Submitted by the co-chairs of the
VMAD Informal Working Group

Informal document **GRVA-19-12**
19th GRVA, 25 June 2024
(For review during the
Troy meeting 20-24 May 2024)
Provisional agenda item 4(d)

Proposal for amendments to the Guidelines and recommendations for ADS safety requirements, assessments and test methods to inform regulatory development

The text reproduced below is based on the document GRVA-18-50. The modifications to that text are indicated in **blue bold** for new characters and ~~blue strikethrough~~ for deleted characters.

I. Proposal

Annex 6., amend to read:

"Annex 6: Track and real-world testing

Track testing

Track testing occurs on a closed-access testing ground that uses real obstacles and obstacle surrogates (e.g., vehicle crash targets, etc.) to assess the safety requirements of an ADS (e.g., human factors, safety system). This testing approach allows for the ADS of physical vehicles to be validated through realistic scenarios by evaluating either sub-systems or the fully assembled system. The external inputs and conditions can be controlled or measured during a test.

Track testing is suitable for assessing the ADS capabilities in nominal scenarios, critical scenarios, and failure scenarios. It can also be used to verify the performance of the vehicles regarding human factors or fall-back in these scenarios. However, operating on test tracks can be resource intensive.

It is recommended that:

(a) track testing be used to assess the performance of ADS in a number of selected important nominal, critical, and failure scenarios, notably given that, unlike real-world testing, track testing can accelerate exposure to known rare events or safety critical scenarios, and in a more controlled and safer environment.

(b) track testing is conducted on a testing ground that is part of, or suitably represents, the ODD of the ADS. This excludes track tests where the objective is to assess compliance with non-ODD or extra-ODD related requirements, e.g. tests verifying that the ADS safely responds to crossing ODD boundaries, where applicable.

(c) a test on public roads that are closed to other road users shall be considered a track test.

(d) real-world variation is included in the test parameters instead of limiting the test parameters to standardised parameters, standardised test objects and standardised test environments. The test parameters should therefore go beyond available standards but should remain within the ODD of the ADS. It is recommended to develop a harmonized method for selecting parameters.

(e) with regard to (d), the test equipment, the test set-up, and the test environment, as well as alterations made to those, are recorded at such a detail that ensures replication of the specific test.

(f) the selection of scenarios to be conducted on a test track is appropriate to the ODD, where possible. Track test environments allow for controllability and assurance that specific parameters that can vary in the ODD can be delivered during physical testing.

(g) the behaviour of the ADS towards other road users is verified on a test track using several scenarios.

(h) with regards to human factors, the human machine interaction is tested with the ADS user under different scenarios to ensure safe use of the ADS.

(i) for track testing a protocol is developed containing minimum requirements that standardise how for the test relevant data are to be collected and analysed (e.g., how the data is recorded, how measurements are derived from the recorded data, and how the measurements are analysed).

(j) to develop the track tests in line with the approach set out in Appendix 1 to this Annex.

It is acknowledged that pass-fail criteria depend on the specific scenario tested, and that the selection of scenarios depends on the ODD of the ADS under test. Moreover, it is acknowledged that a proportion of the required pass-fail criteria are not yet available, and they, or some methods to derive such pass-fail criteria, still need to be developed or parts of them could remain subjective.

As performing assessments in crowded areas could be challenging on test tracks, it is recommended such assessments to be performed in real world tests instead. Such assessments should not cover safety critical scenarios.

Information generated during the track test can be used as additional data to validate the virtual tests by comparing an ADS' performance between a virtual test and a test track on the same scenario. For instance, track testing can be used as an additional tool/method to validate the quality/reliability of the virtual toolchain.

Real-world testing

Real-world testing uses public roads to test the capabilities and compliance with safety requirements (e.g., human factors, safety system) of a vehicle with an automated driving system (ADS) in real-world traffic. It therefore provides an opportunity to validate the safety of the ADS within its true operating environment.

It is acknowledged that also for real-world tests pass-fail criteria depend on the specific scenarios tested and encountered, and that the pre-selection of scenarios depends on the ODD of the ADS under test. Moreover, it is acknowledged that a proportion of the required pass-fail criteria are not yet available, and they, or some methods to derive such pass-fail criteria, still need to be developed or parts of them could remain subjective.

It is recommended that real world testing:

(a) assess ADS in nominal scenarios. It is acknowledged that critical and/or failure scenarios may occur during real-world testing, but they generally should not be tested on purpose. In case such scenario would occur, it shall not be excluded from the assessment;

(b) is done safely. It is therefore recommended, if applicable to the ADS use case, that the test supervisor has the possibility to end the real world test at any point. In addition, it is also recommended that any inappropriate behaviour observed and/or the reason for the forced end is investigated in detail later;

(c) is only conducted if a minimum level of safety of the other road users on public roads and of in-vehicle users of the ADS can be ensured by considering the validation methods of simulation, audit, and track testing as well as the manufacturer's prior real-world testing of the ADS;

(d) is always conducted with other road users. Tests on public roads that are closed to other traffic should be considered as track tests;

(e) be considered for assessing aspects of the ADS performance related to its capability to drive in real traffic conditions, such as:

   i. behavioural competencies;

   ii. interaction with other road users;

   iii. safe and anticipatory behaviour;

   iv. smooth driving;

   v. capability to deal with dense traffic;

   vi. maintaining flow of traffic; and

   vii. being considerate and courteous to other vehicles;

(f) be considered for assessing aspects of the ADS performance at some ODD boundaries (nominal and complex scenarios), i.e. is the system triggering transition demands to the driver when it is supposed to (e.g. end of the ODD, weather conditions). The same testing could be used to confirm the performances related to human factors under these conditions;

(g) be considered for detecting issues that may not be well captured by track tests and simulation, such as perception quality limitation (e.g. due to light conditions, rain, etc.);

(h) be considered for assessing aspects relating to human factors, such as user-initiated deactivation, system-initiated deactivation (not leading to a minimum risk condition), audibility of messages in real world conditions, if applicable to the ADS.

It is furthermore recommended that:

(a) the environment and conditions of the selected test routes reflect the applicable ODD's environment and conditions. In addition, the selected test routes should ensure that the ADS under test is expected to experience complex scenarios;

(b) real world testing is developed in line with the approach set out in Appendix 1 to this Annex.

(c) for real world testing a protocol is developed containing minimum requirements that standardise how for the test relevant data are to be collected and analysed (e.g., how the data is recorded, how measurements are derived from the recorded data, and how the measurements are analysed).

While the ADS is designed to perform the DDT only within the conditions represented by its ODD, it is recommended that real world testing assess the ADS both within its ODD and

outside its ODD (e.g. to determine the ADS's appropriate recognition and response when not in its ODD) on public roads.

Although it may not be possible to encounter all traffic scenarios during a real-world test, the likelihood of covering specific complex scenarios could be increased by selecting a specific type of ODD (e.g., highway) and examining when and where specific elements (e.g., high- or low-density traffic) typically occur.

Specific infractions identified during real-world testing may be reviewed and/or assessed by evaluating the data gathered during that test and any data gathered during additional virtual, track and real-world testing.

Data generated during real-world testing may be used as additional data to validate whether portions of a virtual and/or track-testing environment were modelled properly by comparing an ADS' performance within a simulation and/or track test with its performance in a real-world environment when executing the same test scenario.

It can also be used to support the development of new traffic scenarios for track and virtual testing, allowing for the identification of edge cases and other unanticipated hazardous situations that could challenge the ADS.

The information gathered from real world testing may also support improvements in the hazard and risk analysis and to the design of ADS.

Appendix 1

Introduction

An overview of best practices, procedures, technical resources and tools related to track testing and real world testing showed that numerous test procedures and standards for track testing have been developed and used to assess the safety of vehicles with automated driving systems (e.g., ALKS) and particularly with advanced driver assistance systems, which can serve as input to the to-be-developed track testing methodology. The overview furthermore showed that no test procedure to assess the safety of vehicles with automated driving systems on public roads has been developed yet, with most of the available documentation concerning guidance or specifications on testing (i.e., trails) such vehicles by OEMs during the developmental stages of their systems, or the testing of human drivers.

This ~~annex~~ **appendix** proposes test matrices to support track and real-world testing of ADS and ADS vehicles. This approach recommends the use of one general matrix for physical testing complemented by test matrices designed respectively for track testing and real world testing.

The general matrix for physical testing provides an overview of how the ADS safety requirements could be assessed using track testing, real world testing, or both. The test matrices for track testing and real world testing would differ in design in order to take into account the different settings in which the tests are conducted and to ensure that the strengths of each testing method can be utilized.

The test matrices set out in this annex are illustrative and include indicative rather than definitive criteria.

It is important to note that the ADS is designed to perform the DDT only within the conditions represented by its ODD. Therefore, track testing should be conducted on a testing ground that is part of, or suitably represents, the ODD of the ADS. Real world testing meanwhile may assess the ADS both within its ODD and outside its ODD (e.g. to

determine the ADS's appropriate recognition and response when not in its ODD) on public roads.

General matrix for physical testing

The general matrix would provide a clear overview of the type or types of physical testing to be used for assessing compliance with the applicable safety requirements. ~~The general matrix overviews the type(s) of physical tests suitable for assessing compliance with the ADS safety requirements.~~ The following table illustrates the concept for listing requirements alongside the indication of whether track and/or real-world testing might be suitable for assessment of compliance. The listed requirements are indicative and would be replaced by verifiable criteria defined for the ADS under assessment (see Annex 3 for an approach to defining these criteria based on the high-level ADS safety requirements).

Table 1. Example of the General Matrix for Physical Testing

ADS Safety Requirement	*Track*	*Real World*
1. The ADS should perform the entire Dynamic Driving Task.	Yes	Yes
2. The ADS should control the longitudinal and lateral motion of the vehicle.	Yes	Yes
(…)		
7. The ADS should adapt its behaviour in line with safety risks.	Yes	If encountered
8. The ADS should adapt its behaviour to the surrounding traffic conditions.		Yes
(…)		
30. The ADS should safely manage short-duration ODD exits.	Yes	Yes
31. Pursuant to a collision, the ADS should stop the vehicle and deactivate.	Yes	If encountered
(…)		

~~One very important consideration in applying this matrix is that an ADS (except one at SAE Level 5) is designed to perform the DDT only within its ODD. Except for momentary situations where an ODD element is missing (e.g., an ADS reliant on lane markngs encounters a short stretch of road with obscured markings), an ADS will not perform the DDT outside of its ODD and for safety reasons should not do so. Therefore, track and real world testing of an ADS must occur in a test environment within the ODD of the ADS or one that exactly duplicates all ODD elements.~~

'If encountered' as used in the table above would indicate that real-world testing would not seek to assess the particular requirement but would do so if it occurred during a test. Some situations are clearly undesirable from a safety perspective on public roads. However, given that real-world testing inherently involves uncontrolled parameters, critical traffic situations could organically occur and in this case, the performance with regard to the specific requirement should be assessed. Safety during testing on public roads should also be taken into account, and the assessor or the driver should ensure they can take over the driving task if needed.

Instead of “Yes” or “If encountered”, the table might also be structured to provide more information on the intended objective(s) of the test. For example:

Table 2. Example of alternative structure for the general matrix

ADS Safety Requirement	*Track*	*Real World*
The ADS should respond safely to the cut-in of another vehicle.	Verification of the ADS crash-avoidance response to a dangerous cut in.	Nominal verification that the ADS adapts the vehicle positioning in response to the cut in. Verification of the ADS crash-avoidance response to a dangerous cut in, if encountered.

Matrix for track testing

The following table illustrates an approach combining traffic scenarios, performance requirements, and test specifications into a matrix for conducting track tests. The “scenario” column would cross-reference the testing with the scenario upon which the testing is based, covering the traffic situation, infrastructure elements, objects, ODD elements, etc. The “safety requirement(s)” column would cross-reference the **applicable safety requirements** ~~specifications~~ established for ADS performance under the scenario. The “additional test specification” column would allow for conditions or parameters not described in either the traffic scenario or the safety requirement(s), but are necessary to conduct the track test (e.g. minimum duration of the test).

Table 3. Example of a test matrix for track test

Traffic Scenario	*Safety Requirement(s)*	*Additional Test Specifications*	*Assessment Specification*
Unobstructed travel on a straight path	Safe lateral positioning in a lane of travel	A minimum test duration of 5 minutes	The test shall verify that the ADS does not leave its lane and maintains a stable position inside its ego lane across the speed range within its system boundaries.
Unobstructed travel along a curve	Safe lateral positioning in a lane of travel Adapt to road conditions	A minimum test duration of 5 minutes	The test shall demonstrate that the ADS does not leave its lane and maintains a stable position inside its ego lane across the speed range and different curvatures within its system boundaries.

Cut-in by another vehicle while traveling on a straight path	Respond safely to the cut-in Safe longitudinal positioning relative to a lead vehicle	Scenario with selected parameters to verify the ADS crash-avoidance response to a dangerous cut in per ~~the safety requirements~~ **safety requirement**[39]	The test shall demonstrate that the ADS is capable of avoiding a collision with a vehicle cutting into the lane of the ADS vehicle up to a certain criticality of the cut-in manoeuvre.
ODD exit scenario	ADS detection of ODD boundary Automated response (failed fallback user response or no fallback user)	Test for failed fallback user response	The test shall demonstrate that the ADS is capable of bringing the vehicle to a safe stop, in case of a failed fallback user response.

Matrix for real world testing

The following table illustrates an approach combining performance requirements and traffic situations into a matrix for conducting real-world testing. The "safety requirements" column would specify the verifiable performance requirement(s).

The top rows on the right side set out traffic situations **required** to be encountered during real-world testing. The matrix intentionally uses the term "traffic situation" rather than "traffic scenario" given that real-world traffic cannot be controlled to reproduce predefined scenarios in all cases. **The envisaged descriptions of the situations will be rather general in order to ensure that there is a very high probability of them being encountered during real world testing.** The test route(s)**, therefore,** should be designed to ensure exposure of the ADS **within the ODD** to situations under which the ADS can demonstrate compliance with the safety requirements.

The remaining fields of the matrix describe behavioural competencies defined for the traffic situations per Annex 3. Each behavioural competency summarizes the desired performance in one sentence with a more detailed description to be set out in the testing protocols accompanying the test matrix where necessary. The behavioural competencies correspond to the safety requirement(s) applicable to each traffic situation.

As discussed under the general matrix, the real-world testing matrix allows for "if encountered" assessments. **The "if encountered" may occur in two situations. First, the assessment of safety requirements that are undesirable to be conducted on public roads, but which may nevertheless occur.**[40] **Second, the assessment of safety requirements (during nominal traffic conditions) that cannot be assured (and therefore required) to be encountered during real world testing, but which may occur.**

39 This inclusion assumes the traffic scenario does not prescribe the range of parameters to be selected for the occurrence of a safety-critical situation. If that were to be included in the scenario, this field could be empty.

40 **It should be possible for the assessor to interrupt the test on public roads, should the situation become dangerous.**

An illustration of the first is the example on Row 2.1 of the table on the safe response to a cut-in. The requirement is the assessment of the ADS' response to a (nominal) cut-in of another vehicle during real world testing. The ADS' response to a dangerous cut-in could only be assessed if encountered during real world testing, as signalled by the addition of ', if applicable.'.

An illustration of the second is also the example in Row 2.1 of the table on the safe response to a (nominal) cut-in. This situation is likely but not guaranteed to occur in any or possibly all of the traffic situations listed in the top row of the table. When it does occur it should be assessed.

~~Real-world testing requires assessment of nominal performance but allows for conditional assessment of critical and/or failure performance should such situations occur during the testing. Real-world testing includes assessment of the ADS competency to mitigate safety risks due to external conditions and behaviours of other road users. For example, row 2.1. notes ADS responses to a nominal cut-in by another vehicle as well as the possibility of a dangerous cut-in occurring during the testing.~~

Aspects related to routing (e.g. minimum duration, minimum frequency of a given traffic situation encountered during testing, etc.) would be set out in the accompanying test protocols.

Table 4. Example of a test matrix for real world testing: motorway application

		Traffic Situations				
	Safety Requirements	*Driving on the motorway*	*Merging*	*Lane Change*	*Overtaking*	*Exiting Motorway*
1.1.	Safe lateral positioning in a lane of travel	The ADS demonstrates it does not leave its lane and maintains a stable position inside its ego lane across the speed range within its system boundaries.	The ADS demonstrates it achieves a stable position inside the target lane upon completion of the lane change procedure.	The ADS demonstrates stable positioning inside the target lane upon completion of the lane change procedure.	The ADS demonstrates it achieves a stable position inside the target lane upon completion of the lane change procedure.	The ADS demonstrates it maintains a stable position in the off-ramp lane.
2.1.	Respond safely to the cut-in of another vehicle	The ADS adapts the vehicle positioning in response to the (nominal) cut in. The ADS responds appropriately[41] to a dangerous cut in, if applicable.[42]				
2.2.	Safe longitudinal positioning relative to a lead vehicle	The ADS demonstrates it maintains a safe longitudinal position relative to a lead vehicle.	The ADS demonstrates it maintains a safe longitudinal position relative to a lead vehicle during and upon the completion of the lane change procedure.	The ADS demonstrates it maintains a safe longitudinal position relative to a lead vehicle prior and during the lane change procedure. The ADS demonstrates it maintains a safe longitudinal position relative to a lead vehicle upon the completion of the lane change procedure, if applicable.	The ADS demonstrates it maintains a safe longitudinal position relative to a lead vehicle prior and during the lane change procedure.	The ADS demonstrates it maintains a safe longitudinal position relative to a lead vehicle, if applicable.

[41] What constitutes an 'appropriate response' would then be set out in the testing protocols that accompany the test matrix, sourced from FRAV.
[42] To be determined whether 'If encountered' situations should be included in the matrix itself. Included here, as well as in other parts of the table, as an illustration.

II. Justification

During the 18th session of GRVA, the FRAV/VMAD Integration Document (GRVA-18-50) was adopted. During the presentation guiding the work for this document, it was already announced that during the December meeting of the FRAV/VMAD group it was noted that for Track-/Real World Testing some parts of the text in the original VMAD document (VMAD-34-04/Rev.2) were missing and should be included.

Moreover, during that December meeting some suggestions for improvement of the text were proposed and/or requested. The total set of modifications has been prepared and was presented to and accepted by the FRAV/VMAD group during the session just after the January session of GRVA.

The proposal in this document represents the outcome of this work.

Document VMAD-35-02 provides a colour-coded overview of the amendments included in this proposal, showing in detail which changes refer to:

- VMAD text which had previously been approved by GRVA and WP.29, and was identified as missing in the Integration Document in the December FRAV/VMAD session;
- text modifications resulting from the December FRAV/VMAD session, which could not yet be included in the integration document of FRAV/VMAD submitted to GRVA for its session in January 2024.