges to safety arguments or evidence, thereby validating the safety case. This method emphasizes the explicit representation of arguments and evidence that challenge or undermine safety claims. By identifying challenges to various elements of a SAC, dialectic arguments can be formed. These challenges may manifest as claims that, if true, would dispute existing assertions or as counter-evidence that undermines the argument. The authors suggest that explicit runtime monitoring of the system and its environment for occurrences of counter-evidence to maintain the validity of the safety case [Hawkins and Ryan Conmy (2023)].

Both EA and DA employ iterative approaches; however, our view is that EA provides greater flexibility for developing and resolving doubts.

2.3. *Safety Monitoring and Safety Performance Monitoring*

Safety monitoring and safety performance monitoring are two distinct yet related concepts of safety management. Safety monitoring focuses on continuous observation of risk-related features in the driving environment, enabling dynamic adaptations to driving functions based on residual risk levels or dynamic runtime risk as-

sessments [Haupt and Liggesmeyer (2019)]. In contrast, safety performance monitoring evaluates the long-term effectiveness of Safety Management Systems (SMS) by analyzing safety outcome data, such as incident rates and the success of safety mechanisms. While safety monitoring addresses immediate concerns in a reactive manner, safety performance monitoring adopts a proactive approach over a longer time, aiming to strengthen the overall safety framework of ADS and assure operational safety [Thieme and Utne (2017)].

Safety performance monitoring is facilitated by SPIs, which have long been integral to SMS in the aviation industry. They can be either qualitative or quantitative, and using a mix of both approaches can effectively address various challenges that may arise from relying on a single methodology [Kešel'ová et al. (2021)].

3. Methodology - The Systematic Approach

Our systematic approach involves a 4 step methodology to realize operational safety assurance. It combines different approaches presented in Chapter 2 to facilitate safety performance monitoring, a crucial aspect of operational safety.

STEP 1: Uncertainty elicitation

The first step involves collecting opinions from safety experts (or reviewers) and describing them in terms of decision and confidence values.

Decision and *Confidence* are both represented as qualitative scales with values ranging from 0 to 1. The *Decision* scale indicates the extent to which an expert leans towards a particular conclusion regarding a claim, with values ranging from acceptance to rejection. A value of 1 signifies complete acceptance of the claim, a value of 0 indicates outright rejection, and a value of 0.5 reflects the expert's indecision.

The *Confidence* scale quantifies the degree of information available to the expert that supports the decision made. A value of 1 denotes that the expert possesses sufficient information to justify the decision, whereas a value of 0 indicates a lack of adequate information necessary for justification [Idmessaoud et al. (2022)].

After collecting information from experts, the next step is to plot it on an evaluation matrix shown in Figure 1. The *Decision* and *Confidence* values are then converted into *Belief* (Bel), *Disbelief* (Disb), and *Uncertainty* (Unc) values using the formulae:

$$\begin{aligned} \mathrm{Bel}(x) &= \frac{\mathrm{Conf}(x) - 1}{2} + \mathrm{Dec}(x) \\ \mathrm{Disb}(x) &= \frac{\mathrm{Conf}(x) + 1}{2} - \mathrm{Dec}(x) \\ \mathrm{Unc}(x) &= 1 - \mathrm{Bel}(x) - \mathrm{Disb}(x) \end{aligned} \tag{1}$$

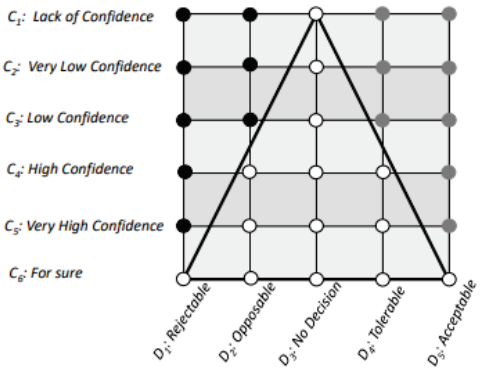

Fig. 1.: Evaluation matrix [Idmessaoud et al. (2022)].

A constraint called "Josang constraint" in the evaluation matrix must be considered while determining the values. The values outside the triangle in the evaluation matrix gives negative belief (black dots) or disbelief (grey dots) which might be invalid. This leads to another set of formulae to adjust the values so that the "Josang constraint" is adhered to [Jøsang (2016)].

When $\mathrm{Dec}(x) > \frac{1+\mathrm{Conf}(x)}{2}$, the Decision values must be set to the following:

$$\mathrm{Dec}(x) = \frac{1 + \mathrm{Conf}(x)}{2} \tag{2}$$

When $\mathrm{Dec}(x) < \frac{1-\mathrm{Conf}(x)}{2}$, the Decision values must be set to the following:

$$\mathrm{Dec}(x) = \frac{1 - \mathrm{Conf}(x)}{2} \tag{3}$$

Bel, Disb, and Unc values are elicited for all the leaf goals in a SAC.

STEP 2: Uncertainty propagation

The second step involves propagating the Bel, Disb, and Unc values upward along the SAC until the top goal. For this process, the type of argument relationship plays a crucial role. The argument relationship types are *Conjunctive argument* (C-arg), *Disjunctive argument* (D-arg), and *Hybrid argument* (H-arg).

In a C-arg, for a conclusion C having two (or more) premises P1 and P2, both premises are together needed to support the conclusion, like a Boolean logic AND gate. In a D-arg, either one of them is required to reach the conclusion C, like a Bollean logic OR gate [Idmessaoud et al. (2022)].

The Bel, Disb, and Unc values are propagated along the SAC using Dempster's rule of combination. Depending on the type of argument in the SAC, their calculation varies. Different formulae used are:

For a D-arg,

$$bel = 1 - \prod_{i=1}^{n} (1 - bel_i) \tag{4}$$

$$disb = \prod_{i=1}^{n} (disb_i) \tag{5}$$

For a C-arg,

$$bel = \prod_{i=1}^{n} (bel_i) \tag{6}$$

$$disb = 1 - \prod_{i=1}^{n} (1 - disb_i) \tag{7}$$

While DST provides a clear mathematical basis for handling uncertainties, this gives the clarity of the uncertainty for the top goals [Wang et al. (2019)].

STEP 3: Claims Resolution

The third step involves resolving each claim based on the uncertainty values determined in the previous steps. First, a certain threshold for uncertainty is defined by a domain expert, which varies and is dependent on various factors, based on the domain, SIL levels, safety criticality of the functions being developed, and corresponding risk scores. The claims, which have uncertainty higher than the defined threshold, have to be resolved. For this purpose, EA is used.

The claims are subjected to EA by questioning its existence with the help of doubts or defeaters. The different types of defeaters are explained in Chapter 2. The EA pattern remains unchanged regardless of whether the defeater is Rebutting, Undermining, or Undercutting. New claims or new evidence can be introduced to counter a defeater. Inference rules combine the purpose of claims with the supporting evidence to address the defeater and describe and update the confidence levels immediately after the defeaters are resolved. This requires efficient combination of claims and evidence. Different combinations of claims and evidence may also give rise to multiple inference rules. When certain evidences or inference rules cannot be further developed, it remains as residual doubt, which is usually left unaddressed or undeveloped. Those that need to be further developed or the ones that have been challenged again, have to be resolved by giving new evidence. Certain claims, whether from the initial SAC or added to counter defeaters, cannot be resolved due to insufficient runtime operational data. For such claims, SPIs are defined accordingly.

STEP 4: SPI Definition

The identified SPIs must be realistic, relevant, and directly linked to safety objectives, regardless of their simplicity or complexity. Generally, a combination of SPIs is needed to clearly reflect safety performance. An SPI is defined as a metric, substantiated by evidence, that utilizes threshold comparisons to validate claims within a SAC. It comprises of a metric-threshold pair that assesses a specific facet of safety. These metrics may pertain to various aspects, including product performance, design quality, process quality, or compliance with operational procedures. A threshold establishes an acceptable target value. A metric alone does not qualify as an SPI, as context is crucial within a SAC [Koopman (2022)].

4. Use Case Analysis

Construction zone assist (CZA) is a feature in ADS designed to enhance vehicle safety and nav-

igation in active construction zones, along the highway and in urban areas. It uses sensors and data analysis to detect construction zones, adjusting vehicle behavior accordingly based on model predictive control (MPC) to ensure safe passage through these environments [Motelay et al. (2023)].

An initial SAC for CZA is created, considering a top goal (G1) as "ADS safely navigates along the construction zones without any collisions or run offs". Two sub-goals G2 and G3 are defined and further broken down into G4 through G9, which are addressed by the solutions (evidence) S1 through S5, as illustrated in Figure 2.

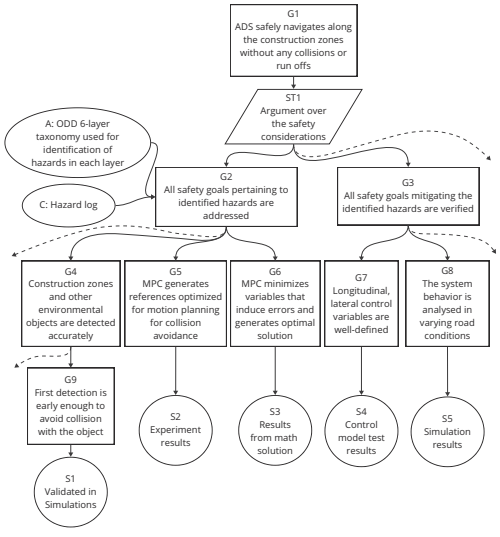

Fig. 2.: Initial Safety Assurance Case for CZA

As described in Chapter 3, uncertainty elicitation is first carried out by safety experts, where they give their opinion about the leaf goals (G9, G5, G6, G7, G8) in the form of *Decision* and *Confidence* (Dec, Conf). This process is carried out using an evaluation matrix where the experts decide on their decision on the leaf goals and the confidence which they support their decision with. This gives a quantified data of the expert opinion in (Dec, Conf) scale, as shown in Table 1.

The Dec and Conf values for leaf goals are then converted into Bel, Disb, Unc using formula (1). These values are then propagated up towards the top goal. G4 has the same set of values as G9

Table 1.: Decision and Confidence values from uncertainty elicitation

Goal	*Decision*	*Confidence*
G9	*0.5*	*0.6*
G5	*0.75*	*0.8*
G6	*1*	*0.8*
G7	*0.5*	*0.4*
G8	*0.75*	*0.6*

Table 2.: Belief, Disbelief and Uncertainty values from DST

Goal	*Belief*	*Disbelief*	*Uncertainty*
G9	*0.3*	*0.3*	*0.4*
G4	*0.3*	*0.3*	*0.4*
G5	*0.65*	*0.15*	*0.2*
G6	*0.8*	*0*	*0.2*
G7	*0.2*	*0.2*	*0.6*
G8	*0.55*	*0.05*	*0.4*
G2	*0.156*	*0.405*	*0.439*
G3	*0.11*	*0.24*	*0.65*
G1	*0.01716*	*0.5478*	*0.43504*

as they are connected as a simple argument. The relationship of G4, G5, and G6 to G2 is that of a conjunctive argument, the values are calculated as per the formulae (6) and (7). The same applies for G7, G8 towards G3. These values propagate towards the top goal as a conjunctive argument using the same formulae. Bel, Disb, Unc values of each goal in SAC is listed in Table 2.

The uncertainty values are individually resolved and based on a tolerance threshold of 0.2. Leaf goals G7, G8, and G9, which have higher uncertainty, are identified for further resolution. Goal G7, which is defined as "Longitudinal, and lateral control variables are well- defined", is resolved by two rebuttal defeaters R7.1 and R7.2. R7.1 is supported by the evidence provided by S4, which is further defeated by an undermining defeater UM7 and is left as a residual undeveloped doubt there. A new goal is added to address R7.2, which requires data from the field. This particular goal requires 2 SPIs, one pertaining to maximum saturation limit and another to minimum satura-

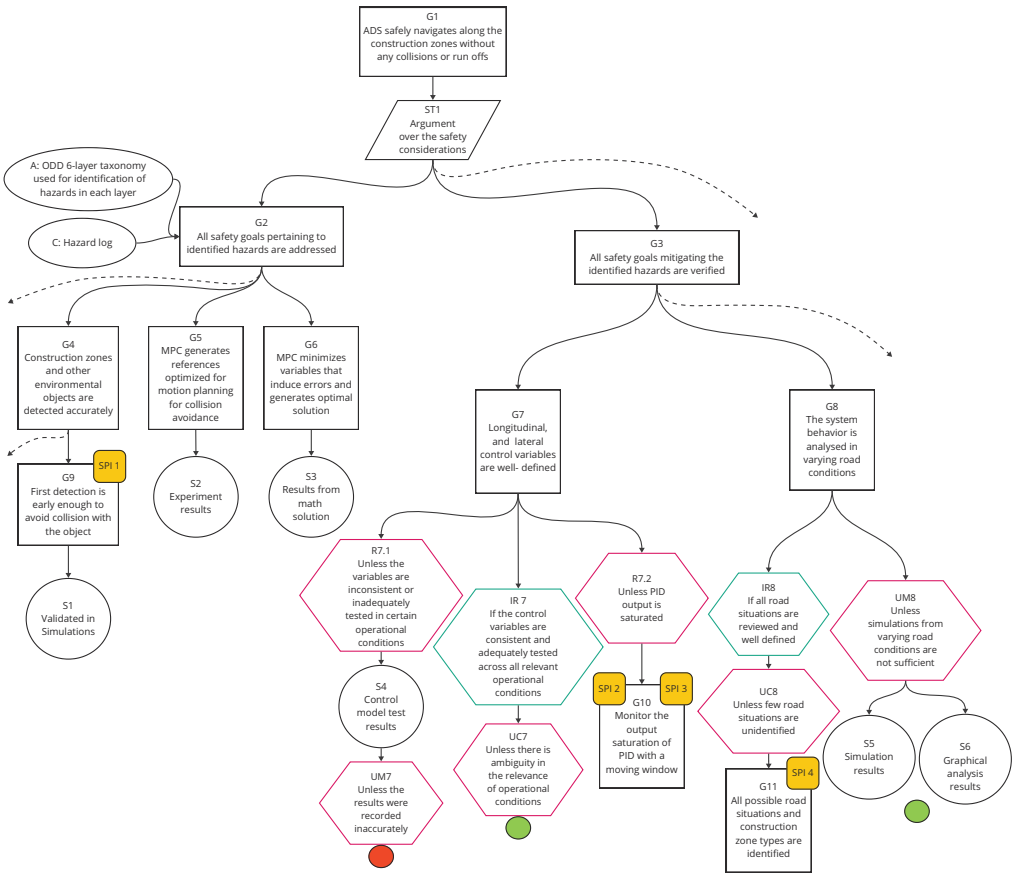

Fig. 3.: Final Safety Assurance Case for CZA

tion limit of PID controller. An ”inference rule” IR7, which is associated to G7, is countered by an undercutting defeater UC7, it is then considered to be acceptable. Similarly, G8 is also addressed using IR8 and UC8 by defining a new goal G11, which an SPI will have to be defined for. A new evidence S6 is added to UM8 to categorize it as acceptable. Goal G9 cannot be addressed with defeaters or counter-arguments but directly with an SPI. The final resulting SAC for CZA is shown in Figure 3.

Lastly, leading SPIs 1, 2, 3 and 4 are defined as follows:

SPI 1: Rate of first detection below TTC_MIN is $<$ than SPI1_THRESH

SPI 2: Rate of PID output saturation $>$ PID_SAT_MAX is $<$ SPI2_THRESH

SPI 3: Rate of PID output saturation $<$ PID_SAT_MIN is $<$ SPI3_THRESH

SPI 4: Rate of EBM_INITIATION without scenario identification is $<$ SPI4_THRESH

These SPIs have to be monitored at runtime for violations to aid Safety performance monitoring.

5. Conclusion and Future Work

This paper presents a novel systematic approach to enhance operational safety assurance for ADS through effective SPM. By integrating DST with EA, we have proposed a systematic methodology for resolving the uncertainties associated with SAC and also identifying SPIs. SPIs utilize real-time field data to continuously refine and validate SACs, ensuring that they remain relevant and effective in addressing the dynamic runtime

uncertainties inherent in ADS operations. The approach helps identify claims that need SPIs in an effective manner and to the best of our knowledge there is no similar approach to holistically address operational safety assurance.

The case study on CZA, exemplifies how this approach can be applied in practice, revealing a structured path to resolving claims and enhancing confidence in safety arguments. In future work, we will focus on representing SPIs and gathering field data for SPI violations and SAC validation. In the event of an SPI violation, a Change Impact Analysis will be performed, supported by Digital Dependability Identities (DDI) technology [Zeller et al. (2023)], to facilitate the update cycle using DevSafeOps.

Acknowledgement

The German Federal Ministry of Education and Research (BMBF) funded this work under grant number 01IS22087Q — AutoDevSafeOps.

References

Asaadi, E., E. Denney, J. Menzies, G. J. Pai, and D. Petroff (2020). Dynamic assurance cases: a pathway to trusted autonomy. *Computer 53*(12), 35–46.

Ayoub, A., J. Chang, O. Sokolsky, and I. Lee (2013). *Assessing the overall sufficiency of safety arguments*. IET.

Cârlan, C., F. Gomez, Y. Mathew, K. Krishna, R. King, P. Gebauer, and B. R. Smith (2024). Dynamic safety cases for frontier ai. *arXiv preprint arXiv:2412.17618*.

Goodenough, J. B., C. B. Weinstock, and A. Z. Klein (2013). Eliminative induction: A basis for arguing system confidence. In *ICSE 2013*, pp. 1161–1164. IEEE.

Goodenough, J. B., C. B. Weinstock, and A. Z. Klein (2015). Eliminative argumentation: A basis for arguing confidence in system properties. *Software Engineering Institute, Carnegie Mellon University, Pittsburgh, PA, Tech. Rep. CMU/SEI-2015-TR-005*.

Gyllenhammar, M., G. R. de Campos, and M. Törngren (2023). Holistic perspectives on safety of automated driving systems-methods for provision of evidence. *Authorea Preprints*.

Haupt, N. B. and P. Liggesmeyer (2019). A runtime safety monitoring approach for adaptable autonomous systems. In *SAFECOMP 2019 Proceedings 38*, pp. 166–177. Springer.

Hawkins, R. and P. Ryan Conmy (2023). Identifying run-time monitoring requirements for autonomous systems through the analysis of safety arguments. In *SAFECOMP 2023*, pp. 11–24. Springer.

Idmessaoud, Y., D. Dubois, and J. Guiochet (2022). Uncertainty elicitation and propagation in gsn models of assurance cases. In *SAFECOMP 2022*, pp. 111–125. Springer.

Jøsang, A. (2016). *Subjective logic*, Volume 3. Springer.

Kešel'ová, M., M. Blišt'anová, P. Hanák, and L. Brůnová (2021). Safety management system in aviation: comparative analysis of safety management system approaches in v4 countries. *Management Systems in Production Engineering 29*(3), 208–214.

Koopman, P. (2022). *How Safe Is Safe Enough?: Measuring and Predicting Autonomous Vehicle Safety* (3 ed.).

Laxman, N., M. Schweizer, and J. Reich (2024). Employing field monitoring and runtime safety verification for cumulative operational safety assurance in a safetydevops process. In *ICSRS 2024*.

Motelay, H., I. Mangini, I. Prando Manzini, V. Moreno Sanches, and J.-V. Zacchi (2023). Construction zone assist. Technical description.

OMG. Structured assurance case metamodel. `https://www.omg.org/spec/SACM/About-SACM/`. Accessed: 2025-01-11.

Shafer, G. (1992). Dempster-shafer theory. *Encyclopedia of artificial intelligence 1*, 330–331.

The British Standards Institution (2022). PAS 1881:2022 Assuring the operational safety of automated vehicles – Specifi cation. Specification.

Thieme, C. A. and I. B. Utne (2017). Safety performance monitoring of autonomous marine systems. *Reliability Engineering & System Safety 159*, 264–275.

Underwriter Laboratories (2022). ANSI/UL 4600 - Standard for Evaluation of Autonomous Products. Standard.

Wang, R., J. Guiochet, G. Motet, and W. Schön (2019). Safety case confidence propagation based on dempster–shafer theory. *International Journal of Approximate Reasoning 107*, 46–64.

Wei, R., S. Foster, H. Mei, F. Yan, R. Yang, I. Habli, C. O'Halloran, N. Tudor, T. Kelly, and Y. Nemouchi (2024). Access: Assurance case centric engineering of safety–critical systems. *Journal of Systems and Software 213*, 112034.

Yager, R. R. (1987). On the dempster-shafer framework and new combination rules. *Information sciences 41*(2), 93–137.

Zeller, M., I. Sorokos, J. Reich, R. Adler, and D. Schneider (2023). Open dependability exchange metamodel: a format to exchange safety information. In *RAMS 2023*, pp. 1–7. IEEE.