Approach to Derive Verifiable Performance

1. Purpose of this document

This document provides an overview on the approach to be used to derive verifiable performance criteria for the certification or, as relevant, for self-certification of ADS, based on the manufacturer/ ADS developer’s description of the Operational Design Domain (ODD) of the ADS.

Operational design domain (ODD) refers to:

> *Operating conditions under which a given driving automation system or feature thereof is specifically designed to function, including, but not limited to, environmental, geographical, and time-of-day restrictions, and/or the requisite presence or absence of certain traffic or roadway characteristics. (SAE J3016)*

Given a specific ODD, it is crucial for the ADS to ensure that:

- it can operate safely within its ODD under conditions reasonably expected in the ODD
- it will be used only within its ODD
- it can monitor whether it is inside/outside its ODD and respond appropriately.

The conditions constituting the ODD in which the ADS was designed to operate will help determine which ADS competencies are required. For example, if an ADS has an ODD which comprises of roads with non-signalised junctions, one of the required behaviour competencies for the ADS in that ODD could potentially be *“unprotected left or right turn”*. However, the same behaviour competency may not be required if the ODD of an ADS is limited to motorways or highways with signalized junctions.

2. Introduction and approach

The concept of “behavioural competencies” is useful in determining the safety of the performance of the Dynamic Driving Task (DDT) by an Automated Driving System (ADS).

The Automated Vehicle Safety Consortium (“AVSC”)) has provided these definitions[1]:

- *Behaviour: Specific goal-oriented actions directed by an engaged ADS in the process of completing the DDT or DDT fallback within the ODD (if applicable) at a variety of timescales.*
- *Behavioural Competency: Expected and measurable capability of an ADS feature operating a vehicle within its ODD.*

Behavioural competencies can be described with different abstraction levels, similarly to functional, logical, and concrete scenarios. Refinement of the competencies from a functional to a more concrete level is possible by following the approach proposed in these guidelines.

[1] AVSC Best Practice for Evaluating Behavioral Competencies for Automated Driving System Dedicated Vehicles (ADS-DVs).

Such competencies track the three broad categories of driving situations that may be encountered in performance of the DDT: nominal, critical, and failure.

Nominal driving situations are those in which behaviour of other road users and the operating conditions of the given ODD are reasonably foreseeable (e.g. other traffic participants operating in line with traffic regulations) and no failures occur that are relevant to the ADS's performance of the DDT.

Critical driving situations are those in which the behaviour of one or more road users (e.g., violating traffic regulations, …) and/or a sudden and not reasonably foreseeable change of the operating conditions of the given ODD (e.g. sudden storm, damaged road infrastructure, …) creates a situation that may result in an immediate risk of collision. In this case, as it is recognised that in some cases the ADS may not be able to avoid a collision, the ADS performance are compared with safety model performance to set the threshold between where avoidance is required and where it is not feasible, but mitigation may be possible.

Failure situations involve those in which the ADS or another vehicle system experiences a fault or failure that removes or reduces the ADS's ability to perform the DDT, such as sensor or computer failure or a failed propulsion system.

Concrete performance requirements depend on the specific situations the ADS encounters, on a reference behaviour that is deemed appropriate for a human driver or a technical system, and on assumptions (e.g. friction values, reaction times) about the behaviour of the vehicle and other road users. Since it is virtually impossible to write a regulation that sets out verifiable criteria for every combination of these variables, this document aims at providing a set of different reference behaviours or safety models together with an overview of the characteristics and required assumptions that can be useful in deriving verifiable performance criteria in some situations. The aim is then to assist those who develop concrete regulations with the selection and parameterization of functions or selection of scalars as pass/fail criteria.

For this, the following is needed:

- An overview of reasonable expectations (which might occur in different ODDs),
- An overview of reference behaviours / safety models that define the boundary between avoidable accidents and mitigation (note that these reference behaviours will not be used for anything else than providing this boundary as a performance criterion).
- A matrix combining suggested reference behaviours / safety models with driving situations.

3. Behavioural Competences Identification

The approach suggests a series of analytical frameworks that could help to derive measurable criteria appropriate for the specific application. These frameworks are divided into:

- ODD Analysis
- Driving Situation Analysis
- OEDR Analysis

3.1.1. ODD analysis

This analysis represents the first step with the aim to identify the characteristics of the ODD. An ODD may consist of stationary physical elements (e.g., physical infrastructure), environmental conditions, dynamic elements (e.g., reasonably expected traffic level and composition, vulnerable road users) and operational constraints to the specific ADS application. Various sources provide useful guidance for precisely determining the elements of a particular ODD and their format definition.[2,3, 4, 5]

As part of this activity, the level of detail of the ODD definition using the ODD attributes will also need to be established.

3.1.2. Driving Situation Analysis

In the driving situation analysis, the behaviours of other road users that are reasonably expected and presence of roadway characteristics in the ODD are explored in more detail by mapping actors with appropriate properties and defining interactions between the objects.

An example of this analysis is given in **[Table 1]**, where static and dynamic behaviours of other objects (including other road users) that the ADS is reasonably expected to encounter within the ODD are described. In the case of vehicles, this includes behaviours such as "acceleration", "deceleration", "cut-in"; for pedestrians, examples of dynamic behaviours include "crossing road", "walking on sidewalk", etc. Some of these behaviours may involve nominal situations (e.g., lead vehicle deceleration at a rate reasonably expected in light of traffic and other circumstances within the bounds of physical limitations[6]) while others may involve critical situations (e.g., sudden cut-ins or unpredictable pedestrian or cyclist behaviour, including behaviours that may violate local traffic laws such as crossing a road outside a designated cross walk).

The behaviour of other road users and the condition of physical objects within the ODD may fall at any point along a continuum of likelihood. For example, deceleration by other vehicles may range from what is expected and reasonable in the traffic circumstances, to unreasonable but somewhat likely rapid deceleration, to extremely unlikely (e.g., a sudden cut-in combined with full braking on a clear high-speed road). The analysis of the ODD and reasonably expected driving situations within the ODD should make distinctions that include an estimate of the likelihood of situations to ensure that the ADS's performance is evaluated based on response to reasonably likely occurrences involving nominal, critical and failure situations but not on the expectation that the ADS will avoid or mitigate the most extremely unlikely occurrences.

---

[2]; E.g., *AVSC Best Practice for Describing an Operational Design Domain: Conceptual Framework and Lexicon*; and *A Framework for Automated Driving System Testable Cases and Scenarios* (NHTSA).

[3] *E.g. BSI PAS 1883:2020 Operational Design Domain (ODD) taxonomy for an automated driving system (ADS) - Specification*

[4] ASAM OpenODD

[5] Road Vehicles — Test scenarios for automated driving systems — Taxonomy for operational design domain

[6] Deceleration of road vehicles is limited by tire-road friction and separating fluid, if any (e.g. wet, ice). It is only in some rare circumstances limited by brake capacity, specifically if the brake torque fades due to hot brakes.

Objects	Events/Interactions
Vehicles (e.g., cars, light trucks, heavy trucks, buses, motorcycles)	**Lead vehicle decelerating (frontal), lead vehicle stopped (frontal), lead vehicle accelerating (frontal), changing lanes (frontal/side), cutting in (adjacent), turning (frontal), encroaching opposing vehicle (frontal/side), encroaching adjacent vehicle (frontal/side), entering roadway (frontal/side), cutting out (frontal)**
Pedestrians	**Crossing road – inside crosswalk (frontal), crossing road – outside crosswalk (frontal),** walking on sidewalk/shoulder
Pedalcyclists	**Riding in lane (frontal)**, riding in adjacent lane (frontal/side), riding in dedicated lane (frontal/side), riding on sidewalk/shoulder, **crossing road – inside crosswalk (frontal/side), crossing road – outside crosswalk (frontal/side)**

Objects	Events/Interactions
Animals[5]	Static in lane (frontal), moving into/out of lane (frontal/side), static/moving in adjacent lane (frontal), static/moving on shoulder
Debris[6]	**Static in lane (frontal)**
Other dynamic objects (e.g., shopping carts)	**Static in lane (frontal/side), moving into/out of lane (frontal/side)**

Objects	Events/Interactions
Traffic signs[7]	Stop, yield, speed limit, crosswalk, railroad crossing, school zone
Traffic signals[7]	Intersection, railroad crossing, school zone
Vehicle signals	Turn signals

Table 1 – Static / Dynamic elements and their properties

3.1.3. Object and Events Detection and Response (OEDR) Analysis: Behaviour competency identification

Once the objects and their reasonably expected behaviours have been identified, it is possible to map the appropriate ADS response, which can be expressed as a behavioural competency. The detailed response is derived from more general and applicable functional requirements, as developed by FRAV. The acceptable ADS response will vary depending on whether the driving situation involves nominal, critical, or failure characteristics.

The outcome of the analysis is a set of behaviour competencies that can be applied to the events characterizing the ODD. **[Table 2]** provides a qualitative example of a matching event – response.

Event	ORUs behaviour (Property)	ODD Static Element	Infrastructure	Weather Conditions	Time of Day	Response	Beahvioural Competence
Car	Riding forward - Decelerating	Straight Road (w/ Lane markings)	Explicit Speed Sign	No precipitation	Day overcast	Follow vehicle, decelerate, stop	The ADS travels with a speed compliant with the road legal limit in its lane. The ADS decelerates to maintain a safe distance following a leading car, travelling in the same lane as ego, decelerating at [] m/ss.
Motorcyclist	Riding forward - Cruising	Curved Road	Implicit Speed Sign	No precipitation	Twilight (dawn., dusk)	Swerving (L/R)	The ADS travels with a speed compliant with the road legal limit in its lane. The ADS keeps safe headway distance from the motorcyclist swerving between lanes.

Table 2 – Example of elementary behaviour competences for given events

The combination of objects, events, and their potential interaction, as a function of the ODD, constitute the set of nominal or critical situations pertinent to the ADS under analysis.

3.2. Nominal Situation Competencies

In these situations, ADS competencies can often be derived by applying traffic laws of the country where the ADS is intended to operate, as well as by applying general safe driving principles for situations not addressed adequately by current traffic laws for human drivers. Examples of such competencies may include adherence to legal requirements to maintain a safe distance from vehicles ahead, provide pedestrians the right of way, obey traffic signs and signals, etc. Of course, some nominal competencies (e.g., safe merging, safely proceeding around road hazards) may not be explicitly articulated or mandated by traffic laws. In some instances, traffic laws may provide wide discretion for the driver to determine the safest response to a particular situation (for example, how to respond to adverse weather conditions). As such not all traffic laws are stated with sufficient specificity to provide a clear basis for defining a competency.

Therefore, an approach to codify rules of the road to provide additional specificity was developed in Paragraph **[6]**. Additionally, application of models involving safe driving behaviour may be needed in addition to reference to codified rules of the road in developing behavioural competencies for nominal driving situations.

3.3. Critical Situation Competencies

The development of these competencies requires analysis of (1) what constitutes such unreasonable behaviour by ORUs and/or a sudden change of the operating conditions that are not reasonably foreseeable and (2) what constitutes an appropriate ADS response to avoid or mitigate the imminent crash. Additionally, it is also important to identify the occurrence of unplanned emergent behaviour in critical situations.

Analysis of the first type may be based on a variety of methodologies, including e.g. IEEE 2846-2022 (which offers guidance on what behaviours by other road users are reasonably foreseeable) and other models of reasonable driving behaviour. Analysis of the second factor may be based on various models of acceptable human driving behaviour in crash imminent situations.

Hazard identification methods (e.g. STPA as mentioned in SAE J3187) which analyse the system design for functional and operational insufficiencies can help identify the occurrence of emergent behaviour which may lead to critical situations.

[PLACEHOLDER [Include a table with example of ORU unreasonable behaviour – running a red traffic light]

Development of behavioural competencies for critical driving situations faces several challenges. No general consensus exists on the appropriate models for the behaviour of ORUs or appropriate responses by the ADS to unreasonable ORU behaviours that make a crash imminent.

3.4. Failure Situation Competencies

FRAV requirements include management of various failure modes. As noted above, failure situations involve those in which the ADS or another vehicle system experiences a fault or failure that removes or reduces the ADS's ability to perform the DDT, such as sensor or computer failure or a failed propulsion system.

In developing the behavioural competencies appropriate for failure situations, the objective is to describe the ability of the ADS to detect and respond safely to specific types of faults and failures. Depending upon the nature and extent of the fault or failure, the responses can include identifying a minor fault for immediate repair after trip completion, responding to a significant fault with restrictions (such as limp-home mode) for the remainder of the trip, or responding to major failures by achieving a minimal risk condition. Communication of the fault or failure condition to vehicle users may also be a desirable ADS behavioural competency.

An example of Failure Competencies is reported in [Table 4].

PLACEHOLDER [Insert table with example of Failure Competences]

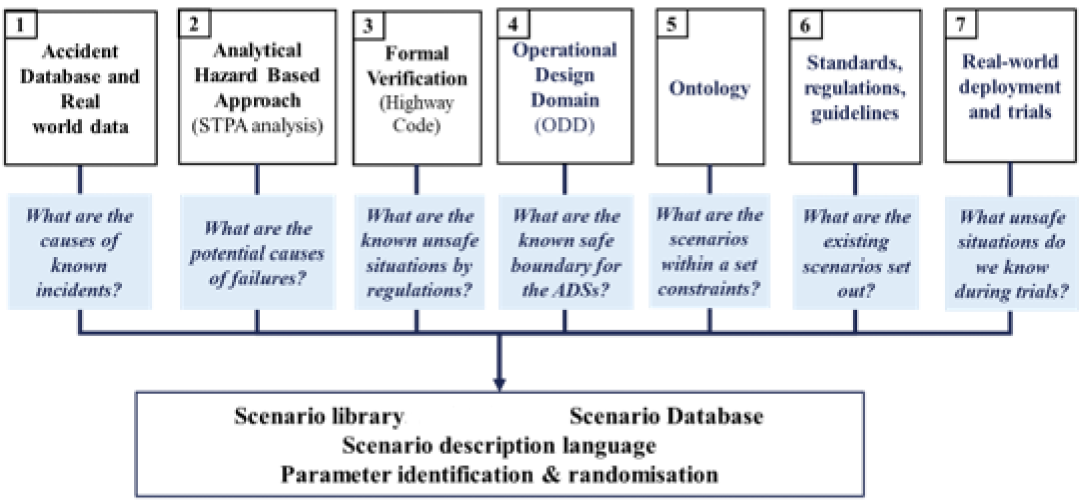
4. Scenario identification

To ensure that the behavioural competences identified in the previous paragraphs are ready to be assessed through the application of simulations or physical testing, ODD-relevant scenarios must be developed. Scenario creation involves use of assumptions concerning the actions of road users that incorporate realistic parameters.

This approach suggests two complementary methodologies to derive reasonably expectable situations which might occur for a given ODD:

- Knowledge-based (e.g. goal-based)
- Data-based

A knowledge-driven scenario generation approach utilizes domain specific (or expert) knowledge to identify hazardous events systematically and create scenarios. A data driven approach utilizes the available data (e.g. accident databases, insurance records) to identify and classify occurring scenarios. [**Figure x]** illustrates various data-based and knowledge-based scenario generation methods.



PLACEHOLDER [Figure 1]

Accident datasets and field data can be analysed to identify accident hotspots and scenario parameters which contribute to causation of accidents carrying high levels of severity.

Knowledge based methods, or other formal techniques can be used to analyse the characteristics of the ADS architecture and identify system failures and hazardous situations [see SAE J3187]. The analysis is then converted into a set of abstract/logical scenarios together with their corresponding pass/fail criteria.

Other knowledge-based methods include the formal analysis approach with the highway code rules for scenario generation. Each of the highway code rules describes a hypothetical driving scenario with the corresponding behaviour and ODD elements. The ODD is a specification set out by the manufacturer of an ADS and it defines the operating conditions within which the ADS can operate safely. Formal models are generated via a model template to create the mathematical representations of those scenarios, collecting the combinations of ODD and behaviour parameters. The analysis reports the manoeuvre parameters that are close of violating the pass criteria and produce scenarios that represent these set of violations. Other knowledge-based methods use formal representation of the ODD and behaviour competencies of the ADS for scenario generation.

Furthermore, the existing scenarios already defined in the standards, regulations or guidelines (Option 6 - KB) can also be utilized for the testing of ADSs, for example the scenarios set out in ISO22737 and NCAP. ISO22737 has been developed for low-speed automated driving systems (LSAD) and the NCAP provides a set of testing scenarios for the safety assurance of vehicles. Option 7 (DB) includes the scenarios that occur during real world trials and deployments. Such scenarios might have not been considered pre-deployment but are key learnings.

4.1. Assumptions: Logical to concrete scenarios

Assumptions concerning the actions of other road users may need to account for cultural differences in driving styles in different geolocations, making it impracticable to harmonise these assumptions across different domains. Therefore, evidence should be provided to support the assumptions made. Existing standards e.g. IEEE 2846-2022 provide a set of assumptions to be considered by ADS safety-related models for an initial set of driving situations. Additionally, several other tools including data collection campaigns performed during the development phase, real-world accident analysis and realistic driving behaviour evaluations, constraint randomisation, Bayesian optimisation besides others can be used to inform values for such assumptions.

5. Application of Rules of Road as Pass criteria and requirements

An approach to define an acceptance criterion related to nominal driving situations is to evaluate the ADS performance against the rules of the road. Furthermore, **[FRAV ADS Safety Topics]** mentions that *"The ADS should comply with traffic rules" (in all driving conditions)*. It is challenging to test against this requirement in the absence of codified rules of the road.

A codified rules of the road also enable the verification of the requirement "the ADS should comply with traffic rules".

Development of functional requirements and test scenarios involves careful description of the behaviours of the road users and the physical circumstances (i.e. the scenery) in which the behaviours occur. If one compares the scope of ODD and the content of current *rules of the road for human drivers* (e.g., UK's Highway Code or the Vienna Convention Chapter 2), a large overlap of scenery aspects and environmental conditions aspects can be observed. It is plausible to follow an ODD based approach and an ODD taxonomy, to model the environmental and scenery aspects of the *rules of the road*. In addition, road user behaviours need to be described. Behaviour can be further divided into ego (vehicle under test) behaviours and actor behaviours. The relevant behavioural competency describes the expected and measurable behaviour of the ego vehicle in the scenario.

Any rule of the road can be classified into two categories:

- Doing some behaviour somewhere
- Not doing some behaviour somewhere

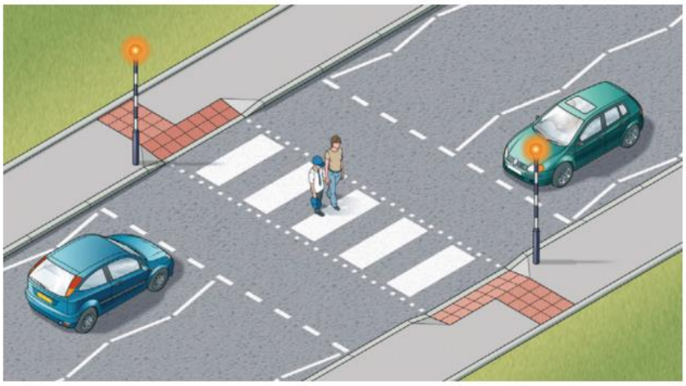
While doing or not doing some behaviour can be defined as part of ADS's behaviour competencies, *"somewhere"* could be considered as part of the "operating condition" or part of the ODD definition. The approach is summarised in **[Figure 2]**.

Current Rules (for human drivers) = *f(Operating condition, Behaviour competency, Assumptions)*

Codified Rule of the Road = *f(operating condition, behaviour competency, driving characteristics)*

Figure 2 – Example of codified Rules of the Road

Example: The UK's Highway Code (HC)[1], Rule 195 states (Zebra crossing): *"As you approach a zebra crossing: look out for pedestrians waiting to cross and be ready to slow down or stop to let them cross; you MUST give way when a pedestrian has moved onto a crossing"*.



Rule 19: Zebra crossings have flashing beacons

Figure 1: Example of zebra crossing from UK's Highway Code:
Source: https://www.gov.uk/guidance/the-highway-code/rules-for-pedestrians-1-to-35#rule19

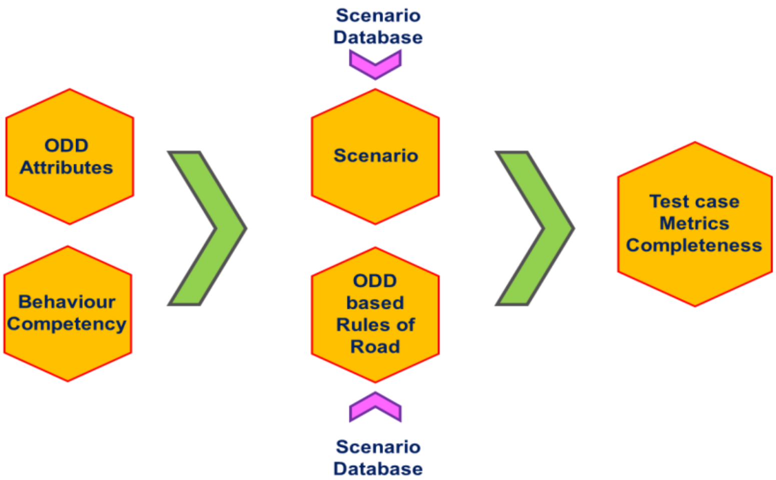
Figure 3 –

From this rule, one can extract the "operating condition or ODD" variables, as well as the behaviour competencies. While *"zebra crossing"* and *"pedestrian"* define the *operating condition*; and *"slow down or stop"* defines the *behaviour competency*, the rule doesn't mention how long should the vehicle be stopped. This is an implicit assumption made by the driver. However, for ADSs, such assumptions will need to be specified. We foresee such assumptions being specific to countries, regions, and cultures.

The proposed process helps makes the "implicit assumptions" in the rules for human drivers into explicit rules. In other words, the proposed process enables to turn "undefined" attributes in the rules of the road (for human drivers) to "defined" attributes in the codified rules of the road.

5.1. Using rules of the road as pass criteria

[Figure 10] illustrates the use of codified rules of the road as a pass criterion for scenario-based testing activities. Every test scenario definition will have ODD and behaviour



competency attributes defined. Every rule of the road will also have ODD and behaviour competency attributes as part of its definition.I Therefore, it is possible to map every scenario to a corresponding rule(s) of the road using ODD and behaviour tags or labels in a scenario **catalogue (VMAD SG1)**.

Figure 4 - Rules of the Road as pass / fail criteria

This approach would allow the test engineer to map each scenario to a corresponding rule (or set of rules). These rules can then serve as the pass criteria during the scenario-based testing approach. This approach can thus enable engineers and authorities to show/assess compliance to traffic rules by making the rules of the road verifiable.

6. Application of Safety Models as pass criteria and requirements for critical situations

The aim of this section is to further specify the methodology to derive the threshold to separate between collisions that have to be avoided and those where only mitigation is required. The aim is NOT to prescribe a specific behaviour of the ADS in any given critical situation; this is only about the outcome. In a mathematical & logical sense, for any given situation, there will be a function depending on variables that partly describe scenario, delivering a Boolean “true” or “false” for whether the collision needs to be avoided, and vice versa for whether mitigation is acceptable:

$$Avoidance[0;1] = f_{safetymodel}(scenario\ variable\ 1, scenario\ variable\ 2, \dots),$$

$$Mitigation[0;1] = 1 - f_{safetymodel}(scenario\ variable\ 1, scenario\ variable\ 2, \dots).$$

It is envisioned that concrete ADS regulations, built using the guidelines as specified here, may contain either a concrete scalar threshold (example: avoid accidents for a driving speed below 42 km/h, see UN R152), or formulate a concrete $f_{safetymodel}$ where all parameters are specified (simplified example from UN R157: when cut-ins of other vehicles occur before a specific TTC, the collision needs to be avoided, the resulting function would be:

$$f_{safetymodel} = [1 \text{ for } TTC_{LaneIntrusion} > (v_{rel}/(2 \cdot 6 m/s^2) + 0.35s); 0 \text{ otherwise}].$$

Choosing appropriate model(s) depends, amongst others, on:

- the balance between risk to the ADS itself vs. risk towards the accident partner (e.g. for pedestrians, it would very likely be acceptable to have a slightly increased risk for the typically belted ADS occupants when the risk for the pedestrian would be significantly reduced, e.g. by earlier or stronger brake intervention),
- the assumed anticipation level *(e.g. is it feasible to anticipate actions of other traffic parameters and start countermeasures earlier, or will it be a simple reaction to faults),*

- the environmental condition parameters. *(e.g. what level of friction is typically available where the ADS are travelling)*,
- the balance between efficiency and acceptable remaining risk *(e.g. passing a pedestrian with no acceptable risk would be possible only with very low speeds, which would render the current sidewalk close to streets infrastructure useless).*

These factors will be different for different situations, or in other words: there would be different $f_{\text{safetymodel},i}$ for different critical situations anticipated to occur in the operational domain of the concrete ADS regulation in pseudo-code:

Example Regulation XXX =

{Situation / parameter range 1, avoidance = $f_{\text{safetymodel},1}$(parameters a,b,c);

address pedestrian accidents in urban areas

Situation / parameter range 2, avoidance = $f_{\text{safetymodel},2}$(parameters d,e,f);

address car-car accidents with cut-in on motorways…}.

The following paragraphs summarize the safety performance models that can be used to assess the behavioural competency of an ADS based on the scenario.

6.1. Review of Safety Models

Based on an initial literature review, safety performance models for the ADS behaviour include:

- Careful & Competent human driver (CC, Ref.: UNECE Reg. 157 Annex 3 §3.3).
  - *This model assumes the characteristics of a typical driver with regard to threat detection, reaction time delay, brake application to identify what a human intervention to e.g. a cut-in maneuver would be.*
- Fuzzy Surrogate Safety Model (FSSM, Ref.. UNECE Reg. 157 Annex 3 §3.4).
- Kinematic Lane Change (K-LC).
- Last-Point-to-Steer (LPS, Ref.: AEBS HDV 03.03).
  - *This model assumes an emergency braking intervention in longitudinal traffic is justified as soon as a collision cannot be avoided by steering (=the last past to steer has passed). Typically the last point to steer for speeds > ca. 30 km/h is later than the last point to brake.*
- Responsibility Sensitive Safety (RSS, Ref.: Shalev-Shwartz et al., 2017).
- Safety Force Field (SFF, Ref.: D.Nister et al., 2019).
- Safety Zone (SZ, Ref.: AEBS HDV 03.03).

- *This model assumes an emergency braking intervention with cross-traffic is justified as soon as the collision partner is no longer able to not enter the path of the ego vehicle AND a collision will occur. In particular for pedestrians, a safety zone of approx. 30 cm is a typical value and this is consistent with the brake distance required for a 5 km/h pedestrian.*

7. Performance Evaluation and Targets

As previously highlighted, nominal situations are considered reasonably foreseeable and preventable for a given ODD and therefore it is expected that the ADS would be capable of handling them without any resulting collision.

On the other hand, failure situations are performed to assess the ADS ability to recognise faults / failures in the system, and respond in compliance with the principles highlighted by FRAV.

For the purpose of defining performance criteria in critical situations, those where others are at fault & behaving unforeseeable & the collision might potentially not be prevented have to be analysed further. In these situations, it is proposed that safety models are used to explore and compare the ADS performance with mathematical formulations to derive what is deemed as preventable or where mitigation strategy is needed.