

Overview of Operational Safety Concepts for Level 4 Automated Driving System Fleets

Safety Hazard Qualitative Risk Scale

A qualitative risk scale is proposed to categorize the identified safety hazards. Each safety hazard is assigned a risk category based on its potential consequences, represented by the end-states of the ESD diagrams developed for each operational phase. Given the scope of the analysis, a detailed breakdown of the consequences under different conditions is not performed at this point (e.g., different speeds may result in different hazard levels in case of a collision).

The proposed multi-dimensional qualitative risk scale is composed of a combination of “relative frequency”, “controllability” and “severity” inspired by the ISO 26262 ASIL risk assessment methodology (International Organization for Standardization, 2018). For this work, a high relative frequency, low controllability, and high severity would result in a high risk.

A conservative approach is used to characterize the consequences represented by the ESD end-states. Conservative risk assessments are generally adopted when the analysis contains significant uncertainties (National Research Council (US) Committee on Risk Assessment of Hazardous Air Pollutants., 1994). There are two main sources of uncertainties in this analysis. The first arises from the project’s scope: the specific conditions in which the risk scenarios take place are not defined (e.g., weather conditions, vehicles’ speed, surrounding traffic information). The second main source of uncertainties refers to the lack of probabilistic or frequency data for complete risk quantification, including hardware, software, and human failures for ADS L4 operating as MaaS.

This approach is a specific strategy employed to address uncertainty and variability for estimating risk that favors one type of error (overestimation) over its converse (underestimation). For instance, any incidental scenario in which a passenger or other road user is involved is categorized as high severity (level 4). Despite the potential overestimation of risks, the proposed scale is satisfactory for describing and categorizing the safety hazards in a hierarchical approach.

The structure of each of these scales is described in the following sections.

Severity scale

The severity is classified on a scale from 1 to 4. The consequences include traffic disruption, property damage-only (PDO), and risk of fatality and injury (passengers and/or other road users). The following considerations are adopted (Table B.1):

- Level 1 corresponds to scenarios in which operation does not lead to any traffic, property, or injury related consequence, e.g., a passenger trip has been successfully completed. Organizational errors and failure to follow procedures are also considered at this level as these do not produce any immediate consequences.
- Level 2 corresponds to scenarios in which the interruption of an ADS vehicle’s operation causes traffic disruptions and any incidents that may occur are so minor as to not result in property damage or injury. It should be noted that some conditions may lead to more or less severe consequences. For instance, multiple vehicles entering MRC close to hospitals

or evacuation routes cause a traffic disruption that may pose a danger to lives, as well as vehicles entering MRC in areas that reduce the road visibility to other road users.

- Level 3 corresponds to scenarios in which the ADS vehicle’s operation has been interrupted or has been involved in an incident. No aggravating factors are present, i.e., no passengers or other road users have been exposed to harm.
- Level 4 corresponds to scenarios in which the ADS vehicle’s operation has been interrupted or has been involved in an incident. This level also covers scenarios where the vehicle is not responsive to remote commands. One or more aggravating factors are present, i.e., passengers or other road users have been exposed to harm.

A conservative approach is taken toward the presence of potential hazards for passengers on board and other road users in the vicinity of the ADS vehicle. As a result, most of the post-incident scenarios are classified as level 4 (fatality and injury) instead of level 3 (PDO), regardless of the severity of the incident itself.

Table B.1: Description of qualitative severity scale.

Consequence	Description	Level	Examples
No incident	The operation occurs as expected. No operational errors that lead to immediate hazards.	1	The vehicle safely completes a trip to the intended destination with no incidents.
			The MOC crew correctly performs expected actions or faulty actions do not lead to hazards (i.e., faulty inspection process does not necessarily lead to a system failure).
Traffic disruption	The vehicle’s operation is interrupted, e.g., a crash does not occur or if it does occur, it is so minor as to not result in property damage and injury. The vehicle achieves MRC and needs to be retrieved by the MOC crew or operates under MR-DDT conditions.	2	The ADS vehicle is dispatched to MOC in MR-DDT condition.
			The vehicle engages MRC and post-incident procedures are initiated. No other road users are involved.
Property damage-only (PDO)	The vehicle is involved in an incident where no passengers or road users are injured.	3	Incident without passengers onboard and no other road users are injured. Post-incident procedures are followed.
			Incident without passengers onboard and no other road users are involved. Post-incident procedures are not followed.
Fatality and Injury	The vehicle is (1) involved in an incident involving injuries or fatalities to ADS occupants and/or other road users, or (2) unresponsive to remote commands with passengers onboard and/or	4	Communication between vehicles and FOC is limited or severed. Vehicle and/or passengers are in an unknown state.
			The vehicle is unreachable or unresponsive to remote commands and fails to autonomously implement DDT fallback actions when required.
			Incidents with or without passengers onboard and/or other road users are involved. Post-incident procedures are followed.
			Incidents with or without passengers onboard and/or other road users are involved. Post-incident procedures are not followed.

	affecting other road users.		FOC is unaware other road users are involved in the incident and does not contact first responders or does not provide them with correct information.
--	-----------------------------	--	---

Controllability scale

According to the Automotive Safety Integrity Level in the ISO 26262 functional safety standard, controllability represents the level of the ability of the driver to avoid harm. However, several challenges have been identified in applying the controllability scale, particularly in the context of automated vehicle operation (De Gelder et al., 2021; Khastgir et al., 2017).

In a MaaS context with L4 ADS when no safety driver onboard, the term can be adapted to represent the ability of the participating agents (the ADS vehicle, FOC operators, and MOC crew members) to avoid harm. This proves a structured approach to categorize scenarios based on how successful these agents are in performing predefined tasks and procedures. It is considered that if the three agents act as expected, they have a higher ability to prevent and mitigate harm, i.e., the operation is designed such that harm can be avoided in most circumstances. Thus, a higher level of controllability is achieved when the ADS vehicle, the FOC operators, and the MOC crew act according to the operational requirements. The agents' actions are categorized as either:

- a) Prevention actions: Actions available to avoid an incident occurring. E.g., a vehicle detects a failure and safely enters MRC (with or without assistance from the FOC).
- b) Mitigation actions: Actions available to mitigate harm after an incident has occurred. E.g., after a vehicle enters MRC, the FOC initiates post-incident procedures.

The controllability is assessed through four levels (Table B.2):

- High (1): High controllability refers to scenarios in which all the participating agents act as expected. This includes scenarios in which the vehicle is rerouted to the MOC due to non-safety critical failures.
- Medium (2): Medium controllability refers to scenarios in which one of the agents does not act as expected. However, other agents may perform additional preventive or mitigative actions. For instance, the MOC fails to detect a vehicle failure during an inspection. However, the ADS and the FOC may detect failure during operation, and the vehicle can perform a DDT fallback before causing an incident.
- Low (3): Low controllability refers to scenarios in which two or more participating agents do not behave as expected. This level refers to scenarios where agents fail to prevent harm, although mitigation actions may still be performed, e.g., the FOC follows the post-incident procedures to recover the vehicle.
- Very low (4): Very low controllability refers to scenarios in which an incident has occurred, and no preventive or mitigative actions are available for the agents to prevent or mitigate consequences. This includes failures to implement safety-related measures during post-incident procedures (i.e., contacting first responders).

Table B.2: Description of qualitative controllability scale.

Controllability	Description	Level	Examples
High	All agents behave as expected.	1	The vehicle safely completes a trip to the intended destination with no incidents.
			The ADS vehicle is dispatched to MOC in MR-DDT condition.
Medium	An agent does not behave as expected and both preventive and mitigative actions may be available.	2	The MOC crew performs less than adequate inspection or maintenance activities. Preventive and mitigative actions are available, i.e., the ADS system may engage MRC if the self-diagnostic module detects a system failure and the FOC operator may engage MR-DDT or MRC if abnormal vehicle behavior is detected.
			The vehicle engages MRC and post-incident procedures are available to mitigate risks.
Low	Two or more agents do not behave as expected and no preventive actions are available. Mitigation actions may still be available.	3	The vehicle engages MRC, but post-incident procedures are not initiated. Mitigation actions are still available, as the FOC operator may initiate post-incident procedures after communicating with passengers and/or first responders.
			The MOC fails to dispatch a recovery team during post-incident procedures. Mitigation actions are still available, as the MOC may dispatch a recovery vehicle after communicating with the FOC operators.
			FOC fails to dispatch a secondary vehicle for passengers to continue trip after a vehicle failure. Mitigation actions are still available, i.e., as the FOC operator may dispatch a secondary vehicle after communicating with passengers.
Very Low	Two or more agents do not behave as expected and no preventive or mitigative actions are available.	4	The vehicle is unreachable or unresponsive to remote commands and fails to autonomously implement DDT fallback actions when required.
			Communication between vehicles and FOC is limited or severed. Vehicle and/or passengers are in unknown state.
			The vehicle engages MRC and post-incident procedures are not followed.
			FOC is unaware other road users are involved in the incident and does not contact first responders or does not provide them with correct information.

Relative frequency scale

As little operational experience has been documented in sufficient depth to retrieve quantitative measures of likelihood or frequency data to characterize the scenarios, the proposed scale is based on the expected relative frequency of the end-state with respect to the initiating event corresponding to each ESD and the events leading to the ESD.

The relative frequency is estimated through:

$$f_{rel} = f_{es} \times f_{ie},$$

where f_{es} represents the relative frequency of an end-state with respect to the other possible end-states stemming from the same initiating event, and f_{ie} represents the relative frequency of the initiating event with respect to a period of ADS vehicle operation.

The relative frequency of an end state is estimated considering the probability of the event that may lead to them. For instance, a successful end state such as “trip successfully completed” is expected to be more frequent than the state concerning an incident and post-incident failures: the path from the IE to the successful end state involves the “success” path of the events, which is expected to have a higher probability than the “failure paths” (e.g., it is expected that the vehicle has a higher probability of functioning as expected than of presenting a critical failure while in operation).

The initiating event relative frequency is categorized as follows:

- High (3): End-states derived from initiating events with expected high relative frequency considering the entire fleet operation. These correspond to a) ADS Vehicle is on-route to destination without passengers, b) ADS Vehicle is on-route to destination with passengers, c) ADS Vehicle is scheduled for passenger pick-up, and d) ADS Vehicle is scheduled for passenger drop-off.
- Medium (2): End-states derived from initiating events with expected medium relative frequency considering the entire fleet operation. These correspond to e) ADS Vehicle is scheduled to arrive at MOC, f) ADS Vehicle is scheduled for pre-shift inspection, and g) ADS Vehicle is scheduled for service maintenance.
- Low (1): End-states derived from initiating events with expected low relative frequency considering the entire fleet operation. These correspond to h) post-incident procedures are initiated.

The end-state relative frequency is categorized as follows:

- High (3): End-states which are expected to regularly occur during the operational phase. This refers to successful end-states indicating a trip has been completed or that inspection and maintenance activities have successfully reflected the state of the vehicle.
- Medium (2): End-states which may occur during the operational phase. This refers to end-states resulting from low-severity vehicle failures and from less than adequate inspection/maintenance procedures.
- Low (1): End-states which are not expected to occur during the operational phase. This refers to end-states resulting from critical vehicle failures and from failures to follow operational procedures during vehicle post-incident management.

This scale is based on modeling assumptions which may overestimate the risk of low-likelihood events. In particular, the likelihood of the end-states resulting from the post-incident procedures operational phase is potentially several orders of magnitude smaller than end-states resulting from the on-route operational phases, which is not captured in the proposed scale ranging between 1-3.

The resulting relative frequency f_{rel} is then categorized into four levels (Table B.3). The description of each relative frequency category is presented in Table B.4. In the event there is data available to quantify both the initiating event frequency and the probability of failure of the ESD events, a new relative frequency scale would need to be developed to adequately reflect each scenarios' risk.

Table B.3: Relative frequency matrix

Initiating Event/End-State Relative Frequency	High	Medium	Low
High	1	1	3
Medium	1	2	3
Low	3	3	4

Table B.4: Description of qualitative relative frequency scale.

Consequence	Level	Examples
High	1	The vehicle safely completes a passenger trip to the intended destination with no incidents. This corresponds to a high relative frequency of the initiating event and end-state.
		The ADS vehicle is dispatched to MOC for inspection and maintenance activities. This corresponds to a high relative frequency of the initiating event and a medium relative frequency of the end-state.
Medium	2	The MOC crew performs less than adequate inspection or maintenance activities. This corresponds to a medium relative frequency of the initiating event and end-state.
Low	3	Incidents with or without passengers onboard and/or other road users are involved. Post-incident procedures are followed. This corresponds to a high relative frequency of the initiating event and a low relative frequency of the end-state.
		The vehicle is unreachable or unresponsive to remote commands and fails to autonomously implement DDT fallback actions when required. This corresponds to a medium relative frequency of the initiating event and a low relative frequency of the end-state.
Very Low	4	Incidents with or without passengers onboard and/or other road users are involved. Post-incident procedures are not followed. This corresponds to a low relative frequency of the initiating event and end-state.
		FOC is unaware other road users are involved in the incident and does not contact first responders or does not provide them with correct information. This corresponds to a low relative frequency of the initiating event and end-state.

End-state Risk Classification

The risk level, assessed through the combination of severity, controllability, and relative frequency, is categorized on a scale 1-5 shown in Table B.5.

- Level 1: Operation proceeds as expected or operational failures do not lead to imminent risks.
- Level 2: Low-level risks. Scenarios where the vehicle operation is interrupted but preventive and mitigative actions are available; or when failures of preventive or mitigative actions do not lead to immediate consequences.
- Level 3: Medium-level risk. Scenarios in which the vehicle’s operation is interrupted and mitigative actions are available; or when failures of mitigative actions do not lead to immediate consequences.

- Level 4: High risk. Scenarios where an incident has occurred, or the vehicle’s operation is interrupted. Mitigative actions have failed or have not been performed, leading to immediate consequences.
- Level 5: Very high risk. Scenarios where the vehicle is at risk of collision, involves passengers or other road users, and mitigative actions have failed or have not been performed and lead to immediate consequences.

Table B.5: Resulting risk matrix.

Controllability	Exposure/Severity	No incident*	Traffic disruption	Danger to property	Danger to life
High	Very Low	1	1	1	2
	Low	1	2	2	3
	Medium	1	2	3	3
	High	1	2	3	4
Medium	Very Low	1	2	2	3
	Low	2	3	3	4
	Medium	2	3	4	4
	High	3	4	4	5
Low	Very Low	1	2	3	3
	Low	2	3	4	4
	Medium	2	4	4	5
	High	3	4	5	5
Very Low	Very Low	2	3	3	4
	Low	3	4	4	5
	Medium	3	4	5	5
	High	4	5	5	5

* Severity Level 1: No incidents correspond to scenarios in which operation does not lead to any traffic, property, or injury related consequence, e.g., a passenger trip has successfully been completed. However, organizational errors and failure to follow procedures are also considered at this level as these do not produce any immediate consequences, e.g., the ADS vehicle has been incorrectly cleared for operation after failing a pre-shift inspection test. For more information, please refer to Appendix B.

Table B.6 presents the characterization of each end-state according to the controllability, severity, relative frequency, and resulting risk category.

Table B.6: Risk Scale – end-state categorization.

High-level scenario	Controllability	Severity	Relative Frequency	Risk Category
ES1: Trip completed successfully	High	No Incident	High	1
ES2: Vehicle arrives at MOC for maintenance	High	Traffic disruption	High	2
EF3: Collision Risk	Very Low	Fatality and Injury	Low	5
ES4: Post-incident procedures are initiated.	Medium	Traffic disruption	Low	2
EF5: Vehicle is stranded	Low	Traffic disruption	Low	3

High-level scenario	Controllability	Severity	Relative Frequency	Risk Category
EF6: Vehicle and passenger are stranded	Low	Fatality and Injury	Low	4
EF7: Passenger at risk	Very Low	Fatality and Injury	Low	5
ES8: ADS Vehicle is on-route to destination with passengers	High	No Incident	High	1
ES9: Vehicle scheduled for pre-shift inspection or corrective maintenance	High	No Incident	High	1
ES10: Vehicle scheduled for service maintenance or system updates	High	No Incident	Medium	1
ES11: Vehicle is stationed at MOC	Medium	No Incident	Medium	2
EF12: Vehicle is unreachable	Very Low	Fatality and Injury	Low	5
ES13: Vehicle cleared for operation	High	No Incident	High	1
EF14: Vehicle incorrectly dispatched for operation	Medium	No Incident	Medium	2
ES15: Vehicle is scheduled for external maintenance	High	No Incident	Medium	1
EF16: Vehicle incorrectly cleared for operation	Medium	No Incident	Medium	2
EF17: Vehicle is not scheduled for external maintenance	Medium	No Incident	Medium	2
EF18: Vehicle passes a faulty inspection	Medium	No Incident	Medium	2
EF19: Passenger is stranded, and vehicle is at risk of collision	Very Low	Fatality and Injury	Low	5
ES20: ADS Vehicle is on-route to destination without passengers	High	No Incident	High	1
ES21: Post-incident procedures are completed	High	Fatality and Injury	Low	2
EF22: Vehicle is not recovered	Low	Fatality and Injury	Very Low	3
EF23: Vehicle not recovered; incident not reported to the MOC	Low	Fatality and Injury	Very Low	3
EF24: Passengers and/or other road users at risk	Very Low	Fatality and Injury	Very Low	4
EF25: Passengers, and/or others at risk; vehicle not recovered	Very Low	Fatality and Injury	Very Low	4
EF26: Passengers and/or others, at risk; incident is not reported	Very Low	Fatality and Injury	Very Low	4
EF27: Passengers are stranded	Medium	Property-damage only	Very Low	2
EF28: Passengers are stranded; vehicle is not recovered	Low	Property-damage only	Very Low	3
EF29: Passenger is stranded; vehicle is not recovered; incident not reported	Low	Property-damage only	Very Low	3
EF30: Other road users at risk	Very Low	Fatality and Injury	Very Low	4
EF31: Other road users at risk; vehicle not recovered	Very Low	Fatality and Injury	Very Low	4
EF32: Other road users at risk; incident is not reported	Very Low	Fatality and Injury	Very Low	4
EF33: Passenger is stranded; other road users at risk	Very Low	Fatality and Injury	Very Low	4

High-level scenario	Controllability	Severity	Relative Frequency	Risk Category
EF34: Passenger is stranded; vehicle is not recovered; other road users at risk	Very Low	Fatality and Injury	Very Low	4
EF35: Passenger is stranded; incident is not reported; other road users at risk	Very Low	Fatality and Injury	Very Low	4
EF36: Vehicle arrives at MOC for maintenance; other road users at risk	Very Low	Fatality and Injury	Very Low	4
EF37: Vehicle and others road users at risk	Very Low	Fatality and Injury	Very Low	4
EF38: Vehicle is stranded; others road users at risk	Very Low	Fatality and Injury	Very Low	4
ES39: Post-incident procedures are completed. No other parties are involved.	High	Property-damage only	Low	2
EF40: Vehicle is not recovered. No other parties are involved.	Low	Property-damage only	Very Low	3
EF41: Incident not reported to the MOC. No other parties are involved.	Low	Property-damage only	Very Low	3

Table B.7 provides the end-states categorized by the corresponding risk level. Given the characteristics of the scenarios leading to each particular end-state, two types of risk aggravating factors are introduced:

- a) Latent effect consequences: End-states that do not lead to any immediate consequences yet may increase the risk of future high-risk consequences.
- b) Incident-related consequences: End-states resulting from scenarios in which an incident has occurred (vehicle is in MRC and post-incident procedures have been initiated).

Table B.7: Risk scale - grouping consequences per levels.

Risk Level	Description	High-level scenarios	Comments
Level 1	No imminent risks.	ES1: Trip completed successfully	
		ES8: ADS Vehicle is on-route to destination with passengers	
		ES9: Vehicle scheduled for pre-shift inspection or corrective maintenance	
		ES10: Vehicle scheduled for service maintenance or system updates	
		ES13: Vehicle cleared for operation	
		ES15: Vehicle is scheduled for external maintenance	
		ES20: ADS Vehicle is on-route to destination without passengers	
Level 2	Low-level risks.	ES2: Vehicle arrives at MOC for maintenance	
		ES4: Post-incident procedures are initiated.	
		ES11: Vehicle is stationed at MOC	latent
		EF14: Vehicle incorrectly dispatched for operation	latent
		EF16: Vehicle incorrectly cleared for operation	latent
		EF17: Vehicle is not scheduled for external maintenance	latent
		EF18: Vehicle passes a faulty inspection	latent
		ES21: Post-incident procedures are completed	incident
		EF27: Passengers are stranded	incident
ES39: Post-incident procedures are completed. No other parties are involved.	incident		
Level 3		EF5: Vehicle is stranded	

Risk Level	Description	High-level scenarios	Comments
	Medium-level risk.	EF22: Vehicle is not recovered	incident
		EF23: Vehicle not recovered; incident not reported to the MOC	incident
		EF28: Passengers are stranded; vehicle is not recovered	incident
		EF29: Passenger is stranded; vehicle is not recovered; incident not reported	incident
		EF30: Other road users at risk	incident
		EF36: Vehicle arrives at MOC for maintenance; other road users at risk	incident
		EF40: Vehicle is not recovered. No other parties are involved.	incident
		EF41: Incident not reported to the MOC. No other parties are involved.	incident
Level 4	High risk.	EF6: Vehicle and passenger are stranded	
		EF24: Passengers and/or other road users at risk	incident
		EF25: Passengers, and/or others at risk; vehicle not recovered	incident
		EF26: Passengers and/or others, at risk; incident is not reported	incident
		EF31: Other road users at risk; vehicle not recovered	incident
		EF32: Other road users at risk; incident is not reported	incident
		EF33: Passenger is stranded; other road users at risk	incident
		EF34: Passenger is stranded; vehicle is not recovered; other road users at risk	incident
		EF35: Passenger is stranded; incident is not reported; other road users at risk	incident
		EF37: Vehicle and others road users at risk	incident
Level 5	Very high risk.	EF3: Collision Risk	
		EF7: Passenger at risk	
		EF12: Vehicle is unreachable	
		EF19: Passenger is stranded, and vehicle is at risk of collision	

*Latent: End-states that do not lead to any immediate consequences yet may increase the risk of future high-risk consequences.

*Incident: End-states resulting from scenarios in which an incident has occurred (vehicle is in MRC and post-incident procedures have been initiated).

Comments

The qualitative risk scale approach is proposed to identify the most critical safety hazards for developing risk mitigation measures. However, a quantitative approach should be developed to determine if certain hazard scenarios present a risk above an acceptable threshold.

Once data is available to quantify both the initiating event frequency and the probability of the ESD pivotal events, a quantitative approach to estimate risk must consider three elements: the scenario (s), the likelihood/frequency of the scenario (f), and the consequences it leads to (C) (Kaplan & Garrick, 1981). The risk may be expressed as:

$$R = s \times f \times C$$

The ESDs and FTs developed in this study are quantitative methods. They can be leveraged for a quantitative risk assessment by estimating the frequency of the initiating events, the probability of the pivotal events, and the impact/consequence level of the end states. To better characterize the consequences, the analysis should specify the conditions in which the scenarios occur (e.g.,

weather conditions, vehicles' speed, and surrounding traffic information). The likelihood estimation should use probabilistic or frequency data, including hardware, software, and human failures. In cases of insufficient data, the analyses can adopt well-established methods such as expert judgment and Bayesian models.

References

- De Gelder, E., Elrofai, H., Saberi, A. K., Paardekooper, J.-P. P., Op den Camp, O., & de Schutter, B. (2021). Risk Quantification for Automated Driving Systems in Real-World Driving Scenarios. *IEEE Access*, *9*, 168953–168970. <https://doi.org/10.1109/ACCESS.2021.3136585>
- International Organization for Standardization. (2018). *ISO 26262:2018, Road vehicles — Functional safety*.
- Kaplan, S., & Garrick, B. J. (1981). On The Quantitative Definition of Risk. *Risk Analysis*, *1*(1), 11–27. <https://doi.org/10.1111/J.1539-6924.1981.TB01350.X>
- Khastgir, S., Birrell, S., Dhadyalla, G., Sivencrona, H., & Jennings, P. (2017). Towards increased reliability by objectification of Hazard Analysis and Risk Assessment (HARA) of automated automotive systems. *Safety Science*, *99*, 166–177. <https://doi.org/10.1016/j.ssci.2017.03.024>
- National Research Council (US) Committee on Risk Assessment of Hazardous Air Pollutants. (1994). *Science and Judgment in Risk Assessment*. Washington (DC): National Academies Press (US); <https://www.ncbi.nlm.nih.gov/books/NBK208270/>