# Specific Intended Uses: Establishing verification, validation and accreditation objectives

Arthur M. Reed, *David Taylor Model Basin, Carderock Division, Naval Surface Warfare Center*, arthur.reed@navy.mil

## ABSTRACT

IMO's implementation of the Second Generation Intact Stability Criteria has put in place a multitiered process by which the adequacy of a vessel's stability can be assessed. The most stringent criteria is Direct Assessment where by a vessel is assessed using a physics-based simulation tool. To be applied to stability assessment, these tools should undergo a formal Verification, Validation and Accreditation (VV&A) to assure that they perform adequately. Before the VV&A can be performed, the problem for which the simulation tool is to be assessed must be defined. This use—the *objectives* of the simulation are defined by the establishment of Specific Intended Uses (SIUs). SIUs will be characterized, and the way in which they are used will be defined.

*Keywords*: Verification, Validation and Accreditation; VV&A; Formal VV&A, Specific Intended Uses, SIU

## 1 INTRODUCTION

Beginning in the early 2000's efforts were initiated to develop performance based stability criteria for commercial vessels with the re-establishment of the intact-stability working group by IMO's Subcommittee on Stability and Load Lines and on Fishing Vessels Safety (SLF) (cf. Francescutto, 2004, 2007). Over time, the terminology to describe the new intact stability criteria evolved from "performance based" to "next generation" to "2nd generation"—the terminology in use today. This entire evolution is described in the introduction to Peters, *et al.* (2011).

The SLF Working Group decided that the second-generation intact stability criteria should be performance-based and address three modes of stability failure (SLF 48/21, paragraph 4.18):

- *Restoring arm variation* problems, such as parametric roll and pure loss of stability;
- *Stability under dead ship condition*, as defined by SOLAS regulation II-1/3-8; and
- *Maneuvering related problems in waves*, such as surf-riding and broaching-to.

Ultimately, a fourth mode of stability failure was added:

- *Excessive accelerations*.

The criteria and processes were first introduced in Belenky, *et al.* (2008). The state-of-the-art in the assessment of vulnerability is presented in detail in Peters, *et al.* (2011) and further summarized in Reed & Zuzick (2015)

The deliberations of the Working Group led to the formulation of the framework for the second generation intact stability criteria, which is described in SLF 50/4/4 and was discussed at the 50th session of SLF in May 2007. The key elements of this framework were the distinction between parametric criteria (the 2008 IS Code) and performance-based criteria, and between probabilistic and deterministic criteria. Special attention was paid to probabilistic criteria; the existence of the *problem of rarity* was recognized for the first time and a definition was offered. Also, due to the rarity of stability failures, the evaluation of the probability of failure with numerical tools was recognized as a significant challenge.

The “Second-generation intact-stability criteria” are based on a two-tiered assessment approach: for a given ship design, each stability-failure mode is evaluated using two levels of vulnerability assessment in the first tier. A vessel that fails to comply with the first- and second-level criteria of the first tier must progress to the second tier where it is examined by means of a direct assessment procedure based on tools and methodologies corresponding to the best state-of-the-art physics-based prediction methods in the field of ship-stability failure prediction.

If decisions regarding the adequacy of a vessel stability-wise, are going to be made based on the predictions of a Modeling and Simulation (M&S) tool, there must be a reasonable assurance that the tool provides acceptably accurate results. The process by which a tool may be determined to be sufficiently accurate is known as Verification, Validation and Accreditation (VV&A).

Reed & Zuzick (2015) quoted “Verification, Validation, and Accreditation are three interrelated but distinct processes that gather and evaluate evidence to determine, based on the M&S’s intended use, the M&S’s capabilities, limitations, and performance relative to the real-world objects it simulates.” Definitions for these three terms are provided below, each followed by a practical commentary relevant to computational tools for predicting dynamic stability.

1. Verification—the process of determining that a M&S’s implementation accurately represents the developer’s conceptual description and specification, i.e., does the code accurately implement the theory that is proposed to model the problem at hand?

2. Validation—the process of determining the degree to which an M&S is an accurate representation of the real world from the perspective of the intended uses of the M&S, i.e., does the theory and the code that implements the theory accurately model the relevant physical problem of interest?

3. Accreditation—the official determination that an M&S, . . . is acceptable for use for a specific purpose, i.e., is the theory and the code that implements it adequate for modeling the physics relevant to a specific platform? In other words, are the theory and code relevant to the type of vessel and failure mode for which it is being accredited?

In the process leading to accreditation by a Flag Administration, VV&A must be a formal process with structure that is prescribed. This structure includes the identification of an Accreditation Authority (AA) and the establishment of accreditation panels; and is described in Reed & Zuzick (2015).

The process of accreditation requires Specific Intended Uses (SIUs)—the objectives against which accreditation occurs, the subject of this paper.

## 2 ROLE OF SIUS IN ACCREDITATION

As just described, accreditation is the process by which a computational tool is certified as being sufficiently accurate and thus acceptable for use in a particular case for a particular vessel or class of vessels. In the IMO context, this would be a vessel of a particular size and proportions, which will have a particular mode of operation. In practice this would also be tied to a particular mode of stability failure, and would be defined as a particular SIU.

SIUs are the statements that define the scope of the problem or simulation that is to be modeled, and for which the M&S will be accredited. In the context of direct assessment under second-generation intact stability, this will need to include a definition of the type of vessel for which the M&S tool is to be accredited—accreditation for small fishing vessels may well not apply to a container carrier; as well as the mode of stability failure that is anticipated to be an issue. There can, and in fact would likely be multiple SIUs for the same VV&A activity.

### *2.1 Example of an SIU*

As stated earlier, the SIU effectively defines the objective of the accreditation. As such, the SIU needs to answer the questions “what” and “why.” The “what” part of the answer will in the case of accreditation have two parts, one part pertaining to the type of vessel, and the other pertaining to the mode of stability failure. An example of this would be the accreditation of a code for predicting parametric roll of a container carrier—container carrier would be the type of vessel and parametric roll would be the mode of stability failure.

The "why" question relates to the way in which the predictions from the code will be used. Will the code be used to determine whether a vessel is susceptible to parametric roll in head seas at 24 kt in a particular sea state, or will it be used to derive a speed polar plots for susceptibility to parametric roll in a series of sea states. The answer to the "why" question serves to define the scope of the effort required in the accreditation process.

To clarify, an example of an SIU is: "The XYZ simulation tool will be used to generate operator guidance polar plots for all applicable speeds and headings against pure loss of stability for RO/PAX vessels in the 11,000–13,000 t displacement range, lengths of 130–150 m, and with beam-to-draft ratios of 4.5 to 5.5. These polar plots will enable the vessel operators to avoid situations where pure loss of stability could be an intact stability issue. The information used to generate the operator guidance polar plots will be developed using numerical data generated by the XYZ simulation tool."

In the example SIU, the answers to the "what" question are RO/PAX vessels in a particular size range with the stability failure mode being pure loss of stability. The answer to the "why" question is to generate operator guidance polar plots for all applicable speeds and headings.

### *2.2 Requirements Flow-Down Table*

The answers to the "what" and "why" questions within the SIU are used to determine what needs to be characterized and analyzed from the perspective of the V&V process. This is accomplished by the development of a Requirements Flow-Down Table. In the Requirements Flow-Down Table, each SIU is decomposed in to several high level requirements (HLRs), which characterize important aspects of the SIU. The HLRs are each further mapped into several detailed-functional requirements (DFRs). A comparison metric and an acceptance criterion are identified for each DFR. Additional clarification is provided by the definition of the comparison metrics and their associated acceptance criteria. HLRs reflect the technical specifications provided by SME-opinion. DFRs provide additional specifications as necessary to more fully-describe each HLR. Requirements Flow-Down Tables are useful tools in high-level assessment of the appropriateness of the proposed accreditation criteria as well as required components of the Accreditation Plan (DoD, 2012).

An example of a Requirements Flow-Down Table, Table 1, is provided for the example SIU given above.

## 3 SUMMARY

With the advent of the second-generation intact stability criteria, IMO has initiated a two-tier performance-based stability assessment process for unconventional hulls with a risk of intact stability failure. If the design fails the first and second level tests of the first tier, it then progresses to the second tier and direct assessment, which requires an accredited physics-based simulation tool.

Accreditation requires that a set of Specific Intended Uses (SIUs) defining the objectives of the accreditation be defined. These SIUs must define what the M&S is to be accredited for (type of vessel and mode of stability failure) and why (the product to be produced by the M&S).

Additionally, the Requirements Flow-Down Table which is used to define comparison metrics and acceptance criteria based on the SIUs are described, and an example is provided.

## REFERENCES

Belenky, V., J. O. de Kat & N. Umeda (2008) Towards Performance-Based Criteria for Intact Stability. *Marine Tech.*, 45(2):101–123.

DoD (2012) Department of Defense, USD (AT&L) MIL-STD-3022 w/Change 1: Department of Defense Standard Practice: Documentation of Verification, Validation, and Accreditation (VV&A) for Models and Simulations, 55 p.

Francescutto, A. (2004) Intact Ship Stability—The Way Ahead. *Marine Tech.*, 41:31–37.

Francescutto, A. (2007) Intact Stability of Ships-Recent Developments and Trends. *Proc. 10th Int'l Symp. Practical Design of Ships and Other Floating Struct* (PRADS '07). Houston, TX, Vol. 1, pp. 487–496.

IMO SLF 48/21 (2005) Report to Maritime Safety Committee, London, UK, 65 p.

IMO SLF 50/4/4 (2007) Framework for the Development of New Generation Criteria for Intact Stability, submitted by Japan, the Netherlands and the United States, London, UK, 6 p.

Peters, W., V. Belenky, C. Bassler, K. Spyrou, N. Umeda, G. Bulian, B. Altmayer (2011) The Second Generation Intact Stability Criteria: An Overview of Development. *Trans. SNAME*, 119:225–264.

Reed, A. M. & A. V. Zuzick (2015) Direct Assessment Will Require Accreditation—What This Means. *Proc. 12th Int'l Conf. Stability of Ships & Ocean Vehicles* (STAB '15), Glasgow, UK, pp. 51–78.

**Table 1** Example Requirements Flow-Down Table.

| High Level Requirements | Detailed Functional Requirement | Comparison Metric | Acceptance Criteria |
|---|---|---|---|
| HLR 1.a<br>Simulation must demonstrate good correlation to model data for ship responses to elemental tests to suggest that underlying physics are sound. | DFR 1.a.1<br>Simulation must demonstrate the ability to successfully predict critical motion values in a large number of Quantitative Accreditation conditions for which model test data is available for comparison. | CM 1.a.1<br>Check-list of quantifiable metrics defining “reasonable” correlation for elemental tests used to inform SME opinion | AC 1.a<br>ARP will vote using SME opinion informed by elemental test comparisons whether to assess subsequent acceptance criteria. |
| | DFR 1.a.2<br>Collective SME judgment shall ultimately decide whether or not this requirement is met (regardless of the code’s ability to meet the suggested quantifiable metrics). | CM 1.a.2<br>SME opinion/judgment | |
| HLR 1.b<br>The simulation and model-scale data must show consistently good correlation ranging from the more simple conditions to the more complex conditions. Good correlation must be demonstrated for the range of operational, environmental, and loading conditions defined in the Quantitative Accreditation scope for which comparison model data are available. | DFR 1.b.1<br>Parameters which characterize the ship’s operating condition relative to the seaway, and identify the corresponding critical motion, must be assessed. | CM 1.b.1<br>Mean values, $\mu$, of achieved speed and heading | AC 1.b.1<br>Differences between mean achieved speed and mean achieved heading for each validation condition must be less than specified amounts. |
| | DFR 1.b.2<br>All comparisons must take into account all known sources of uncertainty (sampling, instrument, condition, etc.). | CM 1.b.2<br>90% uncertainty intervals on the each parameter (model and simulation) | AC 1.b.2<br>The 90% confidence intervals on each parameter value ($\sigma$ and A90%) for a given motion and condition must overlap in order to suggest that the underlying populations (model and simulation) may be the same. |
| | DFR 1.b.3<br>Parameters that are used to define Quantitative Accreditation polar plots risk values and lifetime risk calculation must be assessed. If direct validation of these quantities is not achievable, a sufficient substitute quantity shall instead be assessed. (rare motion metrics) | CM 1.b.3<br>The 90th percentile of peak amplitudes, A90%, of motions (in lieu of exceedance rates of physical limit thresholds which are not expected to be available for validation) | |
| | DFR 1.b.4<br>Parameters that are used to evaluate the Quantitative Accreditation system health must be assessed. (non-rare motion metrics) | CM 1.b.4<br>Mean standard deviation, $\sigma$, of motions | |

**Table 1 (Cont'd)** Example Requirements Flow-Down Table.

| High Level Requirements | Detailed Functional Requirement | Comparison Metric | Acceptance Criteria |
|---|---|---|---|
| HLR 1.c Necessary accuracy of the simulation shall be influenced by an appropriate balance between technical excellence and judiciousness | DFR 1.c<br>Thoughtful engineering judgment shall be applied in the determination of permissible differences between simulation and model test results. | CM 1.c<br>Margin applied to observed sample parameter values (defined in CM 1.b.2 and CM 1.b.3) | AC 1.c<br>The observed values of compared sampled parameters may be deemed acceptable if the difference between the values is less than a specified amount. (margin) |
| HLR 1.d<br>The safety of the ship and sailor must be prioritized and reflected in the criteria established for validation. | DFR 1.d.1<br>Reasonable conservatism on the part of the simulation solution should be endorsed to promote the overall safety of the sailor. | CM 1.d.1<br>Margin applied to observed sample parameter values (defined in CM 1.b.2 and CM 1.b.3) | AC 1.d.1<br>The margin allowed by AC 1.c shall be increased by 50% in the case of over-prediction on the part of the simulation to allow for additional conservatism on the part of the simulation. (additional conservative margin) |
| | DFR 1.d.2<br>Determination of simulation tool success must only be reached using reasonably high-fidelity validation data sets. | CM 1.d.2<br>Combined uncertainty in the comparison, calculated as a function of the 90% uncertainty intervals (CM 1.b.2) on both data sets, model and simulation | AC 1.d.2<br>Successful validation comparisons for both rare and non-rare motions ($\sigma$ and A90%) may only be accepted if the combined uncertainty in both data sets is sufficiently small. |
| HLR 1.e<br>Simulation must be deemed usable for conditions within the current scope of the Quantitative Accreditation for which comparison model test data is not available. | DFR 1.e.1<br>Simulation must demonstrate the ability to successfully produce critical motion values in a large number of Quantitative Accreditation conditions for which model test data is available for comparison. | CM 1.e.1<br>Number of conditions which successfully pass the following criteria: AC.1.b.1 through AC 1.d. | AC 1.e<br>70% of Quantitative Accreditation conditions for which model data are available for comparison must pass criteria (AC 1.a through AC 1.d) for 100% of critical motion parameter values. (rare and non-rare motion assessments calculated independently) |