Submitted by the OPI on Virtual Testing Credibility

ADS-04-10
4th ADS IWG session
8-11 October 2024

Transposition of Virtual Testing Credibility Assessment Requirements[1] into provisions for the forthcoming ADS UNR/GTR

The present document provides a second proposal to transpose the provisions included in the "Guidelines and recommendations for ADS safety requirements, assessments and test methods to inform regulatory development" (GRVA-19-15r1e). In particular the requirements to ensure the credibility of the simulation toolchain(s) that can be used to perform virtual testing as part of the safety and performance assessment of the ADS. The paragraph numbers (not to be considered relevant for the present proposal) reported in the present document as well as the reference text is taken from the aforementioned Guidelines. Amendments to the proposals are recorded in track-changes. Terms and definitions reported in section 3 of the Guidelines are considered valid also for the regulatory text.

[…]

3. Definitions

3.Y *"Virtual Testing"* means a type of testing that uses a simulation toolchain to assess the performance of the ADS

3.33 *"Simulation"* means the imitation of the operation of a real-world process or system over time utilizing a software implementation for some (or all) of the models, tools or test environment.

3.34 "*Simulation toolchain*" means a simulation tool or a combination of simulation tools that are used to support the validation of the ADS safety case and/or of one or more performance/functional/user regulatory requirements

Commented [RD1]: OPEN ITEM: CANADA
Definition of simulation toolchain and link with criticality assessment

4. General requirements

[…]

4.X The manufacturer shall demonstrate that the approach to testing is suitable for the demonstration of the safety case and the compliance with performance/functional requirements.

4.X.1 The manufacturer shall demonstrate that the physical testing (proving ground and/or public road) facilities and environment are suitable for the tests that are being conducted.

4.X.2 The manufacturer shall demonstrate that the simulation toolchain(s) is suitable for conducting virtual tests. The requirements for the simulation toolchain are listed in 5.X.

[…]

5. Requirements/Specifications

5.X. Credibility Framework Requirements

5.X.1 The manufacturer shall describe the intended use(s) of virtual testing and its role in the overall testing strategy.

5.X.2 [RH/UK Alternative] The manufacturer shall demonstrate that the simulation toolchain(s) is suitable to use for virtual testing by demonstrating compliance with the requirement listed in this section.

Commented [RE(2]: OPEN ITEM: the possibility to have more than one virtual toolchain for different intended use with different level of credibility.

[1] From GRVA-19-15r1e: Guidelines and recommendations for ADS safety requirements, assessments and test methods to inform regulatory development

- The manufacturer shall document the capability of the simulation toolchain and explain their claim that it is suitable to undertake the virtual testing.
- The manufacturer shall document the activities and processes that support the claim that the simulation toolchain is suitable to use for virtual testing.
- The manufacturer shall provide evidence that supports their claim that the simulation toolchain is suitable to use for virtual testing.

5.X.2. [JRC Proposal] The manufacturer shall demonstrate that the simulation toolchain(s) is suitable to use for virtual testing by:

- performing a criticality analysis that evaluates the potential risk and consequences of using the simulation toolchain(s) for the assessment of the ADS safety case and functional/user requirements;
- demonstrating that the simulation toolchain(s) fulfils the credibility requirements corresponding to the identified criticality as per the requirements listed in this section.

5.X.3. Simulation Toolchain Management requirements

5.X.3.1. Simulation Toolchain Data Management requirements

5.X.3.1.1. The manufacturer shall manage the data used to develop, verify, validate and update the simulation toolchain(s) throughout its lifetime. The manufacturer shall consider the completeness, accuracy and consistency of this data.

5.X.3.1.2. The manufacturer shall maintain a record of the data used in the validation of the toolchain.

5.X.3.1.3. If the simulation toolchain(s) incorporates or relies upon data/tools from other organizations which are not under the control of the manufacturer, the manufacturer shall demonstrate the measures taken to manage the quality and integrity of that data/tools.

5.X.3.1.4. With regards to input data management and parameters associated with the simulation toolchain(s), the manufacturer shall:

(a) document the data used to develop, verify and validate the simulation toolchain(s) and note important quality characteristics;

(b) provide documentation showing that the data used to develop, verify and validate the simulation toolchain(s) covers the intended functionalities that the virtual testing aims to assess;

(c) document the data and the calibration procedures employed to fit any parameters associated with the simulation toolchain;

(d) explain the reasons for data or parameters changing between releases.

5.X.3.1.5. The manufacturer shall quantify the uncertainty in the simulation toolchain(s) and its outputs that occur because of the quality of the data (e.g. data coverage, signal to noise ratio, and sensors' uncertainty/bias/sampling rate).

5.X.3.1.6. With regards to the data that is produced by the simulation toolchain(s) and its components, the manufacturer shall:

(a) maintain a record of the output from the simulation toolchain(s) during its validation and ensure that they are traceable to the input data that produced them.

(b) document the output data and note any important quality characteristics that can be deduced from analysis of the data, e.g. applying statistical methodologies;

5.X.3.1.7. With regards to the quality of the data that is produced by the simulation toolchain(s) and its components, the manufacturer shall:

(a) ensure it is sufficient to undertake any validation activity;

(b) ensure it is sufficient to allow consistency/sanity check of the simulation toolchain, possibly by exploiting redundant information;

Commented [RH3]: I am not sure how we define the "level of credibility". I appreciate that this will be tied in with Criticality (ISO26262) but we will have to provide some clear guidance around this if we are to use this concept. [NASA-STD-7009A does have some details of "M&S Results Credibility Assessment".

Commented [RE(4R3]: Text changed because we do not have levels of credibility, but the manufacturer shall define its strategy to ensure a credible toolchain based on the criticality level

Commented [RH5R3]: Doesn't the first bullet still mention "criticality level"

Commented [DR(I6R3]: Added to open item list the details on criticality

Commented [RH7]: In practice, the manufacturer will present both the activities and processes that help them to make the **claim** that the toolchain is "credible" or fit for purpose. They will also then provide the evidence that supports the claim. [The assessor will then review the claim and the evidence.]

(c) ensure it is sufficient to justify manufacturer's claims about their safety case.

5.X.3.1.8. With regards to the management of stochastic models, the manufacturer shall:

(a) characterize the variance in the simulation toolchain(s)'s output;

(b) ensure the possibility of a deterministic re-execution of the simulation toolchain.

5.X.3.2. Simulation Competency requirements

5.X.3.2.1. The manufacturer shall document and provide the rationale for their confidence in the competency of:

(a) the personnel that developed the simulation toolchain(s) and its components;

(b) the personnel that assessed the simulation toolchain(s) and its components;

(c) the personnel that used the simulation toolchain(s) to perform the testing with the purpose of validating the system.

5.X.3.2.2. The manufacturer shall have processes and procedures that identify and maintain the skills, knowledge, and experience needed to perform the various activities. The following processes shall be established, maintained and documented.

(a) Process to identify and evaluate the necessary competencies that are required to perform the modelling and simulation activities;

(b) Process for training personnel to be competent to perform the modelling and simulation activities.

5.X.3.2.3. The manufacturer shall maintain records of the personnel in the various teams showing they have received the necessary training and have been deemed competent to perform the modelling and simulation activities assigned to those personnel.

5.X.3.2.4. The manufacturer shall set up suitable arrangements with third-party organisations to ensure that the competency of their personnel is adequate to demonstrate the credibility of the simulation toolchain(s).

5.X.3.2.5. [NEW: Placeholder to ref SMS in case of third-party data/tools providers]

5.X.3.3. Simulation Toolchain Release Management requirements

5.X.3.3.1. The manufacturer shall manage and support the simulation toolchain(s) used for virtual testing throughout its complete lifecycle.

5.X.3.3.2. The manufacturer shall manage and document the simulation toolchain(s) release process. The simulation toolchain(s) release management activity shall include:

(a) a description of the modifications associated with each toolchain(s) release;

(b) a record of any associated software (e.g., specific software product, designations and version) and hardware arrangements (e.g., XiL configuration);

(c) a record of the internal review activities that supported the toolchain(s) acceptance and release.

5.X.4. Simulation Toolchain requirements

5.X.4.1. The manufacturer shall describe the simulation toolchain(s) and identify its scope of applicability, its limitations, assumptions and the sources of uncertainty that can affect results.

5.X.4.1.1. Description of the Simulation Toolchain

5.X.4.1.1.1. The manufacturer shall provide a description of the simulation toolchain(s) and its components.

5.X.4.1.1.2. The manufacturer shall provide a description of the approach adopted in the simulation toolchain(s) validation.

5.X.4.1.1.3. The manufacturer shall provide a description of the acceptance tests and criteria that will be used to determine if the simulation toolchain(s) is considered credible based on the credibility framework.

5.X.4.1.2. Simulation Toolchain Assumptions, known Limitations, and Uncertainty Quantification

5.X.4.1.2.1. The manufacturer shall describe the modelling assumptions and considerations that which guided the design of the toolchain.

5.X.4.1.2.2. The manufacturer shall provide information on:

(a) any assumptions made during the development of the simulation toolchain(s) and its components and the limitations that this places on its scope and applicability;

(b) the rationale for choices made about the level of fidelity of the simulation toolchain(s) and its components.

5.X.4.1.2.3. The manufacturer shall provide justification that the tolerances associated with the simulation toolchain(s) are appropriate and meet the acceptance tests and criteria.

5.X.4.1.2.4. The manufacturer shall provide details of the sources of uncertainty in the simulation toolchain(s) and its components and an assessment of their impact on the results.

5.X.4.1.3. Simulation Toolchain scope.

5.X.4.1.3.1. The manufacturer shall document the scope of the simulation toolchain(s) and identify its limitations. It should refer to the ODD and identify any limitations about its applicability within the ODD.

5.X.4.1.3.2. The manufacturer shall demonstrate how the simulation toolchain(s) imitates the relevant physical phenomena and meets the necessary level of accuracy.

5.X.4.1.3.3. The manufacturer shall demonstrate that the test selection for simulation toolchain(s) validation is sufficient to demonstrate that it will perform effectively within the defined scope.

5.X.4.1.3.4. The manufacturer shall provide a list of tests used for validation and the corresponding parameters and any known limitation.

5.X.4.1.4. Simulation Toolchain Criticality analysis.

5.X.4.1.4.1. The manufacturer shall review the simulation toolchain(s) to assess the criticality of prediction errors and the effect these would have on the manufacturer's claims about their safety case.

5.X.5. Simulation Toolchain Verification requirements

5.X.5.1. The manufacturer shall demonstrate that the simulation toolchain(s) will not exhibit unrealistic behaviour for valid inputs which have not been explicitly tested.

5.X.5.2. Simulation Toolchain Code Verification requirements

5.X.5.2.1. The manufacturer shall document the execution of proper code verification techniques, used in evaluating the simulation toolchain(s) and its components, e.g. static/dynamic code verification, convergence analysis and comparison with exact solutions if applicable.

5.X.5.2.2. The manufacturer shall provide evidence that the input parameter space was sufficiently explored to identify if there are any parameter combinations for which the simulation toolchain(s) shows unstable or unrealistic behaviour.

5.X.5.2.3. The manufacturer shall provide information on any sanity/consistency checking procedures that are used.

5.X.5.3. Simulation Toolchain Calculation Verification requirements

5.X.5.3.1. The manufacturer shall document numerical error estimates (e.g. discretization error, rounding error, iterative procedures, and convergence).

5.X.5.3.2. The manufacturer shall review their analysis and demonstrate that the numerical errors are understood and sufficiently bounded to allow the simulation toolchain(s) to be used for virtual testing.

5.X.5.4. Simulation Toolchain Sensitivity Analysis requirements

5.X.5.4.1. The manufacturer shall provide documentation demonstrating that the input data and parameters that most critically influence the toolchain(s) outputs have been identified by means of appropriate sensitivity analysis techniques.

5.X.5.4.2. The manufacturer shall demonstrate that robust calibration procedures have been adopted for assigning appropriate value(s) to the most critical parameters to ensure that the simulation toolchain imitates the physical system.

5.X.5.4.3. The manufacturer shall demonstrate that sensitivity analysis has been used to identify the critical input data and parameters that needs particular attention in order to characterize the uncertainty of the overall simulation toolchain(s) outputs.

5.X.6. Simulation Toolchain Validation requirements

5.X.6.1. The manufacturer shall quantitatively determine the degree to which the simulation toolchain(s) is an accurate representation of the real-world system by means of a validation analysis.

5.X.6.2. The manufacturer shall provide evidence that the simulation toolchain(s) results are consistent and correlated with the results of the physical tests.

5.X.6.3. The validation shall be performed on a sufficiently representative set of tests in order to substantiate the claims about the capability of the simulation toolchain(s) within its scope.

5.X.6.4. The manufacturer shall define the measures of performance (metrics) that will be used when comparing between the results of physical tests and the output of the simulation toolchain(s).

5.X.6.5. The manufacturer shall use appropriate statistical techniques when comparing the results of the physical tests and the output of the simulation toolchain(s) and its components.

5.X.6.6. The manufacturer shall specify acceptance tests and criteria during the simulation toolchain(s) and its components development activity and will demonstrate that they have been achieved.

5.X.6.7. The manufacturer shall define the methodology and the tests used for the simulation toolchain(s) validation. It should be clear whether the full ODD is within scope of the toolchain or only part of it.

The validation strategy may consist of one or more of the following:

(a) subsystem model validation e.g. environment models, sensor models, and vehicle models;

(b) vehicle system model validation (vehicle dynamics model together with the environment model);

(c) sensor system validation (sensor model together with the environment model);

(d) integrated system validation (sensor model together with the environment model with influences form vehicle model).

5.X.6.8. The manufacturer shall demonstrate that the accuracy criteria defined during the simulation toolchain(s) development have been met.

5.X.6.9. The manufacturer shall provide evidence that the processes related to the validation activity have been followed.

5.X.6.10. The manufacturer shall document their uncertainty characterisation analysis and provide information about how the simulation toolchain(s) should be used and any safety margins that should be applied when it is used for virtual testing.

5.X.6.11. The manufacturer shall demonstrate they have techniques to estimate the simulation toolchain(s)'s critical inputs.

Commented [RH8]: I am not sure what this means? What analysis would the manufacturer produce. What would be considered an acceptable "degree"?

Commented [DR(I9R8]: Suggest to retain original text. Further details in interpretation document

Commented [DR(I10R8]: OPEN ITEM: The level of information specification to be discussed (criticality, validation)

5.X.6.12. The manufacturer shall demonstrate that they have characterised the critical parameters used in the simulation toolchain(s) and its components and where appropriate have identified these as distributions with confidence intervals.

5.X.6.13. The manufacturer shall provide evidence that a proper characterization of the uncertainty of the results of the simulation toolchain(s) and its components, because of any assumptions therein, has been made.

5.X.6.14. The manufacturer shall demonstrate the that they have differentiated between the aleatory and epistemic uncertainties associated with the simulation toolchain(s).

6. Assessment and Test Method

6.X.1. The assessor shall review the manufacturer's credibility framework to determine whether the simulation the toolchain(s) is suitable to undertake virtual testing.

6.X.2. The assessor shall review the documentation and evidence supporting the manufacturer's claims

- A successful outcome of the assessment will be a confirmation that the claims of the manufacturer about the capability of the simulation toolchain(s), including its scope, are correct and that it can be used to perform the virtual testing as part of the ADS assessment.
- The simulation toolchain(s) can only be used to undertake virtual testing once the credibility of the same has been established.

6.X.3. The assessor shall audit the information provided by the manufacturer and may request or carry out additional tests of the simulation toolchain(s) or physical tests. The outcome of the tests shall be reviewed and any concerns or discrepancies shall be raised and reviewed with the manufacturer.

The manufacturer shall provide an explanation of the discrepancies in the results. If the results from the simulation toolchain(s) do not sufficiently replicate the output of physical test or does not have sufficient scope the assessor shall inform the manufacturer.

The manufacturer shall conduct extra validation activity and resubmit their information for further assessment.

6.X.Y. [RH/UK Alternative - TBD] If the assessor is not satisfied with the information provided by the manufacturer or the outcome of the additional tests, then the assessor will document those concerns and inform the manufacturer. The manufacturer will then have the opportunity to revise the documentation and evidence and resubmit. The resubmission should address the concerns raised by the assessor and should also clearly indicate the scope of the changes and whether there are any wider consequences.

Annex X Assessment of Credibility for the simulation toolchain used for virtual testing

I. Introduction, motivation, and scope

1. The use of virtual testing can be beneficial for ADS safety and performance assessment because it provides an opportunity to overcome some of the limitations of real-world testing and allows a substantial increase in the number of scenarios that can be tested. Virtual testing can however produce erroneous but seemingly correct results. This is possible in any situation but particularly when using complex simulation toolchains that are not adequately supported by robust practices that address all aspects of the modelling and simulations process not just validation. Therefore, confidence in the credibility of the toolchain(s) is needed so that it can be used for virtual testing in conjunction with physical testing. A toolchain can be used for virtual testing if its credibility can be demonstrated whilst considering its limitations, assumptions, accuracy and uncertainties.

Commented [RH11]: This is to start a conversation about how the assessor process will work. Documenting concerns, having the manufacturer address them, provide details of what the manufacturer has done, whether the manufacturer has actually made system changes, the impact of those changes , ….

Commented [DR(I12R11]: As it is a new content and quite general on the audit process we might consider to discuss this in the future

Commented [RH13]: Not sure, we want something like "In other words", especially if they say different things.

2. A simulation toolchain can be used for virtual testing if its credibility is established by evaluating its fitness for the intended purpose. It is recommended that credibility is achieved by investigating and assessing five properties of the toolchain:

(a) Capability – what the toolchain can do, and what are the associated risks;

(b) Accuracy – how well does the toolchain reproduce the target results;

(c) Correctness – how sound & robust are the data and the algorithms in the toolchain and its components;

(d) Usability – what competencies are needed and what is the quality of the process that manage its use;

(e) Fit for Purpose – how suitable is the toolchain for the assessment of the ADS within its ODD.

3 Creating an assessment framework is complicated by the diverse features and the variety of toolchains that may be used. A risk-based/informed credibility assessment framework is required that can be applied to any toolchain.

3 Creating an assessment framework is complicated because manufacturers will use a variety of toolchains from different sources and with diverse features. A risk-based/informed assessment framework is required that can be applied to any toolchain.

II. Components of the credibility framework and related documentation requirements

1. (UK Alternative) The credibility framework is a structured way to address all the necessary aspects that are required to produce a toolchain that is fit for purpose. A manufacturer can use this to manage their approach and then to make their own assessment of the suitability of the toolchain. The credibility is established by evaluating all the relevant factors that are considered to be the main contributors to the behaviour of the toolchain and therefore affect its overall suitability. How well each of these factors is addressed indicates the quality of the toolchain, and the comparison between the obtained levels and the required levels provides a qualitative measure of the suitability of the toolchain credibility and fitness for its use in virtual testing.

2. A representation of the relationship among the components of the credibility framework is reported in the following figure.

Graphical representation of the relationships between the components of the credibility framework

Commented [RH14]: We don't appear to discus risk based / informed elsewhere?

Commented [RH15]: Is this about the manufacturer's credibility framework or about the assessor framework.

Commented [RH16]: I have split this to be clear between the manufacturer requirement and the assessor. I am not sure how much we will define the assessor framework.

Commented [RH17]: We don't appear to discus risk based / informed elsewhere?

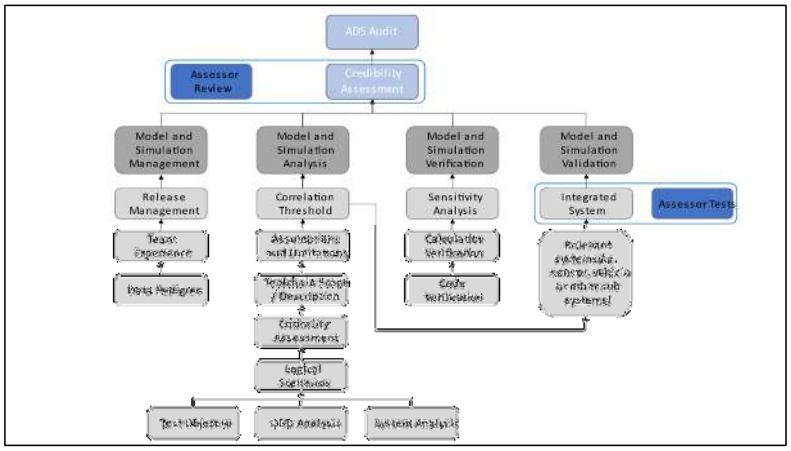
Commented [RH18]: Is this about the manufacturer's credibility framework or about the assessor framework.

Commented [RH19]: I have split this to be clear between the manufacturer requirement and the assessor. I am not sure how much we will define the assessor framework.

Commented [RH20]: This probably needs to be thought about! If there are "required levels" then the "obtained levels" should meet them. The approach described seems to be based on a Maturity Model Framework. I can see it would be useful to have a "qualitative measure" but the outcome will have to be pass / fail. The pass may have "conditions" but the assessment will still have to say it can or can't be used.

Commented [DR(I21]: Picture to be updated

Commented [RD22R21]: RH/UK: May need to edit this diagram if it is included in the final text: titles should match any heading and the use of "Credibility Assessment" may not be the correct.



1. Release management

2. Team's Experience and Expertise.

12. Even though Experience and Expertise (E&E) are already covered in a general sense within the organization, it is important to establish the basis for confidence on the specific experience and expertise for modelling and simulation activities.

13. The credibility of the toolchain and its use for virtual testing depends on the E&E of the personnel involved in its development, verification, validation, deployment and usage. For instance, a proper understanding of the toolchain's limitations and domain of applicability will help to prevent its possible misuse or a misinterpretation of its results.

15. Appropriate management of the E&E of the teams used by the manufacturer is an element of the credibility framework. It can help to ensure that the human element is taken into consideration and the risks associated with this aspect of the process are mitigated and controlled.

17. Experience and Expertise exists at two levels:

19. The manufacturer shall demonstrate how the specific requirements of [this section - Reference] are incorporated into its organisational management system.

19a The independent assessor shall not substitute its judgment for that of the manufacturer regarding the experience and expertise of the organization or its personnel.

3. Data input management

(a) Input data and parameters associated with the toolchain

(i) The manufacturer shall document the data used to develop, verify and validate the toolchain and its components and note important quality characteristics;

(ii) The manufacturer shall provide documentation showing that the data used to develop, verify and validate the toolchain and its components covers the intended functionalities that the toolchain aims at virtualizing;

Commented [RH23]: Can we suggest competency as alternative to E&E. I think that's the generally accepted ISO term is competence.

The Role of Competence in ISO 9001:2015
Competence – The ability of your personnel to apply knowledge and skills to achieve intended results is not just a requirement but a fundamental driver for the success of your QMS. It's about having the right people with the right skills doing the right job at the right time.

Commented [RH24]: This is a bit of commentary that possibly could be incorporated in an earlier paragraph or dropped completely.

Commented [RH25]: None of these are currently requirements, only 17 onward (&16) are requirements. Are there any hidden requirements if not should it be placed elsewhere or condensed.

Commented [RH26]: Not a big fan of this distinction. I think it can come under a single heading and ignore this distinction. I have drafter a version removing the distinction and adding an additional bullet. In an M&SMS the first two bullets would be the part of the process and the last would be the evidence that process was followed.

Commented [RD27]: Not sure this can fit the body of the reg

Commented [RH28]: If kept, this could be clearer about what the assessor should not do!! I don't think it can / should be kept. The assessor should be able to question anything and presumably if they are not happy then would not sign off on the assessment.

Commented [RD29R28]: Happy to remove this

(iii) The manufacturer shall document the data and the calibration procedures employed to fit any parameters associated with the toolchain and its components.

(b) The manufacturer shall quantify the uncertainty in the toolchain resulting from the data quality (e.g. data coverage, signal to noise ratio, and sensors' uncertainty/bias/sampling rate). This will be an input to the final uncertainty analysis of the toolchain.

Commented [RH30]: Left this as "toolchain" only as it should be clear that to get the uncertainty in the toolchain requires an analysis of its components.

4. Toolchain(s) Analysis and Description

5. Criticality assessment

30 The proposed approach to assess criticality is derived from ISO 26262, which requires different levels of qualification for the tools used in the development process. In order to derive how critical the toolchain and its components are the criticality assessment shall consider the following parameters:

(a) The consequences on human safety e.g. severity classes in ISO 26262;

(b) The degree to which the toolchain(s) and its components influence the ADS. A toolchain may be identified as critical but more detailed analysis shows that it is a tool or model within that toolchain that is the key contributor to the criticality.

31. The table below provides an example criticality assessment matrix to demonstrate this analysis. Manufacturers may adjust this matrix to their particular use case.

Commented [RH31]: This may have to go in the interpretation document as its "specifying an approach unless we are mandating it. It could be left in but shown as an example. [30,31]

Commented [RD32R31]: RH/UK: Not sure I understand whether we are planning to retain this. If so we will have to "allow" the manufacturer to claim they have done some analysis and developed the toolchain (or a component) according to a "chosen / determined" level of criticality.

Table 4.
Criticality assessment matrix

Influence on ADS	Significant	N/A			
	Moderate				
	Minor				
	Negligible			N/A	
		Negligible	Minor	Moderate	Significant
		Decision consequence			

32. From the perspective of the criticality assessment, the three possible cases for assessment are:

(i) Toolchains and its components that are clear candidates for following a full credibility assessment.

(ii) Toolchains and its components that may or may not be candidates for following the full credibility assessment at the discretion of the assessor.

(iii) Toolchains and its components that are not required to follow the credibility assessment.

Commented [RH33]: I struggle with this. The manufacturer makes judgments as to what needs to be done and then presents evidence. I don't see the "assessor" being involved in the decision as such. The "assessor" may not think the manufacturer has done a suitable criticality assessment, or maybe during an early stage audit agree what needs to be done. This will also be difficult under GTR!

C. Verification

33. The toolchain(s) verification deals with the analysis of the correct implementation of the conceptual/mathematical models that create and build up the overall toolchain(s). Verification contributes to the toolchain's credibility via providing assurance that the toolchain and all of its components will not exhibit unrealistic

behaviour for a set of inputs which have not been tested. The procedure is based in a multi-step approach described below, which includes code verification, calculation verification and sensitivity analysis.

Commented [RD34]: Moved to 5.X.5.1.

1. Code verification

34. Code verification concerns the execution of tests to demonstrate that no numerical/logical flaws affect the toolchain or its components.

Commented [RH35]: Trying to stay with a single terminology.

2. Calculation verification

38. Calculation verification deals with the estimation of numerical errors affecting the toolchain. ~~The manufacturer shall document numerical error estimates (e.g. discretization error, rounding error, iterative procedures convergence). The manufacturer shall review their analysis and demonstrate that the numerical errors are understood and sufficiently bounded to allow the toolchain to be used for virtual testing.~~

Commented [RD36]: Moved to 5.X.5.3.Y.

38. [RH/UK Alternative] The assessor will review the manufacturer's simulation toolchain documents and evidence to determine whether the manufacturer has made a suitable analysis and has correctly identified the source and estimates of the numerical errors that affect the simulation toolchain. ~~The manufacturer shall document numerical error estimates (e.g. discretization error, rounding error, iterative procedures convergence). The manufacturer shall review their analysis and demonstrate that the numerical errors are understood and sufficiently bounded to allow the toolchain to be used for virtual testing.~~

Commented [RD37]: Moved to 5.X.5.3.Y.

Commented [RH38]: This is an initial "suggestion" for the text that we may have to produce within the assessor section. The problem is that it doesn't "help" the assessor because it doesn't say what to assess against!! The other sections will need something similar.

3. Sensitivity analysis

39. Sensitivity analysis aims to quantify how input data and parameters affect output values and identify which have the greatest impact. The analysis also provides information that is useful in assessing whether the toolchain and its components can continue to satisfy the acceptance tests and criteria when subjected to small variations of the inputs and parameters.

4. Validation

43. The quantitative process of determining the degree to which the toolchain and its components are an accurate representation of the system being emulated. The following elements shall be considered when validating the toolchain:

43. [RH/UK Alternative] ~~The quantitative process of determining the degree to which the toolchain and its components are an accurate representation of the system being emulated.~~ The assessor will consider the following elements when assessing the toolchain:

(a) Measures of Performance (metrics)

Commented [DR(I39]: Moved to 5.X.6.2.

(b) Goodness of Fit measures

Commented [RH40]: This is a poor title. 43 (a) and (b) can probably be combined into a title such as "metrics". Its more about discrete data that can be compared 1to1 and continuos data that has to be compared using some form of statistical analysis.

(c) Validation methodology

(d) Accuracy requirement

The correlation criteria are defined during the toolchain analysis.

(e) Validation scope (the part of the toolchain to be validated)

A toolchain consists of multiple tools, and each tool may use several models. The validation scope includes the toolchain, all tools and models.

Commented [RH41]: Maybe higher up in the section as it is description of what needs to be validated rather than the validation activity.

(f) Internal validation results

(g) Confirmatory Validation

(h) Uncertainty characterisation

This section is concerned with characterizing the variability of the toolchain results. The assessment shall be made up of two phases. In a first phase the information collected from the "toolchain analysis and description" section and the "input data management" section are used to characterise the uncertainty in the input data, in the model parameters and in the various components of the toolchain. Then, by propagating these uncertainties through the toolchain, the overall uncertainty in the toolchain's output can be quantified.

(i) Characterization of the uncertainty in the input data

(ii) Characterization of the uncertainty in the model parameters (following calibration).

(iii) Characterization of the uncertainty in the toolchain structure.

(iv) Characterization of aleatory vs. epistemic uncertainty

Annex 5 - Appendix 2

Documentation structure

1. This section sets out how the above information will be collected and organized in the documentation provided by manufacturer to the relevant authority.
2. The manufacturer shall produce a document ~~(a "simulation handbook")~~ structured using this outline to provide evidence for the topics presented.
3. The documentation shall be part of the toolchain release along with all the appropriate supporting data.
4. The manufacturer shall provide suitable references that allows the relevant parts of the toolchain and the supporting data to be identified.
5. The documentation shall be maintained throughout the whole lifecycle of the toolchain.

Commented [RH42]: Does this imply that TAA or TS receive a copy of the toolchain? I don't really see this as a practical possibility. The toolchain will be embedded into various manufacturer tools, may require licences, might need access to hardware,. ..