Verification and Validation Aspects of Development and Implementation of the Second Generation Intact Stability Criteria

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

ABSTRACT

The IMO Second Generation Intact Stability Criteria, and verification and validation (V&V) are introduced. Then the application of V&V to the Level 1, Level 2 and direct assessment stages of the Second Generation Intact Stability Criteria are discussed. From the perspective of Level 1 and Level 2 verification and validation, the user's only responsibility is to verify that the algorithms for assessing vulnerability to stability failure contained in IMO documentation are implemented correctly. For direct assessment using ship dynamics software for predicting motions in extreme seas, existing well established and documented V&V processes apply. The developers of the algorithms for the Level 1 and Level 2 vulnerability assessments need to validate that their algorithms are consistent across a large range of vessel types and sizes.

The one significant note is that even though, in general, the Level 1 vulnerability assessment can be performed "on the back of an envelope" using a hand calculator, those calculations need to be performed using a spreadsheet program on a personal computer or reliable and consistent verification will be virtually impossible.

Keywords *Verification & Validation, Second Generation Intact Stability Criteria.*

1. INTRODUCTION

For commercial vessels, the classical intact stability criteria is based on the work of Rahola (1939) and is incorporated in the International Code on Intact Stability, the 2008 IS Code (MSC 85/26/Add.1[1]). Similar criteria for naval vessels is provide by Sarchin & Goldberg (1962) and codified in the NATO Naval Ship Code (NATO, 2007a,b) and by a US Navy Design Data Sheet (NAVSEA, 2016). These criteria are prescriptive—that is they are a set of criteria, defined based on empirical data, which are *assumed* to ensure that a vessel meeting the criteria will have adequate static stability. The history of development and the background of the IMO criteria are described by Kobylinski & Kastner (2003); a summary of the origin of these criteria is also available in chapter 3 of the Explanatory Notes to the International Code on Intact Stability (MSC.1/Circ.1281).

Beginning in the early 2000's efforts were initiated to develop performance based stability criteria for commercial vessels with the reestablishment 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 deliberations of the Working Group led to the formulation of the framework for the second-generation intact stability criteria, which is de-

[1] References to IMO documents such as "MSC 85/26/Add.1" appear in the list of references with an "IMO" prefix, *i.e.*, as: IMO MSC 85/26/Add.1. As there is no ambiguity in the names of the IMO citations, the year will be omitted from the citations.

scribed 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.

As the second-generation intact stability criteria are more extensive (deal with multiple stability failure modes) and more complex than the older prescriptive approach to stability, it will be necessary to insure that the algorithms supporting the assessment are consistent and implemented correctly. It is the objective of this paper to provide some insights on these latter two issues.

The paper will begin with a description of the second-generation intact stability criteria process and a definition of Verification, and Validation (V&V). The paper will then discuss V&V for the various levels of the process from both the user's and the algorithm developer's perspective.

2. IMO SECOND GENERATION INTACT STABILITY CRITERIA

The second-generation intact-stability criteria are based on a multi-tiered assessment approach: for a given ship design, each stability-failure mode is evaluated using multiple levels of vulnerability assessment, as necessary. The first two tiers or levels of vulnerability assessment criteria are characterized by different levels of accuracy and computational effort, with the first level being simpler and more conservative than the second.

A ship, which fails to comply with the Level 1 criteria is assessed using the Level 2 criteria. In a case of unacceptable results at the second level, the vessel must then be examined by means of a *direct assessment* procedure based on tools and methodologies corresponding to the best state-of-the-art prediction methods in the field of ship-capsizing prediction. This third-level methodology should capture the physics of capsizing as practically possible.

The three levels of assessment are intended to be of increasing complexity with the Level 1 assessment being a simple "back of the envelope" calculation that should be simple enough that it can be completed for all stability failure modes in a day. The Level 2 assessment is more complex, and might require as much as a week's effort to assess all stability failure modes, and require the use of computational algorithm implemented in a program such as Excel or MathCad—here after referred to as a *spreadsheet*. The third level direct assessment will require the use of serious computing resources and could take a month or more's effort.

The specific IMO rules and regulations are still under development, but the following publications document the current state of the envisioned process for Level 1 and Level 2 of each of the stability failure modes:

- <u>Pure loss of stability</u>: SDC 2WP.4, Annex 1; SDC 3WP.5, Annex 3; SDC 4/5/1/Add.5; SDC 4/5/6
- <u>Parametric Roll:</u> SDC 2WP.4, Annex 2; SDC 3WP.5, Annex 4; SDC 4/5/1/Add.1; SDC 4/5/1/Add.5, SDC 4/5/6
- <u>Dead ship condition:</u> SDC 3WP.5; SDC 3WP.5, Annex 6; SDC 4/5/1/Add.3; SDC 4/5/1/Add.5; SDC 4/5/6
- <u>Surf riding/broaching:</u> SDC 2WP.4, Annex 3; SDC 3WP.5, Annex 5; SDC 4/5/1/Add.2, SDC 4/5/1/Add.5, SDC 4/5/6
- <u>Excessive Acceleration:</u> SDC 3WP.5, Annex 2; SDC 3WP.5, Annex 7; SDC 4/5/1/Add.4; SDC 4/5/6

The procedure for performing direct assessment is described in: SDC 4/WP.4, Annex 1.

3. VERIFICATION AND VALIDATION

Software that is being used for engineering computations, upon which design decisions will be based needs to be correct. The processes by which software is assessed as to it correctness and being adequate for the job is called *verification* and *validation* (V&V)—verification assesses correctness and validation assesses the degree to which it is adequate for the task. Papers and reports by Beck, *et al.* (1996), AIAA (1998), DoD (1998, 2003, 2007, 2012), McCue, *et al.* (2008), ASME (2009), Reed (2009) and Reed & Zuzick (2015) provide different, although consistent, definitions of V&V. The U.S. DoD definitions for these 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 model or simulation 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 a model or simulation is an accurate representation of the real world from the perspective of the intended uses of the model or simulation, *i.e.*, does the theory and the code that implements the theory accurately model the relevant physical problem of interest?

4. V&V FROM THE USER'S PERSPECTIVE

For the Second Generation Intact Stability Criteria, the question of V&V has to cover a broad range of computations/computational tools—from the "back of an envelope" assessment to sophisticated ship dynamics computational tools. As each of the levels of assessment has its own issues, they will be discussed separately, beginning with Direct Assessment, where the computational tools that are traditionally put through the V&V process would be employed.

Direct Assessment

As just stated, the hydrodynamic computational tools for predicting ship dynamics are the types of software for which the V&V processes have been developed. So while these are the most complex software tools that must be put through the V&V process, and the tools for which the most effort will have to be expended, they are the tools for which the process is the most mature. As stated previously, there is an abundance of literature on the subject of formal V&V of software (*cf.*, AIAA 1998; DoD 1998, 2003, 2007, 2012; ASME 2009). Reed & Zuzick (2015) provide a survey of the formal V&V process tailored for the ship stability community.

From the users perspective, it is unlikely that a user will be developing a computational tool for assessing dynamic stability performance in extreme seas; the user will most likely be employing software developed by a third party. Thus, the user will not be responsible for verification of the software, he will have to assume that the software vender has performed that function, and the user will only be responsible for performing validation to assure that the software tool is adequate for predicting the stability failure mode(s) of concern. The Flag Administration, responsible for the vessel being assessed, should have defined the process for *formal validation*.

Level 2 Criteria

For Level 2, the Second Generation Intact Stability Criteria will explicitly provide the user with the algorithm for use in assessing the vulnerability of a ship to each particular stability failure mode. Thus, there should be no requirement for the user to perform *validation* of a spreadsheet that is used to perform the vulnerability calculations. However, it will be necessary to perform *verification* to insure that the calculations are performed correctly.

The issue then becomes one of how best to perform this verification. It would appear that the ideal situation would be to have a series of benchmark cases for each stability failure mode. For each failure mode there would be pairs of cases, one of the pairs being a case that passes the vulnerability test for that mode and one that fails the vulnerability test. For Level 2 algorithms where there are binary decision points within the algorithm, there should be a pair of benchmark cases that will test each branch of the decision tree.

Under these conditions, the user would be required to enter each pair of benchmark data into his spreadsheet and show that the results of each case agree with the expected answer within a specified accuracy, say 2-percent. When a user has performed and passed this level of validation for all five stability failure modes, he could be "certified" by a Flag Administration to use his spreadsheet to assess the vulnerability of his design to stability failure.

Level 1 Criteria

In principle, the Level 1 V&V should be similar in complexity to the Level 2 problem, and have the same approach. However, there is one complication at Level 1. Level 1 vulnerability assessment has been characterized as an assessment that can be carried out on the "back of an envelope" using a hand calculator, but this opens the Level 1 assessment up to a lack of repeatability due to simple calculation errors.

Therefore, it is proposed that, even at Level 1, it be required that the vulnerability assessment for each mode of stability failure be implemented in a spreadsheet. This will vastly reduce the possibility of inadvertent errors due to "hitting the wrong key" on a calculator, and will greatly facilitate verification using he same benchmarking process proposed for Level 2.

5. V&V FROM THE CRITERIA DEVELOPER'S PERSPECTIVE

The developers of the Level 1 and Level 2 intact stability vulnerability criteria are not developing software, so they do not have any responsibility for V&V in the traditional sense. However, they do have responsibility for ensuring that the algorithms that they are developing are consistent—this is a validation function.

What is meant by consistency of algorithms? If the Level 1 and Level 2 algorithms are developed from the same theoretical basis, then the validation can be performed largely at the theory/algorithm basis, but if *not*, then extensive computational testing is required. A *hypothetical* example of a theoretically consistent Level 1 and Level 2 vulnerability assessment would be where the Mathieu equation is used to evaluate the sensitivity to parametric roll, with the Level 1 algorithm using the Mathieu equation without the roll damping term and the Level 2 algorithm using the Mathieu equation with a roll damping term.

In the absence of such a consistent theoretical basis, the validation of the Level 1 and Level 2 algorithms consists of two steps. First, the algorithms must be rational, that is they should not be based on the use of logically inconsistent information and second they must undergo an extensive computational consistency check. To give a *ludicrous example* of a rationality check, a stability failure algorithm based, among other things, on the distance from the earth to the moon would be highly suspect. Someone other than the developer of the algorithm should conduct the *rationality* step of the validation.

The second step, the computational validation, will involve evaluating a large number of vessels of various types and sizes using both the Level 1 and Level 2 algorithms for each mode of stability failure. The metric here is two-fold, first that a vessel in a given loading condition that passes the Level 1 vulnerability test should not fail the Level 2 vulnerability check. And secondly, for those vessels that pass both the Level 1 and Level 2 vulnerability check, the margin at Level 2 should not be smaller than the margin at Level 1—if a vessel passes the Level 1 check by a large margin, it should not pass the Level 2 check by only a small margin, this is admittedly somewhat subjective.

6. CONCLUSIONS

From the perspective of Level 1 and Level 2 verification and validation, the user's only responsibility is to verify that the algorithms for assessing vulnerability to stability failure contained in IMO documentation are implemented correctly. To facilitate this, there needs to be a comprehensive set of benchmark cases that both meet and fail to meet the vulnerability criteria, covering each of the stability failure modes. For direct assessment using ship dynamics software for predicting motions in extreme seas, the well established and documented V&V process of AIAA 1998; DoD 1998, 2003, 2007, 2012; and ASME 2009, *etc*. apply. The developer of the algorithms for the Level 1 and Level 2 vulnerability assessments need to validate that their algorithms are consistent across a large range of vessel types and sizes.

The one significant note is that even though, in general, the Level 1 vulnerability assessment can be performed "on the back of an envelope" using a hand calculator, those calculations need to be performed using a spreadsheet program on a personal computer or reliable and consistent verification will be virtually impossible.

REFERENCES

AIAA (1998) Guide for the verification and validation of computational fluid dynamics simulations. AIAA G-077-1998 Guide, Reston, VA: American Institute of Aeronautics and Astronautics, viii+19 p.

ASME (2009) V&V 20-2009, Standard for verification and validation in computational fluid dynamics and heat transfer. New York, NY: American Society of Mechanical Engineers, x+87 p.

Beck, R. F., A. M. Reed & E. P. Rood (1996) Application of modern numerical methods in marine hydrodynamics. *Trans. SNAME*, 104:519--37, Jersey City, NJ.

DoD (1998) DoD Modeling and simulation (M&S) glossary. DoD 5000.59-M, U. S. Department of Defense, 175 p.

DoD (2003) DoD Modeling and simulation (M&S) verification, validation, and accreditation (VV&A). DoD Instruction 5000.61, U. S. Department of Defense, 10 p.

DoD (2007) DoD Modeling and Simulation (M&S) Management. DoD Directive 5000.59, U. S. Department of Defense, 7 p.

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 MSC 85/26/Add.1 (2008) International code on intact stability, 2008 (2008 IS Code). London, 96 p

IMO MSC.1/Circ.1281 (2008) Explanatory notes to the international code on intact stability, 2008. London, 30 p.

IMO SDC 2-WP.4 (2015) Development of Second Generation Intact Stability Criteria, Development of amendments to Part B of the 2008 IS Code on towing, lifting and anchor handling operations, Report of the working group (Part 1). London, UK, 48 p.

IMO SDC 3/WP.5 (2016) Finalization of Second Generation Intact Stability Criteria, Amendments To Part B of the 2008 IS Code on Towing, Lifting and Anchor Handling Operations Report of the working group (Part 1). London, UK, 141 p.

IMO SDC 4/5/1/Add.1 (2016) Finalization of Second Generation Intact Stability Criteria, Report of the correspondence group (part 2). Submitted by Japan, London, UK, 39 p.

IMO SDC 4/5/1/Add.2 (2016) Finalization of Second Generation Intact Stability Criteria, Report of the correspondence group (part 3). Submitted by Japan, London, UK, 18 p.

IMO SDC 4/5/1/Add.3 (2016) Finalization of Second Generation Intact Stability Criteria, Report of the correspondence group (part 4). Submitted by Japan, London, UK, 24 p.

IMO SDC 4/5/1/Add.4 (2016) Finalization of Second Generation Intact Stability Criteria, Report of the correspondence group (part 5). Submitted by Japan, London, UK, 17 p.

IMO SDC 4/5/1/Add.5 (2016) Finalization of Second Generation Intact Stability Criteria, Report of the correspondence group (part 6). Submitted by Japan, London, UK, 51 p.

IMO SDC 4/5/6 (2016) Finalization of Second Generation Intact Stability Criteria, Draft consolidated explanatory notes for the second generation intact stability criteria. Submitted by the United States, London, UK, 51 p.

IMO SDC 4/WP.4 (2017) Report of the working group (Part 1). London, UK, 21 p.

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.

Kobylinski, L. K. & S. Kastner (2003) *Stability and safety of ships: regulation and operation*. Elsevier, Amsterdam, 454 p.

McCue, L. S., W. R. Story & A. M. Reed (2008) Nonlinear dynamics applied to the validation of computational methods. *Proc. 27th Symp. on Naval Hydro,* Seoul, South Korea., 10 p.

NATO (2007a) Buoyancy, stability and controllability. Chapter III of Naval Ship Code, NATO Naval Armaments Group, Maritime Capability Group 6, Specialist Team on Naval Ship Safety and Classification, Allied Naval Engineering Publication ANEP-77, vii+121 p.

NATO (2007b) Guidance on NSC Chapter III Buoyancy and Stability, Part B: Application. Chapter 3, Guide to the Naval Ship Code, NATO Naval Armaments Group, Maritime Capability Group 6, Specialist Team on Naval Ship Safety and Classification, 91 p.

NAVSEA (2016) Design practices and criteria for U.S. Navy surface ship stability and reserve buoyancy. NAVSEA Technical Publication T9070-AF-DPC-010/079-1. Naval Sea Systems Command. 175 pp.

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.

Rahola, J. (1939) The judging of the stability of ships and the determination of the minimum amount of stability especially considering the vessel navigating Finnish waters. PhD Thesis, Technical University of Finland, Helsinki, viii+232 p.

Reed, A. M. (2009) A Naval Perspective on Ship Stability. *Proc. 10th Int'l Conf. Stability of Ships & Ocean Vehicles* (STAB '09), St. Petersburg, Russia, pp. 21–44.

Reed, A. M. & A. V. Zuzick (2015). Direct assessment will require accreditation—What this means. *Proc. 12th International Conf. Stability of Ships & Ocean Vehicles* (STAB'15). Ed. by S. Secretariat. Glasgow, UK, pp. 49–78.

Sarchin, T. H. & L. L. Goldberg (1962) Stability and buoyancy criteria for U. S. Naval surface ships. *Trans. SNAME*, 72:418–58.