

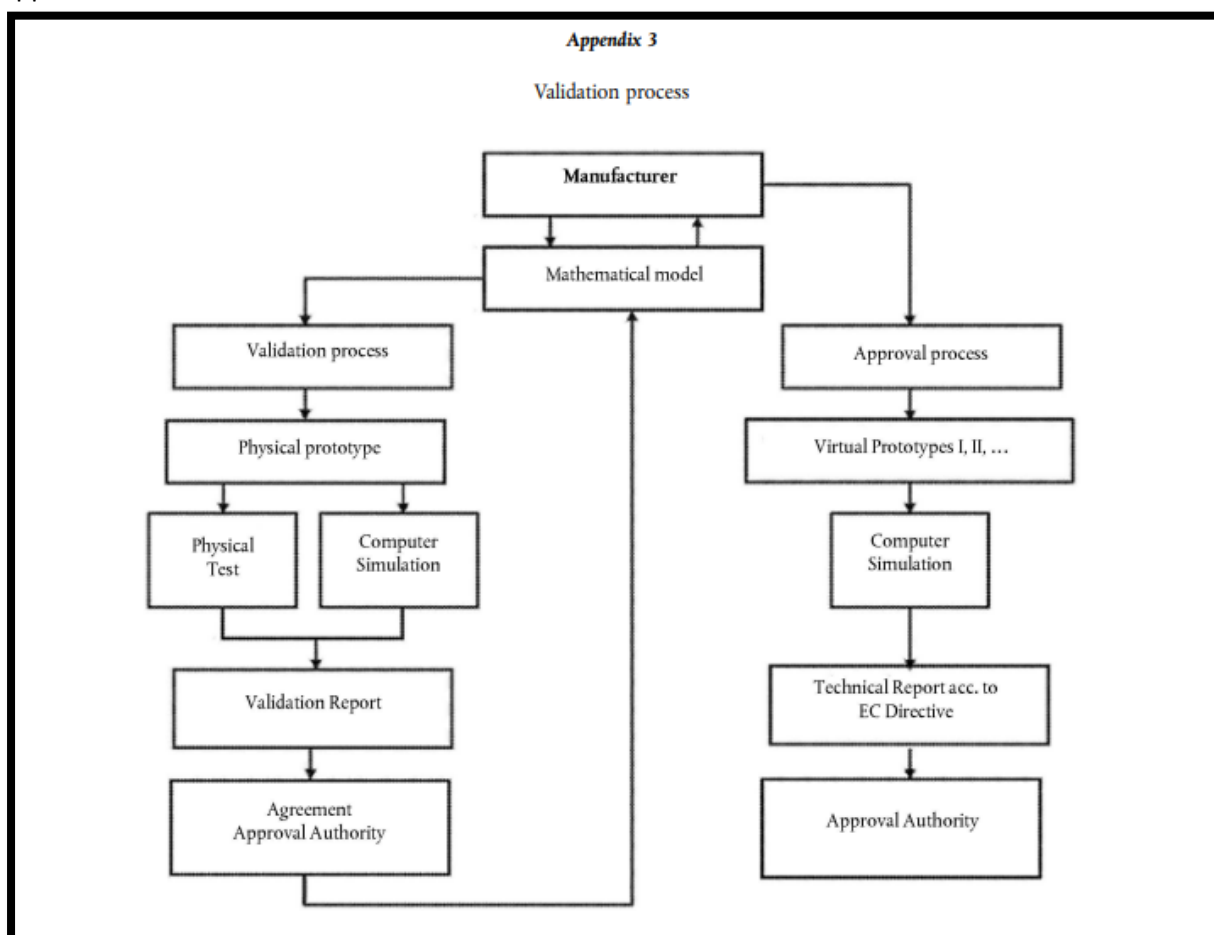
Virtual testing in UN R 152

Introduction

The current GRVA-18 proposal for the use of Virtual Testing in R152 on AEBS does not appear to address all the necessary aspects associated with the specification, development, deployment and use of the simulation tools that constitutes a “Virtual Testing Capability”. GRVA’s VMAD Working Group has been developing a New Assessment Test Methodology (NATM) that includes a description of the requirements associated with a virtual testing capability when it is being used as part of an approval process. A summary of those requirements extracted from the work of the VMAD group is provided in this document.

The complete NATM framework has been developed in the context of an Automated driving System and is more complex than is required for the R152 however the set of requirements associated with the “credibility assessment” of the virtual testing capability are relevant.

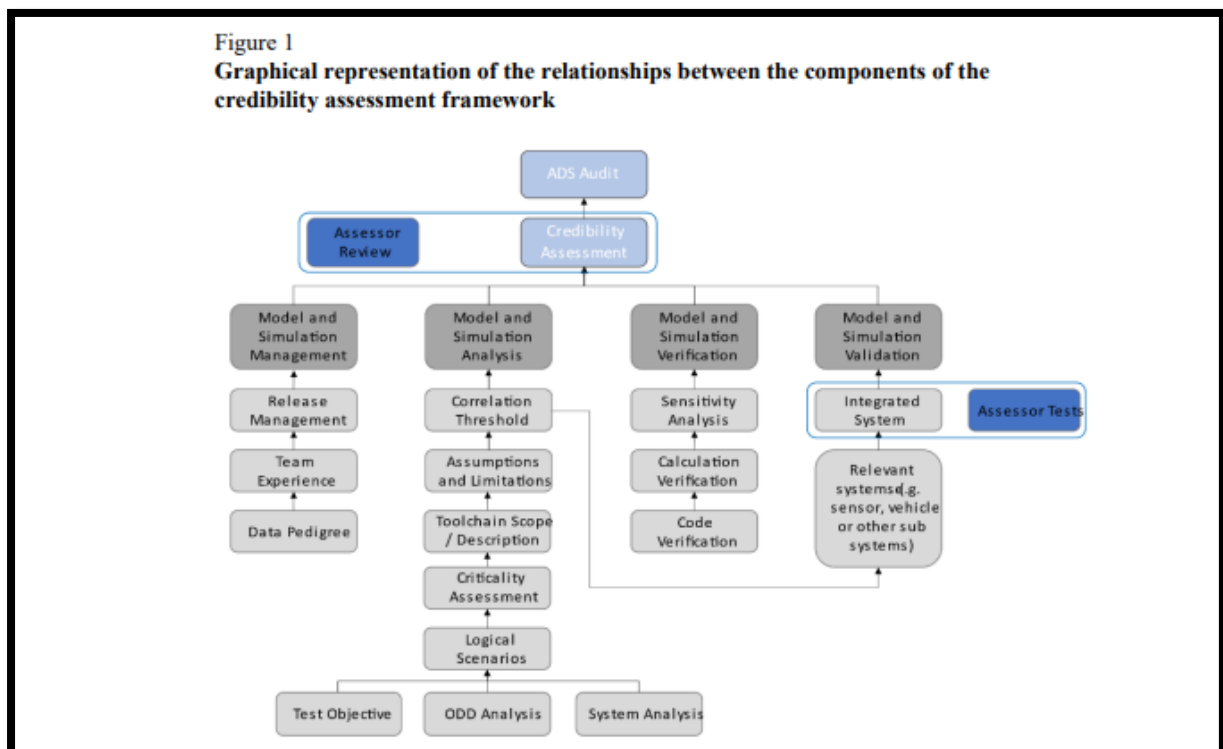
It is also worth noting the generic approach to virtual testing contained in the EU Whole Vehicle Type Approval Regulation – 2018/858 and in particular the following diagram extracted ANNEX VIII Appendix 3.



This shows a clear two stage activity where the simulation toolchain and its components are first developed and assessed before being used. Appendix 1 of the same Annex also provides some “conditions” for the virtual testing method.

In general, it is impossible to formally prove the capability of modelling and simulation (M&S) tools, so an alternative is to adopt a structured approach that builds a body of evidence that creates the necessary level of confidence. This gives the manufacturer reassurance and also allows a third party such as a Technical Service (TS) or Type Approval Authority (TAA) to decide whether the M&S tools are of sufficient maturity to be used for virtual testing.

The UNECE VMAD activity has provided that structure and developed clear guidelines on the requirements around the development and assessment of an M&S capability that is to be used for “virtual testing”. The approach is called a “Credibility Assessment Framework” (the figure below is from NATM Guidelines for Validating Automated Driving System (ADS) – WP.29/2022/58).



The credibility framework shown above covers the various aspects that the manufacturer must consider when producing a virtual testing capability. The manufacturer should develop the processes and documentation that support this and then generate the compliance evidence. The evidence will consist of various documents that show that the processes have been followed and that the results including verification, validation and testing achieve the necessary criteria. This is a complex activity and the details will be agreed with the TS and TAA. The assessment will be an audit process conducted by the TS or TAA of all the information provided by the manufacturer.

The framework is a structured way to describe and report the M&S capability based on the manufacturer’s approach and meeting appropriate criteria. Adopting and following this framework generates confidence in the results because it shows that the manufacturer is following good practice. In other words, the credibility is established by evaluating the factors that are considered to be the main contributors to a successful M&S capability, including, M&S management, team's experience and expertise, description and analysis of the M&S components, data management, verification, validation and uncertainty characterization. Each of these factors contributes to the overall quality and if the assessor decides that the required level has been achieved then the M&S capability is considered credible and fit to use for virtual testing.

Summary of the NATM Credibility Framework Requirements

The following is a summary of the main requirements outlined in the NATM document WP.29-2022-58. It is worth noting that the complete description runs for several pages. The summary follows the outline proposed in the NATM document.

Note: It is possible / probable that some of the aspects may not be applicable to the specific application of AEBS but it is better to allow the manufacturer to claim and give the rationale for any “exemptions” or inapplicability as part of their information pack rather than simply to remove it from the requirement.

Models and Simulation Management

The M&S lifecycle should be monitored, managed and documented. Management activities should be established to support the M&S adopting a work product management approach including:

- **M&S management process**, including a description of modifications to the toolchain and its components, clear designation of the software, acceptance review processes, details of the lifecycle and lifetime support and description of management responsibilities and escalation processes;
- **Release management**, including version control of the toolchain and components used for any approval activities, validation strategy and acceptance criteria, data provenance and traceability, data quality checks covering completeness, accuracy, and consistency;
- **Team's Experience and Expertise (E&E)**, including processes and evidence that the various teams developing, testing and validating and ultimately using the toolchain and its components have appropriate experience and training;
- **Input Data Management**, including traceability of the data used in developing the toolchain and its components and used in the validation activity, evidence that the data is fully representative of the intended scope and functionality of the application, evidence of data quality considerations when it is being used for model development and parameter estimation;
- **Output Data Management**, including records of the scenarios used and the outputs of the validation activity, output data traceability, approach and results from comparisons and correlations, consistency and sanity checks.

Description and analysis of the toolchain and components

The description and analysis should provide a description of the toolchain and its components, identify the applicable parameter space as well as the scope, limitations and the sources of uncertainty that can affect the results. This will include:

- **General description**, including a description of the complete toolchain and its components a clear description of the objectives and metrics;
- **Assumptions, known limitations and uncertainty sources**, including the rationale for the modelling assumptions and hence the limitations, the fidelity required for the toolchain and its components, the tolerance and criteria for the correlation of real and virtual results and information about the sources of uncertainty in the model;
- **Scope**, includes a clear description of the applicability of the toolchain and its components, the accuracy required to emulate the physical phenomena and the fidelity required to do so, the validation scenarios and the corresponding parameter description limitations, acceptance and testing requirements derived from ODD analysis;

- **Criticality assessment**, including the impact of the errors in the toolchain and its components on the safety of the system and any subsequent functional safety requirements.

Verification

Verification deals with the analysis of the correct implementation of the conceptual / mathematical models that create and build up the tools & toolchain. Verification contributes to the credibility by providing assurance that the toolchain and its components will exhibit realistic behaviour for all inputs including those that have not been explicitly assessed. There are several ways to perform verification including:

- **Code Verification** is concerned with activities that try to show that the numerical and logical implementation of the toolchain and its components is correct including, static/dynamic code verification, convergence analysis and comparison with exact solutions, a sufficiently exhaustive exploration of the input parameters domain to identify parameter combinations for which the M&S tools show unstable or unrealistic behaviour and sanity / consistency checking;
- **Calculation verification** deals with the estimation of numerical errors affecting the toolchain and components including, numerical error estimates (e.g. discretization error, rounding error, iterative procedures convergence) and analysis that the errors remain sufficiently bounded;
- **Sensitivity analysis** aims at quantifying how output values are affected by changes in input values and to identify the parameters having the greatest impact on the results. The analysis should include the identification of the most critical parameters influencing the results and a robust calibration procedure for those parameters.

Validation

Validation is the process of determining the degree to which the toolchain and its components are an accurate representation of the real world from the perspective of the intended use. The following should be part of the validation activity:

- **Measures of Performance (metrics)** are defined during the M&S analysis stage and include discrete value analysis, time evolution and analysis of state changes;
- **Goodness of Fit measures** are also used to compare the outputs of the toolchain and its components with physical tests. The results are compared statistically to see if the measures have been achieved;
- **Accuracy requirements** are defined during the M&S analysis and should set the thresholds for the various comparisons. The validation results should show that these have been met.
- **Validation methodology or strategy** is the approach adopted by the manufacturer to show the toolchain and its components are fit for purpose. It includes the choice of scenarios to cover the maximum possible extent of the ODD and validation of subsystem and combinations of subsystems including environment, sensors, vehicle systems, user behaviour, etc.;
- **Validation scope** A toolchain consists of multiple tools, and each tool will use several models. The validation scope includes the appropriate assessment of the toolchain and its components;
- **Internal validation of results** should provide evidence of the validation activity and information related to the processes that were followed, physical tests that were performed and products that were used;

- **Uncertainty characterisation** is concerned with characterizing the expected variability of the virtual toolchain results. The analysis should characterise the uncertainty in the input data, in the model parameters and in the toolchain structure that is collected from the “Description and analysis of the toolchain and components” and the “Input Data Management”. The identified uncertainties should then be propagated through the toolchain and the overall uncertainty of the results quantified and appropriate safety margins established.

French Proposal

The current proposal (GRVA-18-23) does not cover all the requirements identified in the VMAD activity for the credibility framework that have been summarised from the NATM document in the section above. Also, some areas that have been covered have insufficient detail to allow a robust assessment by a Technical Service or Type Approval Authority. The document also does not follow the structure of the NATM document when discussing the various aspects that need to be addressed for the credibility framework.

The following identifies some of those areas that should be addressed in a revised proposal.

- Experience and Expertise (E&E) is touched upon briefly in section 1.1.3e on “Usability”. This makes no mention of the broader issues around E&E throughout the design, development etc. aspect of the M&S capability.
- Physical Testing appears only to be required for the final toolchain (Section 1.2). Physical Testing should be used and evidenced for all the components of the toolchain.
- There is some detail about the overall M&S Management process in section 1.5.7. but this is limited and does not provide a clear structure for the overall process of control and management responsibility.
- Sections 1.5.2 & 1.5.6. mention data input and data management but the NATM document has detailed sections covering both data input and output management.
- There are some references to validity domain in the document but no indication as to how this is derived, e.g., assumptions, limitations and tolerances.
- There is no mention of uncertainty in the document sources and characterisation.
- There is no mention of sensitivity analysis to quantify the effect of variations of input or model parameters.
- There is no mention of criticality assessment that might influence decisions about how stringent any assessment should be.
- There is no reference to a verification activity. Verification and validation are often combined, but it is a separate phase of the M&S development and assurance process and is not generally about comparison with physical tests.
- There is some mention of “accuracy” but limited description (see 1.4.3) or justification for the criteria that are needed for a successful validation, i.e. measures of performance and goodness of fit. These criteria are specific during the analysis of the M&S process to ensure that the final capability is fit for purpose.
- The current document recognises that there is an audit activity but does not provide clear requirements that can be audited. Many of the paragraphs appear to describe an activity and therefore only imply what the audit should assess.
- Section 1.2.1.1 discussed the number of physical tests and is agreement with the Technical Service or Type Approval Authority. This approach does not seem fully in line with the intention of the NATM. The testing that is used to prove the simulation capability should be part of the overall validation strategy. It should be proposed and justified by the manufacturer. As part of the audit this strategy will be reviewed, but it is unlikely that a set

of tests can be uniquely identified it is likely to be iterative and based on the ongoing discussions.

Summary

The current proposal to incorporate virtual testing into R152 does not address all the above aspects of a credible virtual testing capability and does not provide sufficient detail for those that are mentioned. The process of specifying, developing, deploying and managing a virtual testing capability is complex and it is important that all the relevant aspects are addressed in any proposal to ensure that a proper and complete assessment is made during the approval process.